

Grant Agreement ECP-2007-DILI-527003

ARROW

**Evaluation of compliance of the ARROW
workflow with the agreed HLG guidelines on
diligent search**

Deliverable number	<i>D3.2.2.</i>
Dissemination level	<i>Public</i>
Delivery date	<i>1st July 2010</i>
Status	<i>Final</i>
Author(s)	<i>Federation of European Publishers (FEP)</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.

Evaluation of compliance of the ARROW workflow with the agreed HLG guidelines on diligent search

- 1. Introduction 4
- 2. The HLG guidelines 4
- 3. The ARROW workflow 5
- 4. Compliance of ARROW workflow with the agreed HLG guidelines on diligent search 8
 - 4.1 Procedure for due diligence 8
 - 4.1.1 The search is done prior to the use of the work 8
 - 4.1.2 The search is done title by title or work by work; 8
 - 4.1.3 The relevant resources would usually be those of the country of the work’s origin 9
 - 4.1.4 The active search would use the appropriate resources 9
 - 4.1.5 Material consisting of several works 11
 - 4.1.6 The search process should be documented 12
 - 4.2 Measures to prevent future orphan works 12
 - 4.3 Databases of orphan works 13
- 5. Conclusion 15

Executive summary

The objective of this deliverable is to match the main principles of the guidelines on due diligence criteria for orphan works as agreed by the EC i2010 High Level Expert Group on Digital Libraries (HLG) with the ongoing development of the ARROW workflow. ARROW provides libraries with access to selected data sources and tools supporting the identification of the rights status of works and the search for rightholders; as the guidelines on due diligence provide a framework for the design of diligent search procedures regarding rights and rightholders information related to the identification of orphan works, their connection with the ARROW project is highly relevant.

The guidelines, meant to improve conditions for digitisation and online accessibility of cultural works, include a definition of orphan works and lists of due diligence criteria (procedures and of the resources available); they also propose measures to prevent future orphan works and suggest setting up databases with information on orphan works.

ARROW in turn gives access to a network of distributed data providers, as the information needed by the ARROW system to support the diligent search process is stored in the databases of different organisations for their own specific purposes. The design of the workflow formalised this role of “interoperability facilitator” and also illustrates the steps and resources involved in a search procedure (the main actors being libraries, The European Library, Books in Print databases and Reproduction Rights Organisations).

According to the guidelines, the procedure for due diligent search should be based on a number of principles: the search is done prior to the use of the work; the search is done title by title or work by work; the relevant resources would usually be those of the country of the work’s origin. The ARROW workflow complies with all of these principles.

In terms of the resources suggested, ARROW covers a large part of those listed in the guidelines, both common and sector-specific, and definitely the most relevant ones. The ARROW workflow also facilitates compliance with the recommendation to document searches, as the system implements mechanisms to store information, including records of the searches performed.

Furthermore, ARROW enhances the adoption of standards (identifiers and metadata) to foster interoperability and facilitates the access to authority files and the use of authoritative data on personal and organisational names and dates; it thus indirectly supports the adoption of some of the measures put forward in the guidelines to prevent future orphan works.

In sum, the ARROW workflow complies with all of the principles and most of the recommendations of the HLG regarding diligent search. Therefore, it can be considered as a valuable tool for libraries - and other subjects - that need to retrieve information on rights and rightholders for digitisation initiatives.

1. Introduction

The objective of this deliverable is to match the main principles of the guidelines on due diligence criteria for orphan works (hereinafter “the guidelines”) as agreed by the EC i2010 High Level Expert Group on Digital Libraries (HLG)¹ with the ongoing development of the ARROW workflow. The general framework in which the guidelines have been developed has been described in deliverable D.3.1 on the legal framework – First edition².

Arrow provides libraries with access to selected data sources and tools supporting the identification of the rights status of works (in print, out of print, orphan) and the search for rightholders; as the guidelines on due diligence provide a framework for the design of diligent search procedures regarding rights and rightholders information related to the identification of orphan works, their connection with the ARROW project is highly relevant. The deliverable therefore provides a brief overview of the guidelines and a description of the ARROW workflow, and then a detailed illustration of the correspondence of the steps of the ARROW workflow with a number of recommendations put forward in the guidelines.

