



Accessible Registries of Rights Information
and Orphan Works towards Europeana

ARROW

BUSINESS MODEL

February 2011

ARROW Business Model

1. Introduction.....	2
2. Organisational background, mission statement and key strategies.....	2
3. The ARROW value proposition: the services offered	3
3.1 Core value/services	3
3.2 Additional values	6
4. Main ARROW customers	7
4.1 Public institutions.....	7
4.1.1 Libraries	7
4.1.2 Other cultural institutions	10
4.2 Digitisation projects by private entities	10
4.2.1 Rightholders in general	11
4.3 RROs.....	11
5. ARROW use cases	11
5.1. France.....	12
5.2. Germany.....	14
5.3. UK.....	14
6. ARROW’s business architecture	15
6.1 Software	19
6.2 Hardware.....	22
6.3 The hosting and maintenance service	23
6.4 Office infrastructure.....	24
6.5 Personnel.....	24
6.6 The ARROW social infrastructure.....	25
7. Funding model and budget.....	26
7.1 Financing sources.....	26
Public/indirect	26
Private/direct	28
Payments in kind.....	28
7.2 Fee models	29
8. Governance	30

1. Introduction

ARROW (Accessible Registries of Rights Information and Orphan Works towards Europeana) is a project co-funded under the eContent+ programme of the European Commission and managed by a consortium of European national and university libraries, organisations representing authors, publishers and Reproduction Rights Organisations (RROs, collective management organisations in text and image based works).

ARROW aims to establish a system to identify rights, rightholders, rights status of a work, including whether it is orphan or out of print. This will enable libraries as well as other users to obtain information on who are the pertinent rightholders, which are the relevant rights concerned, who owns and administers them and – where possible – how and where they can seek permission to digitise and / or make available the work to user groups. The project also seeks to enhance the interoperability between sources of rights information held by rightholders, RROs and other collective management organisations, agents, libraries and users. Solutions envisaged by the venture include the establishment of systems for the exchange of rights data, the creation of registries of orphan works, information on or the establishment of registries of out of print works, supporting the creation of a network of rights clearance mechanisms.

This document introduces a sustainable mid- and long term business model and governance form for the running of ARROW after the first project phase has been concluded, when ARROW will enter a new phase (currently called ARROW plus) during which the number of participating countries will be expanded to cover progressively all or the majority of the EU Member States and EEA countries and the system set up in the initial phase will be enhanced.¹ It outlines the basic elements of the three main components in the ARROW business model: the ARROW customers; the ARROW value proposal; and the resources/capabilities required to implement the business model. It also outlines some ideas regarding the system's governance.

2. Organisational background, mission statement and key strategies

The features identified as the keys to achieving the project objectives are interoperability, standards deployment and stakeholder involvement. On the latter point, a large and comprehensive network of entities is involved in the project, either as contracting partners, as associated partners/external supporters or otherwise directly involved in and delivering services to the project without having a formal relationship to it (such as certain books in print databases). These include: libraries, publishers associations, reproduction rights organisations (RROs) and other collective management organisations representing authors and/or publishers, Books in Print information source companies,

¹ Special attention shall be dedicated to the countries where the project prototype is already operational and, even more, to those where digitisation plans are under way or at least in preparation.

technology developers and ISBN agencies, both at national and international level. These, and the project itself, are in turn linked to an even broader external network, comprising EDItEUR, ISTC- IA, DOI, ACAP, etc.; the networks of authors associations, publishers associations and RROs; potential key user groups such as libraries and other cultural institutions, in particular those delivering to Europeana.

2.1 Mission statement

ARROW's mission statement is *“to integrate information on rights, rightholders and rights status (thus facilitating their search and retrieval), with a focus on orphan works, therefore building a European wide orphan works registry”*.

The objective of ARROW is to facilitate the identification of rights, rightholders and rights status, and through this also to enable digitisation projects; ARROW is neutral as to who uses its services, as long as the customer pays the established user/service fee. It is therefore open for use by all kinds of players with requirements that can be addressed by ARROW, both public and private, whatever their business models might be.

2.2 Key strategies

The key strategies for the establishment of ARROW are:

- close the information gap through the building of a system to facilitate search for rights, rightholders and rights status and an orphan works registry
- contribute to enhanced cost efficiency for players in the value chain
- serve players in the market equally and be neutral with respect to business models.

Building on these strategies, the paper identifies the key ARROW value proposition – distinguishing between the ARROW “core business” and additional services; the main potential customers of the ARROW system; the key value proposition for ARROW; and the system's business architecture, i.e. the resources and capabilities required to implement ARROW, including costs, funding and revenue streams.

3. The ARROW value proposition: the services offered

3.1 Core value/services

ARROW sets out to offer a custom-oriented and cost and time efficient system to identify (i) rights; (ii) rights status; and (iii) rightholders in copyright based works, to establish a European-wide orphan works registry and to build new/interconnect already established out-of-print registries. A system that enables a cost and time saving identification of these attributes meets expressly stated requirements from ARROW's main customers from the public, non-profit as well as the commercial sectors.

There are numerous initiatives to digitise and make available the world's - including the European - cultural heritage. Digitisation programmes have been put forward by public institutions such as libraries and also by commercial players in the market (most notably Google through its Google Book Search programme, although it is not the only one), as well as by a large number of publishers.

The digitisation and making available of the common cultural heritage has also a strong political support. On a European level this is backed by the European Commission, for instance through the i2010 digital libraries initiative and the establishment of the High Level Expert Group (HLG), which released its final report in February 2010² and the Communication on Europeana next steps; and the European Parliament, including in its own initiative report on Europeana.

Various EU official documents and the statements from those involved in national digitisation projects have clearly pointed out the need to include works that are in copyright in the range of works available to users in digital form. It has further been stressed as a clear requirement that the digitisation and making available of copyright works must not conflict with the publishers' and authors' interests in commercialising them. Consequently, there is a need to identify whether a work is in or out of copyright, whether a work is offered commercially, who the rightholders are and where the rights can be cleared. The ARROW value proposition therefore meets clearly expressed user requirements, as well as those on political levels.

For many potential users, “to facilitate” clearly implies costs savings. Basically, it should allow users to perform the same tasks (i.e., in the case of digitising and making available copyright based works, performing a diligent search for rights, rightholders and right status) while employing less resources. This means that if A is the cost of diligent search without ARROW, and B is the cost for diligent search using ARROW, any price $P < (A - B)$ allows the user to save money. If the total ARROW costs are lower than the sum of the prices paid by its users³, then the system is self-sustainable through revenues generated by its core business.

From the information already available, we know that broadly speaking the costs of diligent search depend on the type of collection that libraries or other users want to digitise and make available.. Average costs depend among other things on the average age of books. Therefore, every estimate risks being imprecise, following the statistical rule that when the variance is high estimates are less good.

The cost structure⁴ should be divided in two parts: The first is related to a part of the search that must be done in any case, unless libraries are planning to make available books that publishers are actively trying to sell (which would be a violation of the fundamental principles of copyright as expressed by the three-step test). This part includes:

- preparation of appropriate metadata records (so either new creation or some manipulation of what already exists)
- search for the “commercial availability” status of the book

² http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg/reports/hlg_final_report09.pdf.

³ As a formula, if total costs $TC \leq \sum P_i$, where $i = 1, 2, \dots, n$ are the users of the service.

⁴ For works not identified as public domain.

- check if the book has been already scanned by other libraries

The second is a proactive search for rightholders. Even if there is a licensing scheme to use orphan works, this search should be done in any case for distribution purposes. It is just a matter of doing it before or after the scanning. Doing it before has two advantages: the first, more important, is that this way copyright principles are respected. The second, more practical, is that if rightholders are found after the scanning, if they disagree with making their works available, then the cost of scanning might be wasted (except for when digitisation is carried out by libraries, which can digitise works for defined purposes under an exception in national legislation, subject to compliance with the three-step test).

