DRM and its risks for long-term archiving

Stefan Hein, German National Library, Frankfurt am Main, Germany

Abstract

This paper puts the subject of Digital Rights and Access Management (DRM) into the context of digital long-term preservation. It examines the risks and challenges for ensuring long-term accessibility and usability of DRM-protected objects on the one hand and for the safeguarding of associated rights on the other hand. The research leading to the results presented in this paper has mainly been undertaken within the EU project APARSEN [1].

Motivation

Digital Rights Management (DRM) includes primarily mechanisms to protect rights like copyright and intellectual property rights of producers and authors of digital content. DRM can be found in digital objects like eBooks, music, movies or video games, principally in any kind of publication. From the archival perspective, the preservation of DRM-protected publications is a new challenge, because the protection mechanisms can be included as a part of the digital object itself and are often accompanied by restrictions in accessing and using the content, and reproducing (copying) the underlying bytes.

Problem

The problem in dealing with DRM lies in two aspects. On the one hand an archiving institution has to deal with the preservation of DRM-protected material. For example access-restricted objects must be viewed as being potentially at risk, as the implementation of future preservation measures can be impeded or even prevented entirely by such restrictions. On the other hand the institution needs to take care of the safeguarding of the associated rights, especially when it comes to using the content. This paper gives an introduction into the subject of DRM for archiving institutions and thus aims to raise the awareness about the associated problems.

Terminology

At first it is necessary to explain and define the terminology of digital rights and DRM.

Digital Rights

Digital rights refer to the 'rights' associated with accessing, using, creating and publishing digital content. The rights can relate to usage permissions as well as access preferences or limitations imposed upon digital content. In this respect, the digital content can be regarded as 'protected material', where the protection is on behalf of the 'creator' or 'owner' of the digital content. These rights can relate to copyright legislation, intellectual property rights or contractual agreements imposed on the content.

DRM

The following definition is given by Renato Iannella [2]:

"Digital Rights Management (DRM) involves the description, layering, analysis, valuation, trading and monitoring of the rights over an enterprise's tangible and intangible assets. DRM covers the digital management of rights - be they rights in a physical manifestation of a work (eg a book), or be they rights in a digital manifestation of a work (eg an ebook)". [2]

The present paper does not extend the meaning to the physical manifestations. It focuses on managing the rights of digital content.

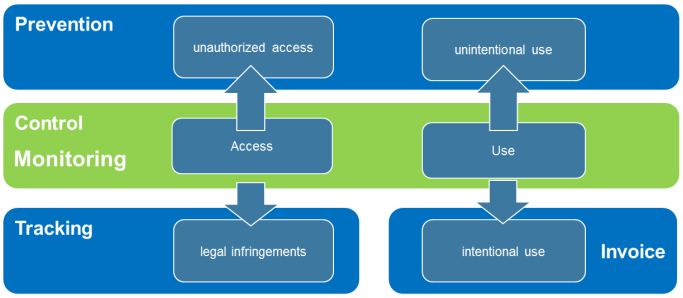

Figure 1. *DRM capabilities*

As illustrated in figure 1, DRM is able to control the access and the use of digital content. If the content owner controls the access, he can prevent unauthorized access and also unintentional use of the content. In the case of Open Access material, DRM could also mean granting access to every person without any limitation of use.

Monitoring features could also be a part of DRM. With these capabilities the content owner is able to track legal infringements - for example in cases of unauthorized access. On the other hand the DRM monitoring technique provides the possibility to invoice the intentional use (e. g. for lending system like online-video rentals).

Approach

The report evaluates the risks that different DRM variants bear for long-term preservation measures. For this purpose, four DRM variants were identified:

1) Data carrier copy protection: With regard to user management, the prevention of copying is a prime example within the context of the entertainment industry. One - albeit unreliable - method is e.g. the deliberate inclusion of errors in the data stream of an audio CD. These errors then prevent conventional CD-ROM drives equipped with error correction systems from reading the data stream, thereby foiling any attempt to copy the music from the CD to another data carrier.

