

Grant Agreement ECP-2007-DILI-527003

ARROW

Correspondence of Arrow infrastructure with Rights Clearance Centres (RCC) and the needs of their users and rightholder requirements

Deliverable number	<i>D3.3.2</i>
Dissemination level	<i>Public</i>
Delivery date	<i>27 September 2010</i>
Status	<i>Final</i>
Author(s)	<i>Federation of European Publishers</i>



eContentplus

This project is funded under the *eContentplus* programme¹, a multiannual Community programme to make digital content in Europe more accessible, usable and exploitable.

¹ OJ L 79, 24.3.2005, p. 1.

Correspondence of Arrow infrastructure with Rights Clearance Centres (RCC) and the needs of their users and rightholder requirements

- EXECUTIVE SUMMARY..... 3**

- 1 INTRODUCTION 4**

- 2 THE HLG KEY PRINCIPLES 4**
 - 2.1 DEFINITION OF RCC 5
 - 2.2 RCC AND THE “DILIGENT SEARCH” FUNCTION..... 5
 - 2.3 RCC AND ORPHAN WORKS 7

- 3 CURRENT EXPERIENCES IN THE FIELD OF RCC..... 7**
 - 3.1 GERMANY 7
 - 3.2 FRANCE 8
 - 3.3 DENMARK 9
 - 3.4 NORWAY..... 9

- 4 ARROW INFRASTRUCTURE AND THE RCCS..... 10**
 - 4.1 GERMANY 10
 - 4.2 FRANCE 12
 - 4.3 DENMARK 13
 - 4.4 NORWAY..... 13
 - 4.5 THE OTHER PILOT COUNTRIES AND POSSIBLE ESTABLISHMENT OF RCCs..... 14
 - 4.5.1 *Spain*..... 14
 - 4.5.2 *United Kingdom*..... 14

- 5 GUIDANCE FOR THE IMPLEMENTATION OF THE “SYSTEM FOR REDIRECTION TO CLEARING MECHANISMS” 15**

- 6 CONCLUSIONS 16**

Executive summary

This deliverable is meant to highlight how the Arrow infrastructure matches the working of rights clearing centres for orphan and out of print works and meets the needs of their users, that is those who seek permission from rightholders or their representatives to use certain works (basically for digitisation and making available), as well as rightholder requirements.

Arrow has been inspired since the beginning by the principles and recommendations of the Copyright Subgroup of the High Level Expert Group on Digital Libraries (HLG); as far as compliance to “emerging” clearance centres is concerned, Arrow also refers to the key principles by the HLG in relation to rights clearance centres (RCC) for both orphan and out of print works.

The Copyright Subgroup of the HLG provided some guidance as to the ideal functioning of RCCs. Arrow, in turn, provides a tool for facilitating diligent search of rights information, rights and rightholders and is therefore a valuable base to enable the creation of RCCs.

The HLG key principles present a definition of RCCs and state their main purposes. Two tasks are crucial: to conduct diligent searches for rights information and rightholders and to grant (or refuse) licenses to users for the works and uses required, or alternatively, redirect to the appropriate rightholder. Though they may be interconnected, the two functions are here analysed as distinct, in order to point out the role of Arrow for the two tasks.

While no such centres have been established yet, Reproduction Rights Organisations (RROs) have been identified as natural candidates to become clearing centres for out of print and orphan works. There are RROs which already include such works in at least some of their licences, for instance, when acting as “one stop shops” for certain secondary institutional uses of portions of copyright works. There are also examples of RROs starting to develop services to assist authors and publishers in granting licences to libraries for the digitisation and making available of out-of-print works, which see RROs as potentially taking up the role of RCCs for orphan and out-of-print works. Compatibility with the HLG principles would normally require that the rights clearance be based on diligent search, for out of print works on voluntary mandating by rightholders and for orphan works on authorisation of the RCC by some legislative measures to grant permissions.

The characteristics of the Arrow infrastructure, especially its role in facilitating diligent searches and building up information on searches and digital availability of works, make it a valuable basis to facilitate the work of RCC. Any clearance centres making use of the Arrow system will be in line with the principles and suggestions of the HLG and any clearance centres built upon the key principles by the HLG could be easily integrated into the Arrow system. Moreover, the very inclusion of RROs in the Arrow workflow constitutes evidence of Arrow’s connection with the already existing clearance centres.

1 Introduction

This deliverable is meant to highlight how the Arrow infrastructure matches the working of emerging clearing centres for orphan and out of print works and meets the needs of their users, that is those who seek permission from rightholders or their representatives to use certain works (basically for digitisation and making available online).

