

BRITISH  
LIBRARY

# Seeking New Landscapes

A rights clearance study in the  
context of mass digitisation of  
140 books published between  
1870 and 2010.

**Barbara Stratton**

**ARROW**

Accessible. Register of Rights Information  
and Orphan Works research. Transparency.

**The author, Barbara Stratton is a UK chartered library and information professional and library copyright consultant.**

Barbara Stratton has asserted her right under the Copyright, Designs and Patents Act 1988 to be identified as the author of this work.

© The British Library Board 2011.

Fig. 21 on page 60 reproduced by kind permission of the author and the publisher from Padfield, Tim: Copyright for Archivists and Record Managers. 4th ed., London: Facet Publishing, 2010.

Appendix 6 is © PLS and ALCS 2011

#### Acknowledgements:

The author acknowledges the contributions made to the execution of this study by the following British Library staff: Emma Wilson, Publisher Licensing Support Officer, Chris Morrison, Copyright Assurance Manager, Benjamin White, Head of Intellectual Property and Rachel Marshall, Licensing Manager.

The British Library would like to thank Tom Rivers for his helpful advice, keen eye and support in the preparation of this study.

This study owes much, including its inspiration, to a previous 2005 US report by Denise Troll Covey from Carnegie Mellon University in the USA entitled *Acquiring copyright permission to digitize and provide open access to books*.



This report is published under a Creative Commons Attribution 2.0 UK: England and Wales (CC-BY 2.0) licence. <http://creativecommons.org/licenses/by/2.0/uk>

Disclaimer: The views expressed in the study are those of the author. It should also be noted that rights clearance is not an exact science. Therefore, at times informed decisions were necessarily made based on the partial evidence available, which in turn formed the basis of subsequent decisions made in the process.

# Contents

Executive Summary	4
Key Findings	5
Introduction	6
The context of this study	7
Methodology	14
About the records in this study	16
Results from this study	33
Conclusions	51
Appendix 1	52
Appendix 2	57
Appendix 3	61
Appendix 4	63
Appendix 5	68
Appendix 6	
ARROW: A service in development to assist 'diligent research'	69
Glossary	74

# Executive Summary

## Background

Libraries, archives and cultural institutions throughout Europe have a responsibility to preserve and provide access to their collections for the public good. In order to achieve these goals in today's Internet enabled society they are striving to digitise their collections and make them available online.

As recently expressed to the European Commission in the Comité des Sages report 'A New Renaissance'\*; the benefits of doing so are great: citizens are able to access material that would otherwise be unavailable to them due to geographic location; economic growth is stimulated through the wide accessibility of cultural goods; and innovative ways of interpreting and building on knowledge and creativity are enabled through technological means.

However there are a number of challenges in mass digitisation of cultural materials that must be addressed if a significant proportion of Europe's culture is to be made available online. Perhaps one of the most critical of these is the process around identification and clearance of rights in material within the last 140 years. Much of this material remains in copyright, meaning it cannot be made available without permission from the rightsholder, although the vast majority of this in-copyright material is no longer in commerce, or was never intended for commercial purposes in the first place. This creates a complex and costly proposition for cultural institutions who find themselves using rights clearance channels and methodology primarily created for commercial negotiations relating to the reuse of individual or very small numbers of copyright works.

This complexity is compounded by the number of 'orphan works' – those copyright works for whom the owner is untraceable – that are held within cultural institutions' collections. As it is not possible to gain the rightsholder's permission to digitise the content it is not possible to digitise the material without effectively breaking the law.

This study was undertaken by the British Library as part of a wider pilot undertaken by the EU funded ARROW\*\* project team. The ARROW system is a network of databases and rights registries designed to enable the identification and rights clearance of works to support mass digitisation throughout Europe. One of its key functions is to support Europeana, the European Commission backed "digital library." The project is a cooperation between libraries and rightsholder groups to find technological solutions to the challenges that clearing rights for mass digitisation presents.

## Aim of the Project

The purpose of the study was to examine the rights clearance process, rather than actually digitise. It used a random selection as possible (see 5.1) of 140 books published between 1870 and 2010 – 10 per decade – from the collection of the British Library. Although the oldest known in-copyright item held at the British Library dates from 1865 it was felt that starting the study in 1870 was sufficiently rigorous to cover the vast majority of in-copyright works. The study sought to determine whether permission could be received to digitise them and to achieve the following goals:

- Identify the copyright status of the material and the proportion of 'orphan works' (those in-copyright works whose owners cannot be identified or located).
- Measure and quantify the level of diligent search currently required to undertake mass digitisation of material from the last 140 years.
- Understand the relationship between commercial activity of library collection items and wider availability.
- Compare the outputs of a 'manual' diligent search for rights holders against the 'automated' ARROW system.

---

\* [http://ec.europa.eu/information\\_society/activities/digital\\_libraries/doc/reflection\\_group/final-report-cd53.pdf](http://ec.europa.eu/information_society/activities/digital_libraries/doc/reflection_group/final-report-cd53.pdf)

\*\* Accessible Registries of Rights Information and Orphan Works towards Europeana: <http://www.arrow-net.eu/>

## Key Findings

- 1 29% of the books were out of copyright, 57% were identified as in copyright and the remaining 14% had an unknown copyright status (presumed to be in copyright for the purposes of digitisation).
- 2 Of the total number of potentially in-copyright works 43% were orphan works, equating to 31% of the total sample. They ranged in date from the 1870s to the 1990s.
- 3 The type of publisher had a large impact on whether works were orphaned with self-published works accounting for 51% of all orphan works in the study.
- 4 Only 21% of the books in the total sample were still in print, almost half of which were in-copyright titles published between 1990 and 2010. Most of the rest were public domain material from the decades around the turn of the 19th and 20th centuries, a considerable proportion of which are available as print on demand rather than warehoused print runs.
- 5 The decade which featured the highest proportion of definitely in-copyright orphan works was the 1980s (50%) which demonstrated that although age may be a factor in whether a work becomes orphaned, even material from the recent past is clearly affected by this issue.
- 6 56.5% of books in the sample were published by non-mainstream publishers such as professional associations, institutions and political organisations.
- 7 Permission to digitise was sought for 73% of the books in the sample. Of these:
  - rightsholders gave permission for just 17% of the books to be digitised;
  - permission was not granted for 26% of the titles;
  - for 26% of the titles no response was received;
  - rightsholder contact details for the remaining 31% of the titles could not be located.
- 8 On average it took 4 hours per book to undertake a “diligent search”. This involved clarifying the copyright status of the work and then identifying rightsholders and requesting permissions.
- 9 In contrast the use of the ARROW system took less than 5 minutes per title to upload the catalogue records and check the results.
- 10 During the study the diligent search and the ARROW system identified the same copyright status for 51% of titles. System improvements made soon after the study closure subsequently brought this figure to 69%. A comparison of the ultimate rights clearance outcome of the works queried via each process showed that in 92% of cases the same result was received.