2) Lightweight DRM: For the purposes of this paper, lightweight DRM (LWDRM) refers to all mechanisms which do not of themselves restrict access to digital objects or their use, but which serve the detection and tracking of legal infringements. This is mostly achieved through the use of marking techniques such as digital watermarks. Digital watermarks may be applied to the digital object in a way which is invisible to the user but which allows the content providers to detect their works e.g. on illegal file-sharing sites. In music files, for instance, this additional

information is embedded in the form of slight, audibly imperceptible frequency modifications [3].

3) Encryption-based password protection: This variant focuses on DRM mechanisms which require no connections to external components (such as authentication servers) during use and which basically manage the access and usage possibilities of objects. The term "access" here signifies the opening of a file object using pre-defined player and display software - even though the act of opening could itself be interpreted as the most basic form of use. Use is therefore always conditional upon having access to the object. An example of this is Adobe's PDF format. It contains functions which render access and usage and it is manageable in a variety of forms (like Print, Edit document, Copy content, Extract pages). This kind of limitation of use is one of the most common DRM variants that libraries such as the German National Library face, primarily in the context of online publications (e.g. eBooks) and dissertations.

4) DRM Systems: This DRM category focuses not only on selected aspects already presented above, but also attempts, by means of a system of diverse components and technologies such as the digital watermarks and encryption methods already examined, to cover all the core DRM areas. The architecture of a DRM system - as illustrated in figure 2 - is outlined by Bill Rosenblatt and consists of the three linked components of content server, licence server and client. The different DRM components can be geographically distributed and communicate via the Internet. This results in a range of dependencies which can affect everything from generation and content through to use. The client, e.g. the media player or the document reader, therefore no longer functions independently as a gateway to the actual content. It is apparent that precisely this interaction between the different components markedly increases the complexity of DRM systems in comparison to the DRM variants already presented [4].

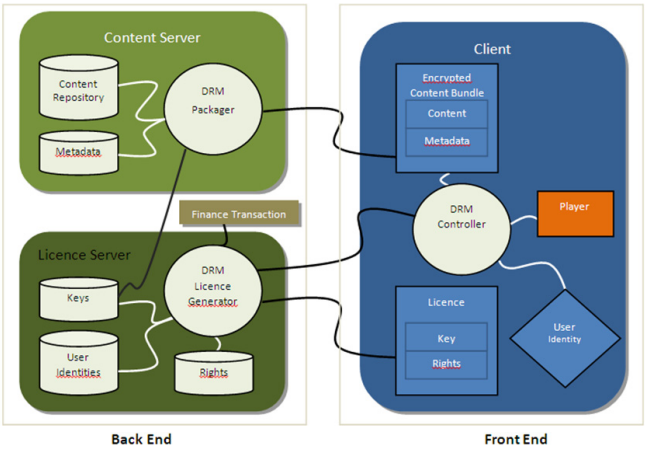

Figure 2. *Architecture of a DRM System (adapted from [4])*

Scale for Long-Term Preservation Risk (LTPR)

To evaluate the risk of different DRM technologies, the following scale (Long-Term-Preservation Risk (LPTR)) is used:

LTPR	Characterization
no risk	No risk for future LTP measures
medium	Possible to use at present (at time of publication) in up-to-date hardware and software environment, current LTP measures restricted, no external dependencies, medium risk for future LTP measures
high	Use and LTP measures already (currently) restricted, high risk for implementation of LTP measures in the future as result of external dependencies

In summary, the higher the LTPR value, the greater the risk in archiving and maintaining the usability of the object concerned. This approach represents an appraisal on the part of the author of this paper that was assessed within the APARSEN project and its work package about DRM [5]. This appraisal also contains a prediction component, meaning that 100% guarantees cannot be offered.

Assessment

In the following, the four DRM variants are evaluated by using the introduced LTPR Scale.