2. The HLG guidelines

The relevance of the guidelines to the ARROW project has already been illustrated in deliverable D.3.2.1. on the guidelines for the definition of orphan works³.

The drafting of the guidelines was suggested in the Final Report by the Copyright Subgroup of the HLG⁴ and they were endorsed by a Memorandum of Understanding on Diligent Search Guidelines for Orphan Works⁵, a form of self regulation subscribed by 27 stakeholders’ organisations representing European rightholders and cultural institutions.

The guidelines aim at improving conditions for digitisation of, and online accessibility to, cultural material by creating mechanisms to facilitate the use of orphan works. They are a voluntary measure and have been designed to be generic because of the diversity of potential solutions and resources at national level.

The guidelines include a definition of orphan works and lists of due diligence criteria, in terms of the procedure and of the resources available for research. They also deal briefly with measures to prevent future orphan works and suggest setting up databases with information on orphan works.

¹ *Sector-specific Guidelines on Due Diligence Criteria for Orphan Works*, elaborated by sector-specific working groups composed by representatives from cultural institutions and the creative sectors, which reported to the HLG and to the European Commission (http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg/orphan/guidelines.pdf).

² D3.1 *Report on legal framework* is available for downloading in the Resources area of the ARROW website (www.arrow-net.eu).

³ D3.2.1 *Guidelines for the Definition of Orphan Works* is available for downloading in the Resources area of the ARROW website (www.arrow-net.eu).

⁴ Pages 14-15 of the Final Report on Digital Preservation, Orphan Works and Out Of Print Works (http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg/reports/copyright/copyright_subgroup_final_report_26508-clean171.pdf).

⁵ Page 1, point 1 and 2 of the Memorandum of Understanding on Diligent Search Guidelines for Orphan Works (http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg/orphan/mou.pdf)

3. The ARROW workflow

The main objective of ARROW system is to provide a comprehensive and easy to use service to support any diligent search model adopted by libraries, by facilitating the identification of rightholders (authors/publishers) and the identification of the rights status works with particular concern to orphan and out-of-print works.

ARROW therefore has been developed as a comprehensive system made up of:

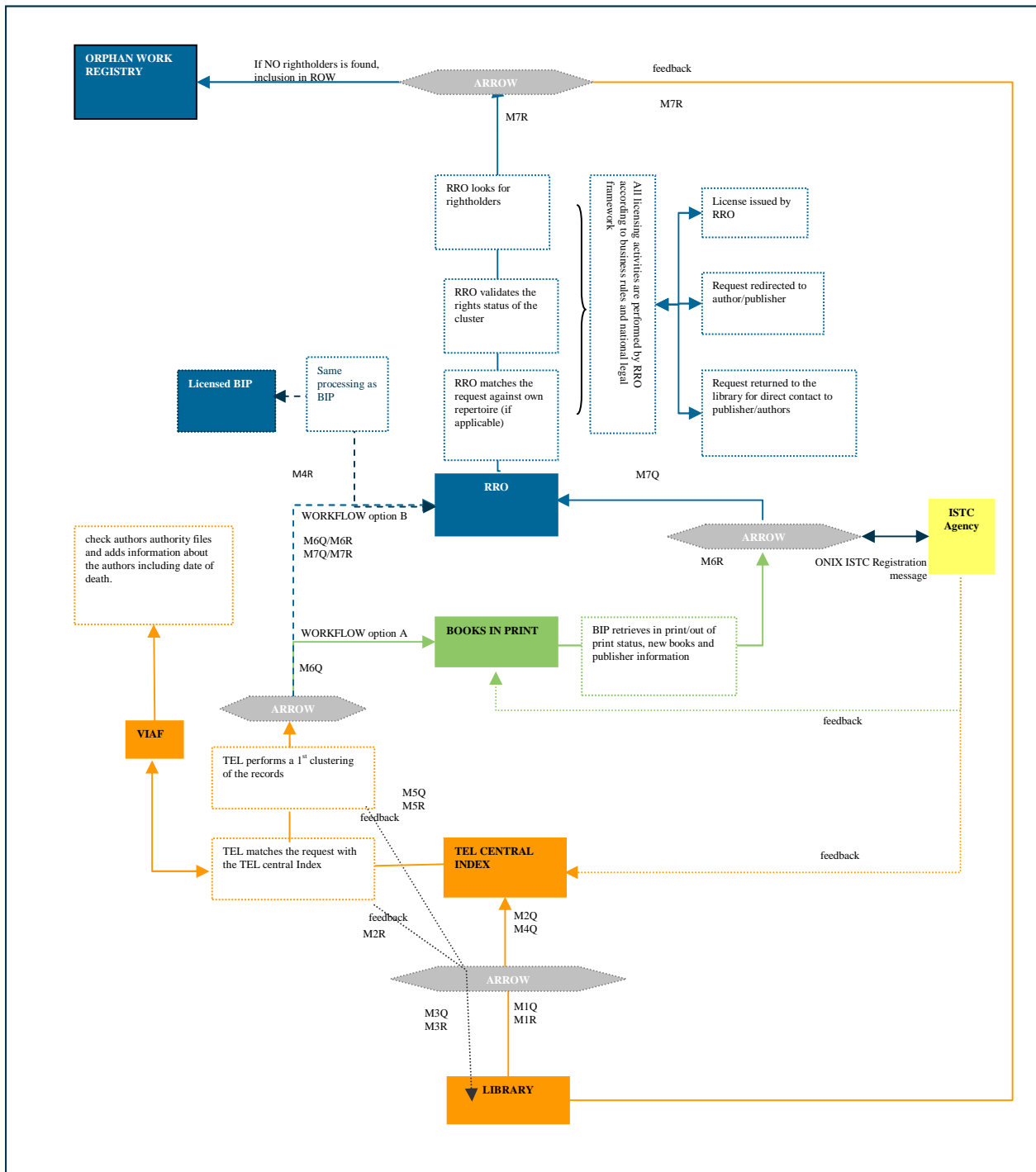
- Distributed Rights Information Infrastructure (RII): a service addressed to libraries to support their diligent search for rightholders, as well as in future perspective to other commercial and non commercial players planning to build a digital collection including copyrighted materials;
- Registry of Orphan Works (ROW): whenever rightholders are unknown or cannot be traced the work will be included in a registry to allow authors, publishers or their representatives to claim their rightholdership.

Information needed by the ARROW system to support the diligent search process is stored in the databases of different organisations for their own specific purposes⁶, therefore the role of ARROW can also be defined as that of “interoperability facilitator” to access and query different IT systems, to retrieve the relevant data, to process and exchange these data with other systems and to add complementary data from other systems.

So since the beginning ARROW was appointed to set up a network of distributed data providers, therefore ARROW is to be considered a network of information sources. This has been formalised in the design of the ARROW system workflow, as represented in the following diagram, which provides a high level description of the Rights Information Infrastructure processes and steps carried out to support a diligent search. The diagram also highlights data providers and information resources connected in the infrastructure and how the results of the diligence search can feed the Registry of Orphan Works

⁶ As also highlighted in the HLG Sector Specific Guidelines on due diligent criteria. For more information on this issue, see §4.1.4 *The active search would use the appropriate resources in this deliverable*. For more information on the ARROW analysis of relevant sources of information to be included in the system, see D5.1 *Analysis of bibliographic resources and clearing mechanisms existing in Europe* and D4.2 *Guidelines for Technical Interoperability* available for downloading in the Resources area of the ARROW website (www.arrow-net.eu)

The ARROW workflow diagram



Starting with the initial request from a library, the resulting “transaction” passes through stages of resource identification, work and manifestation clustering and the identification of related resources, before submission to an RRO for a licensing decision. For the time being, the ARROW workflow runs on a national basis based on the country of publication of the target resource.

The pilot process workflow concludes with the RRO sending its considered response to ARROW: this may involve the grant or denial of a license and/or the provision of additional information to help the library bring the request to complete resolution.