## Conclusions

The study confirmed through analysis of a representative set of titles published within the 140 years between 1870 and 2010, that rights clearance of works on an individual, item by item basis is unworkable in the context of mass digitisation. Mass digitisation potentially involves the copying and making available of millions of copyright works. At 4 hours per book it would take one researcher over 1,000 years to clear the rights in just 500,000 books – a drop in the ocean when compared to the rich collections of Europe's cultural institutions.

Issues of rights clearance are therefore fundamentally important to the creation of publicly accessible collections of content. The current economic climate, in which funding for cultural and knowledge based services is highly limited, makes it even more important to find efficient ways of making public collections available. The material in these collections has, in a great many cases, been put on shelves and essentially forgotten about. To make it widely available in digital form is to increase understanding of our history, our traditions and the world within which we live. To limit this just to items that are clearly in the public domain through the lack of efficient rights clearance mechanisms would mean omitting the 20th century from this understanding. This would be a tragedy.

However, the results from the comparison with the ARROW system were very encouraging. The potential that a single automated diligent search is all that is needed to clear rights – a search for which a user need invest only minimal time in uploading records and reviewing responses – makes mass clearance of rights an achievable goal. ARROW also has the potential to play a significant role in any legislative solution to orphan works, acting as a single registry through which cultural institutions could identify works they have digitised and rightsholders could claim any works to which they own the rights.

In summary the study showed that more efficient ways of clearing rights and providing cultural institutions with legal certainty over their activities are needed to ensure that highly valuable research materials don't remain out of reach of the vast majority of citizens.



# Introduction

This study is based on research commissioned by the British Library as part of the EU-funded ARROW project to examine the diligent search and rights clearance processes required to enable the Library to digitise a selection of holdings as part of a mass digitisation project and make them available to the public online. The study was conducted over an eleven-month period between March 2010 and February 2011 and provides soundly based and quantified evidence about the ease or difficulty of manually identifying rightsholders and obtaining permissions, as well as the time spent on specific rights clearance tasks.

The subject of the study was a computer-generated random sample of catalogue records of books and other monographs<sup>1</sup> held by the Library which had been published in each of the 14 decades from 1870 to 2010. The sample consisted of 140 works, 10 from each decade (a full list of the works included in the sample can be found in Appendix 1).

The methodology was inspired by the research conducted in the Carnegie Mellon University Libraries' *Random Sample Feasibility Study* which Denise Troll Covey reported on in 2005 in her study *Acquiring copyright permission to digitize and provide open access to books*<sup>2</sup>.

This study is part of the European Commission supported **ARROW** (Accessible Registries of Rights Information and Orphan Works towards Europeana) project in which the Library was a UK lead and partner.<sup>3</sup> ARROW is working to find ways to clarify the rights status of literary texts – mainly books – to facilitate their clearance for digitisation and is building an automated European database of rights to aid the mass digitisation of such works. This work will in the first instance support the *Europeana*<sup>4</sup> digital library, the flagship of the European Commission's i2010 Digital Library Initiative, as well as assist in digitisation at a Member State level<sup>5</sup>.

In what follows we describe how the research team worked through the parameters and relevant copyright issues for the study, went about identifying rightsholders and attempted to obtain clearances for digitisation of the works.

---

1 The term 'monograph' is used by librarians to mean a non-serial publication complete in one volume or in a finite number of volumes.

2 Troll Covey, Denise (2005). *Random sample feasibility study*. In *Acquiring Copyright Permission to Digitize and Provide Open Access to Books*. (pp. 11-21) Washington, D.C., Digital Library Federation Council on Library and Information Resources. Troll Covey's sample was taken from the holdings of the Carnegie Mellon University Libraries. The report also covers two other similar studies: 'The fine and rare book study' and 'The million book project study'.

3 ARROW, Accessible Registries of Rights Information and Orphan Works towards Europeana, is a project of a consortium of European national libraries, European and national publishers and collective management organisations, representing publishers and writers. "ARROW aims in particular to support the EC's i2010 Digital Libraries Project by finding ways to identify rightsholders, rights and clarify the rights status of a work including whether it is orphan or out of print."

More information about the project including a list of partners can be found at: <http://www.arrow-net.eu/>

4 "Europeana enables people to explore the digital resources of Europe's museums, libraries, archives and audio-visual collections. It promotes discovery and networking opportunities in a multilingual space where users can engage, share in and be inspired by the rich diversity of Europe's cultural and scientific heritage."

<http://www.europeana.eu/portal/aboutus.html>

5 The Digital Libraries Initiative "was launched in 2005 as part of the Commission's i2010 strategy to boost the digital economy. It builds on work already done by the Commission over a number of years."

[http://ec.europa.eu/information\\_society/activities/digital\\_libraries/background/index\\_en.htm](http://ec.europa.eu/information_society/activities/digital_libraries/background/index_en.htm)

# The context of this study

## 1 The European Digital Library – Europeana

In 2005 the leaders of France, Poland, Germany, Italy, Spain and Hungary wrote a letter to the European Commission that called for the establishment of a European digital library. The following is an extract from that letter:

*“The heritage of European libraries is of a richness and diversity without equal. It expresses the universality of a continent that, throughout its history, has interacted with the rest of the world. Yet, if it is not digitised and made accessible online, this heritage could, in future, not occupy any place in the future knowledge landscape.”*

**Letter to the President of the European Commission by the leaders of France, Poland, Germany, Italy, Spain and Hungary, 28 April 2005<sup>6</sup>**

This call led to the launch of the European Commission's 'i2010 Digital Libraries Initiative' in 2005 which in 2008 established *Europeana* as Europe's digital library. In 2010, the i2010 Digital Libraries Initiative was subsumed as the 'Digital Libraries Initiative' into the Commission's new, and wider, 'Digital Agenda for Europe', described by the Commission as "Europe's strategy for a flourishing digital economy by 2020."<sup>7</sup> The Commission's Digital Agenda and Digital Libraries Initiative<sup>8</sup> work programmes identified that considerable societal and economic benefit is to be derived from digitising the collections of national libraries, archives, museums and other significant institutions in order to make them available online.<sup>9</sup>

