EFG´s Vocabulary and Matching Work

SESSION 4: VOCABULARIES

EFG Workshop on Data Quality and Semantic Interoperability Issues in European Film Archives

Frankfurt am Main, 30 May 2011

Francesca Schulze
(Deutsches Filminstitut – DIF)
f.schulze@deutsches-filminstitut.de
www.europeanfilmgateway.eu

Cofunded by the Community Programme eContent*plus*

Overview of presentation

- EFG Vocabularies
- Matching work
- EFG portal examples (controlled vocabularies)
- Vocabulary Checker
- EFG ingestion workflow (with all cataloguing activities)
- Achievements
- Visions

45 controlled vocabularies

- harmonise heterogeneous and multilingual source metadata (values) collected in EFG
- display values coherently in EFG and Europeana portal
  - ✓ vocabularies established for end users (public)
- express archival data consistently and contextualise them in a reference model
  - ✓ EFG schema: ER model with subject predicate object triples
- compiled value lists for attributes & semantic relationships
- assigned vocabularies with elements of EFG schema (textual documentation)
- translated into 13 European languages
- public version available on EFG project website:
  - ✓ www.europeanfilmgateway.eu/guidelines_and_standards.php

Controlled vocabularies in EFG
What do they cover?

- Formal aspects to describe a film work
  - ✓ form/category, format, provider name, etc.
- Formal aspects to describe a non-film work
  - ✓ document type, format, provider name, etc.
- Types for attributes of different entities
  - ✓ date & name types, title types, activitiy types, etc.
- Spatial coverage
  - ✓ countries, regions
- Semantic relationship types
  - ✓ cast & credits, etc.

➢ no harmonisation of subject terms
➢ no types for non-human characters and roles

Controlled vocabularies in EFG
How were they compiled?

- FIAF Glossary of Filmographic Terms, 2008 (film-related activities)
- ISO Country Codes 3166-1
- AFNOR codes (historical countries)
- Marc Geographic Area Codes (large regions)
- ISO Language Codes 639-1, 639-2
- Marc Relator Codes (non-film activities)
- EBU P/META 2.0 Concept Schemes (language usage)
- EAC Beta (name and date types)
- IANA MIME Media Types
- ➤ Value lists of EFG partner archives

EFG vocabularies:
38 value lists with **1.700** term concepts
Example: "Documentary"

EFG Term	Vocabulary Name	Display Term	Language	Gender
Documentary	Form	Documentary	en	U
Documentary	Form	Dokumentar	da	U
Documentary	Form	Documentaire	fr	U
Documentary	Form	Dokumentarfilm	de	U
Documentary	Form	Dokumentti	fi	U
Documentary	Form	Ντοκιμαντέρ	gr	U
Documentary	Form	Dokumentarfilm	no	U
Documentary	Form	Dokumentární film	cz	U
Documentary	Form	Documentário	pt	U
Documentary	Form	Dokumentinis filmas	lt	U
Documentary	Form	Documentario	it	U
Documentary	Form	Documentaire	nl	U
Documentary	Form	Dokumentumfilm	hu	U

Vocabulary matching:
16 tables with **11.000** matched source values
Example: "Colour"

Data Provider	Source Value	Language	EFG Term	Vocabulary Name
EYE	Zwart-wit	nl	Black & White	Colour
EYE	Kleur	nl	Colour	Colour
EYE	Onbekend	nl	n/a	Colour
EYE	Tinting	nl	Tinted / Toned / Hand coloured	Colour
EYE	Toning	nl	Tinted / Toned / Hand coloured	Colour
EYE	Inkleuring	nl	Tinted / Toned / Hand coloured	Colour

Vocabulary matching

Example: Terms matched to "Assistant camera operator"