The following list describes the categories of actors and sources of information included in the ARROW workflow and their specific role:

- Libraries: they are ARROW's primary "end users". As data providers, they take part to the ARROW infrastructure via The European Library (TEL).
- The European Library (TEL) coordinates access to the resources of European national libraries and the Virtual International Authority File (VIAF) and provides resource identification and clustering⁷ services to ARROW.
- Books in Print agencies (BiP): match incoming requests against their own records and advise ARROW of any matching ISBNs as well as returning ONIX product records containing in print or out of print status, availability and other metadata.
- Reproduction Rights Organisations (RROs): review the consolidated requests from the libraries and the information gathered throughout the ARROW process. On this basis the RROs then proceed to grant or deny licenses, as appropriate, as well as offering further advice to ARROW and the libraries on how best to bring requests to complete resolution. In workflow B, RROs takes over the function of the BiP⁸.

Here is a short description of the main steps in ARROW workflow to provide reference for the analysis carried out in § 4. *Compliance of ARROW workflow with the agreed HLG guidelines on diligent search:*

- **Step 0:** Library creates its user account on the ARROW system, receiving a user name and password and can then connect to the ARROW "secured network"
- **Step 1:** Library submits a query to ARROW requesting a permission to digitise and use a book, namely the copy of the book on shelf in a specific edition/format (paperback, hardcover, large print, etc.) published by a specific publisher, which ARROW validates and acknowledges receipt;
- **Step 2:** ARROW forwards the query to The European Library for the matching process;
- **Step 3:** The European Library processes the request, identifying the book(s) by matching the metadata query with the records present in its Central Index and returns the results to ARROW;
- **Step 4:** ARROW forwards the results to the Requesting library for validation;
- **Step 5:** the requesting library validates the matching results, choosing the closest match
- **Step 6:** ARROW asks The European Library to perform a clustering process for each exact match identified;
- **Step 7:** The European Library performs the clustering process and returns the cluster(s) to ARROW;
- **Step 8:** ARROW forwards the results to the Requesting library for validation (optional);
- **Step 9:** the library validates the clusters (optional)
- **Step 10:** ARROW forwards the validated clusters to the BIP to match each manifestation, define the publishing status of each work, and provide publishing information;
- **Step 11:** BIP returns the results to ARROW;
- **Step 12:** ARROW forwards the information gathered to the RRO to match with its own records, advise upon available licences and provide information about the rights holders;
- **Step 13:** RRO returns the results to ARROW, proposing or denying a license (or stating that no license is required) and offering further information about rightsholders;
- **Step 14:** ARROW sends feedback to the Requesting library about the rights status of each work (in print/out of print/public domain or orphan), the offer or denial of a license, details of an alternative RRO to contact for a licence where applicable and rightsholder information if available.
- **Step 15:** in case the work is considered orphan, ARROW registers the work and associated information in the Registry of Orphan Works

⁷ In the ARROW context, a cluster is defined as group of manifestations (books) that share the same underlying intellectual work.

⁸ United Kingdom and Spain have adopted Workflow B.

4. Compliance of ARROW workflow with the agreed HLG guidelines on diligent search

This section evaluates in detail the matching between the ARROW workflow and the principles and recommendations contained in the guidelines on diligent search. Every principle and every recommendation are examined and compliance of steps of the ARROW workflow is highlighted.

4.1 Procedure for due diligence

As recommended by the HLG⁹, the procedure and methodology of the search and further matters of orphan works is to be established by the cultural institution in compliance with the applicable guidelines. Similar cultural institutions are encouraged to cooperate and publish search procedures and methodologies.

The HLG guidelines firstly set a number of principles for identifying and/or locating the rightholders:

- the search is done prior to the use of the work;
- the search is done title by title or work by work;
- the relevant resources would usually be those of the country of the work's origin.

Moreover, the HLG guidelines define other principles specifying the procedure and methodology to follow:

- the active search would use the appropriate resources (listed in the guidelines);
- material consisting of several works (whether including “embedded works” or based on “underlying works”), should be searched first on the basis of the guidelines in the sector of the ‘main material and where the rightholders of these works cannot thereby be identified and/or located, the guidelines for the sector of the embedded and/or underlying work must also be followed’;
- the search process should be documented.