Europeana now includes world famous and diverse objects such as *Philosophiae Naturalis Principia Mathematica* – Issac Newton's treatise on the laws of motion, drawings by Leonardo da Vinci, and Johannes Vermeer's painting, *Girl With A Pearl Earring*. At its launch *Europeana* received over 13 million hits per hour on its first day causing the website to crash.<sup>10</sup> At the time of writing, *Europeana*'s collection had reached 15 million items.<sup>11</sup> Based at the Koninklijke Bibliotheek, the Royal Library of the Netherlands, it is actually not a digital library in itself but a gateway or portal to the content contributed by Europe's major cultural institutions, led by Europe's national libraries, archives and museums. Some 1,500 institutions have already contributed to *Europeana* including the Rijksmuseum in Amsterdam, the British Library in London, and the Louvre in Paris as well as many smaller institutions. Many European cultural institutions are contemplating, or are already engaged in, mass digitisation of entire collections for *Europeana*, some having won funding from the European Commission's ICT Policy Support Programme.<sup>12</sup>

:

---

<sup>6</sup> Passage translated from the original French.

<sup>7</sup> [http://ec.europa.eu/information\\_society/activities/digital\\_libraries/doc/letter\\_1/index\\_en.htm](http://ec.europa.eu/information_society/activities/digital_libraries/doc/letter_1/index_en.htm)

<sup>8</sup> [http://ec.europa.eu/information\\_society/digital-agenda/index\\_en.htm](http://ec.europa.eu/information_society/digital-agenda/index_en.htm)

<sup>9</sup> [http://ec.europa.eu/information\\_society/activities/digital\\_libraries/index\\_en.htm](http://ec.europa.eu/information_society/activities/digital_libraries/index_en.htm)

<sup>10</sup> European Commission (05/05/2010). Digital Agenda: Commission welcomes European Parliament's support.

(Ref. IP/10/538)

<http://europa.eu/rapid/pressReleasesAction.do?reference=IP/10/538&format=HTML&aged=0&language=EN&guiLanguage=en>

<sup>11</sup> Casey, Zoe., 'Europeana crashes to earth.' *European Voice* 21.11.08

<http://www.europeanvoice.com/article/2008/11/europeana-crashes-to-earth/63159.aspx>

<sup>12</sup> <http://www.europeana.eu/portal/aboutus.html>

<sup>13</sup> [http://ec.europa.eu/information\\_society/activities/econtentplus/about/index\\_en.htm](http://ec.europa.eu/information_society/activities/econtentplus/about/index_en.htm). In 2009 the ICT Policy Support Programme replaced the Commission's eContentplus Programme. Examples of Digital Libraries Programme projects funded by the Commission's eContentplus and the ICT Policy Support Programme can be seen at:

[http://ec.europa.eu/information\\_society/activities/econtentplus/projects/index\\_en.htm](http://ec.europa.eu/information_society/activities/econtentplus/projects/index_en.htm)

The Commission's view is that it is important to digitise Europe's cultural and scientific collections: to enable "all Europeans to access Europe's collective memory and use it for education, work, leisure and creativity", to "contribute to Europe's competitiveness", to present "Europe's rich and diverse heritage on the Internet" and to protect "cultural assets from irretrievable loss." The Digital Agenda for Europe's Digital Libraries Initiative would not just have cultural value, but also help promote the growth of education, technology markets, media and tourism in Europe. The Commission recognised "digitisation and digital preservation are knowledge-intensive activities that are likely to grow considerably in the coming years"<sup>13</sup> offering potential for innovation and job creation within these industries.

The Commission also reported that "high-quality digital content is a key driver for large scale industrial activities" and this view is borne out by evidence provided to the Commission's 'Comité des Sages'<sup>14</sup> hearing in October 2010, which recognised that there are significant opportunities for technological innovations particularly through the growing market for applications or 'apps' to provide access to online information content while on the move and the potential for convergence of access to information via mobile devices such as smart phones and tablet PCs.

In addition to the economic opportunities offered by access to this information for technology companies, it is also worth noting that an increasing number of museums and libraries have also been providing direct access to their content via mobile apps thus opening up more channels through which to explore cultural content. Two examples of such apps are the US National Library of Medicine's 'Gallery of Mobile Apps and Sites', and the British Library's 'Treasures App'.<sup>15</sup>

## 1.1 Orphan Works and the European Digital Libraries Initiative

The Digital Libraries Initiative is driving the agenda for finding legislative and voluntary solutions to facilitate the digitisation of orphan works – copyright works for which the rightsholder(s) cannot be located or identified after a diligent search. However, five years after the Digital Libraries Initiative commenced in the European Union, only three countries have an appropriate code or legislation to allow the digitisation and putting online of orphan works – and even in these countries availability is only limited to their own jurisdictional borders. Denmark and Finland rely upon statutory extended collective licensing arrangements for this, and Hungary (where the Patent Office is empowered by statute to grant a licence for the use of an orphan work if the user demonstrates that a reasonable search has been made) have introduced a statute backed solution similar to the Canadian system which provides for a statutory supplemental licensing scheme operated by a central authority.<sup>16</sup> Licensing schemes from collective management organisations (CMOs), however, arguably do not solve the problem of unpublished orphan works (including unpublished letters, diaries and a range of other works on paper) because such authors are generally not represented by collecting societies whose role is mainly represent creators and entrepreneurs who make their works commercially available. This means that much material held in libraries, archives and other cultural institutions, often due to its unique nature of high interest to researchers cannot be digitised and made available without the institution potentially risking infringing copyright.

---

<sup>13</sup> European Commission (24 August 2006). Communication on the Digitisation and Online Accessibility of Cultural Material and Digital Preservation of (C(2006) 3808 final).

