19 / 20 / 21 novembre 08

*Cinéma et audiovisuel :
quelles mémoires numériques pour l'Europe ?*

DES PROJETS EN EUROPE, DES PROJETS EUROPÉENS?

EFG – A gateway to Europe's film heritage

Georg ECKES,
Deutsches Filminstitut – DIF

Thanks for the invitation, it is a great pleasure. I work as the project coordinator of the project EFG (European film Gateway)[1] at the German Film Institute (Deutsches Filminstitut – DIF). The project is supported by the eContentplus Programme of the European Commission, and it is one of the projects that aim at providing content from the archives, libraries and museums to the "European Digital Library" Europeana[2]. The Europeana prototype already launched yesterday, but in order to make it richer and get more collections into Europeana, these so-called aggregator projects have been put in place. EFG is a so-called Aggregator project for the film domain for Europeana.

EFG is not a digitization project; it is rather about access to already existing digitized collections. It is also not a video-on-demand project, since it has a scope that goes beyond the moving images: It also deals with film-related material, and it is designed as a kind of showcase for the archives to show what they have, and if possible to show it for free to the general public. I would like to quote Elisabeth Niggemann, the Chair of CENL (Conference of European National Librarians), who describes Europeana's objectives as follows: "For the users it is not important whether the sources of knowledge and experience are kept by archives, libraries or museums, but to get access to the sources they want, and to be able to use them across types of sources and sectors of institutions". In a nutshell, this is the philosophy behind Europeana. Europeana is about providing access to cultural heritage across the domains, across media types, across countries, across languages and across institutions. Film archives and the film domain are included in the concept of Europeana.

There are several projects grouped around Europeana that provide content to it. For example for the library domain there is "The European Library" (TEL)[3], certain national digital libraries that can provide content to Europeana, a museum aggregator called Athena[4], which is a project that will start soon, a portal for European National Archives called EPA[5], and also from the audiovisual domain there is the EUscreen project[6] which will be a follow-up project to VideoActive[7].

---

[1] http://www.europeanfilmgateway.eu

[2] http://www.europeana.eu

[3] http://www.theeuropeanlibrary.org

[4] http://www.athenaeurope.org

[5] http://www.europeanarchivesgateway.eu

[6] http://ec.europa.eu/information_society/events/cf/document.cfm?doc_id=9107

[7] http://www.videoactive.eu/VideoActive/Home.do

The European Film Gateway project complements the activities of these other activities by aggregating content from the film archives.

Why aggregators? Why not deliver it directly? Of course that is possible and it should also be done. But during the work on the Europeana prototype, it has proven very difficult to homogenize and make interoperable the content and especially the cataloguing information that comes from the different domains and institutions. By now Europeana has integrated more than 140-150 institutions, and according to the Europeana office, it proved to be not scalable to integrate metadata individually on an institution-by-institution basis. Hence, Europeana stresses the need for domain-specific aggregators to do some work beforehand, in terms of homogenizing and cleaning data. This is what projects like the European Film Gateway do or will do. In short, the project is about cleaning up the film archives' own backyard. If you are into cataloguing and databases or heritage or film institutions you know that there is a lot to be done.