The principles and recommendations of the Copyright Subgroup of the High Level Expert Group on Digital Libraries (HLG) provide some guidance as to key principles for the functioning of rights clearance centres (RCC) for both orphan and out-of-print works. Arrow, in turn, provides a tool for facilitating diligent search of rights information, right status and rightholders as well as on information on orphan works and rights clearance and is therefore a valuable tool for existing RCCs and for the creation of new ones. While no such centres, i.e. with the task of addressing specifically and chiefly the clearance of rights to use orphan and or out-of-print works on a large scale, have been established yet, Reproduction Rights Organisations (RROs) have been identified as natural candidates to become clearing centres for out of print (provided there is a demand, on a voluntary basis, with the agreement of rightholders) and orphan works.

The deliverable therefore elaborates a definition of RCC based on the principles of the HLG (§ 2) and analyses the current tentative experiences in some European countries (§ 3). Subsequently, it demonstrates the correspondence of the Arrow infrastructure with the principles and suggestions of the HLG through a practical analysis on how Arrow is going to work in the countries involved; the very inclusion of RROs in the Arrow workflow shows Arrow's connection with the already existing clearance centres (§ 4).

Finally, the deliverable provides practical guidance for the implementation of the part of the Arrow infrastructure defined as "System for redirection to clearing mechanisms" (§ 5).

2 The HLG key principles

The relevance of the principles to the Arrow project has already been illustrated in Deliverable 3.3.1 on the clearance mechanisms for out of print works.¹ The key principles are part of the annexes to the Final Report on Digital Preservation, Orphan Works, and Out-of-Print Works by the Copyright Subgroup of the HLG.² There are two sets of principles, one for orphan works and one for out of print works; both outline some principles on the clearance of rights and the purpose of rights clearing centres, and provide definitions accordingly.

¹ D3.3.1. *Guidelines for clearance mechanisms for out of print works* is available for downloading in the Resources area of the Arrow website (www.arrow-net.eu).

² The Report can be found at: http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg/reports/copyright/copyright_subgroup_final_report_26508-clean171.pdf. The annexes are Annex 6: Recommended Key Principles for Rights Clearance Centres and Databases for Orphan Works (http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg/reports/copyright/copyright_subgroup_final_report_26508-annex6-ow-270508.pdf) and Annex 7: Recommended Key Principles for Rights Clearance Centres and Databases for Out of Print Works (http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg/reports/copyright/copyright_subgroup_final_report_26508-annex7-opw-270508.pdf). The documents are also accessible through the Arrow website, at <http://www.arrow-net.eu/resources/useful-information.html>.

2.1 Definition of RCC

The Final HLG Report generically suggests the creation of RCC as possible part of the solution of the orphan works issue. As to out of print works, the Report recommends *“that each Member State considers encouraging the establishment of national Rights Clearance Centres (RCC). These clearance centres could act as national portals for clearance of rights in respect of out-of-print works unless the proposed user finds it simpler to contact the rightholder directly. Existing CMOs such as RROs could run the portals. The rightholders may also opt for other solutions. Subject to the mandate from the rightholder, the RCC may:*

- a. Grant the permission and offer a licensing agreement;*
- b. Redirect the request to the pertinent rightholder;*
- c. Refuse permission (which may e.g. be the case if the CMO does not have the mandate to grant the permission).*

*The RCC will not encompass all rightholders and all works. It would, however, be expected to represent a substantial portion of them.”*³

Key elements of the RCC tasks, for our purposes, are:

- the identification of RCCs as *“portals”* that users can use to solicit a licence;
- such *“portals”* are envisaged as *“national”*: there are in fact strong reasons to foresee *“national”* development of RCCs; however, *“The national rights clearance centres are foreseen to be interoperable to allow a European-wide solution” (Ibid)*;
- RROs are mentioned as natural candidate to be the RCC in this field;
- possible answers that users can receive are clearly pre-figured out: a RCC can (a) offer the license; (b) re-direct the users to the rightholder; (c) refuse the license.

From the Arrow viewpoint this last element is important since it defines a possible technical specification of the system: the Arrow will be compliant with the HLG principle on this matter if it enables any RCC to communicate any and all those possible answers to users. In the current situation, with RROs as natural candidates for the RCC role, Arrow’s compliance with the HLG principles is largely facilitated, as national RROs are fully integrated in the system workflow providing the user with a set of possible pre-determined answers in line with what the HLG recommends (see the cases presented in this paper, § 4).

