

CultLib - cultura libera

A Digital Repository of Free Cultural Objects

We preserve cultural objects in digital form
and make them freely available,
today and for future generations.

Version: 1.2
Date: 9. September 2009
Author: Dr. sc. math. Hartwig Thomas
Translator: Anne Smith

On the possession of ideas

If nature has made any one thing less susceptible than all others of exclusive property, it is the action of the thinking power called an idea, which an individual may exclusively possess as long as he keeps it to himself; but the moment it is divulged, it forces itself into the possession of every one, and the receiver cannot dispossess himself of it. Its peculiar character, too, is that no one possesses the less, because every other possesses the whole of it. He who receives an idea from me, receives instruction himself without lessening mine; as he who lights his taper at mine, receives light without darkening me.

That ideas should freely spread from one to another over the globe, for the moral and mutual instruction of man, and improvement of his condition, seems to have been peculiarly and benevolently designed by nature, when she made them, like fire, expansible over all space, without lessening their density at any point, and like the air in which we breathe, move, and have our physical being, incapable of confinement or exclusive appropriation.

*Thomas Jefferson,
letter to Issac McPherson 13. August 1813*

Summary

The project *CultLib* of the association *Digitale Allmend* (Digital Commons) has as its goal the founding of a Swiss trust *Pro Cultura Libera* for the creation and operation of a digital repository of free cultural objects. For this purpose it will collect works, objects of art and culture in the widest sense, which deserve being readily accessible over time and make them available to the general public without cost. Such works and collections will be made accessible and catalogued and stored in formats suitable for long-term preservation in a digital repository, called *CultLib* for short. They will be published under permanent, citable addresses for the benefit of the general public.

The rapid developments of the last two decades have led to the odd situation that although there has never been so much communication as today, the cultural activities of the human race will not be documented as comprehensively in libraries, archives, and museums in the long term as those of previous generations. The establishment of the *CultLib* repository is designed to counteract this “digital amnesia”.

The two main causes of this spreading loss of history are developments in technology and the current copyright law.

The *CultLib* repository will accept cultural works which can be distributed without problems. This means that they will be presented free to the general public by the legal owners of the rights, for example by means of a *Creative Commons* license. Users will not need to fear any campaigns against the copying and utilization of them.

In contrast to many ephemeral collections of free works of culture in “Web 2.0” *CultLib* is interested in long-term preservation. This means, therefore, that upon delivery the texts, pictures, recordings, moving-picture sequences must be processed, that is supplemented by meta-information (XMP, Dublin Core) and converted to suitable formats for long-term storage.

The labor involved in this process necessitates that the acceptance of new works in the repository be coupled with an evaluation of their cultural significance. Such a review process will hinder the *CultLib* repository from becoming a hodge-podge of trendy contributions.

Table of Contents

1 Digital Amnesia and the Increasing Loss of History.....	5
1.1 Two Reasons for the Digital Amnesia.....	5
1.2 Loss of History as a Result of Copyright Laws.....	6
1.3 The Ephemeral Nature of Many Internet Collections.....	6
1.4 Against the Loss of History and Memory.....	7
1.5 CultLib, a Project of the Digitale Allmend (Digital Commons).....	7
2 A Repository of Free Cultural Objects.....	8
2.1 Culture.....	8
2.2 The Concept of a Work.....	9
2.3 Free Access.....	9
2.4 Making Content Accessible.....	10
2.5 Duration of Preservation.....	10
2.6 Repository.....	11
2.6.1 Access for People and for Programs.....	11
2.6.2 Basis for Exhibitions of Cultural Agencies.....	11
3 Technical Aspects.....	12
3.1 File Formats.....	12
3.2 Long-term Identification.....	12
3.3 Metadata und Catalog.....	13
3.4 Open Systems.....	13
4 Organization.....	14
5 Operation and Finances.....	15
6 Conclusion.....	16

1 Digital Amnesia and the Increasing Loss of History

For centuries human civilizations have set up libraries, archives, art collections and museums in which books, music, paintings and other cultural objects are preserved and available to the general public. The interaction with cultural objects of earlier times is considered to be valuable in development of humans and their cultures and is promoted in schools and universities.