Another reason why aggregators make sense is that with a portal such as Europeana you inevitably reduce information to some degree. To circumvent this, we are going to build a portal specifically for film archives and cinémathèques, so that if a user finds something in Europeana from the film archives, he can go to the EFG site for more information, then again gets to the individual site of the archive for even more detailed information.
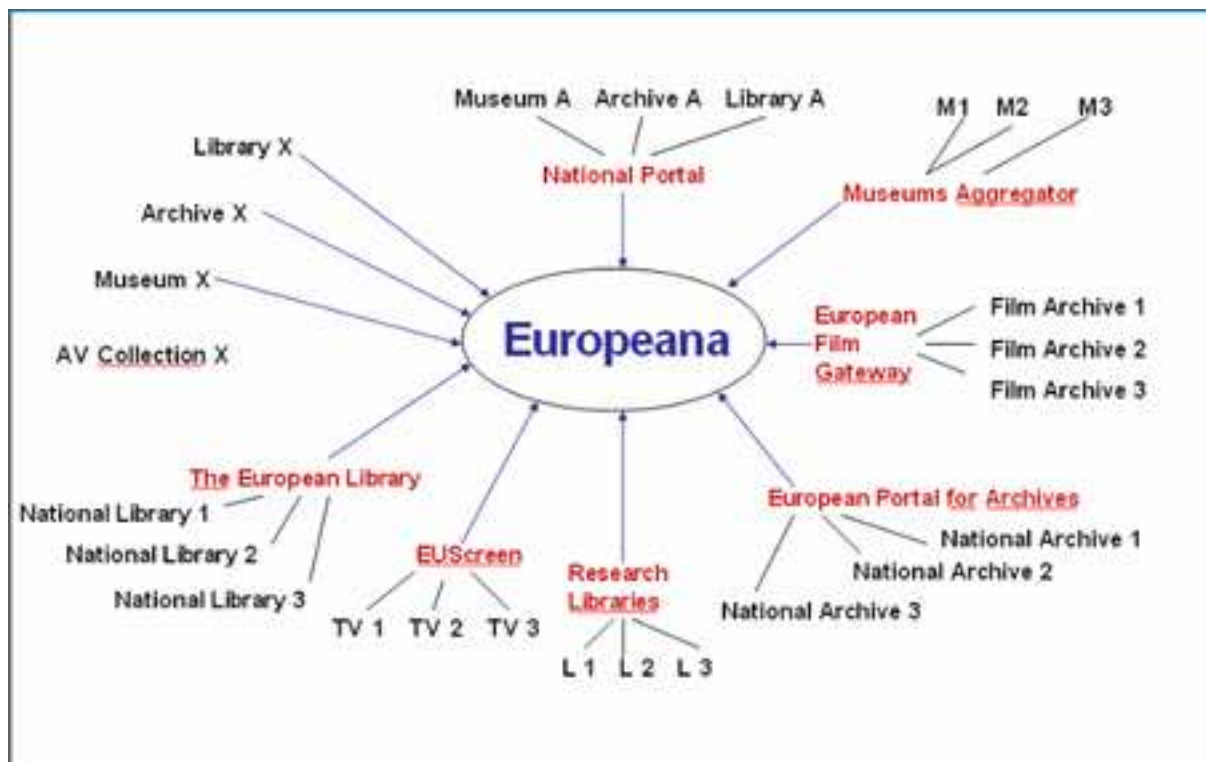
Why EFG?

There is a growing number of digitization projects and digital repositories from film institutions. It is getting increasingly difficult to get an overview of this. We need a registry of collections and items as well as film-specific authority files on a trans-institutional and trans-national level. We need common interoperability standards for exchanging information between film heritage institutions and especially for exchanging film-related metadata across the barriers of countries and languages.

Of course when publishing content on the World Wide Web there are also several IPR issues such as how to deal with works of film that are presumed "orphan", how to identify rights owners and how to come to agreements concerning world-wide internet access to films which are currently out of distribution. In our experience many film archives still lack experience with these issues, and this one thing we are planning to tackle in EFG.

What is EFG supposed to achieve?

It is about building a digital showcase for collections of film archives and cinémathèques; this includes moving images, images, text material and sound material: Collections that have been digitized in many local or national digitization projects. EFG does not have a selection stance, it relies on the decisions which are made in these individual projects in the institutions. Therefore, it is just about building a framework, an information space, to give access to what has been generated in those individual projects.

It is designed as a central access point to federated digital collections and repositories. EFG is about copying the metadata and providing free public access through a direct link to the digital object which remains on the individual institution's site; it is not about copying content. We are also planning to build authority files for films works and persons to be able to provide highly reliable filmographic information coming from Europe's film heritage institutions; information coming from several existing projects and databases that you can already find on the web, for example the Cinéressources from the Cinémathèque.