[http://ec.europa.eu/information\\_society/activities/digital\\_libraries/doc/recommendation/comm\\_recomm/en.pdf](http://ec.europa.eu/information_society/activities/digital_libraries/doc/recommendation/comm_recomm/en.pdf)

<sup>14</sup> [http://ec.europa.eu/information\\_society/activities/digital\\_libraries/comite\\_des\\_sages/index\\_en.htm](http://ec.europa.eu/information_society/activities/digital_libraries/comite_des_sages/index_en.htm)

<sup>15</sup> <http://www.nlm.nih.gov/mobile> and <http://www.bl.uk/app>, See also the 'M-Libraries' international wiki list of library mobile apps at <http://www.libsuccess.org/index.php?title=M-Libraries>

<sup>16</sup> In Canada, if the copyright owner is not located after a reasonable search, the user may apply for a licence from the Copyright Board of Canada. Lifshitz-Goldberg, Yael (2010): Orphan works. WIPO Seminar – May 2010: lecture summary. [http://www.wipo.int/meetings/en/doc\\_details.jsp?doc\\_id=143252](http://www.wipo.int/meetings/en/doc_details.jsp?doc_id=143252)



The Commission's High Level Group on Digital Libraries (operative 2006-09) was tasked with making recommendations<sup>17</sup> on a range of issues arising from the Commission's i2010 Digital Libraries Initiative, one of which was orphan works. A 'Memorandum of Understanding on Diligent Search Guidelines for Orphan Works'<sup>18</sup> was included in the recommendation as well as a model licence for out of print works. This MOU did not include agreement on the criteria needed for mass digitisation but did recognise the need for legislation. Likewise, the Comité des Sages' recent report, *The New Renaissance*,<sup>19</sup> has called for a "European legal instrument for orphan works as soon as possible". It recommends that "some form of registration" is necessary to avoid creating future orphan works, and that cultural institutions should have access to "collective licensing solutions" enabling cultural institutions "backed by legislation, to digitise and bring out of distribution works online, if rights holders and commercial providers do not do so."

In response the Commission published its Draft Directive on orphan works in May 2011, which is intended to provide a framework that will allow Member States to make national provisions concerning orphan works that are legally valid across borders.<sup>20</sup> However, the Draft Directive as currently proposed is unlikely to have a positive impact on mass digitisation projects since it requires a 'diligent search' to be made for each individual work, which, as will be demonstrated by this study is likely to still be beyond the means and resources of most cultural and educational institutions (See Section 11 below). It also excludes unpublished works, sound recordings, stand alone artistic works and in certain instances film also.

## 2 Constraints affecting the mass digitisation and online availability of the UK's cultural collections

Many UK libraries and archives make works available to readers from all over the world for the purposes of learning, research and cultural benefit and where possible they do so online. However, the ability to maximise the benefits of the digital environment is restricted by the current copyright framework in the UK, which inhibits delivery of publicly available online access to much of the cultural repertoire that is in copyright:

- In the UK there is no straightforward licensing solution available for the mass digitisation of in-copyright cultural artefacts (i.e. mostly late 19th and 20th century works);
- Unpublished works (a large part of archive collections) are in copyright in the UK until 2039;<sup>21</sup>
- Orphan works may not be digitised or otherwise exploited beyond the provisions of the UK's copyright exceptions,<sup>22</sup> without the permission of the individual rightsholders;
- A single work may have more than one rightsholder but if just one of multiple rightsholders cannot be located, the work is rendered orphaned.

---

<sup>17</sup> European Commission, High Level Expert Group on Digital Libraries (2009): Final report. Digital libraries: recommendations and challenges for the future. See Annex 2 of report (p.16) for further links to the MOU and Diligent Search Guidelines:

[http://ec.europa.eu/information\\_society/activities/digital\\_libraries/doc/hleg/reports/hlg\\_final\\_report09.pdf](http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg/reports/hlg_final_report09.pdf)

<sup>18</sup> [http://ec.europa.eu/information\\_society/activities/digital\\_libraries/doc/hleg/orphan/mou.pdf](http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg/orphan/mou.pdf)

<sup>19</sup> European Commission, Comité des Sages (2011). The new renaissance: report of the 'Comité Des Sages' reflection group on bringing Europe's cultural heritage online. [http://ec.europa.eu/information\\_society/activities/digital\\_libraries/doc/refgroup/final\\_report\\_cds.pdf](http://ec.europa.eu/information_society/activities/digital_libraries/doc/refgroup/final_report_cds.pdf)

<sup>20</sup> [http://ec.europa.eu/internal\\_market/copyright/orphan\\_works\\_en.htm#directive](http://ec.europa.eu/internal_market/copyright/orphan_works_en.htm#directive)

<sup>21</sup> Copyright Designs and Patents Act 1988 (known as 'CDPA') c. 48 Sch. 1 para 12.

<sup>22</sup> CDPA Chapter III Acts Permitted in relation to Copyright Works, ss. 28-76. <http://www.legislation.gov.uk/ukpga/1988/48/contents> and <http://www.legislation.gov.uk/ukpga/1988/48/part/1/chapter/III>

## 2.1 The transactional cost of obtaining permissions

A 2009 study by the Collections Trust and the Joint Information Systems Committee (JISC) found that the average proportion of orphan works estimated to be in collections across the UK public sector was 5-10%; this proportion was found to be higher in archives since they hold unique unpublished papers, records and audiovisual materials, many of which are more likely to be orphaned.<sup>23</sup> The three Carnegie Mellon University Libraries studies in the US between 1999 and 2001 reported on by Denise Troll Covey,<sup>24</sup> found between 23% and 31% of works orphaned. A British Library project to digitise more than 4,000 hours of sound recordings and make them available online for researchers included the digitisation of 220 oral history recordings of jazz musicians and promoters made by the Library itself in the late 1980s. This required seeking 200 permissions for this sub-set of recordings alone. Of these, 53 rightsholders (26.5%) could not be traced and 13 more (6.5%) had died with no heirs or estates found to approach, meaning 33% of these recordings were orphaned. A further 27 (13.5%) did not respond to the Library's enquiries.<sup>25</sup>

A European Commission report by Anna Vuopala published in 2010,<sup>26</sup> assessed the cost for rights clearances of orphan works in digitisation projects covering text, artistic works, sound recordings and audiovisual materials that had been conducted by 22 European libraries and archives. It found that the costs of clearing rights, including the time and effort to obtain licences from rightsholders, were "costly and cumbersome" for institutions.<sup>27</sup> Vuopala writes "In fact the amount of time and effort to obtain licenses to digitise works has overwhelmed many of them. Data shows that the older the work and the less economic value it has, the more it costs to clear rights to use it."<sup>28</sup> She goes on to give two examples:

- The UK National Archives spent two years and £35,000 clearing copyright to digitise 1,114 old wills and make them available online. It was able to locate the rightsholders and obtain permission "in less than half of the cases."<sup>29</sup> In this project "Approval was granted for 597 documents. Permission was rejected for 45 documents and laid conditional on a payment in 77 cases. Right holders for 385 documents were not identified."<sup>30</sup>
- An Austrian university library digitised 200,000 doctoral dissertations from 1925-1988 under a national copyright exception but was unable to make it available online due to the "disproportionately high transaction costs involved in clearing the rights for them." Digitisation cost €150,000 but the library estimated that the rights clearance transaction costs "would be 20-50 times higher than [the] cost of digitisation."<sup>31</sup>

---

<sup>23</sup> Korn, Naomi (2009). In *From The Cold: an assessment of the scope of 'orphan works' and its impact on the delivery of services to the public*, JISC and the Collections Trust. <http://www.jisc.ac.uk/media/documents/publications/infromthecoldv1.pdf>

<sup>24</sup> Troll Covey, Denise (2005). *Acquiring Copyright Permission to Digitize and Provide Open Access to Books*, Washington, D.C., Digital Library Federation Council on Library and Information Resources.

<sup>25</sup> Response from the British Library to the Independent Review of Intellectual Property and Growth. March 2011. <http://pressandpolicy.bl.uk/imagelibrary/downloadMedia.ashx?MediaDetailsID=887>

<sup>26</sup> Vuopala, Anna (2010). *Assessment of the orphan works issue and costs for rights clearance*, European Commission. [http://ec.europa.eu/information\\_society/activities/digital\\_libraries/doc/reports\\_orphan/anna\\_report.pdf](http://ec.europa.eu/information_society/activities/digital_libraries/doc/reports_orphan/anna_report.pdf)

<sup>27</sup> Vuopala, Anna (2010). *Op. cit.* p.5

<sup>28</sup> *Ibid* p.5

<sup>29</sup> *Ibid* p.5

<sup>30</sup> *Ibid.* pp.39-40

<sup>31</sup> *Ibid.* p5

## 2.2 The loss of 'Clause 43'

The examples in paragraph 2.1 above demonstrate the time and cost burdens of making a diligent search, even within smaller digitisation projects. In the context of the mass digitisation programmes as envisaged for Europeana, the diligent search becomes an impossibility particularly given the lack of adequate mechanism in the UK to seek copyright clearance and permissions.

However at one point, there was hope for a UK solution. The Digital Economy Bill (now the Digital Economy Act 2010),<sup>32</sup> introduced by the then Labour government in the House of Lords in the 2009–10 session of Parliament, had included a clause (Clause 42 subsequently Clause 43)<sup>33</sup> to allow the use of orphan works. Clause 43 had proposed important flexibilities for the handling of orphan works by cultural institutions which could be legislated for at national level without the need for a European Directive, by making provision for both a statute backed licensing scheme through collecting societies and, where appropriate, for special licences for individual orphan works issued by the relevant Minister. This clause was supported by all the UK's national museums, galleries, libraries, the university sector, the BBC and Consumer Focus. However, vigorous opposition to Clause 43 from a certain group within the photographic sector created an extensive debate. In March, a general election was called and the by then controversial clause was dropped from the bill to enable the outgoing government to push the legislation through. It received Royal Assent on 8 April, four days before Parliament was dissolved.

This study illustrates the impact of the loss of Clause 43: the copyright status of a considerable proportion of the works in the study's sample is unknown and a significant number were found to be orphaned.

## 2.3 The role of collecting societies in clearing rights for mass digitisation

Collecting societies<sup>34</sup> are organisations which exist to represent rightsholders and enable large scale licensing of copyright material through economies of scale. Throughout Europe and the rest of the world there are a number of these organisations providing rights clearance on a territory and copyright work category basis (e.g. sound recording rights in the UK, or rights in literary works in Germany).

Collecting societies usually maintain their own discrete databases of creators and their creative works which they match against usage data received from licensees in order to collect and distribute royalties for the reproduction, performance and communication of copyright works. In the pre-Internet age this model worked well as it was easy and logical to delimit the usage of copyright works by territory – i.e. where reproduction took place, or where a licensee was economically or legally resident. In addition the number of potential licensees was limited to those who had the means to distribute or otherwise exploit copyright works on a wide basis. However over the last two decades the availability of digital technology and the wide spread use of the Internet have led to a situation in which distinctions of territoriality have become less easy to apply and the volume and complexity of copyright works' usage has increased dramatically.

---

<sup>32</sup> Digital Economy Act (2010 c.24).

<sup>33</sup> House of Commons Bill 89 09-10 (Brought from the Lords) 17.03.s10 <http://services.parliament.uk/bills/2009-10/digitaleconomyhl/documents.html>

Clause 43: <http://www.publications.parliament.uk/pa/cm200910/cmbills/089/10089.49-55.html#i901>

<sup>34</sup> Collecting societies are also known as collective management organisations or CMOs. Those collecting societies who administer licensing for reproduction of copyright material are called Reproduction Rights Organisations or RROs.

Collecting societies have historically sought to work with each other where possible to identify efficiencies such as offering licences that contain bundles of rights and linking their databases and processes.<sup>35</sup> In order to address the challenges of the digital licensing landscape they have also been working towards more efficient and joined up ways of working through cooperating with other collecting societies as well as other parties who hold relevant databases.<sup>36</sup>

However many who wish to use copyright works still find it difficult to identify rightsholders and acquire the necessary rights. This ranges from large scale services wishing to launch copyright content services on a multi-territory basis to individuals and organisations wanting to use specific pieces of content on a smaller scale.<sup>37</sup> This is because there are often multiple copyright elements and owners, sometimes within just one piece of content let alone large collections, for which permission to use is administered in different ways by organisations in different jurisdictions and operating under different constitutions.<sup>38</sup> Users of copyright content therefore need to liaise with multiple rights organisations, creating significant inefficiencies. Whilst there have been numerous calls from government, industry and consumers for a 'one stop shop' which would allow users of all types of copyright works to identify rightsholders and acquire the relevant rights in a single place, this has not yet been realised.

Projects such as ARROW, which bring together rightsholder organisations with libraries and other information repositories, bring the concept of the one stop shop closer to realisation. ARROW could help insulate users of copyright works from the complexities and inefficiencies inherent in the current licensing regime by acting as a single point of access.

