

Request for Proposals: Data Platform Modernization of *Treasures from the Film Archives* (FIAF)

1 General Introduction

1.1 Overview

The database *Treasures from the Film Archives* contains unique information about silent-era film holdings in film archives around the world. The database was conceived in 1994 by the Fédération International des Archives du Film (International Federation of Film Archives, FIAF) as a tool to aid the work of preservation, research and film exchange between archives. The database currently provides filmographic (title, director, year, etc.) and item-level (condition, format, generation, etc.) information on over 60,000 works, describing silent shorts and features, including fiction and non-fiction films, from many of the world's major film archives.

This request for proposals contains mandatory (**must**) and optional (**should**) functionality and tasks. All functionality in proposed solutions must be described, designed, implemented and carried out in a way that the core tasks of the database can be completed, even if optional functionality is not implemented.

1.2 Goals

The ambition of the FIAF Cataloguing and Documentation Commission is to modernize the technical infrastructure that delivers *Treasures*, to achieve these key aims:

- Streamline the data lifecycle and remove manual processes wherever possible:
 - data preparation by partner archives
 - data delivery by partner archives
 - data ingest, disambiguation, and cleanup by FIAF editorial staff
 - data exchange from FIAF to commercial partners ProQuest and Ovid, and to FIAF's restricted website
- Reduce FIAF staff manual labor updating the dataset with new contributions, ideally creating an interface for FIAF affiliates to contribute data directly. This should enable contributors to import or enter and update their own data while following a FIAF template or data mapping. The proposed *Treasures* infrastructure should provide for vocabulary

(authority) term disambiguation and data management with minimum possible human intervention (although a certain degree of human analysis may always be required)

- Reduce the staff labor cost and overhead of hosting the infrastructure, ideally by delivering a technology stack that can be hosted either within existing FIAF servers, or by an alternative low-cost cloud provider (Heroku, AWS, etc.)

Modernization must include the setup and implementation of a new database with a new data model, using an open-source database management system (currently: FileMaker).

A report commissioned by FIAF in 2020 is appended to this request. It provides in-depth details about the existing *Treasures* infrastructure, data model, workflows, and user profiles.

Also appended to this request is a spreadsheet outlining the data exchange format that data contributors will be asked to follow when submitting data in the future.

2 Current Infrastructure

Following is a general description of the current *Treasures* infrastructure. Please refer to the appended report for more detailed information.

2.1 Ingest, Backend

Treasures currently uses a proprietary FileMaker system, and the updating process is semi-manual. Paid contract staff to support this system include a FileMaker consultant to develop database functionality and run import scripts, as well as a data editor to manually confirm authority term disambiguation and decide on each new data point to be merged with existing records or added as a new record (for more information see report, chapter on “Data Structure Information”, which will be provided on expression of interest).

2.2. Front end

Via the [FIAF website](#) (only accessible to logged-in users working for FIAF affiliate archives) users can enter a search term in a search box, which returns results in list form. With another click, a detailed view of the metadata of a film title can be viewed. Note that the FIAF website is only one search interface. Others are offered by contract via Ovid and ProQuest.

Using an advanced search, you can search specifically by film title, country, contributing archive, and director. It is also possible to narrow the search by year.

Simple search and advanced search can also be combined.

The result contains the following entries on the details page:



METROPOLIS (G, Fritz Lang, 1927/1925/1926)