The HLG definition also suggests that one of the ways to authorise digitisation and making available of works could be *“Collectively via Rights Clearance Centre administered by a CMO such as a RRO. Depending on the mandate, the license offered by the CMO may either be offered on a transactional basis (i.e. case by case) or offered as a repertoire licence. A “repertoire licence” means that the library/archive through the licence is (a) granted preauthorisation to digitise and make available the works that the Rights Clearance Centre has in its repertoire (b) normally at a standardised set of conditions.”*⁴

In terms of the Arrow specification, this means that the system should enable the handling of information both at title by title level and at repertoire level.

2.2 RCC and the “diligent search” function

In the HLG key principles (set out in the annexes to the Final Report), the RCC tasks are defined also with respect to *“diligent search”*. Two sets of principles have been elaborated, for orphan works and for

³ See note 2.

out of print works, which have some differences related to the procedures recommended to deal with those kinds of works. According to the principles on orphan works, RCCs are seen as part of a network of resources that “*shall, inter alia (...)*

1. enable users, institutional as well as commercial ones, to conduct an online search, in order to discover whether:

- a. the identity and/or the location of a rightholder to a given work required by the user is/are known;*
- b. information is available on previous diligent search on the rightholder to the work to enable further search or engage in a diligent search for the rightholder(s) to the work;*
- c. the work is already registered as available digitally, e.g. from another library/archive/ museum or other user who has already sought and obtained clearance through a default licence or, when applicable, on some mechanism based on legislation.*

2. enable users, institutional as well as commercial ones, to proceed from discovery of digitisation status to:

- a. digitise the work, subject to the relevant terms and conditions established, and, when applicable, subsequently making the digitisation accessible to users on the conditions established in the licensing agreement; or*
- b. enable access to a work that has been digitised and made available digitally from another.”*

For out of print works the envisaged network of resources should:

“1. enable cultural institutions such as libraries, archives and museums to conduct an online search in order to discover whether:

- a. the work has been declared out-of-print and is already registered as available digitally, e.g. from another library which has already sought and obtained permission to digitise the work;*
- b. the work has been declared out-of-print and no permission has been granted to make it available online*

2. enable cultural institutions such as libraries, archives and museums to proceed from discovery of digitisation status to:

- a. seek permission to digitise the work, either directly from the rightholder(s) or through the procedures established by the centralised Rights Clearance Centre (RCC);*
- b. digitise the work subject to the relevant terms and conditions established, and, when applicable, subsequently make the digitisation accessible to its users on the conditions established in the licensing agreement; or*
- c. seek access to a work that has been digitised and made available digitally from another source, e.g. another national or foreign library, archive or museum, or another authorised user.”*

It should be clear that there are two distinct functions: *right clearance* and *diligent search*. The HLG report suggests that there is a link between the two, and that the RCC – whose nature is primary related to right clearance – could also have a role in the management of rights information, and thus in diligent search.

The Arrow system deals with right information only, not right clearance. However, one of the key pieces of right information that are managed is about the identification of the entity entitled to clear rights, for a particular title and for a particular use (typically, digitising and making available on the internet). It thus purports to provide user information on where rights could be cleared and interoperability. The Arrow system is also conceived to “redirect” the user to this entity (e.g. an RRO), so as to solicit and potentially obtain the license desired, including when the entity is not already directly integrated in the system workflow.

⁴ See note 2.

The Arrow system being specifically designed to deal with the information management aspect, it can very well complement the work of a RCC, whose primary task is the clearance of rights. Arrow can support RCCs in accordance with the HLG sets of principles, also taking into account their differences, as illustrated in relation with the cases presented in this paper (§ 4).

2.3 RCC and Orphan works

The HLG also sets out a number of principles upon which the working of RCC should be based. With regard to orphan works, the RCC is foreseen to be an authoritative source for the clearance of rights to digitise and/or make available orphan works in accordance with the national legal regime.

Also in this case, the HLG guidelines provide that the RCC may operate either as a pure clearing centre of orphan works that are determined outside its sphere, or be also engaged in the diligent search for the rightholder. The RCC shall clarify whether it is tasked with performing both tasks and, if not, which of these tasks it performs.

Finally, in the case of orphan works, the Arrow system is designed to support any known operational model for RCCs, being able to manage different “roles” independently from the entity playing each role (rights clearance, diligent search or both).⁵

3 Current experiences in the field of RCC

A few projects exist and others are being planned which envisage a role for RROs in handling authorisations to digitise and make available works, both orphan and out of print.⁶

3.1 Germany

In Germany, the Börsenverein (German publishers and booksellers association), the German National Library (DNB) and VG Wort (German RRO for text works) have agreed on a scheme for the digitisation of books in the DNB’s stocks which are out of print (i.e. no editions of the work are actively commercialised in any form by the rightholder) and have been published before a cut-off date (1965 in the first implementation) . The aim is to enable the DNB to begin digitising its holdings, starting with older works, with a view to including them in the German Digital Library. All institutions in Germany which wish to digitise those works and make them available on a non-commercial basis will be able to use this scheme for works:

- i. which fall in the category defined in the agreement, so out of print at work level and published before the defined date;
- ii. which the rightholder(s) concerned did not withdraw.