1. fotografassistent
2. Ass. Cameraman
2. foto-ass
3. Ass. Cameraman
Assistant camera
Assistant Cinematographer
Assisterende fotograf
B-foto
Camera assistant
C-foto, Danmark
First assistant camera
Foto praktikant
foto-ass
Foto-ass i Nicaragua
Fotograf - motorcykel
Fotograf 2. assistent
Fotografass.
Fotografisk assistance
Kameraass.
Kameraassistent
Multicam technician
Second assistant camera
Suppl. kamera
B-foto, 2. unit
B-Fotograf
Kamera, 2.unit
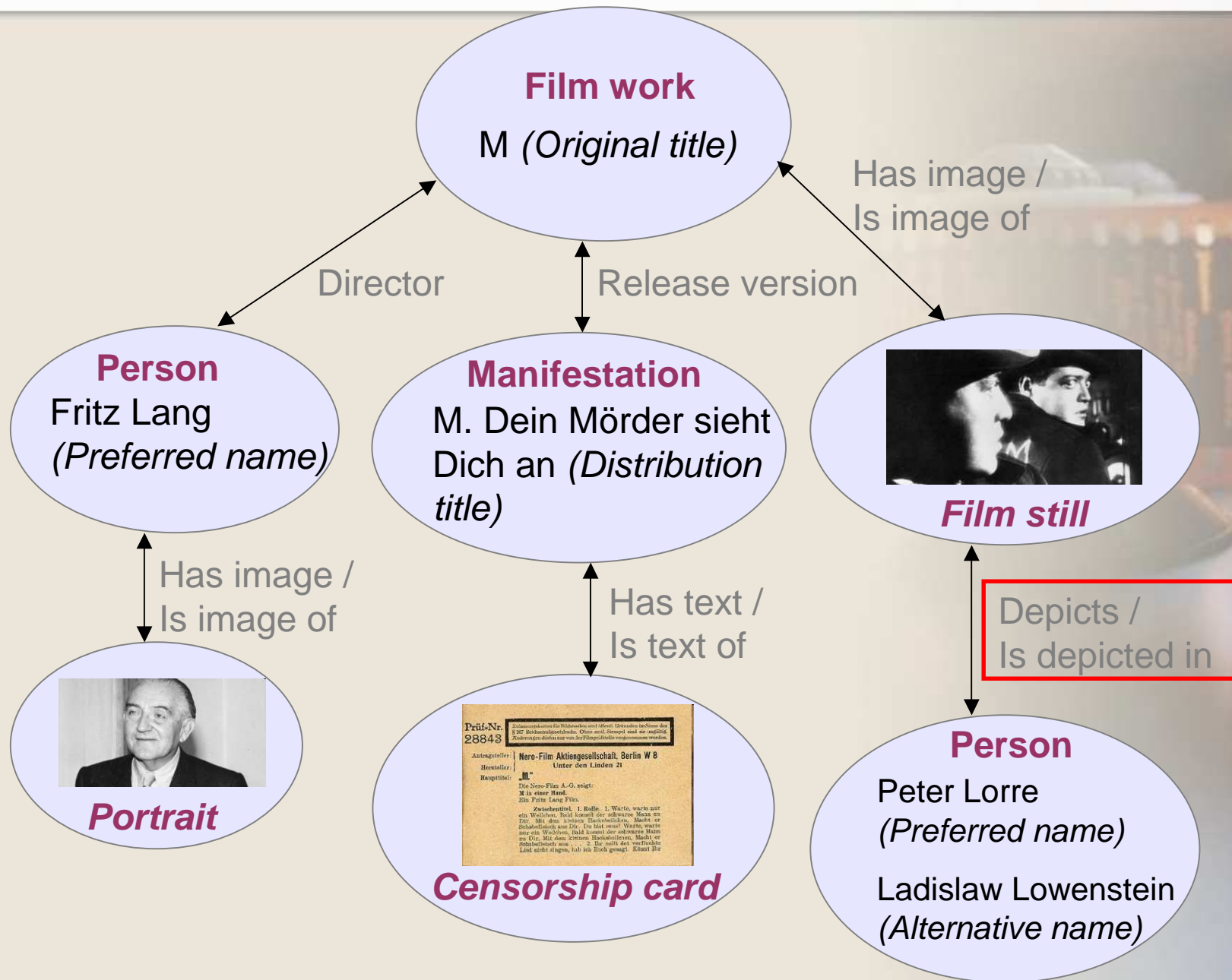
Heterogenous source values