Film details

Full Film Title	METROPOLIS (G, Fritz Lang, 1927/1925/1926)
Film Title	METROPOLIS SCHICKSAL EINER MENSCHHEIT IM JAHRE 2000
Film Country	G
Film Director	Lang, Fritz
Film Year	1927
Serie	
Prod. Company	Universum-Film AG (Ufa)
Producer	
Cast	Helm, Brigitte Fröhlich, Gustav Abel, Alfred Klein-Rohden (Klein-Rogge), Rudolf Loos, Theodor George, Heinrich Rasp, Fritz
Writer	Von Harbou, Thea Lang, Fritz
Photography	Freund, Karl Rittau, Günther
Credits	Moroder, Giorgio
Archive	Bulgarska Nacionalna Filmoteka (Sofia) [BGSB] Bundesarchiv-Filmarchiv (Berlin) [DEBB] Cinémathèque Royale de Belgique (Brussels) [BEBR] Cinematca Brasileira (Sao Paulo) [BRSC] Cinematca do Museu de Arte Moderna (Rio de Janeiro) [BRRC] Svenska Filminstitutet (Stockholm) [SESS] National Film and Sound Archive (Canberra) [AUCF] Deutsches Filminstitut - DIF (Frankfurt) [DEFD] Filmmuseum München (Munich) [DEMF] George Eastman Museum (Rochester) [USRG] Gosfilmofond of Russia (Moscow) [RUMG] Museum of Modern Art - Department of Film (New York) [USNM] Library and Archives Canada (Ottawa) [CAOL] BFI National Archive (London) [GBLB] Deutsche Kinemathek / Museum für Film und Fernsehen (Berlin) [DEBD] Fondazione Centro Sperimentale di Cinematografia - Cineteca Nazionale (Rome) [ITRF] Fondazione Cineteca Italiana (Milan) [ITMF] Arhiva Nacionala de Filme - Cinematca Romana (Bucharest) [ROBA] Cinémathèque Québécoise (Montreal) [CAMC] Cineteca del Friuli (Gemona) [ITGC] Cinematca Nacional de Angola (Luanda) [AOLC] Nga Taonga Sound & Vision / The New Zealand Archive of Film, Television and Sound (Wellington) [NZWN] Norwegian Film Institute (Oslo) [NOON] UCLA Film & Television Archive (Los Angeles) [USLU] UC Berkeley Art Museum & Pacific Film Archive (BAMPFA) (Berkeley) [USBU] Filmoteca Española (Madrid) [ESMF] Harvard Film Archive (Cambridge) [USCH] Choson Minjuju'i Inmingonghwaguk Kugga Yonghwa Munhongo (Pyongyang) [KPPC] Danish Film Institute (Copenhagen) [DKCD] Museo Nazionale del Cinema - Fondazione Maria Adriana Prolo (Turin) [ITTM] Jugoslovenska Kinoteka (Belgrade) [RSBJ] Lobster Films (Paris) [FRPL] Academy Film Archive (Los Angeles) [USLA] Filmarchiv Austria (Vienna) [ATVF] Centre National du Cinéma et de l'Image Animée (Paris / Bois d'Arcy) [FRPC]

Access	16 mm: ITRF, ITGC, AOLC, USLU, USBU, CAOL, DEBD, USCH, ITTM, AUCF 35 mm: ITRF, CAMC, NOON, ESMF, DEBD, AUCF Format unspecified: NZWN, GBLB Laserdisc: USLU 1/2" video: USLU 16 mm acetate positive: DEFD 35 mm acetate positive: DEFD
Non access	35 mm nitrate positive: USLU Format unspecified: USRG, DKCD 9.5 mm positive: DEBD 16 mm dupe negative: DEBD 16 mm positive: DEBD
Note	USCH: 2 access copies Inclusion of a title in this database does not guarantee its availability nor completeness. Users should contact individual archives for more information.
NFPF	
N ^o (id)	24740

3 Functional Requirements

3.1 Registration process

FIAF maintains a registration system for current archive affiliates which provides access to restricted areas on the FIAF website, including access to the *Treasures* database. Ideally, the proposed system should be able to use the existing registry to authorize access for the three user profiles described below.

The system must address the requirements of three different user roles:

- a) End user (search, filter, see search results)
- b) Data contributor (manage, import, export own holdings)
- c) Editor (manage, import, export all holdings; administer user roles and value lists)

End user functions can also be used by data contributors and editors. Data contributor functions can also be used by editors.

3.2 User profiles

3.2.1 End user

End users are employees of FIAF-affiliated film archives who are authorized to view holdings in the *Treasures* system (see registration procedure above).

Users must have the ability to:

- 1) Search for film works and items, filter search results;
- 2) View and browse a results list of film works; and
- 3) Access a detail page with full information about a work and its items.

Typically, users will search for film title, year of release, country of production, director(s), production company, producer, cast. End users must be able to filter search results by these criteria: year of release, country of production, holding archive(s), sound/silent, film format, carrier type.