4.1.1 The search is done prior to the use of the work

From a legal perspective it is important to carry out a search before using the work because otherwise the user risks, inter alia, infringing copyright rules as the copyright holder might not agree to the said use. Prior consent to use protected works forms the basis of international copyright legislation. Article 9 of the Berne Convention says that “*Authors of literary and artistic works protected by this Convention shall have the exclusive right of authorising the reproductions of these works, in any manner or form*”.

With regard to this first principle, ARROW can definitely be consulted before any use of a work, as searches only require to submit a defined set of metadata, specified accordingly to libraries' cataloguing practices, and that can be easily exported from libraries' existing records. Moreover, searching through ARROW also requires the library to specify what usage the library is planning to do with the work (digitisation and/or associated exploitations), clearly implying that the search is to be performed prior to the use of the work. Finally ARROW can be searched also as iterative process, prior to any further additional use of a work that had been already searched for. (Ref. to ARROW workflow Step 1).

4.1.2 The search is done title by title or work by work:

According to the guidelines on orphan works, agreed by the High Level Group on digital libraries, diligent search would generally have to be carried out on a title by title basis, based on available data. This will ensure that the search is sufficiently accurate, as the identification of all relevant rightholders

⁹ The European Digital Library Initiative, *Sector Specific Guidelines on due diligence criteria for Orphan Works, Joint Report*.
http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg/orphan/guidelines.pdf.

of a work (not limited to the first author, but also including others like editors, illustrators, authors of commentaries and prefaces, etc.) is necessary to correctly establish the rights status of the work.

To facilitate the use of the ARROW system in large scale digitisation programmes, a library can search ARROW submitting in a single process a list of book records and specification of what rights are requested to the ARROW system. Regardless of how many requests are made at the same time (how many books or titles are contained in a single submission), the ARROW system splits the submission at item (book/title) level and the search is performed by the system title by title. (Ref. to ARROW workflow Step 1). Moreover, for each book (title) submitted by the library for searching, ARROW retrieves information regarding the underlying work, by means of work data extraction, matching and clustering procedures and algorithms, thus ensuring that the search is performed work by work as well. (Ref. to Step 3, 7 and 10 in the ARROW workflow)

4.1.3 The relevant resources would usually be those of the country of the work's origin

From a legal point of view the principle of the country of origin can be justified in the sense that copyright protection is territorial in nature. Thanks to international Conventions, protection is extended from the country of origin (country of first publication, Article 5 (4) a of the Berne Convention) to the rest of contracting countries. However it is more likely to find the relevant sources in the country of origin where the protection is initially granted. The country of origin criterion applies when the country of origin of the publication or author is known. In these circumstances there should not be major difficulties in focusing on the country concerned to conduct the diligent search and decide on the authorisation of the digitisation and making available. Moreover, in relation to books and other textual works the country of origin is almost always known and, in case it is not known, the language may also offer an indication.

The ARROW system has been conceived from the beginning to be able to work on a country basis; this means that starting from the country of the work's origin, which is in principle that of publication, all the appropriate sources of information in that country are connected in the ARROW workflow. For this reason the "country of publication" of the book the library is searching is one of the key information (metadata) ARROW collects and process, along with the "language of text", in order to identify relevant sources of information to be queried in the workflow, from Step 1 of the workflow on.

4.1.4 The active search would use the appropriate resources

The HLG guidelines suggest a number of relevant resources available for research, both common and sector-specific, including the European Digital Library, legal deposit institutions/national libraries, published biographical resources for authors, published indexes and catalogues from library holdings, collective rights management organisations and (sector-specific) lists of books in print.

As anticipated, the above mentioned categories of sources of information have been included in the conceptual model of the ARROW workflow and then implemented on a country basis.

First of all, the HLG recommends including in the due diligent search some information sources that can be grouped under the common characteristic of belonging to the "library domain": catalogues of national libraries, the European Digital Library, indexes of library holdings, biographical resources for authors, etc.