✓ 60 % terms from free-text fields

✓ 40 % terms from controlled vocabularies

    (diverse vocabulary lists at archives)

EFG vocabularies:
7 lists with 210 semantic relationship concepts
Example: "Director"

EFG Term	Vocabulary Name	Domain	Display Term	Display Term Inverse	Range	Language	Gender
Director	AgentRelation	AVCreation	Regisseur	Regisseur	Person	nl	U
Director	AgentRelation	AVCreation	Režisér	Režisér	Person	cz	M
Director	AgentRelation	AVCreation	Režisérka	Režisérka	Person	cz	F
Director	AgentRelation	AVCreation	Rendezı	Rendezı	Person	hu	U
Director	AgentRelation	AVCreation	Regissør	Regissør	Person	no	U
Director	AgentRelation	AVCreation	Instruktør	Instruktør	Person	da	U
Director	AgentRelation	AVCreation	Ohjaaja	Ohjaaja	Person	fi	U
Director	AgentRelation	AVCreation	Realizador	Realizador	Person	pt	M
Director	AgentRelation	AVCreation	Realizadora	Realizadora	Person	pt	F
Director	AgentRelation	AVCreation	Režisierius	Režisierius	Person	lt	M
Director	AgentRelation	AVCreation	Režisier	Režisier	Person	lt	F
Director	AgentRelation	AVCreation	Réalisateur	Réalisateur	Person	fr	M
Director	AgentRelation	AVCreation	Réalisatrice	Réalisatrice	Person	fr	F
Director	AgentRelation	AVCreation	Regisseur	Regisseur	Person	de	M
Director	AgentRelation	AVCreation	Regisseurin	Regisseurin	Person	de	F
Director	AgentRelation	AVCreation	Σκηνοθέτης	Σκηνοθέτης	Person	gr	U

Film work
M *(Original title)*

Director

Release version

Has image /
Is image of

Person
Fritz Lang
(Preferred name)

Manifestation
M. Dein Mörder sieht Dich an *(Distribution title)*

Film still

Has image /
Is image of

Has text /
Is text of

Depicts /
Is depicted in

Portrait

Censorship card

Person
Peter Lorre
(Preferred name)

Ladislaw Lowenstein
(Alternative name)

**Detail page:
VIDEO**

RELATED NAMES

Dreyer, Carl Th. | Screenplay
Jul, Christen | Screenplay
Maté, Rudolph | Director of photography
Taldy, Tonka | Editor
Zeller, Wolfgang | Music composer

> More

RELATED TITLES

No matches

◀ Back to search results
🖨 Printer-friendly version
👤 Send to friend

View on Danish National Filmography

Vampyr

Other title(s):	Vampyr (L'etrange aventure de David Gray) Allan Grays sælsomme hændelser Vampyr (Die seltsame Geschichte des Allan Gray)
Genre:	Feature film
Country:	DE FR
Year:	1932
Runtime:	75 min
Description:	En ung mand, Allan Gray, ankommer til en kro ude på landet. Samme aften får han besøg af en ukendt mand, der beder ham om at hjælpe med at frelse en ung pige, der er i fare. Hun viser sig senere at være den ukendte mands datter. Kort efter bliver den ukendte mand, der bor på et stort gods, myrdet, og den unge mand må selv prøve at komme den unge pige til undsætning. Derfor giver han blod til en transfusion, forhindrer den syge kvinde i at tage gift, befrier hendes søster, der er bundet til en seng, og drager til sidst bort med hende. Desuden er det lykkedes ham at standse den forbandelse, der hvilede over egnen, ved sammen med tjeneren at drive en jernpæl gennem en gammel kvindelig vampyrs hjerte. Derefter sørger tjeneren for, at hendes medskyldige, lægen, bliver kvalt - i mel (fra "Carl Th. Dreyer født Nilsson" af Martin Drouzy, 1982). A young man, Allan Gray, arrives at a countryside inn. At the very same evening an unknown man pays him a visit and asks him for help to save a young girl who is in great danger. As it turns out the girl is the unknown man's sister. Shortly after the visit the unknown man - resident at a great manor - is murdered, and Allan Gray has to come to the young girls rescue all alone. He gives blood to a transfusion, prevents the sick woman from taking poison, sets her sister who is tied to a bed free, and finally sets out with her. Furthermore he has succeeded in putting an end to the curse that has been upon the locality; together with the servant he has driven an iron pile into the heart of an old female vampire. The servant makes sure that her accomplice, the doctor, suffocates - in flour.
Provider:	Det Danske Filminstitut
Production company:	Film-Production Carl Dreyer
Director:	Dreyer, Carl Th.
Colour:	Black & White
Sound:	Without sound