Although the vast majority of artistic works are now generated and communicated in a digital form, one hardly finds any serious collections of such works in this form whose contents are stored in a long-term manner. There is the danger that fewer works from the last twenty years will be preserved for future generations than from earlier decades. Attention has been drawn to this danger on international¹, European², as well as national (Swiss)³ levels.

1.1 Two Reasons for the Digital Amnesia

There are two reasons which are primarily responsible for this development. On the one hand the traditional institutions associated with preservation of cultural artifacts could not keep up with the turbulent technical developments in computer technology and communication; at the same time those creating digital cultural artifacts were so fascinated by the new possibilities that they tried out everything and fought ideological battles in regard to computers without concerning themselves with the preservation of their works. On the other hand international copyright law - almost unnoticed by the general public - had completely changed the regulation of ownership of "intellectual property".

Until recently it was not profitable for those in the software and services business to offer long-term, stable solutions. With proprietary, undocumented formats one could keep out the competition out. The users were not prepared to use open software with documented, standardized formats. The annual order for office software is granted to suppliers of proprietary software without a call for tenders as required by law. Large private companies behave no differently.

Formerly only the unauthorized commercial utilization of the works of others - which caused material harm to the owner of the rights - was illegal; free communication and further use of the content of the cultural objects was normal and was one of the basic public rights, of which schools, churches, universities, radio and television stations took full advantage. The new copyright law declared every expression of a person as his or her absolute private property and categorically forbade any kind of non-sanctioned use of the material. This development of copyright law to the complete privatization of "intellectual property" took place throughout the world between 1985 and 1995.

1 UNESCO Charter on the Preservation of the Digital Heritage
http://portal.unesco.org/ci/en/ev.php-URL_ID=13367

2 EU-Projekt PLANETS (Preservation and Long-term Access via NETworked Services) http://www.planets-project.eu/docs/comms/Planets_Project_Brochure.pdf

3 Memopolitik, Bericht des Bundesamtes für Kultur
<http://www.bak.admin.ch/themen/kulturpolitik/02082/index.html?lang=de>

In Switzerland the new copyright law became effective in 1992, the WIPO-copyright contract (World Intellectual Property Organization), came into effect on 1 July 2008. The above-mentioned international and national analyses identify the modern copyright law as one fundamental cause of digital amnesia.

1.2 Loss of History as a Result of Copyright Laws

These two factors have led to the current widespread and increasing lack of history of the younger generations. In 1974 in East Germany Wolf Biermann's song "Soldat, Soldat" (Soldier, Soldier) was played on cassettes clandestinely copied and passed on to others and everybody learned it by heart, sang it in the streets; in Switzerland Franz Hohler sang songs by Biermann, Vian and the Beatles without worrying about copyright questions. Now such works may not be used freely by other artists until 70 to 150 years after their creation.

There is a general insecurity concerning the rights in regard to digital data. The total privatization of intellectual property has led to demands for penalties of a fantastic dimension, that bears no relationship to the actual damage. Therefore those required to make decisions increasingly avoid all risks and forego using data from digital media in order to avoid possible law suits.

Many cultural objects today have been abandoned. That means that nobody knows any more to whom the rights belong, but copyright law and its executors forbid any publication with the permission of the owner of the rights or his or her heirs.

Thus today's youth learns little about the events, books, films, and the music of the sixties, seventies or eighties. And one only hears of the happenings in the last twenty years which the large media concerns judge to be profitable, despite the fact that the worldwide distribution of information was never so simple nor so cheap as today.

It is foreseeable that today's copyright laws, in spite of all the international contracts, will not be able to withstand the normative forces of reality. But until a world-wide correction of the total privatization of all intellectual activity becomes political reality, one will only be able to make those works accessible to the general public, which have been made available by the owner of the rights or whose authors have been dead for more than 70 years.