3.2.2 Data contributor

Data contributors are usually employees of film archives who are responsible for uploading, entering or editing data about their holdings in the *Treasures* system. They will be the most intensive users of the proposed system.

The following subsections describe required functionality for data contributors:

3.2.2.1 Import into *Treasures*

Contributing institutions must be able to import data representing multiple items linked to multiple film works in a single batch. For more information about the exchange format, see Appendix 2 (*Exchange format field list*)

Data import must be possible via a standardized file format, ideally CSV (alternatively as XML).

The import functionality must meet the following requirements:

- 1) Columns in data import files must map logically to fields in *Treasures* destination database;
- 2) The system must be able to validate (and/or reject) data upon import against data rules, formats, and value lists stipulated in the data exchange format; and
- 3) The system must provide failure / error reporting with sufficient detail for a data contributor to troubleshoot the reasons for a failed import (e.g. identifying row and column of failed entry)

The system should normalize imported data where required. For example:

- Date formats should be normalized to ISO-8601 upon import;
- Vocabulary term variants (such as genre; “Romantic comedy” → “Comedy”) should be normalized to preferred terms upon import (for example via a simple thesaurus)

The system must be able to link item records with filmographic work records. The system should have a convenient means of identifying matches to existing filmographic work records, or to

create new work records as needed. This matching process could happen either before, during, or after import.

Apart from importing into Treasures, it must be possible for a data contributor to manually create and change item and work records through a user interface. This applies to archives with small holdings or with limited technical capacity to generate exports. See also 3.2.2.3 *Manage Holdings*.

3.2.2.2 Export from *Treasures*

Each data contributor must have the possibility to export their own organisation's holdings, i.e.: All work/item records of items contributed by the respective archive.

See also 3.3.2 *Data Exports*.

3.2.2.3 Manage holdings

The system must enable data contributors to directly create and edit individual records via a web-based user interface. This must support contributions from archives that are unable to create and import bulk datasets.

The following permissions for data contributors must be implemented:

- can insert, update, delete only self-contributed items;
- can insert works;
- can delete only works for which no items exist in the database from other data contributors.

Data contributors must be able to search Treasures by all common data points such as title, date, agent name, country of production, content description. They should also be able to search for works held in specific archives or multiple archives. The system should allow sorting of results to enable contributors to navigate complex result sets.

3.2.3 Editor

Editors are usually employees of or contractors working for FIAF. They administer the *Treasures* database, provide service and help to data contributors and end users, export and submit Treasures data to other publishers and, in general, take on responsibility for the correctness and integrity for the *Treasures* data and system.

The following subsections describe required functionality for editors:

3.2.3.1 Manage and edit all holdings / records; Data management

Editors must be able to edit (insert, update and delete) all works and all items from all contributors from an easy-to-use user interface.

Editors must be able to administer value lists / controlled vocabularies in the Treasures system and database.

3.2.3.2 Administration of users and archives

Editors must be able to add and remove users, and to change user roles. Editors must be able to add, delete and edit records for contributing archives, i.e.: Names, abbreviations and further information about film archives contributing data to *Treasures*.

3.2.3.3 Export

Export functionality for editors is basically the same as for data contributors, with the exception that an editor must be able to export ALL data stored in the *Treasures* database.

The system must support complete data export, or export of subsets based on admin-stated constraints. For example (not an exhaustive list), an editor may export:

- all records *created* within a date range
- all records *modified* within a date range
- all work records and all item records
- work records associated with item records created or modified within a date range
- all item records from a specific archive
- work or item records created or edited by specific contributors

See also 3.3.2 *Data Exports*.

3.2.3.4 Reporting

Editors must have access to system-level reports. The system must log data updates to work and item records, including account details of users who modify data. These logs should be accessible only to editors.

The system must enable an audit of deleted and modified records.

3.3 Data requirements

3.3.1 Data model

The existing data model uses a flat structure—it summarizes unique items held in each archive per filmographic work, with one row per work per archive. This model must be replaced by a more normalized data model meeting the following minimum requirements:

- Separate tables for works, items, persons, corporate bodies, contributing archives and possibly other main entities
- Separate table or tables for value lists

In this model, a “work” record refers to the most generalized information about a film—director, country, year, etc. An “item” record refers to a specific instance of that work held by an individual film archive. Please refer to the appended report for more detail about this conceptual model and the FIAF film cataloging guidelines.