Result page: filtering and thumbnail display

MY EFG
REGISTER | LOGIN

EFG *beta*
european film gateway

ABOUT EFG

king

Search >

Find films, photos and texts from the collections of Europe's film archives

> HOME   > SEARCH RESULTS: "KING"

FILTER

Provider
☑ All
☐ Cinecittà Luce S.p.A. (79)
☐ Deutsches Filminstitut - DIF (57)
☐ Det Danske Filminstitut (35)
☐ FernUniversität Hagen (12)
☐ Filmarchiv Austria (7)
☐ La Cinémathèque Française (5)
☐ Národní filmový archiv (5)
☐ Magyar Nemzeti Filmarchívum (1)

Year

Language
☑ All
☐ DE (9)
☐ IT (5)
☐ CS (4)
☐ EN (4)

Media
☑ All
☐ image (131)
☐ video (41)
☐ text (29)

APPLY

VIDEOS more videos   1 2 3 4 5 ... 11 >   41 Results

Nulla di fatto per Jannilli e Jimmy Ki...
Newsreel
1952
Cinecittà Luce S.p.A.

Kong Haakons Besøg paa 'Herluf Trolle'..
DK/Dokumentarfilm
DK 1906
Det Danske Filminstitut

"Divi" in Italia: Tyrone Power a S. Ma...
Newsreel
1948
Cinecittà Luce S.p.A.

La spogliarellista danese Katy King in...
Newsreel
1961
Cinecittà Luce S.p.A.

IMAGES more images   1 2 3 4 5 ... 33 >   131 Results

King's Palace Delhi
Drawing
La Cinémathèque Française

King Kong und die weiße Frau
Picture
Filmarchiv Austria

King Kong und die weiße Frau
Picture
Filmarchiv Austria

King Kong und die weiße Frau
Picture
Filmarchiv Austria

TEXTS more texts   1 2 3 4 5 ... 8 >   29 Results

Die Fabel von King-Kong
Press material
1933
Deutsches Filminstitut - DIF

King Kong
Press material
1933
Filmarchiv Austria

King Kong
Press material
FernUniversität Hagen

King Kong
Title list
1933
Filmarchiv Austria

Result page: TEXT

EFG beta
european film gateway

ABOUT EFG

Find films, photos and texts from the collections
of Europe's film archives

Search >

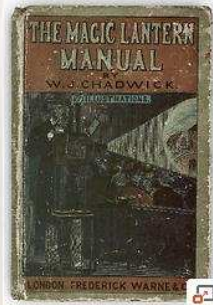> HOME > SEARCH RESULTS: "LANTERN" > THE MAGIC LANTERN MANUAL

RELATED NAMES

Chadwick, William Isaac | Author(s)

RELATED TITLES

No matches

◀ Back to search results
♥ Social networks
🖨 Printer-friendly version
👤 Send to friend

View on cineressource.net

The magic lantern manual

Creator: Chadwick, William Isaac
Date created: 01.01.1885
Description: 154 p. : front., fig., couv. ill. en coul. ; 19 cm
Keywords: Lanterne magique
Provenance: La Cinémathèque Française
Provider: La Cinémathèque Française
Rights: William Isaac Chadwick
Document type: Book
Language: en

FURTHER RESULTS   1 2 3 4 >   13 Results

The magic lantern manual	How to make lantern slides	The magic lantern manual	The magic lantern ; or world as it goe...
Book	Book	Book	Book
1878	1898	1885	1822
La Cinémathèque Française	La Cinémathèque Française	La Cinémathèque Française	La Cinémathèque Française