1.3 The Ephemeral Nature of Many Internet Collections

To be sure, there are a multitude of collections of cultural objects today in the world-wide web, from Flickr to Youtube. The seriousness of most of these initiatives, however, must be called into question, as anybody and everybody can load contents of their choice onto the net. Some of them, like Wikipedia, have established an internal evaluation procedure in order to brake the tendency of valuable collections of knowledge being reduced to an accumulation of private interests. Others have turned into gigantic garbage dumps of family photos. Neither the permanence of the address nor the long-term life of the data formats is guaranteed. A useful means of accessing the materials is often lacking. It is hardly likely in thirty years time that there will be a RealPlayer or a FlashPlayer with which one can play these proprietary formats. When works released on Flickr suddenly disappear again from one day to the next,

or appear again as works protected by copyright with ban on copying, the reliability of the open licensing cannot be considered to be reliable.

1.4 Against the Loss of History and Memory

The *CultLib* project wants to resist this increasing trend towards the abandonment of history, towards digital amnesia and at the same time meet the challenges posed by today's computer technology.

For this it is primarily dependent on the initiative and support of private individuals, as state institutions up until now have proved to be unsuited for the purpose of attaining this goal.

1.5 *CultLib*, a Project of the Digitale Allmend (Digital Commons)

The Swiss association, Digitale Allmend (Digital Commons) has set itself the goal of promoting the distribution of *Creative Commons* licenses. In an e-mail discussion of the Digitale Allmend on 27 July 2007, one of its members, Philippe Perreaux, suggested that a database should be set up, in which users could enter information about publications with a *Creative Commons* license which are freely available to the general public. This database would be freely accessible on the Internet, so that users would be able to find free works more readily and so that the creators themselves might achieve a greater degree of renown by means of this catalog.

The project *CultLib* is the result of the stimulating discussion with the members of the Digitale Allmend which led to the further development of this idea as exhibited in the project outlined here.

2 A Repository of Free Cultural Objects

In the *CultLib* repository collections of free cultural objects are to receive long-term storage and made accessible to the general public. This chapter will define concretely what is meant by this statement.

2.1 Culture

It can hardly be denied that cultural objects in digital form are rarely collected systematically and made publicly accessible. This is in part due to the fact that culture usually is a non-commissioned, subsidized production with negligible profits. In contrast to software, the sciences or journalism there are no individuals or institutions who are specifically interested commercially in the long-term preservation of culture in digital form.

Accordingly there are comprehensive, well-managed collections of open software and respectable initiatives in regard towards free accessibility in the sciences and useful newspaper archives with free access.

CultLib is not intended as a competitor of these collections in related fields, but rather as a partner interested in cooperation.

Although there are very complex, modern interactive works of medial art, *CultLib* will limit the components of the works in its repository for the present formally to digital

- Texts,
- Images,
- Tone and
- Video sequences.

With texts we mean typographically produced pages, which may contain images, but no recordings or video sequences.

With images we understand picture files of any sort and any resolution, i.e. photographs, scanned documents, image data from scientific instruments such as X-rays or satellite pictures.

With tone sequences we mean tone data of any provenance or quality, whether they be traditional music, bird songs or computer generated sounds.

With video sequences we mean any type of moving images, independent of their screen format and resolution.

A work - like a website for example - can also consist of a inseparable compilation of such data files, which only together represent the work as a whole.

The fact that the great majority of cultural objects exists in one of these forms justifies this provisional restriction. As we assume that it is improbable that today's programs will be runnable in environments such as will be available in thirty years time, we do not see any viable possibility at the moment for long-term preservation of computer games or works of medial art.

2.2 The Concept of a Work

Not every picture uploaded on Flickr is a work which is worthy of long-term preservation. *CultLib* does not intend to serve as a photo album for families nor to be at the disposition of those carrying out criminal or inhuman activities.

Therefore a review process will have to be set up, in which a jury will be given the task of deciding which works will be given a place in the repository. Just as in similar processes of peer review in the sciences and software engineering, this jury should and will be able to consult external experts in the field for this evaluation. In each case care must be taken that the evaluatory decision be made in a context free of religious, political and ideological prejudices.