There is little empirical evidence of diligent search costs when conducted with the current existing tools (that is, “before ARROW”). Examples include the following:

- The German National Library (DNB) conducted an experiment on a limited sample of titles, and found that the average time spent in search is 52 minutes. The cost depends of course on labour costs in the specific context.
- The Italian Publishers Association (AIE) has estimated the costs for a service to search for Italian rightholders for claiming books within the Google Settlement. AIE planned to charge around 9 euro per title when a rightholder was found, with no charge if the rightholder was not found, so – the average costs per submitted title would be less than that.

All of these costs are for the full service. It is difficult to estimate the share of the two parts of the service separately. However, a rough estimate can be that 50% of the cost depends on the first phase, and 50% for the second. So, when considering the actual rightholder search only, the estimated costs would be reduced by 50%.

ARROW aims at fully automating the first part of the search, whilst for the second it depends on the comprehensiveness of the RRO databases. So, initially, it will contribute to cost reduction more for the first part of the search flow. In the long run, ARROW is designed in a way to foster the enrichment of data so as to automate also the second part of the service, which initially will require some human intervention.

Costs generally have to be assessed against potential benefits. From the library viewpoint, even without considering legal costs for liability, the benefits include:

- not digitising books that rightholders will later ask to remove, once again saving the cost of digitisation;
- the possibility to negotiate with rightholders (once found before scanning) the permission to make their works available; in this case, the cost saving will depend on the level of licensing fees agreed with the collecting society in charge.
- not digitising books already digitised by other libraries;

All kinds of users can enjoy such benefits, plus in general the consistent time savings expected from the system. In addition, the value proposition of ARROW also includes

the fact that easier handling of the rights to digitise and make available can improve the legal certainty of digitisation projects and therefore enhance their possibilities to provide access to works.

In conclusion: to have a diligent search is both more respectful of the copyright principle (which rather than an obstacle is a facilitator of the digitisation process), with particular emphasis on moral rights, and often more convenient from the economic point of view of the digitiser.

The capacity of this cost saving model to actually predict the sustainability of the ARROW system in the mid- to long term depends nonetheless on a series of necessary assumptions. The main assumption is that the European Commission and the EU Member States will follow up the political statements on digitisation programmes and on the building up of Europeana as a portal to provide access to the European cultural access which includes works in copyright with concrete digitisation programmes. It is also presumed that the European Commission and the EU Member States will encourage and facilitate national digitisation programmes through providing the necessary funding.

Those engaged in or planning to carry out digitisation programmes shall be informed of the opportunity presented by ARROW. As ARROW will need to recover the costs of its services, it is required that users of the ARROW system pay directly or indirectly for the use of it (with the possible differences that will be explored later).

3.2 Additional values

The RII consists of a number of fragmented tools (e.g. clustering techniques, ISTC applications, ISNI applications, etc.) that can be used also for purposes different than “facilitating rightholder search”.

For example: clustering the different editions of the same work is something needed for the functioning of ARROW, but it can also be useful for other applications, within the ISTC framework. Services in this field will be offered against remuneration.

The ARROW infrastructure will provide values in various other ways, such as offering access to information through a central registry, including in particular a Registry of Orphan Works; closing the information gap through information based on data from high quality sources; enabling a larger portion of works to be digitised and made available legally; assisting to local registries and building interfaces with the ARROW system.

The improvement of rights information retrieval will also enable publishers or RROs to issue licences for digitisation / online availability / print on demand (reprint) or e-books on demand for certain out of print books; remuneration for those services could be paid directly via ARROW. ARROW could also facilitate negotiations between publishers and authors on the reuse of works.

ARROW will also offer the opportunity for users to access data on commercial availability for all the editions of a textual work, linked with information on rights status, conveniently and at reduced costs.

4. Main ARROW customers

Several kinds of players can have a need for a system like ARROW for their digitisation initiatives, their motives being preservation and making available of cultural heritage for libraries, and profit for commercial entities, either linked directly to sales or to the possibility to offer a wider range of results to users performing internet searches.⁵

In the short and middle term, provided that certain conditions are met (in particular, the availability of funds for digitisation programmes), libraries and other public institutions are assumed to be the main users of ARROW. Digitising and making available the cultural heritage is principally the duty and task of public institutions and that is also where we currently find the majority of the digitisation projects which involve copyrighted works. Digitisation programmes are therefore also chiefly publicly funded initiatives.

Also private bodies have launched plans to digitise in-copyright works, such as Google and Microsoft. Others, which are entering the e-book market, can have an interest in doing so. In general, this concerns all those who want to clear rights to make use of copyright works regardless of their status including out-of-commerce and orphan works. RROs may also have a certain need to verify information that may be available through ARROW.

In its first phase of operations ARROW is going to serve stakeholders' agreements at national level (see use-cases description below), with libraries, publishers and CMOs setting the rules in a way that balances the interests of all the parties.

Chapters 4.1 to 4.5 describe the general design of the business model highlighting the value that in principle the system is able to offer to the main players of the digitisation value chain. Chapter 5 describes how this general approach is going to operate through the first use cases emerging in France, Germany and UK.

4.1 Public institutions

4.1.1 Libraries

Libraries are expected to be the main users of the services that ARROW will offer, at least in the start-up phase. Libraries hold vast collections of books and are requested to digitise and make them available also through Europeana. Europeana aims at being one of the main resource centres for education and research. The objective is to digitise and

⁵ The range of players that can have an interest in using ARROW is thoroughly analysed in D3.6 *Report on business models – Edition 2*, available for downloading in the Resources area of the ARROW website (www.ARROW-net.eu). Here we provide a synthetic outline, focusing more on ARROW's offer than on other players' plans.

make available 25 million in 2014⁶, from the 2010 level of 14 million objects.⁷ It is a declared aim that this should include in-copyright works⁸ with a particular focus on out-of-commerce works, assumed to represent 90% of the national libraries' collections. This creates potentially a substantial market for a service that can assist libraries in identifying whether the prospective millions of new items are in or out of print and provide information on where the rights to digitise and make available in-copyright works can be cleared, and whether the work is in or out of commerce or orphan.

There is a strong political will to see a steady increase of the cultural material digitised and made available through Europeana, and there is a declared intention for this to encompass works under copyright. A number of official documents by the European Institutions (European Commission, Council, European Parliament) supported the creation of Europeana and regularly call for enhancing its content, often focusing on the need for the inclusion of in-copyright material and for solving the issue of orphan and out of print works. Such documents also underline the importance of respecting copyright legislation in the process and the need for adequate funding for Europeana.⁹ Nonetheless, while they mention books explicitly, they do not set precise goals for books as a separate category. Moreover, it is finally up to individual libraries, and especially national ones, to establish and carry out digitisation plans, within the framework of national programmes

⁶ Outline Business Plan for Europeana as a service of the EDL Foundation, November 2008; available at: http://version1.europeana.eu/c/document_library/get_file?uuid=0c6c6078-8026-4297-9367-dd6d14b73c2e&groupId=10602.

⁷ Of which around 1.2 million books.

⁸ See, for instance, the European Parliament report on “Europeana - next steps” by MEP Helga Trüpel, http://www.europarl.europa.eu/meetdocs/2009_2014/documents/cult/pr/793/793669/793669en.pdf, statement 12

⁹ See, for example: Communication from the Commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions “i2010 – A European Information Society for growth and employment” - COM(2005)229, 1 June 2005 (<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2005:0229:FIN:EN:PDF>); Communication from the Commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions “i2010: Digital Libraries” - COM(2005)465, 30 September 2005 (<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2005:0465:FIN:EN:PDF>); Commission Recommendation of 24 August 2006 on the digitisation and online accessibility of cultural material and digital preservation (2006/585/EC) (<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:236:0028:0030:EN:PDF>); Council Conclusions on the Digitisation and Online Accessibility of Cultural Material, and Digital Preservation (2006/C 297/01) (<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2006:297:0001:0005:EN:PDF>); European Parliament resolution of 27 September 2007 on “i2010: towards a European digital library” (2006/2040(INI)) (<http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP/TEXT+TA+P6-TA-2007-0416+0+DOC+XML+V0//EN>); Communication from the Commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions “Europe’s cultural heritage at the click of a mouse - Progress on the digitisation and online accessibility of cultural material and digital preservation across the EU” - COM(2008)513, 11 August 2008 (<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2008:0513:FIN:EN:PDF>); Council conclusions of 20 November 2008 on the European digital library Europeana (2008/C 319/07) (<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2008:319:0018:0019:EN:PDF>); Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions - Europeana: next steps - COM(2009)440, 28 August 2009 (<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2009:0440:FIN:EN:PDF>); European Parliament Report of 5 May 2010 on “Europeana - the next steps” (2009/2158(INI)) (<http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP/NONSGML+REPORT+A7-2010-0028+0+DOC+PDF+V0//EN&language=EN>); Council Conclusions of 10 May 2010 on Europeana: next steps (2010/C 137/07) (<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2010:137:0019:0021:EN:PDF>).

and objectives. It is therefore more indicated to refer directly to other documents about digitisation, especially those elaborated by or through libraries, to have a somewhat more precise idea of the magnitude that can be expected from digitisation initiatives.