1) Data carrier copy protection, LTPR = medium

Data carrier migration is a key LTP measure, meaning that the prevention of all activities aimed at separating the data stream from the carrier should be regarded as risky. The data carrier copy protection currently prevents copying of, e.g., audio CDs. If the data stream cannot be separated from the data carrier, this carries a high risk for future LTP measures because the necessary players and/or software may no longer be available. Use is, however, possible at present with common player devices (e.g. hi-fi CD players). Based on the principle of "what you can hear/see, you can copy", this permits LTP measures to be performed, albeit with restrictions e.g. in the form of loss of quality (digital-analogue conversion).

2) Lightweight DRM, LTPR = no risk

Lightweight DRM involves no restrictions on access or use; the data stream is therefore accessible and the content usable at all times. The marking of digital objects therefore poses no risk for use or LTP measures.

3) Encryption-based password protection, LTPR = medium

Access to the data stream and use of the content is predicated upon knowing the password. The password must be saved separately and linked to the actual content. The user must be given the password when access is granted. If only limited usage rights, such as text extraction, are granted yet the content can still be displayed, it can no longer be predicted with any certainty whether the conversion tool will require precisely this feature in the future. The execution of current and future LTP measures therefore carries risks.

4) DRM system, LTPR = high

Given that access to and use of the content is restricted similar to the "password protection with encryption" variant, objects protected by DRM systems also carry the same risks. A further problem factor is the existence of an external license server, and connection to it is a precondition for encryption. Even today, use may be impaired or prevented entirely in the event of the content provider going out of business, network problems etc.

Recommendations

The recommendations on the handling of DRM protected material and digital rights are based on the results of two DRM studies and four concrete user scenarios of the National Libraries of Austria, Germany, the Netherlands, and the British Library [5].

Restrictively, it needs to be added that there are only few truly reliable practical experiences beyond prototypical experiments with the execution of preservation actions on DRM protected materials. Because DRM - as part of the content - emerged on the market only a couple of years ago, there was little need to migrate or emulate this content to prevent it from obsolescence. However, the authors of this paper are convinced that the consideration of the compiled recommendations will facilitate the long term preservation of DRM protected materials and the protection of associated rights.

The compiled recommendations are most of all prophylactic in nature. Under "prophylactic measures", we will in the following understand measures that are taken before the actual archiving process, during or at least shortly after the ingest process. The goal of these measures is to recognize potential threats for the execution of future preservation actions early and, if possible, to remove them with current means.

General Recommendations

a) Keep the technical design simple - Keep the variations in type of roles, processes and rights as simple as possible. Don´t give external parties (e.g. publishers and other rights holders´) and internal parties a lot of choices. Select a limited number of variations out of which they can choose one that offers the best fit. This could mean that you will be implementing a variation that offers less than what might be possible in theory. But less makes it more manageable and affordable. Start simple and slowly expand in a controlled manner.

The system should be fully scalable and flexible. This could be achieved through standardization of processes, with all DRM components linking to common data held in centralized repositories and machine-readable databases. This automated system should be balanced against business processes.

b) DRM and Rights Policy – One of the studies that were mentioned above showcased as a best practice the definition of an institutional DRM and Rights Policy. The policy defines how DRM protected materials and their associated rights are treated. The policy should also define which DRM variants or restrictions are accepted or not. Already the process of discussing and defining such a policy creates awareness on all levels and introduces transparency. When published, the Rights Policy establishes confidence for publishers and content creators (rights holders) and can sensitize users to respect the rights of the digital objects that they use.

The Rights Policy should also contain rules for changing and adding usage rights for the purpose of auditing. Usage rights definitions should be simplified and streamlined.

c) Collaboration between rights holders and archives - The DRM and Rights Policy mentioned as measure b) could be negotiated with a publisher or another content creator before they submit their content, if the resources of the preserving institution allow for individual arrangements and the benefits justify the effort. This could, for example, be the case for big publishing houses. If the preserving institution can guarantee appropriate DRM on the objects in their archive, then rights holders will be much more inclined to deposit the digital objects free of DRM. A good example for that is the agreement on digital publications between the National Library of the Netherlands, the Dutch Publishers Association and the International Association of STM Publishers.