2.3 Free Access

Current copyright laws make the management of a repository of digital cultural works in the sense of a museum or a library impossible, as soon as it contains works which cannot be made freely accessible to interested users. As long as the film and music industry are putting all of their efforts into making their works unavailable by means of Digital Rights Management, an institution like *CultLib* can only concern itself with accessible works.

CultLib will have its domicile in Switzerland - at least at the beginning - and be subject to the Swiss understanding of international copyright laws.

Therefore those works which are published under a *Creative Commons* license are especially suitable for being accepted in the repository.

The American expert in constitutional law, Lawrence Lessig, became well-known through his work "Code" on questions concerning copyright law and the Internet. In his book *Creative Commons* he transfers the concepts of Public Domain Software to cultural objects. He created the *Creative Commons* licenses which are made use of by artists all over the world in order to make their works available to the general public for free copying.

Many artists who are aware that their creative activities are seriously threatened by the problems associated with the total privatization of all "immaterial property", have made their works available to the general public for free copying.

The international licensing system of the *Creative Commons* (CC) allows the creators also to subject the free works to certain restrictions (no unauthorized commercial utilization, no alterations that were not permitted by the creator, or alterations only if the right of free copying is applied to derived works).

Whereas the CC-Licenses have the advantage that they meet the requirements of national as well as international law, it is intended that all works may be accepted for storage in the *CultLib* repository which may be placed at the disposal of anonymous users for free utilization. Among these are copies of works whose creators have been dead over 70 years; works paid by the (Swiss) government or subsidized works, which according to federal law about public access to government administration are freely accessible; or all works whose legal creators have made them freely available, even if they have not done so formally with a CC-license (URG (Swiss Copyright Law) Art.

10/1): “The creator has the exclusive right of determining whether, when and how the work will be utilized.”).

2.4 Making Content Accessible

Long-term storage demands that the preserved works be processed in order to make them accessible. It is necessary that the users clearly be informed about the context of the creation of the work.

We know from the boulevard press that the meaning of a picture can be completely distorted, if the context is missing in which it was created.

Each file of every work must therefore be supplemented by metadata before it can be accepted in the repository. As long as reasonable, practical metadata standards exist, these should be used for the technical realization of the processing. At the moment the format XMP with the “Dublin Core” for the metadata concerning the contents and EXIF for those concerning the technical ones seem to be the ones that are most suitable.

The metadata concerning the contents therefore will include: identifier, creator, title, subject, description, rights, date, publisher, contributor, format, language, type and coverage.

The metadata concerning the technical aspects will include: resolution, color palette, pictures/second, speed of the photo (sometimes determined automatically by the camera), place of the shot (sometimes automatically determined by the camera by means of GPS), ...

The metadata are to be embedded intrinsically within the files. The repository will set up a database (catalog) based on the the embedded metadata and will make this available to users as search criteria.

Each file is to be given a world-wide unique ID (e.g, GUID). This will make it also possible to identify two copies of the same file in different repositories as identical on the basis of the address.

2.5 Duration of Preservation

In order to fulfill the demands for long-term preservation, the data must normally be processed upon delivery for acceptance in the repository.

Only a few formats for each data type (text, image, tone, video) will be suitable for long-term preservation in the repository . These must fulfill the criteria of free utilization and clear definition. Internationally standardized formats are most suitable for long-term preservation (for example, PDF/A, ODF, OOXML with restrictions, JPEG, PNG, MP3, WAV, MPEG 2, ...). The processing of the incoming data consists of converting it to one of the standard formats of the *CultLib* repository.

A further demand placed on the data formats and their processing is the possibility of embedding XMP meta data.

Finally, part and parcel of long-term preservation is a citable, permanent address at which the work may be found.

2.6 Repository

A necessary last restriction concerns the term repository. It lays emphasis on the storage of the works. Their general accessibility results in the fact that they may be presented in diverse ways to their users.