EFG Backend Tools

➤ Vocabulary Checker

➤ Authority File Manager

➤ Metadata Editor

Vocabulary Checker

➢ Data cleaning on EFG level
➢ Tool used to verify vocabulary matchings

- matching intellectual work by partner archives
- no vocabulary management tool
- vocabularies and matching tables are managed outsite the EFG System

Vocabulary Checker

EFG european film gateway

EFG Rotten Checker

Welcome franca.debole | EFG Tools | Logout

Last Update: *2011-01-24T13:10:49+01:00*

Record XML

Invalid fields are **highlighted**

```
<efg:efgEntity>
  <efg:award>
    <efg:identifier scheme="CP_CATEGORY_ID"> DIF_award_0019FB69CE854B46BCC7D2764E442FDA </efg:identifier>
    <efg:recordSource>
      <efg:sourceID> 0019FB69CE854B46BCC7D2764E442FDA </efg:sourceID>
      <efg:provider schemeID="Institution acronym" id="DIF">Deutsches Filminstitut - DIF</efg:provider>
    </efg:recordSource>
    <efg:name> Prädikat: wertvoll </efg:name>
    <efg:sponsor> FBW </efg:sponsor>
    <efg:date> 2004-08-01 </efg:date>
    <efg:relAvManifestation>
      <efg:identifier scheme="CP_CATEGORY_ID"> DIF_avCreation_2BAB82FFFFF44B6DBD3FB5178DFA1D5D </efg:identifier>
      <efg:title> Greenhorn </efg:title>
      <efg:type>film</efg:type>
    </efg:relAvManifestation>
  </efg:award>
</efg:efgEntity>
```

This value is invalid ✕

Allowed values for *AVManifestationRelation* vocabulary are:

▷ Alternative version
▷ Archive version
▷ Censored version
▷ n/a
▷ Original version
▷ Other version
▷ Partial version
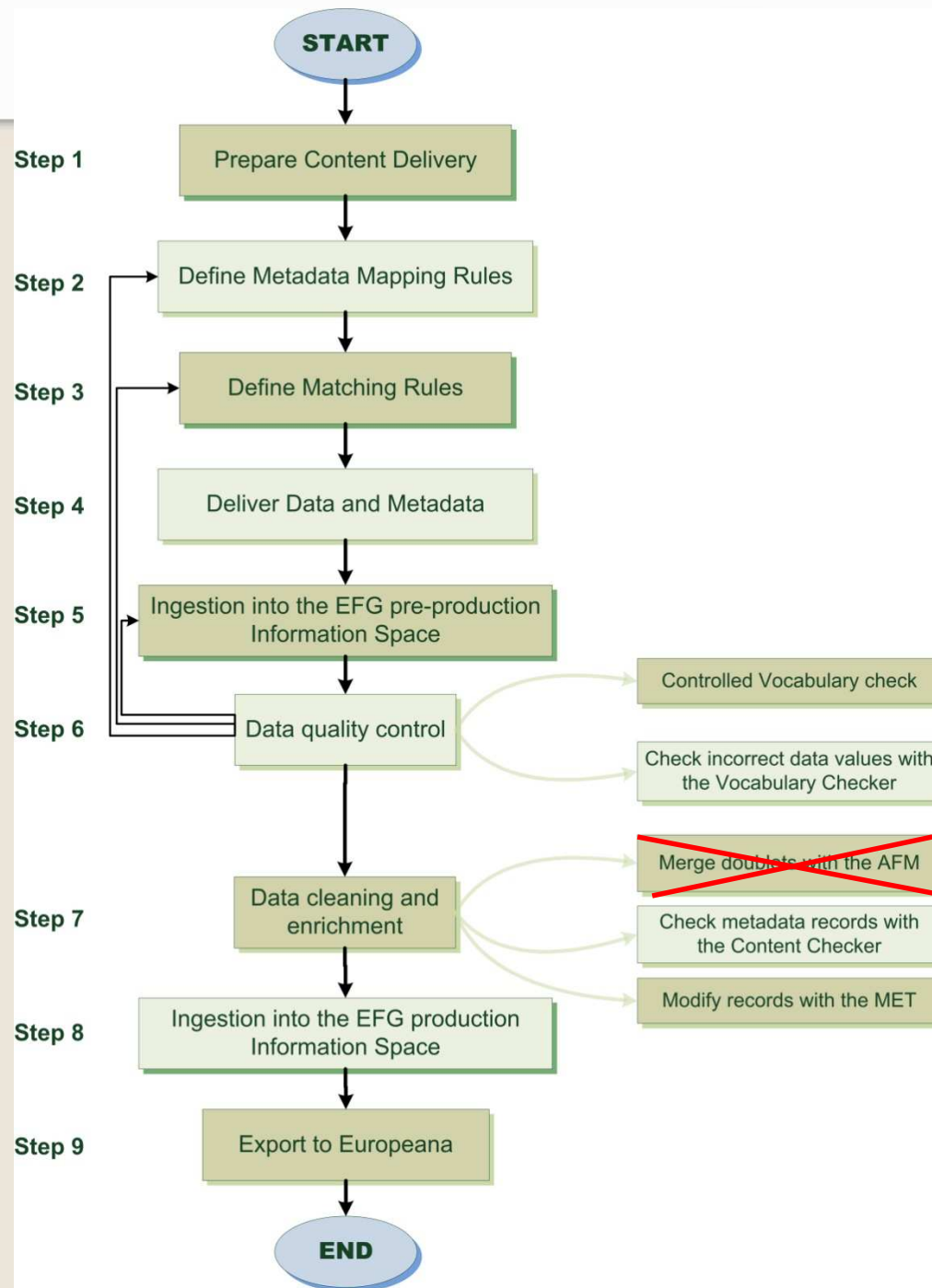▷ Release version
▷ Short version
▷ TV adaptation