The EFG consortium consists of 20 partner organisations including 14 archive partners from Germany, Italy, Denmark, the Netherlands, Czech Republic, Austria, Cinémathèque Française, Greece, Hungary, Norway, Portugal, Lithuania and Switzerland.

The remaining partners are mainly responsible for technical aspects, such as our main technology provider in Pisa. The EDL foundation which is behind Europeana is also a partner in this project, as well as the Association des Cinémathèques Européenes.

*What are the main challenges in this project?*

There are two main challenges: The interoperability of metadata and issues resulting from the fact that most of the archives do not hold the rights to most of the films they preserve.

Interoperability of metadata

In the film domain we have cataloguing rules like the FIAF cataloguing rules but they are currently under revision and they are not consistently applied, or not applied at all at the moment, which makes it difficult to aggregate information. Furthermore, there is no common standard for exchange of data between film archives. This results in a wide variety of data record structures which actually hamper the aggregation of information on a trans-institutional level. Controlled information such as authority files and controlled vocabularies can only be found within institutions; in most cases they are not used across institutions at the moment, which also hampers the data exchange. A further very serious problem can be seen in the varying quality and depth of indexing across the institutions. And last but not least, common metadata exchange protocols such as the Open Archives Initiative's Protocol for Metadata Harvesting (OAI-PMH) is rarely implemented at the moment, although it is gaining ground. In fact, we are aiming at using OAI-PMH as the main means of data exchange in EFG.

There are two kinds of information which are being aggregated in this project:

- *Authority files*
- *Object descriptions and location of objects*

o
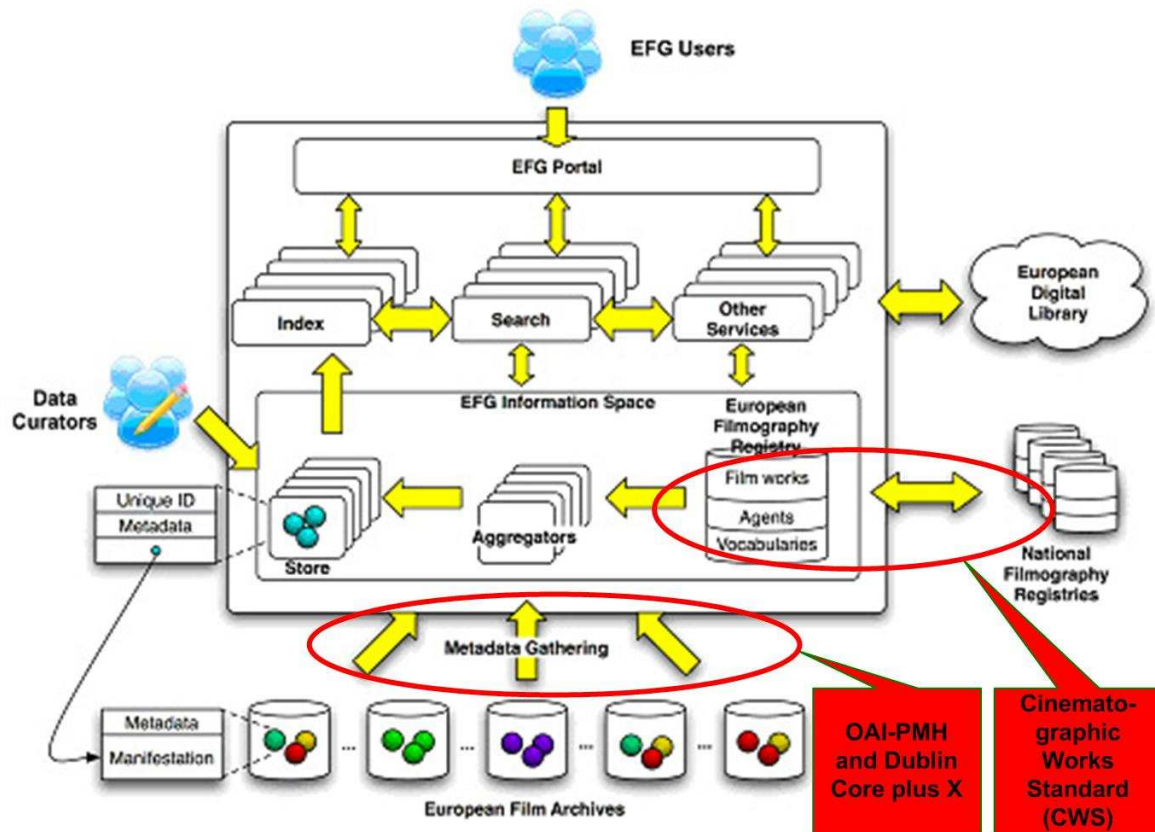
At the bottom of the chart you can see the actual digitized collections and objects from the institutions, to be harvested probably via OAI-PMH into the EFG information space, probably using a simple exchange format such as Dublin Core plus several other fields or entities. This is something we are working on at the moment.