Although a digitisation projects by libraries may not be linked to Europeana; still, they can have a need to find rights information and clear rights. Such initiatives are relatively common, often limited in scale, mostly driven by thematic considerations (focusing on a specific topic, a territory, an event, etc.). University and research libraries with special collections are particularly likely to have similar plans.

A broad indication can be derived by the Numeric Study realised in 2007-2009 to measure the progress of the digitisation of Europe's cultural heritage. According to the Study Report¹⁰, released in May 2009, about half of the European national libraries had plans to digitise their collections, while only 3.5% of the digitisation work had been carried out; the plans involved around two thirds of all the holdings of the libraries. This means undoubtedly a very large number of items to digitise, although precise figures were not available. Still, overall figures for all kinds of institutions (national libraries, public libraries, higher education libraries and others) pointed at the existence of over 220 million books and serials that could be digitised, while only some 4 millions had been already. Even taking into account that for many libraries about half of their collections are not included in digitisation plans, the potential is very high. The study estimates that some 90 million books (rare and not) should enter digitisation plans by libraries.

Somewhat more precise information, though more limited in scope, can be obtained by the report on digitisation in European national libraries that was published as part of the EDL (European Digital Library) Project in February 2008¹¹. According to the report, digitisation by national libraries was seriously underdeveloped and issues such as difficulties with copyrighted works, but especially the costs of digitisation, were holding back most initiatives. However, the overall picture of digitisation projects and programmes reported by national libraries envisaged to pass from some 280,000 digitised books in 2006 to more than 1.5 million in 2012. This is clearly a quite low-end scenario, which can be considered as a minimum basis (and it only regards national libraries).

Thus, all in all, in the next few years digitisation programmes by libraries could involve between hundreds of thousands and millions of books per year; a large number of those are likely to be out-of-print, and some will turn up to be orphan. The main unknown quantity is whether there will be enough funds to actually carry out the planned initiatives (and to undertake more, in order to make significant progress and match somehow the intentions expressed in many political declarations), as well as whether any new initiatives will involve works under copyright and therefore need for information on rights and rightholders and possibly for clearance mechanisms for digitisation and making available.

¹⁰ Available at: http://cordis.europa.eu/fp7/ict/telearn-digicult/publications_en.html.

¹¹ EDL Report on Digitisation in European National Libraries 2006-2012, available at: <http://www.theeuropeanlibrary.org/portal/organisation/cooperation/archive/edlproject/downloads/EDL-D3%201%20final.pdf>.

As far as budgets are concerned, the EDL report mentions that only a few libraries have an explicit and ongoing budget for digitisation, of between 100,000 and 200,000 € per year. The Numeric study estimated that all in all libraries had a budget available for digitisation of some 33 million € (of which 25 belonged to national libraries). If compared to the cost of digitising a book, these figures can be put into perspective: estimates in the EDL study mention from a minimum of 0.08 to 0.15 € per page (16 to 30 € for an average 200-page book) to as much as 1.30 € per page (260 € per 200 pages) considering all costs related - adding metadata, OCR (Optical character recognition), selection and preparation of materials, quality assurance, hard and software, storage and preservation, etc. The Numeric study reported a median cost among libraries ranging from 0.10 € per page (higher education libraries) to 0.70 € per page (public libraries), with national libraries in between at 0.48 € (respectively 20, 160 and 96 € per 200 pages). Budgets so far allocated seem therefore insufficient for real mass digitisation projects.

Even more uncertainty marks the possible inclusion of works in-copyright in the digitisation plans of libraries. Upon enquiring with the national libraries of some of the largest countries in the EU (UK, Germany, Spain, France), the first finding is that where there are mass digitisation programmes, these do not include copyrighted works. Public domain works are given the priority, for various reasons, ranging from the urgency of preserving the oldest works to the possible difficulties in handling copyright related aspects. ARROW is set to provide help both by reducing the cost of digitisation and by facilitating the inclusion of in-copyright works in digitisation initiatives, while allowing the treatment of large volumes of requests.

4.1.2 Other cultural institutions

Other cultural institutions, while not being focused on books, hold book collections, and may undertake their digitisation at some stage, autonomously or in the framework of the larger national / European plans. These include museums, archives and similar entities; though on a different scale, similar considerations apply to them as to libraries.

4.2 Digitisation projects by private entities

The interest of private players can be to offer content commercially or to provide access in the form of search results; in both cases, the availability of a very large pool of resources is paramount, hence the push for digitisation.

As already clarified, ARROW is designed to be business neutral. Any player in the e-book market, commercial or non-commercial, will potentially benefit from the use of ARROW. In addition to serving public institutions such as libraries, it may also be used by any private company interested in the digitisation of books, rights information and rights clearance including search engines such as Google, as well as others in the e-book or e-value chain such as Amazon, Microsoft, the Gutenberg Project, the World Digital Library, the Internet Archive, publishers and so on.

Such companies may also be involved in projects on digitising and making cultural heritage available to the public through public-private partnerships, which are becoming

increasingly more important in the field of digitisation due to lack of public funding being offered towards library digitisation projects. Just as libraries, any private company wishing to scan and use books needs to search (diligently) for rights, creators and publishers, rights status and rights clearance centres. This is the core service offered by ARROW.

4.2.1 Rightholders in general

Though the direct value is for those undertaking digitisation initiatives, rightholders have an interest in seeing ARROW work properly. However, the direct financial interest that can justify paying for the core ARROW services is limited. Diligent search is a user duty. The publisher community calls for proper implementation of that, but individual publishers have no incentive to pay for it unless they use it themselves to clear rights (which in some cases can happen). Some revenue could, nonetheless, be expected from requests for information on the status and clearance of rights by commercial entities for the reuse of works which are orphan or out of print.

4.3 RROs

RROs are expected to administer rights in relation to orphan works; assist authors and publishers in the licensing of out of print works and alternate format production of works to make them accessible to visually impaired persons; license digitisation projects, etc. This may also entail some limited use of ARROW to identify right status and rightholders, when users address not only licensing but also information requests to RROs.

If so, ARROW will represent a cost saving factor for RROs who have to undertake diligent searches. In case of licensing schemes, RROs would use a part of ARROW's revenue stream to cover the cost of administering such schemes and hold money in escrow to be paid out to reappearing rightholders to orphan works.

Benefits to authors and publishers via RROs could include: (i) licensing of digitisation and/or making available of copyright works; (ii) licensing of out-of-print works; (iii) administration of orphan works registries; (iii) administration of out-of-print registries; (iv) carrying out diligent search for rights, rightholders and rights status; (v) providing information based on the data in registries. Users in general will benefit from the RRO indemnity for the use of orphan works in case such a scheme is put in place.

In sum, the value proposition of ARROW for RROs consists in delivering an interoperable system to manage rights information and allowing the maintenance of orphan works registries.

5. ARROW use cases

The first occasions to assess the reliability of the business model described in Chapter 4 are three use cases that are emerging at the beginning of 2011 in France, Germany and UK, i.e. three out of the four countries where the system has been piloted and validated. They are three digitisation programmes planned on two cases (France and Germany) on a

massive scale and with new legal support, in the third case, a niche programme, planned by an individual charity organisation in a particular field.

In all the three cases the ARROW consortium is involved in the design of the programmes' implementation, to assess if the system is able to provide real value to them and – as a consequence – to assess the business model.