In addition to measure b) it could also be helpful to create awareness of the risks of DRM by training the content creators and publishers. This could be done by individual discussions, group seminars, webinars or presentations at relevant conferences or book fairs – perhaps also by referring to this paper.

Recommendations for the handling of DRM-protected objects

d) DRM detection - As a basis for any further treatment, the detection of DRM mechanisms in archival objects is required. Such mechanisms can be detected with manual checks, either of each single object or as sample checks as part of the quality protection. The responsible person can, for example, check if the object can be deployed with the respective viewers / players. Potential access or usage restriction can thereby easily be found. If sample checks are conducted, it must be recognized that a certain amount of DRM protected objects will be ingested. When large volumes are ingested, it is preferable to use automated mass processing applications, i.e., software tools, for these checks. The Open Source *File Information Tool Set (FITS)* [6] deploys a range of recognized analysis tools like JHOVE. These tools provide, at least for common formats like PDF and Microsoft Word, an initial indication if DRM is used. The results of these tools can be used for risk assessment, for example by defining a LTPR or an Ingest Level [7].

e) Measures when DRM is detected - The detection of DRM is only sensible if a pre-defined measure or at least any kind of reaction follows suit. On the basis of the LTPR concepts presented above, the following measures are conceivable:

LTPR = no risk - The data object does not contain any DRM, or, respectively, the contained DRM mechanisms like watermarks do not harm the execution of long term preservation actions. Consequently, the object can be ingested into the long-term archive.

LTPR = medium - A data object with an associated DRM mechanism should not be archived without further analysis. It is recommendable to request a version of the object from the data provider that is free of DRM. If this should not be possible, the conversion into a format or a data carrier that is free of DRM can be considered. If the legal circumstances allow for it, the "digital-to-analogue conversion" could be an option, even if a lossy one. If such measures towards normalization are of greater complexity and require a more thorough preparation, it is recommendable to archive the object, but to record the kind of DRM mechanisms as

part of the technical metadata (see measure f). Because it is possible to use objects with a medium LTPR with current hardware and software, the objects should be normalized as soon as possible after ingest.

LTPR = high - Because objects of this category can even currently only be used with restrictions and will certainly result in restrictions during normalization of preservation actions, these objects should not be archived and DRM free versions should be requested from the content providers.

The Ingest Level Concept that is in use at the German National Library, for example, leads to rejection of all objects with any kind or DRM [7]. It is, however, not always an option to reject DRM protected objects, respectively, to request DRM free versions, especially when the producer cannot be identified anymore. Furthermore, not every content provider is immediately willing to provide its objects without DRM to the preservation institution.

In these cases, it can only be attempted to create awareness for the problem on the side of the producer / content provider. If there is a legal mandate, the preservation institution can use it as an argument. Also the guarantee that the rights will be protected via an institutional access management, so that no disadvantages result from DRM free objects for the content provider, can assist the argumentation. It will, however, imply additional effort for the preservation institution elsewhere, namely in the implementation of such an access management.

If the request for DRM free versions turns out unsuccessful, the measures f) and g) remain.

f) Documentation and Archiving of DRM - If there is no alternative to archiving the object with DRM protection, it is recommended to document it as detailed as possible in the Data Management Functional Entity (see OAIS) at the data level [8]. "As detailed as possible" means to provide all possible details concerning the DRM mechanism used, for example, the kind of usage restrictions.

The documentation in Data Management puts the preservation institution in the position to conduct certain measures later, for example, a normalization or later DRM removal (see measure g). Moreover, the capturing of DRM information in a database enables the creation of a comprehensive statistical basis that allows for reliable statements about the quality of the data holdings and for estimations about the portion of protected objects.