According to the outline of the initiative, VG Wort will act as the RCC, thus confirming the forecast that RROs are the natural candidate to be RCCs. It will in fact provide a centralised point, a “one stop shop” for handling authorisations for digitisation and making available, based on a broad agreement (not compulsory) in line with the national legislation.

⁵ For more information, see D6.2 *Registry of Orphan Works management system*, which will be available in the Resources area of the Arrow website (www.arrow-net.eu).

⁶ A short description is given here; more details can be found in deliverable D3.5 *Report on legal framework – Second Edition*, available for downloading in the Resources area of the Arrow website (www.arrow-net.eu).

The use of Arrow⁷ will allow the conduct of an online search to clarify the status of a work (in particular to check the out of print status at work level); it will also provide information - if available - on previous diligent searches and the possible availability of a digital version of the work.

In terms of the two tasks identified above, the search here is simplified for books published before the agreed cut-off date, since it is sufficient to determine the out of print status of the work concerned. However, the management of right information is much less trivial than it appears at first glance, for two reasons:

- The determination of out of print status at work level implies the capacity to identify all the editions of the same work that have been published over the time, which is one of the key functions of the Arrow "Rights Information Infrastructure".
- The management of the system for withdrawal of works or rightholders also requires management of data at work level. It is important to be sure, in fact, that when a rightholder withdraws one or more works, all the editions of that work(s) are correctly withdrawn in the system.

3.2 France

Another experiment is foreseen in France, although still at an earlier stage.

Orphan works - The High Copyright Council (Conseil de la Propriété Littéraire et Artistique - CSPLA, which includes relevant cultural sectors, users and public administrations) officially recommended a modification of the French intellectual property law in order to establish compulsory collective management for orphan works in the written and visual art sectors, in particular for their digitisation and making available online.⁸

This mechanism should allow Collective Management Organisations (CMO) accredited by the Ministry of Culture to provide for relevant authorisations, while granting legal security to users. The accreditation by the Ministry would be conditioned by usual criteria and by the participation to a common portal aiming at facilitating access and updating information on orphan works. The criteria to define orphan works and therefore the requirements to meet for a "proved and serious" search could be fixed by a committee with an equal representation of rightholders, users and the government.

A working group of the text and image based sector led by the French RRO CFC (Centre français d'exploitation du droit de copie) explored since 2007 concrete solutions to support searches. It suggested the following two-level scheme:

- As a first step, the user would have to perform elementary searches to limit the number of orphan works involved by its request. Arrow would clearly fit at this point, helping to make the search.
- For the remaining works whose rightholders would remain unfound, the metadata should be published on the internet on a dedicated website for a period of time to allow the rightholders to claim their rights. This phase can be managed using the tools developed within Arrow for the Registry of Orphan Works.⁹
- If the work is not claimed, then the user would receive a non-exclusive permission limited in time and with the possibility for it to be renewed to digitise and to give online access to the work, against

⁷ See also below, for more details on the correspondence of Arrow with the principles on RCC.

⁸ Opinion of the special CSPLA Committee on orphan works adopted on April 10, 2008, available at <http://www.cspla.culture.gouv.fr/CONTENU/avisoo08.pdf>.

⁹ See note 5.

a remuneration to be negotiated with the CMO and to be based on the compensation for the digitisation and for the effective use of the work.

- The CMO would therefore only endorse the research undertaken by the user by checking that it is meeting the official criteria fixed by the committee including rightholders, users and the government.
- The user would provide the CMO with declarations of consultations (title by title) in order to determine whether complementary searches would need to be undertaken, as a second step, for the most consulted works.

In this case, the diligent search would be under the responsibility of the user, while the RCC role is limited to the clearance of rights for the digitisation and making available orphan works. Despite the consensus among the members of the CSPLA since 2008, the orphan works bill has not been prepared yet, due to the current discussions on out-of-print works, which may require the coordination of both systems.

Out of print works - The French Ministry of Culture and French publishers and authors are currently considering the possibility of an agreement on out-of-print works from the 20th century. This project proposed by the Ministry would be based on four fundamental principles:

- The digitisation of all the collections by the French National Library at the expense of the State.
- An overall agreement on a massive set of works, going beyond the negotiation on a work by work basis, but with flexible mechanisms to include and withdraw works.
- A secure agreement legally binding the three parties (Ministry/ publishers/ authors)
- A model of dissemination and commercial exploitation of works with mechanisms for income distribution to be defined.