Bibliographic data from the catalogues of Europe's national libraries is one of the key data sources in ARROW. Originally, it was envisaged that the ARROW system would query each of the national library catalogues separately. However, bibliographic data are already aggregated through The European

Library¹⁰, which provides an integrated access to the online catalogues of European national libraries (which in most cases are also in charge of receiving legal deposits). For this reason, it has been decided to query TEL as the single access point to the resources of national libraries in Europe. In future perspective The European Library is playing the role of main aggregator for library domain for Europeana, thus enlarging the coverage also to other types of libraries. The interaction with The European Library includes, inter alia, TEL to get additional information on authors from VIAF (Virtual International Authority File), including all name variants, dates of birth and death, nationality of each author or other contributor (translator, editor, illustrator, etc.). (Ref to Step 2, 3, 4, 5, 6 and 7 in the ARROW workflow).

The inclusion of The European Library in the ARROW workflow therefore covers a relevant part of the common resources indicated by the HLG to identify and/or locate the rightholders and ensures that scaling the ARROW system to all the European Union member countries would not be an issue.

Secondly, the HLG agreed that for the implementation of the orphan works definition, the relevant sources for the due diligent search for rightholders would include specifically the books in print database in the country of the digitising institution and the one in the country of origin of the works when different from the country of origin of the digitising library.

As a matter of fact, the rights status of the work (in public domain or copyrighted, in print or out of print), is one key information needed to support due diligent search: on the basis of this information, terms and conditions for the use of a work may vary according to the national legal scenario and agreements in place between rightholders and their representatives and agents.

In order to know the in print/out of print status of a work, ARROW needs to know the in print/out of print status of each single manifestation belonging to the work. Books in Print databases are the data source that ARROW queries to retrieve that information.

In the ARROW workflow, Books in Print are included either via direct connection or, in case for example the Books in Print database has been already integrated within an RRO system¹¹, via connection to the source of information providing also books in print related information. This part of the workflow is essential to distinguish between an in print and an out of print work and covers the part on sector-specific resources specified by the HLG. (Ref. to Step 10 and 11 in the ARROW workflow)

As to the other source indicated by the guidelines as sector-specific for books, the Public Lending Right authority databases (where available), the recommendation is implemented by ARROW inasmuch as the RRO or other organisation involved is also the PLR administrator or responsible for the distribution of PLR remuneration.

The guidelines for the definition of orphan works also add that it is essential to check the published sources on the history of legal bodies and agencies administering relevant standards and identifiers including their databases (ISBN agency for books, which in several cases correspond with the BiP database holders). The ARROW workflow therefore complies with the above requirements when involving the BiP in the process. Moreover a standard structure for Publisher and Imprint data

¹⁰ Launched as an operational service in March 2005, The European Library provides a single point of access to the bibliographic and digital collections of the National Libraries of Europe. In spring 2010, 46 from the 48 national libraries in Europe have made their collections available in The European Library. The 7 national libraries, who are partners in the ARROW project, are all full-members of The European Library network. For more information about The European Library please refer to: <http://www.theeuropeanlibrary.org/>

¹¹ This option has been adopted for the ARROW workflow in the United Kingdom and in Spain.

exchange has been proposed for elaboration and implementation to refine the ARROW workflow and allow an interoperable search of publishers as rightholders.

Thirdly, according to HLEG guidelines, one of the resources to check for diligent search is collective management organisations such as collecting societies and Reproduction Rights Organisations (RROs) in the relevant countries.

ARROW complies with this requisite since it involves RROs. Whenever separate RROs exist for different categories of rightholders, as for example authors and publishers, all the relevant organisations – independently or networked among each other – are included in the ARROW workflow. RROs have the task to match the information received from ARROW against their repertoires and if rightholders can be traced or contacted the RRO carries on the process. In this case either the RRO has a mandate from rightholders to license the usage, or it is not authorised to license the usage; in any case, it passes the information to ARROW.