5.1. France

In France the Ministry of Culture and publishers and authors representatives are finalising an agreement to facilitate the right clearance on books published in the twentieth century to serve a broad plan for digitisation of the French cultural heritage. The main characteristics envisaged for this plan are the following:

- The National library and the publishers association collaborate on the programme objectives and the agreement intended to cover rights clearance, through a sophisticated model where libraries will be in charge of a large scale digitisation plan to make available on line currently out of commerce books at work level.
- The introduction of a compulsory collective management in the French legislation will allow the clearance of rights on out of commerce works published in the twentieth century in France. This will not require reaching agreements with all the rightholders on a title by title basis, since the entitled RRO will be able to grant licenses also for non mandating rightholders
- However, individual rightholders maintain a first option in exploiting their works online and, in any case, the right of opting out from the agreement, withdrawing all or part of their books from the licensing scheme. This implies the need of a massive notification of rightholders, possibly on an individual basis, so as to facilitate rightholders in exercising the options.

There is broad awareness in the stakeholders' community that the implementation of such a sophisticated agreement requires an infrastructure for the management of data flows. The key elements in the agreement are the identification of out of commerce *works* using data related to commercial availability of *books*¹².

In France a discussion has been initiated to assess if ARROW can be the best solution for serving the implementation of the programme. The discussion was essentially aimed at answering this question: what is the value that ARROW offers in comparison with realistic alternatives?

The ARROW consortium proposed adopting this approach because it is the only one consistent in the long term. The analysis started in November 2010 and has not been completed yet, as it depends on the timeframe of the programme itself, which is still under discussion. The initial findings are the following:

¹² Since the out of commerce status, as defined by the HLEG, is related to the decision of rightholders not to further exploit a work, commercial availability can be seen as a proxy of the occurrence of such decision.

- The flexible ARROW architecture fits the general requirements of the agreement. Since the current version of the system has been implemented before knowing the terms of the agreement (which in part are still unknown), ARROW provides the general infrastructure that now it will be necessary to tailor to be used in this specific context. Technical analysis comparing the expected requirements and ARROW functionalities shows that the “customisation” requires minor implementations. In most of the cases, it will be necessary only to set some parameters in the algorithms that have been already developed. In a few cases some additional features should be implemented to answer very specific needs, but the integration within the ARROW workflow does not create problems, thanks to the modularity of the architecture.
- Since the data necessary to manage the programme are dispersed in several sources, the integration that ARROW provides of those sources (and additional ones, if required) is a value because otherwise the sources would need to be queried separately, so increasing the need for manual handling of data (and therefore costs)
- The identification of the out of commerce status at work level is a key functionality of the ARROW system. This fits the first main requirement of the programme, as far as only out of commerce works will enter the licensing scheme. The integration of this functionality into a single workflow including some form of rightholders search, allows the management of also any notification programme that the agreement would envisage in a coherent sequence.
- The trans-national nature of the ARROW system might be a value for two reasons: (i) it allows, if required, the determination of the commercial availability of works not just in the French market but on trans-national basis, which would be coherent with the intention of making the works available online without territorial limitations; (ii) it enables informing foreign rightholders, possibly through the RRO network, for example in the case of a translated work, where the original author has an important stake, or for books initially published in France but then re-published by a foreign publisher.

This has consequences for the ARROW business model. Lessons learned up to now, in comparison with the general assumptions described in Chapter 4, are:

- The importance of data management in the large scale digitisation plans means it will be necessary to allocate an appropriate budget to this phase. This is true also in programmes that exploit simplified schemes for rights clearance. Therefore, the assumption made in our business plan in this respect seems to be confirmed.
- The European coverage of the system can be a value to improve management of a national initiative, because of the inevitable trans-national nature of making works available online without territorial limitations.
- As for the first uses of the system, the costs of the management of a complex European system may be too high if they have to be covered by only one country, but becomes very competitive when it is possible to exploit the economies of scope and scale deriving from the European (and possibly international) scale itself. Therefore, si can also be assumed that there will be a need of general public support up to the moment when the demand for the service will come from a significant number of countries seems to be confirmed.

5.2. Germany

The German use case has many similarities with the French. Rightholders' associations and the national library agreed on a scheme for simplified right clearance for books that contain out of commerce works and were published in Germany up to 1965. The right to withdraw own works from the scheme is granted to all rightholders.

Again, though with some differences (the most relevant is the cut off date, but it is not the only one) the main requirements emerging from the agreement are the correct determination of the out of commerce status at work level and the support to the information flow between the library and individual rightholders to increase full acceptance of the programme, though a formal notification programme at individual level is not provided by the agreement.

The dialogue between ARROW and the working group set in the stakeholders' community to design the programme produced the same results as in France, so increasing the evidence of the assessment.

Particular characteristics in this case are:

- The level of integration of the German, Austrian and Swiss markets suggested the integration also of the information flows between those countries, which is a feature that ARROW can implement.
- There is a specific request to integrate in the workflow also some rights information related to images embedded in books, which is possible thanks to the collaboration of the CMO specialised in this area in Germany.

Both requirements will be approached during the ARROW Plus project, which includes the incorporation of Austrian data sources in the system and a pilot related to images.

An additional lesson learned analysing this case is that the system is flexible enough to manage in parallel different requirements coming from different national environments while maintaining a European scope and a single basic infrastructure.

5.3. UK

The UK case is different from the others. A private charity – Wellcome Trust – is planning a digitisation programme in a specific scientific area (“Modern Genetics and its Foundations”) and needs to implement it without having any new legal support similar to that planned in France or Germany.

To assess the feasibility of a large digitisation plan including also copyright works, in terms of time and costs to undertake a diligent search, the Wellcome Trust has set up a pilot programme comprising around 1,400 monographs from the library holdings related to genetics and inheritance, published between 1850 and 1990. Results from this pilot will be used to orient the future digitisation policy of the organisation. Works that are found to be still commercially available will be excluded from the digitisation programme.

So, the rightholders search should be complete and licenses must be gathered by individual rightholders, using the intermediation provided by the local RROs. It is to be

noted that, in this case, in the absence of any legislation about orphan works, the information provided by the system is valuable for the user only when a rightholder is found ready to grant the license.

This case was the most typical before the recent developments in some countries, and thus it is not surprising if ARROW also fits the requirements of such a programme. However, there are important lessons learned, also in this case, from the viewpoint of the long term sustainability of the system.

- When the collection to be digitised is defined according to a certain area of interest (a discipline, in this case) rather than a country of publication (as it is for programmes focused on the cultural heritage of a country), the trans-national nature of the service required is even more important. From the business viewpoint, this means that the value of the system depends on the level of coverage of most important data sources in different countries. Again, since the integration of data sources (and the licensing for the use of their data¹³) is expensive, reaching the break even point is realistic only if and when the demand for the service is massive, which again confirms our statement about the need for a transitional public support for the service.
- The nature of those plans includes also the need to integrate non-European data sources, and in particular the US ones. The ARROW system is open to this integration, and – for instance – an interest in collaborating has been expressed by Copyright Clearing Center, the US RRO. As books have, for centuries, been the main vehicle of spreading culture across boundaries (so that good libraries also have foreign books in their collections), and as making available on the Internet is a new way to reinforce this, we need to adopt a global vision of the problem. The distributed model that characterises ARROW is the most suitable to have this approach.

6. ARROW's business architecture

Behind the service it is going to provide and its immediate manifestation - the online user interface that allows searches for rights information - the ARROW system also consists of a number of concrete elements. It comprises common and specific pieces of software, hardware equipment, a series of functionalities and therefore a number of people to maintain the entire infrastructure and provide the related services. All of these elements in turn constitute ARROW's business architecture, a description of the system components that involve costs and contractual relationships.