The second type of information, the authority files, comes from the national filmographic registries, so this is information about film works, not exactly about digital objects. They are to be aggregated in a European filmographic registry for film works and for persons.

CWS is an abbreviation for Cinematographic Works Standard, this is the standard that results from the standardization initiative. It is currently under development and we will be finished soon. If you are interested in it you can look at filmstandards.org *and will find news on the development of the CWS.[8] This standard is supposed to provide a framework for the unambiguous identification of film works that are catalogued under different distribution titles; the same applies to persons. This is one thing that Serge Bromberg pointed to yesterday when he asked the question "why do you restrain yourself in the Midas project and only aggregate information about non-fiction, non-feature films? Why don't you also take information about feature films and aggregate them?" Because you have the same feature films catalogued or registered in many different cataloguing systems and many different institutions in many different countries. If you have no proper identifier for them and no proper rules for identifying the original title or for example the German or French distribution title, you get search results like this:*

Search results (4)

Specify Search »

Title	
HURDES LAS TIERRA SIN PAN	Spain
HURDES, LAS (Tierra sin pan) (Terre sans Pain)	Spain
Las Hurdes	Spain
LAS HURDES O TIERRA SIN PAN	Spain

This is a search result from the Midas project;[9] when you find this you do not know whether these entries refer to the same cinematographic work or to different ones. You simply cannot find out what version of this film is the one you are searching for when the data is just thrown into one big repository.

Hence, we need a framework for the identification of film works and conversions that are associated with film works. In the case of "Las Hurdes" you have the original mute version directed in 1933 (Original Spanish title: "Las Hurdes, tierra sin pan"), and based on this work you have another film work from 1936 which is the sound version in French (French distribution title: "Terre sans pain"; Spanish distribution title: "Las Hurdes o tierra sin pan"; English distribution title: "Land without bread"), and you have additional credits to this film work for the music, musical score or narrator. Associated with this new work you have different versions in different languages. When this is summarized properly under one film work record, you do not get hundreds of search results when searching for a film title, since the film can be properly identified in an aggregated database.

This work of associating different film versions with one single record can not be done in a fully automated way. It has to be done by humans because it is an intellectual decision to state that this film belongs to this work, or this person has the same name but came from different databases. Because of this, there is a lot of indexing work in the EFG project. However, we are not aiming at building a full European filmographic registry because it is impossible within the scope of a project like this. We are aiming at a kind of nucleus of a European Filmography.

IPR issues

The film archives assembled in EFG have the same problems as many other archives:

- *Copyright usually expires seventy years after the death of the creator, which leaves very few film works freely available in the public domain. I think the percentage of films in European archives that the archives do not hold the rights to is over 80%.*

- *The creators are often very hard to find, and it is a very difficult to find out whether a film can be seen as an "orphan" or not.*

What are we going to do about this? First of all we are going to do a survey about copyright regulations in the European countries concerning the use of public domain works, of copyrighted works and orphan works, especially with regards to film archives. The resulting report will be public, and will also be available on the EFG website. This survey is being carried out by the Nederlands Filmmuseum. The findings from this survey will provide several general guidelines for film archives, basically about what to do and how to proceed when researching cleaning and negotiating rights towards Internet access, into a kind of checklist for checking: "What do I have to have done before being able to publish material on the World Wide Web?"