### 3 A brief explanation of United Kingdom copyright term

The current UK copyright law is the Copyright, Designs and Patents Act (1988 c.48) known as 'CDPA',<sup>39</sup> as amended by a raft of secondary legislation or Statutory Instruments.<sup>40</sup> The most important provisions with regard to this study concern those for copyright term in literary, dramatic, musical and artistic works.<sup>41</sup> The CDPA is now compliant with the European Term Directive 1993 (codified in 2006),<sup>42</sup> which has harmonised copyright terms in the member states of the EEA (European Economic Area) to the basic rule of the author's life plus 70 years.

Before the CDPA came into force, literary, dramatic, musical and artistic works that had not been made available to the public during their authors' lifetimes remained in copyright in perpetuity until this occurred, whereupon their copyright terms became finite. In these cases copyright endured for a further 50 years from the end of the year in which the event occurred. 'Made available to the public' includes a work or copies of a work being sold or offered for sale or published. In the case of literary, dramatic, and musical works it also includes public performance or broadcast and, for artistic works, exhibition in public or inclusion in a film showed in public.

---

<sup>35</sup> A good example of this is within the UK where the CLA, who license on behalf of the Publisher Licensing Society and the Authors' Licensing and Collecting Society, have an arrangement with DACS, who represent artists and designers. This enables the licensing of reproductions of illustrations in printed material along with the literary content.

<sup>36</sup> An example of this is the development of ISNI – the International Standard Naming Identifier – <http://www.isni.org/> which will assist with the identification of public parties involved in the creation and exploitation of media content.

<sup>37</sup> An example of the ways in which this has created challenges for collective licensing can be demonstrated by the European Commission's decision to introduce greater efficiency to the music licensing market in Europe through encouraging greater competition between CMOs. This is demonstrated by the 'Statement of Objections' IP/04/586, 03/05/2004:

<http://europa.eu/rapid/pressReleasesAction.do?reference=IP/04/586&format=HTML&aged=1&language=EN&guiLanguage=en>

<sup>38</sup> For example Germany has a highly regulated collecting society arrangement in which their CMOs' terms of reference are laid down in legislation.

<sup>39</sup> The UK Intellectual Property Office makes available an unofficial copy of the copyright provisions in CDPA (a long and complex statute) at <http://www.ipo.gov.uk/pro-types/pro-copy/c-law/c-law-acts.htm>. The full official copy of the Act is at <http://www.legislation.gov.uk/ukpga/1988/48/contents>

<sup>40</sup> Official copies of Regulations amending CDPA are at <http://legislation.gov.uk>. They are directly accessible via links from <http://www.ipo.gov.uk/pro-types/pro-copy/c-law/c-law-statutory.htm> The most relevant Regulations to this study are The Duration of Copyright and Rights in Performances Regulations 1995 (S.I. 1995 no. 3297) which implements the Term Directive 93/98/EEC (now codified as Directive 2006/116/EC), apart from Article 4 – Publication Right, in the UK. The Copyright and Related Rights Regulations 1996 (S.I. 1996 No. 2967) implements inter alia Article 4 of Directive 93/98/EEC.

<sup>41</sup> CDPA s.12 <http://www.legislation.gov.uk/ukpga/1988/48/section/12>

<sup>42</sup> The Term Directive 93/98/EEC (now codified as Directive 2006/116/EC). <http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32006L0116:EN:NOT>

### 3.1 Transitional provisions in the CDPA for copyright term

When the CDPA came into force in August 1989, a number of transitional provisions were introduced, the most prominent for libraries and archives being those concerning photographs and unpublished works. Copyright duration may vary depending on the date of creation or publication and therefore the simple rule of life plus 70 cannot always be relied on. Figure 21 in Appendix 2 sets this out in more detail.

The CDPA replaced the indefinite term provisions for works still unpublished on 1 August 1989 with a copyright term of 50 years from 1 January 1989 to 31 December 2039. However, on the implementation of the Term Directive in UK law in December 1995, authors who died on or after 1 January 1969 retain copyright in their unpublished works for the full term introduced by the Directive of life plus 70 years. Where the author had died on an earlier date, the extension of term to life plus 70 years did not produce a longer copyright term in his or her unpublished works than had originally been provided in the CDPA. The term in these cases still expires on 31 December 2039.

### 3.2 The European Term Directive 1993

The Term Directive was implemented in the UK on 1 December 1995 and took effect on 1 January 1996,<sup>43</sup> generally extending copyright and moral rights subsisting on 31 December 1995 by 20 years and reviving copyright in some works already made available to the public in which copyright had already expired under the UK's former term of author's life plus 50 years. However, the revival of rights did not entitle rightsholders to any rights of action with regard to uses made of the revived works during the time they were in the public domain. The specifics of the Directive's application to UK copyright term is set out in more detail in Appendix 2:2(v).

Classes of works covered in this study in which copyright was extended or revived include published literary, dramatic and musical works whose authors had died between 1922–44 and artistic works, including photographs, by known authors who died between 1922–44. This includes photographs of known authorship taken before 1 July 1912.<sup>44</sup>

It is worth noting that since the implementation of the European Information Society Directive 2001 in the UK<sup>45</sup> in 2003, 'making available to the public' additionally now includes communicating a work to the public which covers its use via the Internet.

---

<sup>43</sup> The Duration of Copyright and Rights in Performances Regulations 1995 (S.I. 1995 no. 3297) implements the Term Directive 93/98/EEC (now codified as Directive 2006/116/EC), apart from Article 4 – Publication Right, in the UK. The Copyright and Related Rights Regulations 1996 (S.I. 1996 No. 2967) implements *inter alia* Article 4 of Directive 93/98/EEC.

<sup>44</sup> Prior to implementation of the Term Directive it could be assumed that copyright in all pre-1912 photographs had expired since both the 1911 and 1956 Copyright Acts (1911 ss.3, 21; 1956 Sch. 2, Sch. 7 paras 2,16) gave them a copyright term of 50 years after creation.

<sup>45</sup> Information Society Directive 2001/29/EC implemented by the Copyright and Related Rights Regulations 2003 (S.I. 2003 no. 2498).