At this point, it would also be conceivable to renounce any further DRM specific measures and to limit the attention to bit stream preservation and to the protection of the DRM mechanism. However, from the point of view of the author of this paper, this is not recommended. Especially on the example of DRM systems, it becomes obvious that the reproduction or emulation of all external dependencies, in particular of the individual backend components of a DRM system, will hardly be possible. Even the option "password protection with encryption" involves the danger that the password is lost sometime in the future. The password needs to be archived and kept accessible together with the preserved content. In the case of copy protection it needs to be taken into account that current hardware that can deal with the protection measures will most likely not be available in the future. So it is highly doubtful if, for example, a copy protected audio CD will be readable in a future device at all, independently of the robustness of the data carrier itself.

g) DRM removal - If the legislation allows it for memory institutions, the removal or bypassing of DRM protective measures during the ingest process could be a feasible step, e.g., as part of the normalization of archival content.

There are, however, a couple of critical points that need attention:

- The technical realization of this strategy needs a thorough examination of each of the data file type dependent DRM protection measures in order to identify or develop suitable tools. Therewith, it is a relatively laborious strategy. Sometimes it needs to be checked if these tools can, under consideration of national or European legislation, be legally acquired and used.
- Moreover, it needs to be clarified if the removal of DRM protection measures constitutes a migration (especially in terms of normalization). In particular, the question rises whether this touches upon the authenticity of the object. Quality checks need to ensure for every scenario that all significant properties are unchanged after the removal of DRM protection.
- The manipulation of content makes checksums unusable. This is critical if these checksums were meant to be used for the assessment of the data's integrity, especially if the author is no longer reachable to confirm the authenticity of the content.
- It needs to be taken into account that the removal of a password encryption is possible only to a limited extent. If the password encryption is robust and the length of the password is sufficient, it is almost impossible to crack a password with a *brute force* attack in justifiable time.
- The tools utilized for DRM removal need maintenance and support and, potentially, additions.
- The removal of DRM protection mechanisms can be CPU intensive and time consuming. Thereby, it influences directly the complete processing time of an object.
- The removal of password encryption does not necessarily create an object that is free of DRM or change its legal status at the same time. If, for example, a PDF document can be accessed after password removal, it needs to be converted into a version that is free of DRM. At the same time the archiving institutions need to take care of the safeguarding of the associated rights which were managed and controlled by the removed DRM mechanisms.

Even if a series of arguments seems to speak against the suggested approach, the APARSEN study has shown that the removal of DRM during migration is already applied for example for video games [5].

h) Analysing the existing data stock - One of the findings of the APARSEN DRM survey is that 60% of the respondents have no concrete plan to analyse their already archived objects for risks that could come up with DRM mechanisms in the future [5]. Unknown or undocumented DRM protection could be a problem which is not solvable. In the worst case, access to the object is forbidden by DRM protection and no one knows how the mechanism works or what requirements are necessary to gain access or to provide specific usage functionalities. Therefore it is advisable to analyse already archived objects or at least their generated technical metadata to detect any restrictions in time. The analysis and the following steps could be supported with the measures that are presented in this section.

Recommendations for the protection of digital rights

i) Detection of Rights Information - The associated rights are by no means always clear or documented with the archival object. In order that digital rights can be protected, they must in cases of doubt need to be detected and documented. As a first approach, preservation institutions can of course contact the content provider. Beyond that, tools like the *Public Domain Calculator* can help to identify rights [9].

j) Documentation and Application of Rights - If the rights are known, it is necessary to document them appropriately. Here, recognized standards like so called Rights Expression Languages (REL) should be used to do this in a digital form. These languages use mostly the commonly known XML standard in the same way as other metadata standards like Dublin Core or the Metadata Encoding and Transmission Standard (METS) of the Library of Congress do. Concrete examples for RELs, which are discussed in [5], are:

- Open Digital Rights Language (ODRL),
- METSRights,
- eXtensible Rights Markup Language (XrML),
- CopyrightMD.