In case rightholders can not be traced or contacted, the work is marked as “probably orphan” and the RRO notifies to ARROW the inclusion in the Orphan Works Registry. (Ref. to Step 12, 13, 14 and 15 of the ARROW workflow)

Finally, the joint involvement of BiPs and RROs in the ARROW workflow allows effectively collecting the information whether a work has been declared out of print and facilitates the licensing in case the RRO has the mandate from the rightholders. This complies with the main principles for out of print works that have been agreed in the guidelines by the HLG¹²; (i) a work is out of print if the rightholder concerned has declared it not to be commercially available, (ii) rightholders should be in a position to decide whether a work which has been defined as out of print should be licensed to be made available.

4.1.5 Material consisting of several works

According to the HLG, material consisting of several works (whether including “embedded works” or based on “underlying works”), should be searched first on the basis of the guidelines in the sector of the ‘main material and where the rightholders of these works cannot thereby be identified and/or located, the guidelines for the sector of the embedded and/or underlying work must also be followed’. This recommendation, mainly based on practical considerations, is nonetheless in line with the existing legal principles and the international and European legal framework

The ARROW system, as conceived at the time of writing, is focussed mainly on textual material and books in particular and therefore the search is done in the text-books sector, however for books including other types of works embedded in a textual work – as images and photographs – the ARROW system performs also a search for rightholders of the embedded work (mainly contributors such as illustrators and photographers), whenever they are recorded in libraries’ catalogue records or in the other sources of information connected in the workflow. Further developments are foreseen for a future stage of the project, to include in the ARROW workflow also appropriate sources of information for embedded works as visual material, thus allowing an enhanced search.

¹² Page 17 of the Final Report on Digital Preservation, Orphan Works and Out Of Print Works
(http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg/reports/copyright/copyright_subgroup_final_report_26508-clean171.pdf)

4.1.6 The search process should be documented

In order to ensure legal certainty and prove that the diligent search has been conducted according to the requirements, it is advisable to gather enough evidence in this respect. The documentation of the search process will provide this evidence.

The guidelines mention that the search should be documented: this could easily be done through ARROW as the system implements mechanisms to store information, including records of the searches performed along the whole workflow. The use of standard and unique identifiers to identify and track works and transaction allows an easy retrieval and display of all the results obtained in the different steps of the workflow.

This storage of information complies also with another recommendation of the guidelines, according to which a statement that rightholders could not be traced (including such information on the rightholder(s) as available and appropriate) is considered good practice. The results of searches performed via ARROW can be in fact recorded and also connected to individual works to facilitate future queries; in particular, in case a work is marked as probably orphan at the end of Arrow workflow, relevant data documenting the search will be stored in the Orphan Works Registry (see also § 4. 3)

Furthermore, insofar as these requirements directed also and especially towards the subjects who intends to carry out digitisation activities, such subjects can refer to ARROW as the main resource used; ARROW lists the resources which are searched (Library, TEL, VIAF, BiP, RRO; sometimes authors' and publishers' organisations' databases) as a part of its documentation. ARROW does not deal instead with the suggestion by the guidelines to make announcements of the search in a number of media (such as websites, relevant publications, social and professional networks, mass media); this is left to the library or other entity performing the search.

4.2 Measures to prevent future orphan works

The HLG guidelines define orphan works for each of the sectors interested; for text materials, the definition is as follows: "An orphan work is a work protected by copyright but the current owner is unknown or untraceable by diligent search. The current owner of the copyright might be the author or other creator, some other first owner if the rights (such as the author's employer -when applicable) or a publisher) or any rightholder who is presumed to be the rightholder according to the legislation or contractual agreement or any successor of the first owner."¹³

As stated in the guidelines, a *"work becoming orphaned is essentially an issue of missing or incomplete information"*¹⁴. Therefore the guidelines recommend measures to improve the availability of information on works, rightholders and rights:

- Use of electronic and other identifiers, including the name of the author
- Creation, use and maintenance of metadata on the work and its rightholders, for example through the automatic metadata creation as images are created
- Enhancing the value of standard identifiers (such as ISAN, ISBN, ISRC, ISMN, ISWC) and registries by recording the devolution of rights to second and further rightholders (as the French and Spanish audiovisual public registries already do);
- Naming relevant rightholders in/on works and their packaging, covers etc
- Keeping a record of relevant rightholders by users as far as practical
- Recording authors' death dates in name authority files of national bibliographies

¹³ The European Digital Library Initiative, *Sector Specific Guidelines on due diligence criteria for Orphan Works, Joint Report*: http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg/orphan/guidelines.pdf.