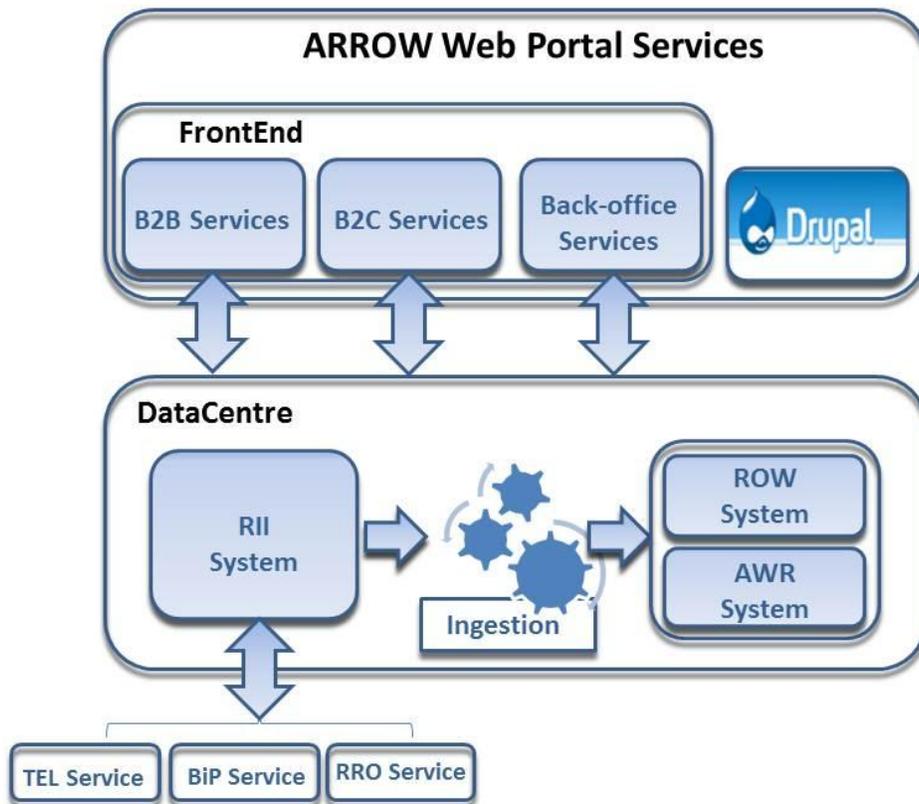
The ARROW System is a comprehensive service to support any diligent search model adopted by libraries, by facilitating the identification of rightholders (authors/publishers) and the identification of the rights status of works with particular regard to orphan and out-of-print works.

¹³ To give an idea of the scale of the problem, to serve the Google Settlement Agreement announced two years ago (and still not operative), which is global in its nature, Google declared that they spend annually around 2.5 million dollars just in licenses for using data. The Settlement included a simplified scheme for clearance, so it does not include proper diligent search for rightholders.

The ARROW System is made up of the following components:

- ARROW Web Portal Services
- The Rights Information Infrastructure (RII)
- The ARROW Work Registry (AWR)
- The Registry of Orphan Works (ROW).

The figure below shows a schematic representation of the ARROWARROW system. The results and the information collected during the RII workflow form the basis for the AWR and therefore for the ROW which is a subset of the above mentioned AWR as described in the following paragraphs.



The **Rights Information Infrastructure (RII)** is at the heart of the ARROW system. The RII is the backbone and the engine that enables ARROW to query and retrieve information from a multiplicity of data providers, in multiple formats, to make the formats interoperable, to process this information and take decisions on the successive elaboration and finally to exchange information according to a planned workflow. Building on the RII, the ARROW System receives a request for permission to digitise and use a manifestation of a work (for instance a book) from a library and after querying the data providers included in the workflow and elaborating the gathered results, provides information on the work rights status.

To simplify the complexity of the system, the workflow can be divided into three main processes corresponding to the three domains involved, each made up of further processes that contribute to the output. Each process is supported by a well-defined set of ARROW messages¹⁴.

- The **first main process** takes place in the library domain and involves The European Library (TEL) as main actor and the Virtual International Authority File (VIAF) as source for authors' information (supported by messages M2 and M4); this phase also involves the main clustering activity.
- The **second main process** takes place in the Books in Print domain and involves BIP organisations or databases in each of the countries included in the ARROW system; it adds further information to the output obtained from the previous process in the library domain (supported by message M6).
- The **third main process** takes place in the Reproduction Rights Organisation domain and involves RROs organisations or databases in each of the countries included in the ARROW system; it adds further information to the output obtained from the previous **process** in the library and BIP domains (supported by message M7).

As a result of the abovementioned three processes, at the end of the ARROW workflow, the following pieces of information are retrieved in the message exchange:

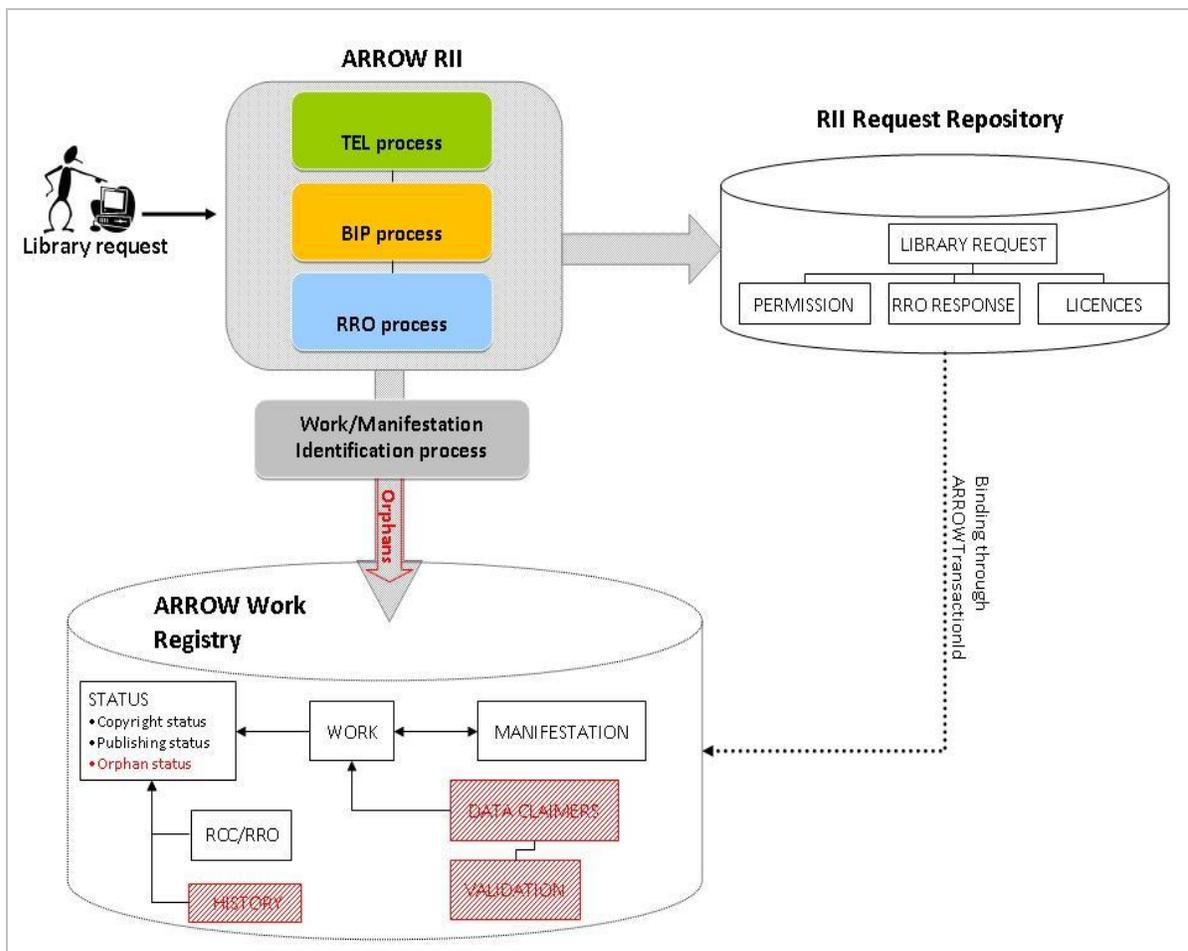
- Work information
- Manifestation information
- Relation between each manifestation and the work they belong to
- Relation between works
- Authors and other contributors information
- Relation between each identified author and the work they have contributed to
- Relation between each piece of information (work, manifestation, author) and the reference source that provided that information (TEL, VIAF, BIPs, RROs)
- A set of so called ARROW Assertions on each work: Copyright Status, Publishing Status and Orphan Status

¹⁴ For a detailed description of the ARROW message suite, see D4.3.2 *Specification for metadata messaging formats* available for downloading in the Resources area of the ARROW website (www.ARROW-net.eu).