---

[9] http://www.filmarchives-online.eu

To summarize rights issues in this project, of course regarding moving images we want to have the full film, but in many cases that is not realistic. In this project we have to be prepared to compromise, giving access to parts of films, giving access to films in reduced streaming quality and so on. These are all compromises that can be made with rights holders or exploiters. It can also be a means of collaboration with rights owners and exploiters. For example, we can offer to someone who has a video-on-demand service, trailers from this service to be accessible via EFG and in Europeana so that people can find them more easily. This is a means of providing access to moving images, in a collaborative way together with rights holders and exploiters. We have to ask rights holders for permission, and we offer cooperation with rights holders, we are interested in co-operations such as public-private partnerships here.

A third point is that I suggest archives give access to a film which is considered orphan, after you have performed a diligent search for the rights owner, a search which has been properly documented.

Schedule

The project just started in September 2008 and so I cannot yet report any results to date. It will run for three years until August 2011. As far as the metadata part is concerned, we are planning to have an interoperability schema by May 2009, and we plan to launch the portal by June 2010. If you would like more information and news about the project, consult our website www.europeanfilmgateway.eu.

Joël DAIRE

Merci Georg d'avoir tenu le défi de présenter ce projet très lourd et très complexe en trente minutes chrono.

Pour les institutions qui participent, ce projet est un vrai défi du fait de sa complexité et de la durée que nous avons pour atteindre les objectifs. Si nous y arrivons, et nous l'espérons tous, sur les deux questions majeures sur lesquelles Georg a insisté aujourd'hui, nous prendrions tout d'un coup dix ans d'avance, et ce serait une victoire formidable.

Un intervenant de la salle (1)

J'ai deux questions juridiques qui s'adressent aux deux intervenants. Par rapport à l'allongement des droits voisins qui touchent les droits producteurs, interprètes, etc., à 90 ans et qui, actuellement est agité au niveau européen.

Luis FERRÃO

Je crois que vous faites référence à la proposition de directive pour étendre le droit des artistes et interprètes et des producteurs de phonogramme de 50 ans à 95 ans. Cette proposition est sur la table du Conseil du Parlement, en processus de codécision. Elle fait l'objet de discussions qui s'annoncent encore relativement longues. Ce sujet n'est pas très pacifique. Cette proposition ne vise pas les droits des producteurs de films. Il s'agit de droits voisins des artistes et interprètes et concerne surtout le domaine musical même si l'objet peut évoluer de sorte à couvrir aussi l'audiovisuel. Tout cela reste ouvert, de même que la durée de 95 ans. On ne peut pas dire grand-chose à ce stade, puisque l'issue du débat est impossible à prévoir. Cela peut déboucher soit sur une extension telle que proposée ou soit sur une extension moins large, soit sur rien du tout.

Un intervenant de la salle (1)

L'autre question est sur ce que vous avez cité sur la chaîne de l'évolution des droits des ventes, etc. Envisage-t-on une publicité de la propriété? C'est bien de trouver l'ayant droit, mais en est-on sûr? Quelle est la sécurité juridique de ces droits?

Luis FERRÃO

Le problème de la chaîne des droits a été identifié par le sous-groupe de droits d'auteur. Il existe des normes identifiantes comme ISAN pour l'audiovisuel, mais ces identifiants ne suivent pas la chaîne des droits alors que la réalité est dynamique. Un décalage existe forcément, à partir d'un certain moment, entre ce que l'identifiant dit et la réalité des droits. C'est pourquoi nous avons proposé de développer des métadonnées, des identifiants qui soient eux-mêmes dynamiques, qui suivent la chaîne des droits. Tout cela a fait l'objet de projets de recherche et autres. Cela fait aussi partie du projet Arrow que j'ai cité, qui vient de démarrer et qui implique des ayants droit de différents secteurs, y compris le secteur culturel, des éditeurs, etc., pour développer ces normes interopérables, mais qui reflète le dynamisme de la chaîne des droits.

* * * * * *

Georg ECKES