2.6.1 Access for People and for Programs

In addition to a standard presentation, the “permanent exhibition” of the contents of *CultLib* on the website *www.cultlib.ch* of *CultLib*, the works will be made available for inclusion for other uses, primarily through a service-based architecture by means of a web-service interface. The collections in *CultLib* will be separately introduced and described, for example by the editors or sponsors. These descriptions are naturally free works also which will be stored in the repository.

2.6.2 Basis for Exhibitions of Cultural Agencies

Whereas the contents of *CultLib* will be grouped primarily according to their origins or presented in collections, third-party cultural agents will be able to assemble their own collections as exhibitions and present them on their own websites.

3 Technical Aspects

Most of today's libraries, archives, art collections and museums have only recently turned their attention to the problems of long-term preservation of digital cultural objects as is documented by the above-mentioned international, European and national (Swiss) initiatives.

The greatest challenge for these institutes is dealing with the technical aspects of long-term preservation of digital content. Their employees on all levels, due to their upbringing and training usually have little knowledge of technology, and often have great difficulties coming to terms with the new media. The technical experts whom they call in, on the other hand, have little idea of the demands placed on a larger cultural institutions for long-term preservation and are for the most part too one-sidedly fixed on specific hardware and software solutions.

3.1 File Formats

As already mentioned above, works must be converted into one of the formats suitable for long-term preservation and the embedding of metadata in order to be integrated in *CultLib*.

Only a small number of internationally standardized formats will be allowed here. Should one of these formats become obsolete in the future, then at least a uniform migration into a newer one can be carried out.

For each format a procedure for this processing must be established with which the incoming data files may be converted into the format chosen for preservation and be tested for its conformity. In support of this process *CultLib* will make user-friendly, secure and efficient tools available which will make the entry and alteration of metadata, as well as the format conversion possible.

3.2 Long-term Identification

Whereas a URL (Universal Resource Locator) represents a "location" in the internet, a unequivocal, permanent address is needed for *CultLib*. Although the URL's have recently been renamed URI's (Universal Resource Identifier), they still denote just a "location". This is evident from the fact that it is impossible on the basis of two URI's to determine whether the resources they refer to are identical.

Only with a long-term, permanent identification can resources be citable, independent of changes in providers and alterations in the IP-address or in the Domain Name System.

The relation of resources to one another (i.e., references to a previous version, to the work form which this portion was removed, ...) must be contained in the descriptions, just as this is the case by new editions of a book today.

Only by means of unequivocal identification can works (i.e., descriptions, scientific analyses, reviews, ...) make clear reference to other works.

3.3 Metadata und Catalog

CultLib will fill the database, which in this digital collection will assume the function of a catalog, primarily from the embedded metadata of the works. In addition, the catalog will establish the structure (“Tectonics”) of the repository (Collection - Work - File). As the work will be understood as an independent unit, which also has an existence outside of the collection, its metadata will contain no reference to the collection from which it originated.

The relationship between a collection and a work is primarily realized in the catalog. A description of the collection can, however, refer to the works contained within it. The relationship between a work and its file components is more tight. The metadata of the work must reference all files which comprise it. (A work is, for example, to be realized as a ZIP file with a standardized XMP metadata record.)

Whoever removes a part or parts from a work (for example, pictures from a text) will be legally required by the regulations of *CultLib* to add to the metadata of the component a reference to the original work, or to leave existing metadata unchanged. In cases where this is technically possible, embedded files of a complete work shall be furnished with embedded metadata that refer to the work.

Catalogue inquiries will not only be possible on the *CultLib* website, but will be available also via Web-Service interfaces in standardized form for other websites and platforms.

3.4 Open Systems

A repository of free works which is based on software that is not free would be a contradictory combination and already on the basis of that destined to fail. The software that is used to implement *CultLib* must therefore be capable of fulfilling the requirement that it be based as far as possible on open sources. Where proprietary tools are used, their effects must correspond to clear, testable standards.