The initial library request, including the permission request, the information gathered and inferred during the TEL and BIP processes and the RRO answer are stored in the RII repository.

The ARROW Work Registry (AWR) stores and maintains all these pieces of information for every request processed by ARROW. The Registry of Orphan Works (ROW) is based on a subset of the AWR, respecting specific criteria that will be made publicly available to specific categories of users for specific purposes.

The figure below provides a high level overview of the **ARROW Work Registry (AWR)** and its relation with the RII.



At the end of the ARROW RII workflow, the RII comes out with two different outputs: the first one constitutes the basis for the RII repository, while the second one those for the AWR. The AWR stores all the relevant pieces of information collected by the RII workflow in a structured way that allows the retrieval and use of that information in the framework of ARROW services:

- Work metadata
- Manifestation metadata
- Authors and other contributors metadata
- The ARROW Assertions on each work
- Reference source (TEL, VIAF, BIPs, RROs) of work metadata, manifestation metadata, authors and contributors metadata

The implementation of the Rights Information Infrastructure, and in particular the ARROW Work Registry (AWR), constitutes the foundation for the **Registry of Orphan Works (ROW)**. The ROW core database in other words can be seen as a view of the AWR, result of the RII workflow. Each library request instances a new ARROW workflow. The RII workflow stores all the responses obtained from the data providers and also extracts and stores all the necessary metadata about works and related manifestations to be stored in the AWR, as explained above. Being a subset of the AWR, the ROW starts ‘empty’ and gets populated by digitisation requests being processed through the ARROW system in an automatic way whenever the output of the RII process indicates that the work could be an orphan.

Since the ARROW System runs at European level and the RII elaborates requests from different European countries, ARROW is designed from the beginning to support both the setup of a centralised infrastructure for the ROW and the set up and integration of other orphan work databases on a country basis (National ROWs). Under this perspective ROW functionalities are designed to consider country specific requirements to enable system scalability and interoperability with National ROWs.

6.1 Software

The architecture of the ARROW system comprises a set of software components, each one accomplishing a different task, such as asynchronous message management, data storage, workflow management and services for querying data.

A first list of such components is presented here in detail and will be further revised at the end of the ARROW project when all the necessary analysis and implementations have been provided.

The main components of the ARROW Web Portal Services are hosted in Cineca and are listed in the table below:

- FrontEnd ARROWPublicService Client
- FrontEnd ARROWReviewService Client
- FrontEnd Presentation Layer component
- FrontEnd Service
- FrontEnd Data Access Object
- FrontEnd Authentication
- Drupal CMS

The Right Information Infrastructure is made up of the following components:

- RII Authentication / Authorization
- RII Web Service::Frontend
- RII Web Service::Provider
- RII Queue
- RII Asynch Listener
- RII DataCentre::Workflow Engine
- RII DataCentre: Activity
- RII Data Access Object (DAO):
- RII Data Model
- RII Connector Manager
- RII Client
- RII Polling

A list follows of other ARROW RII components external to the ARROW system (TEL Service, BIP Service, RRO Service) which were conceived, designed and implemented during the ARROW project with the aim to serve the ARROW system and hence are part of the results achieved in the project.

The main components of The European Library technical infrastructure that are used for ARROW and are hosted in TEL are:

- The European Library MARC central index
- Work Matching Pre-processor
- Similarity Search Engine
- Metadata Repository
- Matching and Clustering Engine
- VIAF Connector
- TEL ARROW Connector

The main components of BIPs and RROs data providers are:

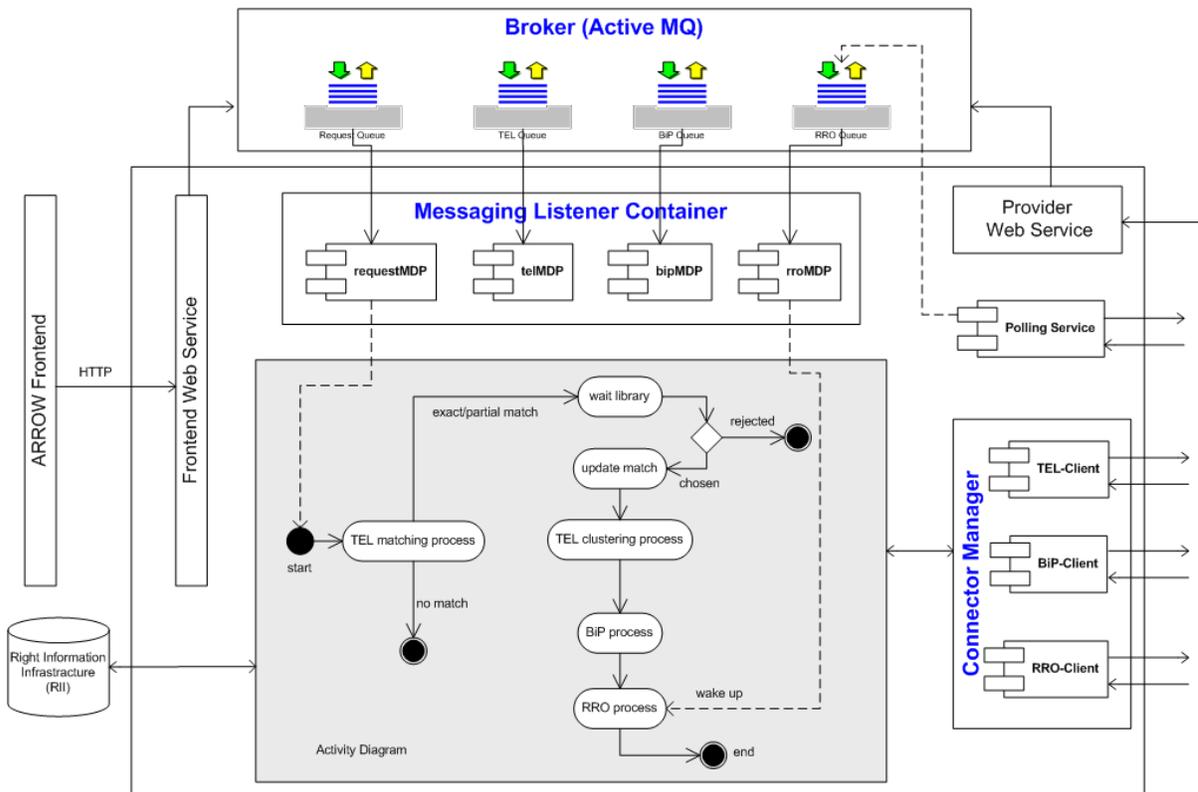
- VLB web services enhancements
- Standard BIP query WS
- Standard RRO query WS
- Standard BIP response WS client
- Standard RRO response WS client
- BiP or RRO Internal ARROW workflow manager

The ARROW Work Registry and the Registry of Orphan Work are made up of the following components

ARROWARROW

- AWR/ROW Web (B2C)
- AWR/ROW Web Service (B2B)
- AWR/ROW Service
 - AWR/ROW feeding
 - Identification handler
 - History manager
 - Claiming manager
- AWR/ROW Model
- AWR/ROW DAO

The figure below provides a more detailed overview of all the main RII software components and relationships between them. It shows a simplified Datacentre architecture containing some of the components described previously as well as the current workflow represented by the activity diagram highlighted in grey.



The ARROW Datacentre has the main task of managing the whole workflow, retrieving and sending data from and to data providers. It stores such data and all the data generated by the system's activities. The relational DBMS used for the ARROW repository is Oracle.

Besides that, ARROW runs also a website, which provides information about the project; partners upload news documents on it. This as well has to be hosted and managed, including from the editorial point of view, which involves time and costs.

Bibliographic records from Europe's national libraries are a key data source for ARROW. This data is used to both identify the specific edition of a book that a library wishes to digitise and to identify all other editions of the same book that potentially share, totally, or in part, the same intellectual work, for further processing in the ARROW workflow.