Rights information should be preserved along with the other representation information and the bibliographical information of records. It is also possible to use the PREMIS Data Dictionary to describe rights information or embed other Rights Expression Languages like the ones shown here.

Independently of the selected implementation, it needs to be ensured that the access of archived objects is always organized according to applicable law. If the information needed to ensure this in the access systems originates from Data Management or any other system like the library's catalogue or an own Rights Management solution, does not matter.

The preservation of rights information needs to meet the same standards as the preservation of the archival content itself. That means that if rights information is stored in data bases, future access needs to be ensured, and the data model needs to be interpretable and usable in the future, too.

The system for the management and preservation of rights information needs to account for changes in the rights information. If the rights owner, for example, withdraws some rights that were previously granted, the rights information needs to be updated accordingly. The update needs to be respected, of course, by the access function. It is further important to be able to manage and change rights for a whole set of content, not only one by one. The logging of changes or audit trail should not be forgotten (which employee changed which value). Managing rights metadata should be administered centrally for example by using a Rights Management System. Keep the amount of rights metadata as low as possible to limit the maintenance burden. The high-level principles described in the user scenario of the British Library require that the data should be live, reliable and reusable [5].

k) Inform about digital Rights - The users are not always aware of the opportunity of rights infringements when using digital materials. Beyond the suggested DRM and Rights Policies, it can be helpful to display some information about copyright and the limits of fair use immediately before the user accesses the requested archival information. This will help to raise the user's awareness concerning digital rights.

l) Storage in the Archival Package - In addition to option j) (Documentation and Archiving of DRM), it is conceivable to store rights information within the Archival Package. In Open Source Software products, for example, it is common use already to include usage licences like the GPL [10] as a text file into the software package. It is, however, recommendable not to use this information for evaluation during access and use, but to apply a procedure like it is described in e). The reasons for this become clear quickly, given that the package needs to be downloaded and unpacked each single time before the rights information become visible. To regulate access, however, the rights information is needed before the access package is submitted to the user. If rights information is stored and managed in a database system, it can be controlled and maintained more easily and efficiently and it can also be integrated into access functions more easily. Consequently, the inclusion into the Archival Package is sensible for the case that the system used in scenario j) is damaged or destroyed. Ideally, the rights information that is stored in the Archival Package can then be transferred automatically to the new or repaired system, using standardized RELs like PREMIS or METSRights.

Conclusion

From a memory institution's point of view, a distinction must be made between digital rights and the DRM techniques. For memory institutions, safeguarding the protective rights of their archived assets is essential, and therefore they either fall back to already existing mechanisms, for example their retrieval systems or their own internal rights management system. This approach requires that the digital archive is a durable trusted archive and that the owner of the objects trusts the repository. At this point the funding of the model can be problematic, as the repository has to finance the archive infrastructure, but does not have the authority to provide access. In this situation, public funds have to sustain the archive's infrastructure and the (commercial) publishers can exploit their assets without worrying about the durability of the assets. Even if the latter prefers a type of 'all-in-one' solution in the end, the demand to process, manage, and archive rights and rights information properly within a system will always be present. In order that such a solution can be an "open" DRM system or solution for public institutions, it is important to use openly standardized components and open metadata standards like the ones presented above. It is also essential to invest in training and qualification, because only a skilled and competent personnel is able to operate a DRM system accordingly and take care of the preservation of the content and the associated digital rights.

Through the integration of proprietary rights control mechanisms as an integral component of digital objects, a new problem has arisen regarding long-term archiving. The main cause of this problem has been that access and restrictions of use could hinder the preservation of the object. If access to the content is already blocked, the problems involved in executing LTP measures are clearly apparent. Preservation measures without access to the actual content are not viable. Technical or other types of metadata (e.g. bibliographic) can only – if at all – be extracted to a limited extent from protected files. According to OAIS, however, these data need to be incorporated in the data management and are essential for meaningful preservation planning and the execution of preservation actions [8]. The encrypted content could also conceal malware (viruses, Trojans)

which could enter the archive and remain undiscovered by virus scanners.