The plans on out of print works are still too vague to allow indepth analysing but they could nonetheless open ways for the establishment of RCC and the use of Arrow's services. In particular, the same elements seen for the German case is likely to be relevant: by definition an agreement on out of print works requires the determination of the out of print status at work level and the management of withdrawal and inclusion of works and/or rightholders also requires a system for dealing with right information.

3.3 Denmark

In Denmark, the copyright act was amended recently to make it possible for institutions and rights holders to enter agreements about ie. mass digitization based on a collective licence model,

An agreement based on the new provision has to be approved by the Ministry of Culture; Copydan, as the Danish institution that can enter collective licence agreements on behalf of Danish rightholders, would be the likeliest candidate as a RCC. In any case, rightholders always have the right to prohibit the use of their works. However, for the time being, due to the financial situation in Denmark mass digitisation projects are not given any governmental priority.

3.4 Norway

The current experience in Norway involves as well a RCC, but it is to some extent of a different nature from the others mentioned. In May 2009, Kopinor (the Norwegian RRO) and the National Library of Norway signed an agreement regarding a pilot project on the digitisation and making available of books. The Bokhylla ("Bookshelf") project allows the National Library to make books published in Norway in the 1790ies, 1890ies and 1990ies – approximately 50,000 books protected by copyright – available on the Internet under an Extended Collective Licence with a right for rightholders to withdraw

works. Most titles published in the 1990ies and some titles published in the 1890ies are assumed to be in copyright. The books that are made accessible under this project cannot be printed or downloaded but are displayed page by page in j-peg format for the reading on the screen and only to Norwegian IP-addresses.

This initiative, based on ECL, has a number of important limitations of the permitted uses of digitised in-copyright works: the making available is limited to books published in Norway, and downloads and printing are not allowed. The rightholders are offered a right to withdraw works at any stage. From the Arrow viewpoint, it is relevant that no diligent search is envisaged beyond identifying the titles that belong to the decades covered by the project and those which the rightholders have withdrawn.

4 Arrow infrastructure and the RCCs

Arrow can be a valuable instrument in the work of a RCC, by facilitating diligent search for rights information and rightholders and making it also easier to obtain information on the availability of licences and thus to clear rights for the use of works.

Considering the principles and suggestions by the Copyright Subgroup of the HLG, we can assess the value of Arrow in respect of RCC through considering the two functions highlighted above: the *diligent search* and the *rights clearance*.

The compliance of the Arrow infrastructure (the RII - Rights Information Infrastructure) with the HLG guidelines has been already analysed in the Deliverable 3.2.2.¹⁰ It would be appropriate to expand the analysis to examine further the role of right clearance in relation to the granting of licences for digitising and making available written works.

As far as RROs act as the RCC, as anticipated by the HLG and confirmed by the early experiences described above, the first element to be emphasised is that the RRO should be prepared to integrate its database in the Arrow workflow (some already are) as one of the sources to be queried to provide users with information. Arrow deals with such information, and it is technically irrelevant for the system if the RRO communicate that the RCC able to provide a licence is the RRO itself or any other entity. This means that the neutrality of Arrow is confirmed also with respect to the range of tasks that any particular RCC would assume and thus in respect to the level of engagement of the RCC in the diligent search.

To illustrate the possible role of Arrow, it is useful to analyse it for the different cases. In this analysis we also considered countries where the discussion about the RCCs in this domain is less developed than in the cases illustrated in the previous chapter.

4.1 Germany

The role of Arrow in the emerging scheme is to facilitate the identification of the status of a work as being out-of-print and the management of the withdrawal of works by individual rightholders.

The piloting of the Arrow system in Germany started up before the agreement on out-of-print works with a cut-off date; however, the current workflow already provides the relevant information necessary for the decision making within the framework of the agreement.

¹⁰ *Evaluation of compliance of the ARROW workflow with the agreed HLG guidelines on diligent search*, available at www.arrow-net.eu.

The first two steps of the Arrow workflow (querying TEL index and Books in Print) are targeted to:

- i. precisely identifying the book concerned starting from the metadata record provided by the requesting library;
- ii. clustering all editions of the same work published over time;
- iii. determining the in print / out of print status of the work.