# Methodology

## 3.3 Case study: The effect of the UK's long copyright term

The consequence of a long copyright term is illustrated by the example of Jessie Saxby, one of the Shetland Islands' most famous authors.<sup>46</sup> Jessie Saxby died, aged 98, on 27 December 1940. In her lifetime she had written over 150 works, beginning at the age of 17 when, as Jessie Edmonston, she published some short stories in 1859 in *The Scottish Annual* ('The sea-girt home') and in the same year in the 1 October issue of the *Family Herald* ('Steen: a Shetland story'). Under the UK copyright term, these 1859 works thus remained in copyright up to 31 December 2010, 151 years after publication.

The current most likely candidate for the longest copyright protection of a UK published work is Dame Sarah Mair (1846–1941), who at the age of 19 in 1865 published 'On Making a Noise in the World' and other articles in *The Attempt*, the magazine of the Ladies' Edinburgh Debating Society, that she had founded.<sup>47</sup> These early works will enter the public domain at the end of 2011, 146 years after publication.

## 4 Procedure

The study was split into four main phases. Firstly a list of 140 titles was drawn from as random a sample of British Library collection items as possible. The next phase involved searching for the rights status of each book – in copyright, out of copyright or copyright status unknown – then determining who the relevant rightsholders were. This was followed by a period of attempting to locate and contact each of the rightsholders identified to ask for permission to digitise the material. The final phase involved collating the results and drawing conclusions.

### 4.1. Parameters for the study

- The search only referred to MARC<sup>48</sup> records, not the physical book in order that the conditions of running an automatic search through the ARROW database could be replicated, so far as possible, in the manual search process, thus enabling a true comparison of the results.
- The study referred to the 'Memorandum of Understanding on Diligent Search Guidelines for Orphan Works' launched in 2008 by the European Commission's High Level Expert Group on Digital Libraries (HLEG).<sup>49</sup> This MOU was an agreement to observe a set of diligent search guidelines to locate rightsholders before declaring that a work was orphaned if that search was unsuccessful. However as the report acknowledges "No agreement was reached by the stakeholders on the possible need for specific due diligence criteria for mass digitisation" and the guidelines are not considered practical by cultural bodies in the context of mass digitisation.

---

<sup>46</sup> Smith, Brian (2004): 'Saxby, Jessie Margaret Edmondston (1842–1940)', Oxford Dictionary of National Biography, Oxford University Press.

<sup>47</sup> Shepley, Nigel (2004): 'Mair, Dame Sarah Elizabeth Siddons (1846–1941)', Oxford Dictionary of National Biography, Oxford University Press.

<sup>48</sup> MACHine-Readable Cataloguing <http://www.loc.gov/marc/faq.html#definition>

<sup>49</sup> i2010 Digital Libraries Initiative High Level Expert Group (2009): Final Report – Digital libraries: recommendations and challenges for the future. Section 2.1.2, p4. [http://ec.europa.eu/information\\_society/activities/digital\\_libraries/doc/hleg/reports/hlg\\_final\\_report09.pdf](http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg/reports/hlg_final_report09.pdf)

- The decision was taken to limit the reference sources consulted to identify the rights status to the extensive resources already available in the Library reading rooms or those available free of charge online. The Diligent Search Guidelines accompanying the MOU mention probate records “depending on the country and the works in question” at the end of a long list of information sources.<sup>50</sup> However, the Guidelines are generally not considered practicable for use with mass digitisation projects, the conditions of which the study proposed to replicate.
- This meant that, in this instance, enquiries into authors’ wills were not pursued. To have done so would have incurred additional costs in terms of time for searching, fees would have had to be paid to obtain copies of the wills from the Probate Registry or to access older wills that might be identified in the National Probate Calendar (1861-1941) on the Ancestry UK Database.<sup>51</sup> In the case of older wills, they could well have led to extensive additional searches down two or three generations to try to identify the current rightsholders and their location, with the risk that this route might still not have led anywhere. In this respect the study did not fully replicate the furthest extremes of diligent searching since searching wills would have been utterly impracticable in the case of a mass digitisation project.
- Reference sources were also interrogated to locate publishers or other intermediaries that might lead to the rightsholders’ location (see Fig. 6, Section 6.6 below for a list of resources used).

## 4.2 Process

- The searches were to be conducted as far as possible within the above parameters. The information sources used and the findings were to be recorded. Additionally administration and project meetings were to take place on a regular basis according to the contracted time available from the consultant and allocated by the Library staff over the 11 month period. This also allowed time in the last four months of the study for rightsholders and intermediaries to be contacted and for responses to be received from them.
- The copyright status of the works needed to be established by finding death dates for rightsholders named in the MARC records, where these were not already recorded. Where none could be found and the copyright status was therefore unknown, it was assumed that the work still needed clearance. In order to draw comparisons with the ARROW system it was decided to only seek to clear rights in illustrations and other third party contributions in the works where the rightsholders were named in the MARC record as responsible for material in the work. As with the search to clear the authors’ rights this would be done without inspection of the book itself.
- Works would be regarded as orphaned if the rightsholder(s) could not be located or identified after a diligent search (see Section 8 for more detail on this).
- The ‘in print’ status of the works in the sample would be established to see if it had any correlation to the willingness of rightsholders to grant permission.
- Rightsholders would be contacted by the most effective means available (email, web form, fax, letter or telephone) at least three times over a three month period during the contact phase of the study. A further month was allowed to wrap up any ongoing negotiations. Onward referrals from initial contacts would also be followed up as far as possible within the time available.
- The study would close on 11 February 2011 (the ARROW project itself was to close on 28 February 2011). Lack of response from rightsholders, responses received after the closing date, or any negotiations not concluded by that date would, for the purposes of the study be treated as permission not granted.

---

<sup>50</sup> 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](http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg/orphan/guidelines.pdf) and Appendix  
[http://ec.europa.eu/information\\_society/activities/digital\\_libraries/doc/hleg/orphan/appendix.pdf](http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg/orphan/appendix.pdf)

<sup>51</sup> <http://www.ancestry.co.uk>

# About the records used in this study

## 5 Selection of the random sample catalogue records for the study

A random sample of 1,000 MARC 21 records per decade of books produced between 1870 and 2010 (14,000 records in total) was drawn from the British Library catalogues by interrogating Aleph, the administrative interface to the British Library Integrated Catalogue. Aleph selected records at random across the 10 Dewey top-level (000–900) classified disciplines.<sup>52</sup> The sheer volume of records in the catalogue meant it was impossible to extract the random sample from a full set of catalogue records. Because Dewey classifications were only routinely used for cataloguing purposes from the 1970s it was also not possible to filter items based on Dewey classification. Probably reflecting the nature of many national library collections it was interesting to note that the majority of the initial 1,000 results pulled back for each decade was so-called “grey literature” – material not published by commercial publishers.