Management and control of usage should be analyzed in more detail and should consider questions like the following:

What effect do restrictions have on the duration and frequency of use? It should be apparent that time restrictions are basically impracticable for LTP measures. It is very difficult to define at present when an LTP measure, e.g. format conversion, should be conducted. Restrictions on the frequency of use would equate to restricting the number of uses of LTP measures. The question remains unanswered whether e.g. analysis tools for preparation or post-processing (e.g. quality assurance) constitute an incidence of use and therefore reduce the number of uses.

What happens when usage rights expire later or are withdrawn? DRM gives rights holders the possibility to withdraw usage rights retroactively. Such withdrawal can affect all copies of a work currently in circulation. Naturally, institutions dedicated to safeguarding the cultural heritage will find it difficult to reconcile this situation with their responsibilities.

Considering the composition of usage rights as described by Rosenblatt, and given the uncertain nature of the future, the impression remains that all restrictions imposed upon reproduction rights, transport rights and rights to create derivative works pose risks for long-term preservation [2].

The author of this report hopes that the proposed catalogue of recommendations has provided type of 'first aid' support for this problem to all affected institutions.

References

[1] "About APARSEN", 2014. [Online] Available: http://www.alliancepermanentaccess.org/index.php/aparsen/ [Accessed 13 03 2014].

[2] R. Iannella, "Open Digital Rights Language (ODRL) Version 1.1," 2002. [Online]. Available: http://www.w3.org/TR/odrl/. [Accessed 25 11 2013].

[3] R. Grimm and C. Neubauer, "LWDRM - An Alternative Rights Management System," 2004. [Online]. Available: http://waste.informatik.hu-berlin.de/Grassmuck/drm/Folien-Grimm-Neubauer-eng.pdf. [Accessed 25 11 2013].

[4] B. Rosenblatt, "Enterprise Digital Rights Managament," 14 July 2005. [Online]. Available: http://www.giantstepsmts.com/Authentica-RMS%20Whitepaper.pdf, pg. 5. [Accessed 06 03 2014].

[5] K. Kaur, S. Hein, S. Schrimpf, M. Ras and M. Holzmayer, "Report in DRM preservation", 2014. [Online]. Available: http://www.alliancepermanentaccess.org/wp-content/plugins/download-monitor/download.php?id=D31.1+Report+on+DRM+preservation. [Accessed 06 03 2014].

[6] "File Information Tool Set", 2013. [Online] Available: http://projects.iq.harvard.edu/fits.[Accessed 06 03 2014].

[7] Schmitt, K., & Hein, S., "Risk Management for Digital Long-Term Preservation Services", from IPRES 2013 : proceedings / of the 10th International Conference on Preservation of Digital Objects: http://purl.pt/24107/1/iPres2013_PDF/Risk%20Management%20for%20Digital%20Long-Term%20Preservation%20Services.pdf. [Accessed 06 03 2014].

[8] CCSDS, "Reference Model for an Open Archival Information System (OAIS)," June 2012. [Online]. Available: public.ccsds.org/publications/archive/650x0m2.pdf. [Accessed 26 11 2013].

[9] *public domain calculation*. (n.d), from Europeana Connect: http://outofcopyright.eu/index.html. [Accessed 06 03 2014].

[10] "General Public Licence", 2007. [Online] Available: http://www.gnu.org/licenses/gpl.html .[Accessed 06 03 2014].

The research leading to these results has received funding from the European Community's Seventh Framework Programme FP7/2007-2013 – ICT-2009.4.1: Digital Libraries and Digital Preservation– under grant agreement No 269977.

Author Biography

Stefan Hein is a software developer in the context of processing digital objects and their digital preservation at the German National Library since 2010. He graduated with a diploma in computer science at the Humboldt University at Berlin. The current main focus of his work is the conception and development of workflows around the different import and long-term preservation processes for digital publications.