In the current pilot, this information is used to start a third phase, namely to look for rightholders within the VG Wort database, and possibly providing information about the licences that VG Wort is able to grant to the requesting library. VG Wort provides such information by extracting it from its IT system, and sending it to Arrow in a standardised message form. VG Wort is currently answering following the general scheme described in the HLG guidelines, and thus:

- a. In a number of cases, VG Wort will provide information through Arrow whether it is authorised to grant a licence. This is limited at the stage, since the actual procedure for collecting ad hoc mandates has not been finalised.
- b. In other cases VG Wort will inform the user that it may contact the appropriate rightholder so as to ask for the requested authorisation.
- c. In a third set of cases it will refuse the licence. This may be for various reasons, which have been defined within the work of message standardisation, and include the cases of lack of capacity to locate the rightholder, positive decision of rightholder not to grant that particular use, etc.

The structure corresponds thus to the “specification” provided by the HLG (see § 1), of which it is a concrete implementation.

Because of this structure, as far as the Arrow functioning is concerned, the implementation of the German simplified out-of-print works agreement would imply the following procedures:

- when the date of publication is *posterior* to the defined cut-off date, the above described procedure will remain the same;
- when the date of publication is *anterior* to the defined cut off date, VG Wort will:
 - check – using the information that Arrow collected in the first phases – whether the work is in print; if the case, it refuses the licence (unless the rightholder authorised it independently from the general agreement);
 - if the work is out of print, check if any of the records in the cluster of editions that has been provided by Arrow for that book has been subject to withdrawal by the rightholder; if the case, it refuses the licence;
 - if none of the two cases above occurs, VG Wort will grant the licence according to the general conditions of the agreement.

This information can already be managed by the Arrow system, so that VG Wort can easily provide the answers according to the described decision tree as described above.

The value provided by Arrow is in the supply of clustering coming from the combination of information resident in library catalogues, Books in Print databases and a standardised, consistent methodology to integrate this information in the RRO database. This will allow to avoid two problems that are typical in these cases:

- To wrongly consider a book out of print because many editions of that work are in fact out of print but others are in print that are not correctly clustered with the previous. It is to be noted that this type of error is more frequent for old books, which may have had several editions over time while

usually only one, the latest, is still commercialised. So, it is sufficient to miss one out of several editions to cause the error.

- To wrongly match the request coming from the library and the withdrawal of works coming from the rightholder. The two sides of the chain use different metadata records to interact with the system, and the matching may be difficult. Typically, the library uses its metadata record that was formed many years ago, possibly employing methodologies that are not in place now, and that may contain errors of different types. The rightholder may compile a form at the moment of the claiming, or – if it is a publisher – provide an ONIX record referred to a number of books. The Arrow process is designed to maximise matching efficiency between records with different origins, written in different formats. Furthermore, as for the determination of the out of print status, the clustering function is also important to be sure that if a rightholder withdraws a certain edition of a work, the same work is not digitised when a library asks for permission referring to another edition¹¹.

The capacity to identify works (and not just books) already digitised by other libraries is a long term value that is to be considered. When used for multiple digitisation programs, the system provides valuable information to avoid duplications and to reduce costs.

4.2 France

The situation in France is similar to the one in Germany: the Arrow pilot is designed independently of the implementation of the current discussion about orphan works and out of print works solutions, since the elements of the possible legislation and voluntary agreements have not been fixed yet (and a technical implementation requires precise terms to fix specifications); however, the implementation is scalable towards the envisaged solutions as known to the Arrow national group, which however includes all of the organisations participating in the discussion.

As for out of print works, the model for interaction could be similar to the one in Germany: as far as any standard licence applies to certain category of books, the Arrow system can be used to check if a particular title belongs to that category. Furthermore, in analogy with the German case, Arrow can be used to facilitate the management of works that may have been withdrawn (and, if provided, included) by rightholders. The current structure of the system supports the relevant solutions allowing the RCC to receive requests from users and to provide answers in a standardised way.

Concerning orphan works, one crucial difference between the French and German potential models seems to be in the role of the RCC in the diligent search function. According to the discussions so far, the role of CFC in determining whether a work is orphan could be limited to checking whether the user followed the required procedure. However, when the user employs Arrow to carry out the search, CFC will also be one of the sources of information for the search, its rightholder and work database being integrated in the Arrow workflow.

One aspect to be noted for the potential use of Arrow in France is in the last phase of the possible process for managing orphan works. As said above (§ 3.2) the discussions on orphan works have pointed to the need of publishing in an internet portal metadata of works whose rightholders remain unlocated, which could constitute the basis of a local Registry of orphan works (ROW)¹². This function could be supported by the technology under development in Arrow for the “ROW management

¹¹ Both types of errors were very frequent for old books in the Google Settlement database when analysed (August 2009) for benchmarking. Example of the analysis is in the AIE study on Italian best sellers authors of XX century, where errors of one of the two cases were found in 81% of the works considered. See AIE, *Analysis of commercial availability of important Italian authors' works in the Google Settlement Database*, http://dx.medra.org/10.1389/090907_aie_hearing.