[Back]

Powered by D-NET

Co-funded by the Community programme eContent*plus*

EFG is solely responsible for the content of this site, which does not represent the opinion of the Community.
The Community is not responsible for any use that may be made of the information contained on this site.

EFG ingestion workflow with all cataloguing activities

START

Step 1 — Prepare Content Delivery

Step 2 — Define Metadata Mapping Rules

Step 3 — Define Matching Rules

Step 4 — Deliver Data and Metadata

Step 5 — Ingestion into the EFG pre-production Information Space

Step 6 — Data quality control
- Controlled Vocabulary check
- Check incorrect data values with the Vocabulary Checker

Step 7 — Data cleaning and enrichment
- Merge doublets with the AFM
- Check metadata records with the Content Checker
- Modify records with the MET

Step 8 — Ingestion into the EFG production Information Space

Step 9 — Export to Europeana

END

Achievements

- Compiled film-relevant controlled vocabularies in 13 languages
  - ➤ Require further refinements for needs of filmographers
- Harmonised statements that can usually be found in filmographic databases (e.g. persons, film works, corporations)

- Contextualised film archival data through semantic relationships (esp. cast & credits, events, etc.)

- EFG vocabularies can be used in other contexts (available on: EFG project website)

- Harmonised heterogenous and multilingual source values in a common filmographic database (EFG XML records)

- Display values consistently in EFG and Europeana portal

Outlook: What could be done after EFG I?

Administrative

➢ Apply vocabulary management tool to update and maintain EFG vocabularies

➢ Represent EFG vocabularies in XML and link them to EFG metadata schema
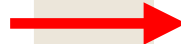
EFG portal

➢ Provide vocabularies in different languages in EFG portal

➢ Exploit EFG vocabularies for the portal to satisfy user needs and apply them for a proper retrieval (media- and document-specific filters, categories as drop-downs in advanced search)

➢ Find solutions for subject indexing, esp. World War I → in possible EFG II project

Other uses

➢ Discuss EFG vocabularies among the film archival community

➢ Further develop vocabularies for needs of filmographers

Stay informed and involved!

- ➢ EFG Vocabularies

- ➢ Data Provider Handbook

  - ✓ Data Cleaning and Enrichment Guidelines

- ➢ Reports on WP 3 work (Outcomes)

http://filmstandards.org

www.europeanfilmgateway.eu

Thanks for your feedback!