3.3.2 Data exports

Format for export will be the same as for import (*see Treasures exchange format*). User must be able to choose CSV/TSV or XML and, optionally, JSON.

Optionally, the system should expose a REST API which receives search queries of the common data points over HTTPS, and delivers XML or JSON to permitted clients over HTTPS, to enable integration with publishers’ data supply infrastructures or other integrations.

3.3.3 Data validation

Value list and data type validations as foreseen for batch upload must also be applied to the manual inserting of data (e.g., countries lookup, ISO date format enforced, syntax of names enforced, etc.).

System should enable the locking of specific records or fields within records, to prevent any change suggestions for entire records or for specific fields within records.

4 Other Requirements

4.1 Data migration

The successful candidate must plan and execute migration of data from the existing *Treasures* database into the new data model. Migration must be complete where logically possible. Where

existing data cannot be conformed to the new data model, legacy data must be retained separately so that it can be manually entered in the new system afterwards, as necessary.

The data migration concept must sufficiently consider the transition from a flat, work-based data structure to a relational work-item structure. It must include suggestions and/or aids for how to normalize existing data (especially work/item) efficiently.

4.2 Open source components

Applications and technologies used must be open source only. Exceptions to this rule are only possible with explicit consent by the customer (FIAF). If proprietary components or applications are used, the bidder needs to make this explicit in the proposal.

4.3 Data and system security

Effective measures must be taken to prevent data integrity and system security infringement. Obviously, this is valid for the whole system and all its components. However, it is of special importance with regard to upload functionality. For example, security measures must include prevention of SQL injections and other known threats.

4.4 Integration into existing technical infrastructure

- Delivered product must be able to export data to existing vendor infrastructure (ProQuest, Ovid) as appropriate
- Delivered product must integrate with FIAF website user registration and authorisation

4.4 Start-up/launch

The system must be delivered in production state, together with sufficient information for an IT administrator to deploy, install and configure all of its components. Successful bidder will be required to supply technical support on request.

4.5 Documentation

The successful candidate must supply English-language documentation about the system sufficient to explain in detail the following:

- Installation/setup/configuration
- Use
- Technical dependencies
- Graphical ER-Schema of new Treasures database

4.6 Project management and quality control

The successful candidate will assume responsibility for the following project management requirements:

- Plan and organise kick-off meeting and other meetings (including agenda, minutes)
- Assure quality of delivered product
- Use issue tracking/reporting system

5 Required Information from Bidder

Proposals must be submitted as a single PDF document via email to Rutger Penne, treasures@fiafnet.org.

In order to be eligible for evaluation, submitted proposals must meet the following criteria:

- a. All text must be in English.
- b. Pricing must be in €. Only fixed price offers are acceptable.
- c. Presentation and description of at least one reference project similar to the proposed project in nature and size completed not before 01.01.2015, including
 - i. Start and end dates
 - ii. Name of client
 - iii. Contact information (e-mail, phone) to client
- d. Alternatively, proposals must include a portfolio of comparable work.
- e. A draft project plan with milestones must be included, with development extending up to a maximum of six months, followed by up to three months of testing. Milestones must detail time (relative to kick-off) and nature of the achievements as well as efforts required (in days). Milestones must provide enough detail to be able to track project progress and identify dependencies. Mandatory milestones are:
 - i. Kick-off
 - ii. Approval of concept
 - iii. Delivery and integration of system for inspection
 - iv. Approval
- f. A description of how the functional requirements will be met, with reference to the description above (approx. 4-6 pages). This includes a full description of all software applications and frameworks to be used.
- g. Describe availability for support contract upon completion of the proposed project

Appendix 1: report commissioned by FIAF in 2020

Appendix 2: exchange format field list

6 Timeline Until Kick-off

2021-03-08	Publication of request for proposals (<i>this document</i>)
2021-04-05	Deadline for bidders' questions (to treasures@fiafnet.org)
2021-05-03	Submission of final proposals
2021-05-17	Applicants notified by FIAF, Award of contract
2021-06-14	or later (TBD) Kick-off meeting