¹² See note 5.

system”, which is designed precisely to meet this type of requirement, including the possibility of management of a national ROW in an independent way while maintaining full interoperability with the other national ROWs using the same infrastructure.

The value offered by the Arrow ROW technology is in the immediate integration of the ROW with the Arrow right information infrastructure (RII). When saying that “metadata about a work” should be published on the internet for a period of time to allow rightholders to claim their rights, there are two questions to be answered: “What metadata?” and “Metadata referred to what?”. Both have been at the very heart of the Arrow work since the beginning. The Arrow ROW model is designed to handle:

- Metadata records properly referring to works (so not to the individual books that contain a certain work), expressed in a standardised way. Furthermore, every work is planned to be identified using a standard numbering scheme (the ISTC).
- Metadata records to all the manifestations (books) including the same work. These will be also stored and showed to the public, so to allow rightholders to claim at the same time all books concerned and thus minimise the risk of having the system consider by mistake as orphan a book that has been claimed in another edition by the rightholder.

The use of standard tools to identify works also facilitates the management of relations between translations of the same work¹³. This is not fully developed in the first release of the system, because of limitations of the sources of information currently available, but the system is scalable to enable such management. At that point, the rightholder will be enabled to claim different editions of his/her works published in different languages, without searching multiple databases.

4.3 Denmark

The case of Denmark could under certain circumstances be similar in its key elements to the German example, so that the value that Arrow can offer could also be very similar. Denmark is not part of the current Arrow piloting; however, Copydan, the Danish RRO, is an associate partner/external supporter of the project. Meetings between some national stakeholders and Arrow representatives have been held at different levels in order to examine the potential contribution from Arrow to the implementation of national digitisation projects. While Denmark is following the Arrow project with interest, its value in the collective licence model has to be further discussed.

4.4 Norway

The Norwegian Bookshelf pilot comprises rights clearance under an Extended Collective Licence (ECL), thus encompassing orphan and out-of-print works. As determined by the rightholders search has been confined to identifying works published in the decades comprised by the project and rightholders who did not want to be included and/or works which have been withdrawn.

Prior diligent search is not a prerequisite for rights administration under this scheme neither for actual search of individual rightholders nor for the more limited purpose of identifying the category of works

¹³ The term work is used here as in the ISTC standard, to refer to “a distinct, abstract intellectual or artistic creation predominantly comprising a combination of words, whose existence is revealed (i.e. “published”) or intended to be revealed, through one or more textual manifestations” (see the ISTC User Manual, at http://www.istc-international.org/html/multimedia/pdfs/ISTC_User_Manual_2010v1.2.pdf). The clustering algorithms used in Arrow allow grouping editions belonging to the same “principal work”, i.e. including also editions that are not part of the same work, but are linked to it. For example, a new edition, or an illustrated edition of the same work is grouped in the Arrow cluster, though the difference between the two cases is tracked. However, since the algorithm operates on metadata records that are expressed in a certain language, it is very difficult – at the current stage – that a translation of a book is properly clustered together with the original version, unless the library catalogue included the original title as an instance within the MARC record for the translated edition.

as in the previous cases; the project in fact includes also works commercially available, as long as published in the decades involved. As a consequence, Arrow has not been used. Some potential use can be envisaged for the administration of rights, for instance to identify works that have been withdrawn. However, at the present conditions, Arrow provides very limited value to the management of such a scheme, though stakeholders are considering possible uses in case of any evolution of the system.

4.5 The other pilot countries and possible establishment of RCCs

The Arrow system is currently piloted in two more countries: Spain and the UK. In those countries the model for the creation of RCCs for the management of orphan works and/or out of print works is still under construction. In the Arrow pilot, the issue is being discussed along the lines of the two other pilot countries, France and Germany.

In the combination between the two functions of “diligent search” and “rights clearance”, and in the absence of any simplified scheme, the RROs in the respective countries tend to work towards covering both.

4.5.1 Spain

The Spanish RRO, Cedro (Centro Español de Derechos Reprográficos), which is fully involved in the Arrow pilot as an associate project partner, recently launched a service for searching rightholders of specific works¹⁴. When a library queries the Arrow system, it is possible that Cedro is already able to provide a final answer, since it already knows the right status and rightholder(s) of the work and either is able to grant the authorisation to scan and make available the book or is able to refuse it, according to the rightholders’ will.