ARROW uses The European Library¹⁵ (TEL), an operational service which aggregates the catalogues and digital collections from Europe's national libraries, as the single data source for library-domain bibliographic data. The European Library team is also responsible for the work metadata extraction, matching and clustering components of the ARROW system.

The European Library's central index is core to The European Library's operational service and is part of the standard technical architecture.

The other components, work matching pre-processor, similarity search engine, work metadata repository, matching and clustering engine and the TEL / ARROW connector, have been developed specifically by The European Library team for ARROW. The maintenance of these components should be included in the operational costs of the ARROW system.

The most significant effort required by The European Library team is the integration of new countries into the ARROW workflow. This requires pre-processing and indexing of the work metadata, plus the fine-tuning of the matching and clustering engine to ensure the best results in line with the local cataloguing practices of the project. Until now, the costs of undertaking this work have been covered by the ARROW project. However, the additional work required to integrate new countries will need to be considered for the operational service. Periodically, the re-clustering of the national library catalogues will be required. It is anticipated that this activity can be included in the on-going maintenance costs of the operational system.

Once its form and function have been established, the Registry of Orphan Works that is foreseen as one of the outcomes of the work of ARROW will potentially have some software requirements as well.

6.2 Hardware

The entire ARROW system is hosted on CINECA sites, with the exception of the ARROW RII components external to the ARROW system (TEL Service, BIP Service, RRO Service). Cineca provides both the staging and production environment of the ARROW system and of the corresponding databases. Cineca maintains the services, this means the provision of following services: hosting and maintenance, Ordinary maintenance, Integrity and Confidentiality of Data through the CINECA Data Centre, Network infrastructure.

¹⁵ For more information about The European Library please refer to: <http://www.theeuropeanlibrary.org/>.

The ARROW hardware infrastructure is made of an application component and a database component.

Application-presentation tier:

The ARROW services run on a Linux cluster composed by two servers in configuration of high availability, in order to keep services running even in case of malfunctions, (currently hosted at CINECA) that CINECA specifically designed for high performance and data critical services.

Database tier:

The Database infrastructure is composed by two P595 IBM. The databases are hosted in configuration of high availability on Logical partition of the previous mentioned infrastructure. The files of the databases are on Storage Area network.

6.3 The hosting and maintenance service

The hosting service currently offered by CINECA includes:

- Hardware, operating system, Oracle software licenses
- Hardware and operating systems installation and management
- Oracle database installation, management and hosting in high availability environment
- Application software installation and management
- Systems and infrastructure tuning and customization
- Systems monitoring services
- Daily data saving and restoring (including incremental backup and disaster recovery)
- System assistance and network management from 8.00 to 19.00 CET on working days
- Security assessment and monitoring, including server and service monitoring

The software licences included are:

- Oracle 11 (Enterprise Edition) licence
- Operating system licence

The European Library incurs charges for the hardware and hosting of the ARROW-specific components of The European Library technical infrastructure. These costs will need to be included in the operational costs for the ARROW service. The European Library central index is part of The European Library operational service and will not incur additional costs for ARROW. The European Library uses open source technology to provide the ARROW service. There will therefore be no licensing costs needed.

Integrity and Confidentiality of Data through the CINECA Data Centre

The hosting and maintenance services also entail other activities (all currently performed by CINECA through its storage and network infrastructure and through procedures followed by its personnel) related to the integrity and confidentiality of data:

- Data backup and archiving (two different services: backup/restore and archive/retrieve).

Network infrastructure

The CINECA Data Centre that provides the services has a network infrastructure that allows benefiting of different types of connectivity: Local area network connectivity, Internet connectivity and Firewall and VPN (Virtual Private Network) services. All net equipment (routers, switches, firewalls) is configured as clusters for high availability in order to guarantee connectivity also in case of hardware and software failures or maintenance activities.

6.4 Office infrastructure

Once it becomes a stable, up and running entity, ARROW will need some kind of normal operative office structure. The size of this structure will depend on the form taken by the relationship of ARROW with its current service providers (see below), and in particular on the balance between the internalisation and outsourcing of its activities; the more tasks are managed directly by ARROW (especially in the technological domain), the larger the structure will need to be. Its costs will depend on the size and location of the office; the overall level of activity of the system could also have an impact on this variable.

6.5 Personnel

To carry out its envisaged activities and tasks, ARROW will need to employ personnel; while for the time being all the work is done by the project partners and a few external subcontractors, once ARROW becomes an independent entity it will have to be endowed with its own human resources. Personnel costs will be a regular cost category; however, the exact level of personnel required will depend on the degree of outsourcing of ARROW's activities and on its overall level of activity.

On a short and medium term, apart from the running of the technical infrastructure itself, other tasks required to be carried out to run ARROW ensue from the following activities:

- Manual intervention in some phases of the work-flow (to be identified during the validation)
- Day-to-day administration, office management
- Help desk for users of the system
- Legal services
- Communication, marketing

Not all of these, however, have to be internalised, as the need for some services (legal, marketing) might be limited.

Regarding the help desk, being the potential number of users of the system quite large, it's possible to envisage that ARROW will have to handle large amount of requests and different types of questions. Some persons can be responsible of the first-level help desk

in order to answer the most commonly asked questions, or provide resolutions that often belong in an FAQ base. If necessary an issue tracking system will be used to allow a logging process to take place at the onset of a call. If the issue isn't resolved at the first-level, it is escalated to a second, where more difficult calls are handled, which often can deal with software specific needs, such as updates and bug-fixes. In these cases it can require specialised technical assistance.

6.6 The ARROW social infrastructure

The ARROW infrastructure will be organised as a federated rather than a centralised system. It will be a network of resources, accessible from a single access point. Because of its network nature, the ARROW system does and will consist also of a set of relationships with other players.

This ARROW “social infrastructure” is also to be organised in order to avoid vulnerability, in particular when important elements of the ARROW system are based on relationships with third parties. ARROW needs the involvement of such key players in its governance and to design stable contractual agreements with third parties that are crucial to provide the service (e.g. The European Library, VIAF, Books in print providers, RROs).

The relationship with other entities is foreseen to take the form of a network of contractual links, which will constitute yet another cost category. The first such relationship to be settled will be related to the legal form ARROW will take as an independent entity. The company will be set up as a foundation or a company limited by guarantee (Ltd) depending on where the headquarters will be located. Some set up costs will incur.

As for the core activities of ARROW, a phase of the workflow is managed through The European Library. The relationship between ARROW and The European Library (TEL) will therefore be formalised in order to ensure long term sustainability. TEL bears an important part of the software and hardware costs, which shall be compensated once the system is up and running.

In a second phase, the relations with data providers are to be addressed, too as ARROW bases its work on the collection and connection of information from a number of sources, namely:

- bibliographic and authority data: national libraries (through TEL and VIAF)
- commercial availability and publishers' details: Books in Print catalogues
- rightholders' details and licence information: RROs repertoires.

Agreements with BiP, RROs and VIAF are therefore to be considered. ARROW will have to acquire licences on third party data, for instance for the BiP catalogues. In some cases this kind of relationship could be settled as an exchange of services (mutually granted access to information).

ARROW will use the books-in-print database services of VLB in Germany, Electre in France, Dilve in Spain and Nielsen in the UK. As the system expands to further countries, costs of other BiP databases will be possibly need to be taken into account.

Other services will be provided by external companies - including existing partners - through outsourcing, such as the hosting and maintenance of the ARROW system. These are currently provided by CINECA, which could keep offering them on a contractual basis.¹⁶

The same applies to the services related to the marketing and communication for promoting the adoption of ARROW as an instrument for diligent search; these are foreseen to be outsourced.

Software and any other kind of licences needed will be part of the ARROW social network as well, and so will the choices of the contractual relationship with the personnel that will be employed to manage the system and all the related activities.

Intellectual property

The working of ARROW will produce elements on which issues of intellectual property might arise. These will include non-original databases, works or metadata newly created within the ARROW process is reserved and other pieces of software (like the aforementioned clustering algorithm). For all of these an IP policy is currently being devised and will be managed as appropriate.