4 Organization

A trust, *Pro Cultura Libera*, is to be founded as the organizational core of *CultLib*.

There will then be a trust according to Swiss law under the name of *Pro Cultura Libera* which has the goal of making culture in digital form freely available over the longterm to the general public. The trust will present itself on its website www.cultlib.ch.

This trust is responsible for the management of the repository and the selection of the works. It can authorize external providers with the management and consult with external experts in relation to the selection of works.

Pro Cultura Libera is ideologically, politically, and otherwise neutral. It will take part in the local, national and international exchange of free culture as far as its means enable it to do so. It will endeavor to expand its knowledge about its collections and promote the use of its services. In order to achieve these goals, it will work together with other organizations and institutions both in Switzerland and abroad.

The Project *CultLib* of the association *Digital Allmend* has the realization of the trust as its goal. It is energetically led by

- Hartwig Thomas*, computer programmer, CEO of *Enter AG*, Zurich; Projects in the field of longterm preservation for the Swiss Federal Archives and for the EU Project PLANETS (Preservation and Long-term Access through Networked Services).

- Bruno Jehle*, president of the company Vision Information Transaction AG, Aarau; longterm experience in the field of digital image processing, printing, pre-press and as Internet provider; currently involved with the establishment of the *BJ Institute* in Hyderabad, India and Aarau.

5 Operation and Finances

CultLib, like museums, libraries, art collections, and archives, can only exist by means of contributions from culturally interested sponsors who with these contributions create a monument which will last longer than bronze. Neither those that store works in the repository nor the user of free works should pay for the operation.

The largest portion of the costs will be caused by the evaluation and preparation of the contents. The technical infrastructure is primarily an initial investment which later will no longer be of great significance.

As administrative organizations charge set fees for the copying and use of free culture, they are required (URG Art. 10/1), to repay any monies collected against the will of the creator of free works.

Sponsors can either promote the *CultLib* generally oder a special collection that is to be included in the repository. They will be mentioned in the metadata and in the presentation layer.

The software development of the *CultLib* infrastructure can be used as a technical basis for interested private and public institutions for projects of a similar nature (company archive, private archives, ...). The *CultLib* developers can provide support for such users.

The file format processing for longterm archiving can be offered to interested private and public institutions as a service for projects of as similar nature. The *CultLib* processors can be contracted to offer work or training for such users.

6 Conclusion

The initiative *CultLib* of the association *Digitale Allmend* has the goal of setting up and operating a digital repository of free cultural works.

For this purpose a Swiss trust by the name of *Pro Cultura Libera* shall be set up. This will collect cultural works, which will be available to the general public for free use, and whose intrinsic value makes it worthwhile for them to be available for a long period of time. Such works and collections will be made accessible, catalogued, and stored in appropriate formats for longterm preservation in a digital repository - known as *CultLib* - and published under permanent, citable addresses available to the general public.

The goal of *CultLib* is to work against digital amnesia and the loss of history of future generations which is primarily a result of the turbulent technical development of new forms of information and communication, as well as the modern copyright law. For twenty years now, the latter has left citizens feeling insecure with the total privatization of the ownership of any human expression and increasingly hinders cultural development.

Pro Cultura Libera does not take any position in religious, political or ideological questions. It takes part in the local, national and international exchange of free culture within its means. It endeavors to make its contents known and promotes the use of its services. In order to reach its goal it works together with other organisations and institutions in Switzerland and abroad.

Pro Cultura Libera is promoted by

- *Members of the trust*, for whom the preservation of culture and intellectual exchange is important,
- *Artists*, who make their works available to the general public for free utilization,
- *Contributors*, who finance the inclusion of works and collections in the repository,
- *Curators and scholars*, who with exhibitions and interpretations create a path for the general public through the stored works,
- *Companies and private individuals* interested in long-term digital storage, who make use of the service provided by the *Pro Cultura Libera* specialists in the processing of such material
- *Institutions and organizations* which make use of the technology of *CultLib* and exchange contents with *Pro Cultura Libera*, and
- all *users* of free cultural objects.