¹⁴ Ibid.

ARROW indirectly supports the adoption of the above mentioned measures: by the setting up of a pan-European information infrastructure connecting relevant source of information, ARROW enhances the adoption of standards (identifiers and metadata) to foster interoperability and increases the awareness of the importance of creation, use and maintenance of correct and updated metadata. Likewise, it facilitates the access to authority files and the use of authoritative data on personal and organisational names and dates.

Finally, the potential connection between the ARROW workflow and the registration of ISTC numbers (International Standard Text Code) can contribute to the measures suggested by the guidelines which could help to prevent future orphan works, as these include the use of electronic and other identifiers and enhancing the value of standard identifiers.

4.3 Databases of orphan works

The High Level Expert Group also developed a set of key principles for orphan and out of print work databases and right clearance centres. These principles establish that the overall solution for orphan works consists, inter alia, of databases of orphan works to facilitate users in their search, which are needed irrespective of any legislative solution. Furthermore, the document specifies that the rationale of the database is to provide assistance to users in their search endeavours. As part of the process and upon certification of the claimant, the information on the rightholders and the consequences, etc. for the license will be published in the orphan works database.

The Final Report by the HLG explicitly recommends that Databases and Rights Clearance Centres for orphan and out-of-print works are established¹⁵. It also maintains that ARROW is a first step in this direction. One of its annexes¹⁶ provides recommended key principles for clearance centres and databases for orphan works; according to those, the combined use of databases and clearance centres should enable users, institutional as well as commercial ones, to conduct an online search, in order to discover the identity and location of rightholders, find available information on previous diligent searches or find whether the work is already registered as available digitally.

The guidelines suggest specifically developing databases which contain information on titles, rightholders and any metadata which are useful to identify and find rightholders. The ARROW workflow thus further matches the guidelines' documentation requirement and the one focused on databases on orphan works as it foresees that the information retrieved will feed into an Orphan Works Registry. Data stored in the Registry and possibly in the system records also matches the suggestions and recommendations of the Final Report.

In fact the ARROW workflow is necessary to gather all the information to deduce the status of the work to which the library's book belongs. In particular Step 14 in the ARROW workflow foresees that the RRO declares that the work is "probably orphan" if the rightholders can not be traced following the previous steps. In this situation, the RRO will check in the ROW (Orphan Work Registry) if the work has already been recorded as orphan. If that is not the case, the RRO will record the work as orphan.

¹⁵ Page 5, i2010: Digital Libraries High Level Expert Group – Copyright Subgroup, Final Report *"Digital Libraries: Recommendations and Challenges for the Future"*
(http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg/reports/hlg_final_report09.pdf).

¹⁶ Page 2, Annex 6: i2010 Digital Libraries Copyright subgroup's 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/copyrigh-annex6.pdf).

In this perspective, the ARROW workflow for Rights information Infrastructure constitutes the fundament for the Registry of Orphan Works, as the registry is likely to start “empty” and will be populated by digitisation requests being processed through the ARROW system in an automatic way, in the case the output of the process indicates that the work can be an “Orphan” The Registry of Orphan Works will also be available for searches and for rightholder to claim their rights ownership individually or through a collective representative organisation or agent.¹⁷

¹⁷ At the time of writing, the Registry of Orphan Work is still being set up and management tools for rightholders are currently under evaluation. More information will be published in D6.2 *Registry of Orphan Works management System*.

5. Conclusion

As this analysis has shown, the ARROW workflow complies with all of the principles and most of the recommendations of the HLG regarding diligent search. Therefore, ARROW can be considered as a valuable tool for libraries - and other subjects - that need to retrieve information on rights and rightholders for digitisation initiatives.

The legitimacy conferred to the guidelines by their wide backing by representatives of the rightholders and of cultural institutions position ARROW as an ideal support to search under any kind of legal framework regarding the digitisation and making available of orphan works.