ARROW will not hold ownership of the original databases or data used either during or after the project. These databases and data will remain the property of the information providers and they will only be used, both by the ARROW partners and the end users for the exclusive purposes of the ARROW search.

7. Funding model and budget

7.1 Financing sources

Public/indirect

As a project under the eContentPlus programme ARROW is financed 50% by the European Commission and managed by a consortium of partners. Several original partners, together with a number of new ones, have applied for a 2.5 year extension of the project for the broad implementation of ARROW throughout Europe (ARROW plus), after which the system will need a stable and sustainable flow of revenues to cover its costs and a suitable governance model.

While a revenue flow can be envisaged, coming both from the core ARROW services and from its additional value, there is also a rationale for a sustained public funding to the

¹⁶ CINECA's services to keep the ARROW system up and running mainly consist of the hosting and maintenance of the Rights Information Infrastructure, ARROW Work Registry, Registry of Orphan Works and ARROW websites (project, RII and ROW), and the inclusion of new countries.

system. The ARROW services will enhance the EU's cultural policy by facilitating digitisation initiatives and offering solutions to issues related to out-of-print and orphan works and contribute to the realisation of declared EU goals. Equally important, a large portion of the projects that are expected to request the use of the services offered by ARROW are public institution ones.

The political declarations and action plans not only call for enhancing digitisation initiatives throughout Europe, including of copyrighted works - mainly to feed into Europeana, they also explicitly call for support to mechanisms that can facilitate digitisation of out-of-print and orphan works¹⁷ (and for the ARROW project itself as part of this category: "*actions at European level - such as the ARROW project in which rightholders and cultural institutions together address the creation of databases of orphan works - should be backed up by national efforts*"¹⁸).

The aim of the consortium partners is that ARROW becomes a self sustainable system as soon as possible after the ARROW plus best practice project period. Investments made during the ARROW and ARROW project plus periods may be considered as sunk costs. A stable funding provided by the public sector will be required in the introductory phase and the first years of regular operation following the project periods. It is proposed that this be in the form of a flat rate of € 0.0011 per inhabitant payable on the basis of population by the EU Member States, EEA countries and Switzerland against free use of ARROW for their publicly owned institutions. Because of the relatively moderate costs involved and on the basis of cost efficiency the European Commission will be asked to consider whether to meet or integrate the costs of the Member States. Most of the funding and all of it at the end of the introductory phase will be payment for services offered by ARROW.

Also, the use of ARROW depends on the number of digitisation projects in the EU Member States. It is required that the Commission and EU Member States follow up their declarations with concrete plans to initiate such projects which will then again make use of the ARROW services.

In the long run the European Commission and the EU Member States are expected to continue their contributions to the financing of ARROW through payment for the use of services as long as and to the extent to which their institutions make use of them. Economies of scale makes it cost efficient for EU Member States and the Commission to pay directly for the services required by institutions that they own rather than leaving this to the individual institutions.

¹⁷ See, for example, the Council conclusions of 20 November 2008 on the European digital library Europeana, which invites the Member States to "*establish mechanisms to facilitate digitisation and online access to orphan works and to out-of-print and out-of-distribution works, while fully respecting right holders' rights and interests*".

¹⁸ "Europe's cultural heritage at the click of a mouse - Progress on the digitisation and online accessibility of cultural material and digital preservation across the EU", Communication from the European Commission COM(2008)513, 11/8/2008.

Governments which plan to fund national digitisation programmes should be advised to create a separate budget line for “diligent search” in general or to fund ARROW directly in order to support the diligent search in institutions they own. EU Member States, EEA countries and Switzerland could pay an annual fee to ARROW to meet the costs by public institutions and allow them to make use of defined ARROW services (search for rights, rightholders, rights status and in the orphan works registries) for free. The fee should be calculated on the basis of the estimated proportion of the total requests to ARROW by public institutions divided between the countries on the basis of population.

As previously said, a yearly contribution from the European Commission on behalf of all EU Member States would be a cost-efficient solution. The Commission is requested to contribute to 50% of the cost of the establishment of ARROW and its operation for the first five years of the operation. In a longer perspective, co-funding by the Commission may be required to maintain cross-European interoperability. Furthermore, some countries (in particular where the national market is too small to justify the investment) risk to be excluded from a service provided only on commercial or quasi-commercial basis and could therefore be in need of some European funding.

Private/direct

Another main source of funding for ARROW will be direct payments by the users of fees per use of its services, first and foremost for making searches through the ARROW infrastructure. This is likely to apply to any private users of ARROW.

ARROW is “business neutral”. This does, however, not mean that every user will always be offered the same tariffs for the use of its services. Different rates may apply depending on factors such as type of use; economy of scale, degree of public service offered, etc. Frequent and high volume users will also be offered access through subscription models including in the form of a monthly or yearly fee. The fee payable by public institutions will also consider any payment made by the EC or national governments. The payment model will allow national public institutions to use some or all of the core ARROW services for free where the EC, national governments or both in combination subscribe to the use of ARROW and pay the subscription fee on their behalf.

Payments in kind

The usage of the ARROW system will be against payment, directly or indirectly, by its users. Payment can also be in kind. ARROW is going to rely on third party databases for its works, some of which are usually accessed at a fee. A good proportion of the data providers are likely to be in turn users of the ARROW services (national libraries in particular, directly and via TEL) or to somehow benefit from the system operations, both in terms of its core services (for example RROs, in case of increased activity) and its additional value (clustering and facilitating ISTC registration, for example, can be of interest for Books in Print databases). When considering licences on third party data, ARROW could take into account these elements in the fee setting.

7.2 Fee models

There are two basic models for the direct payment of ARROW services by its users: by actual use and by subscription. In its purest form, the former would link the amount to be paid to the number of queries performed simply by multiplying this number by a given individual fee. As the current project is intended to be replaced by a best practice project aiming at a broad implementation of the ARROW system the current project partners have decided to await further experience before fixing the tariff per use. It is, however, assumed that the fee would be small compared to the total costs of digitisation and also smaller than the cost of diligent search through other means under most circumstances. It is also foreseeable to have different tariffs for queries that result in finding a rightholder or not, for example X_A and X_B (with the latter possibly zero).

A subscription would be a yearly sum that users would pay to access the services of ARROW, to some extent independently from the actual usage. Its level, however, could be set more easily and adequately by establishing different tariffs according to the potential use of the system by the customer. Thus, a possible scheme would be:

Number of queries allowed	Subscription fee
Up to Q_1	X_1 (possibly free)
Between Q_i and Q_{i+1}	X_{i+1} (we can set as many layers as appropriate, with $i=1, 3, \dots, n-1$; n identifies the maximum number of queries after which the top tariff would apply)
Unlimited queries	X_{MAX}

The fees could be calculated both in advance, giving the user the possibility to make up to a certain number of queries per period, or be calculated afterwards; in this case, there could be a fixed fee depending on the total number of queries made and a variable fee related to the number of queries resulting in the identification of a rightholder. Finally, both possibilities could be offered, for example a subscription scheme coupled with payments per number of queries for usage below a certain threshold.

The contributions by Member States envisaged in the previous section would constitute a form of subscription; again, this could be more or less linked to the potential usage by national institutions, for example in case of proportionality with the size of library collections or of population (in this case seen as a proxy for the benefit that a country would have in the improving of digitisation initiatives). Such fees could be reduced proportionally after a certain period, according to the progress in digitisation.

The system could also foresee that the payment for the ARROW services, in the case of private users, be made as a percentage of the amounts paid to rightholders in case a licensing deal is made (for successful queries resulting in a deal), or as a percentage of the licence fee paid to a RRO or RCC in general in case a mechanism is in place for licensing orphan works (should the result of the query be a presumption that a work is orphan).

As the current ARROW project is likely to be replaced by a best practice project on the broad implementation of the ARROW system it is appropriate to gather the additional experience that a new project period will enable before finalising the outline of a budget for ARROW as an ongoing company.

8. Governance

The ARROW governance model will of course reflect the chosen model of funding. A new legal entity will be established to be the “owner” and manager of the ARROW system; it will be a not-for-profit entity which should be comprised of the existing partners and be open to new members as well. The entity will be set up as a foundation or a company limited by guarantee (Ltd), depending on where the headquarters are located.

- END -