However, it often happens, at the present stage, that Cedro does not hold all the necessary information to provide a final answer, either because it knows the rightholder but has not received any mandate and thus need to ask for it, or because it does not know the rightholder yet. In these cases, a further action by Cedro is required, and Cedro offers the user to exploit the new service, in particular for the second case. The Arrow workflow has been implemented to allow Cedro to communicate to the users this opportunity within the normal Arrow process.

In the long term, the Arrow structure is designed to support the Cedro functioning as a RCC for orphan and out-of-print works as well as any other solution that stakeholders decided to implement.

4.5.2 United Kingdom

Currently, also in the UK pilot, when a library asks for information about rights status and rightholders for a particular work, and this information is not held in the RRO repertoire¹⁵, a further search might be conducted by CLA (Copyright Licensing Agency), the UK RRO, in collaboration with the Publishers Licensing Society (PLS) and/or the Authors Licensing and Collecting Society (ALCS) and then stored in the system for reuse.

Again, in a case where the value provided to the user depends on the capacity of delivering a positive result, the “diligent search” function is seen in conjunction with providing rights clearance, and thus is

¹⁴ See www.conlicencia.com/Conlicencia/OtrosServicios/LocalizacionTitulares.html, where the service named “Localización de Titulares” is described.

¹⁵ The peculiarity of the British situation is that this function involves a network including the RRO in proper sense (CLA) and the two entities that represent publishers (PLS) and authors (ALCS). The Arrow pilot made it possible to create an interoperable environment among the three so that from the user viewpoint the answer comes from one single point.

covered by the same entity(ies) that are expected to grant the license. How this may be developed in the future is difficult to say, as it depends on the evolution of the legal framework and stakeholders' agreements.

5 Guidance for the implementation of the “System for redirection to clearing mechanisms”

The two functions discussed in this document in relation to RCC – the diligent search and right clearance – were conceived in the project proposal as being separate tasks, and thus it was planned to create a “redirection” function to bridge the gap between the two.

The analysis here conducted shows that in practice the two functions are integrated, and the development of the Arrow system has taken into consideration this reality. Therefore, the management of the information about the licences that users can acquire is integrated in the Arrow workflow.

As a conclusion, users will normally be directed to the RCC for rights clearance; no further redirection has been built in at this stage but will be considered if it is later discovered that this is necessary. When the RCC makes use of Arrow the rights clearance process will be integrated with the Arrow workflow, at least when the RCC role is carried out by a RRO whose repertoire databases are integrated in the Arrow system.

Also when the RCC is not a RRO the Arrow system is applicable as it is designed to manage “roles” and “functions” independently from the organisations that are authorised and/or mandated to function as a RCC. It might be possible that one single organisation plays two or more roles (for example, a RRO involved both in the search for rights, rightholders and rights status as well as in the right clearance) or, alternatively, that each of the functions is managed by a distinct organisation. The system allows to integrate any emerging RCC within the existing Arrow workflow without requiring additional implementations.

Following the HLG principles, RCC are meant to enable to conduct an online search to find out the right status of a work, identify and/or locate rightholders and find information on previous similar searches and the possible registration of a work as already digitally available. To facilitate the management of rights information is actually the primary goal of Arrow, which includes ascertaining a work's right status and possibly identifying and locating rightholders. Arrow will also retain information about the searches performed, thus satisfying the other criteria; the creation of a Registry of Orphan Works (ROW) will further contribute with respect to this.

RCC are also supposed to enable cultural and other institutions to seek permission to digitise or make available a work (either directly from rightholders or through a RCC), digitise the work subject to the relevant terms and conditions and seek access to a work that has already been digitised. The Arrow system foresees the possibility for users to obtain direct information about the availability of licences, or even to be proposed a licence, in both cases by an RRO. And there is always the possibility to be redirected toward an individual rightholder for negotiation. The recording of past searches will help identifying works that have already been digitised and the relative conditions.

Not only is the Arrow system, then, practically fit for the emerging and potential clearing centres; it is also fully compatible with the principles of the HLG and can be considered in general as an adequate support for the creation and functioning of RCC according to various possible schemes.

6 Conclusions

As this analysis has shown, the Arrow infrastructure complies with the principles and recommendations of the HLG Report regarding RCCs and is compatible with the current implementations of these recommendations. It can therefore be a valuable support for the carrying out of the tasks of such centres and possibly the establishment of new ones, both because of the way it will facilitate and improve the diligent search for rights information and rightholders and because of the records it will keep of the searches performed and of their results (including the creation of a Registry of Orphan Works).

The current experiences confirm the role of RROs as natural candidates for the task, and demonstrate as well compatibility with a system like Arrow, since RRO information is integrated within the Arrow workflow.