The selection of 14,000 catalogue records was then filtered in a spreadsheet to exclude reprints, non-English language works, official government publications and foreign language materials so as to provide a base selection of English language publications from a variety of publishers across a broad spread of subjects. For instance, if official government publications had been retained in the selection their sheer number would have severely distorted the nature of the small sample.

A random sampling procedure was then run on the remaining records in the spreadsheet to produce a sample of 140 records comprising ten English language monographs per decade. Even at this stage, given the large amounts of grey literature that was being returned from a random electronic selection of the library's collection, the 140 records selected had to be reworked to ensure that the final list contained a sufficient amount of monographs produced by commercial publishers to ensure that the comparison with the ARROW system produced some meaningful results. The reason behind this was that the databases of collecting societies generally comprise commercially produced material rather than grey literature.

Mass digitisation projects in libraries and archives often seek to digitise large collections without applying detailed selection criteria, since any selection process requiring manual intervention is expensive. Other than the removal of the material referred to above from the sample, the intention was to try to obtain as random a sample as possible to mirror the random nature of materials and kinds of issues that might be encountered in a mass digitisation project.

---

<sup>52</sup> The Dewey Decimal Classification System (DDC), developed in 1876 by Melvil Dewey, is the most commonly used library classification scheme in the world. Library classification schemes arrange library stock on the shelves in a specific order to make it easier to locate particular items or to browse in specific subject areas. For further information see <http://www.oclc.org/dewey/> and [http://en.wikipedia.org/wiki/Dewey\\_Decimal\\_Classification](http://en.wikipedia.org/wiki/Dewey_Decimal_Classification). Dewey classes are described at [https://secure.wikimedia.org/wikipedia/en/wiki/List\\_of\\_Dewey\\_Decimal\\_classes](https://secure.wikimedia.org/wikipedia/en/wiki/List_of_Dewey_Decimal_classes)

## 5.1 Non-UK English language publications

The works in the final sample turned out to be predominantly published in the UK but a small percentage (16%) were published elsewhere and included 25 English language publications from Canada and India (both dominions of the former British Empire), Ireland (an integral part of the UK until 1921) and the United States (see Fig. 1). Most of the non-UK works in the sample were also published by non-UK, English language publishers. There were six exceptions: two were published by Oxford University Press, the first in India in 1932 at a time when the country was still part of the British Empire, and the second, an edition published simultaneously in the UK and India in 1998; the third was published by Penguin in Baltimore in 1969; and three books that were published in 2008–2009 in India by Dorling Kindersley, part of Penguin and in turn member of the Pearson global publishing group. All six books were retained in the sample since non-UK English language publications are well represented in the British Library collections and therefore would be part of any mass digitisation activities.

Under the Berne Convention and the Universal Copyright Convention 'national treatment' rules apply to foreign publications, protecting them in the same way as national works. However, where the copyright term in the home nation (i.e. where the work is to be copied) is greater than that in the foreign nation (i.e. where the work was published), the lower foreign term is applied (see Appendix 2 para 2.2(iv)). This meant that the copyright laws of these four countries had to be taken into account in order to establish whether or not they were in the public domain (see Appendix 3).

**Fig. 1: Distribution by decade and country of non-UK published works**

DECADE	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s	Total
Canada		1						1							2
India		1	1	1	1	1	1							3	9
Ireland	1	1					1								3
USA					1	2		2	1	2				2	10
Total	1	3	1	1	2	3	2	3	1	2	0	0	0	5	25

### 5.3 Records from the 2000s decade

Nine of the 10 records from the 2000s decade appeared to be ‘pre-cataloguing records’ i.e. records created when the book is on order in advance of it being fully catalogued and accessioned by the Library, a process which may necessarily be delayed given the vast scope of legal deposit and new acquisitions. Not only did these records not contain an assigned British Library shelfmark (an indicator that the book was accessioned by the Library), but there were not yet any entries for them in the public facing catalogue of the British Library – the Integrated Catalogue. These records also lacked place of publication data since this is generally verified by cataloguers on sight of the physical book.

Despite this, for the purposes of this study, which was only using MARC records and not referring to the physical books they described it was decided that all records should remain in the sample as there was still sufficient information in these records, such as the book’s ISBN, to establish rights information via publishers’ websites without any material difference in the outcomes.

### 5.4 Third party rights in embedded artistic works

The MARC records for nine books in the sample credited named persons or studios (that were not otherwise author(s), editors or translators of the work) with third party rights in embedded artistic works such as illustrations (see Fig. 2). Section 6.1 sets out the structure of a MARC record in more detail however illustrators and other named third party rightsholders are usually noted in the 245 or 500 field.

It was decided that wherever possible these clearances were best followed up via the publisher, or via the author acting as intermediary in the first instance, as they were most likely to have contact details or some knowledge of where the rights for the particular work were likely to sit, and then follow those leads if appropriate. However, one 1934 work included a cartoon on the back cover by a well known cartoonist who died in 1970. This artistic work needed to be cleared separately as the rest of the publication was in the public domain. In this case the artist’s heirs were easy to find since his daughter maintained a website for his works, but negotiations were still in train when the study was closed on 11 February 2011.

The other eight artists’ names were also queried against the membership database of the Design and Artists Copyright Society (DACS), a collecting society which manages rights for a wide spectrum of visual artists, via the interface on the DACS website<sup>53</sup> but no matches could be found. Library staff also phoned DACS to ask if there was any other way to check for the illustrators but, despite DACS’ helpfulness in searching their own records, this also identified no further matches.

Fig. 2: Numbers of works with named third-party rightsholders in embedded artistic works by decade

1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s	Total	% whole sample rounded
0	0	0	1	2	0	1	1	1	0	1	1	1	0	9	6

<sup>53</sup> <http://www.dacs.org.uk/index.php?m=3&s=3&c=9>



## 5.5 Types of publisher

Although the rights more often than not sit with the authors, it is the publisher which, even if it no longer has rights in the work, is the best route to tracing the author and any other relevant rightsholders. However this approach is not always straightforward. Although a publisher may have held some rights, in older works these are likely to have reverted back to the author over the course of time and digital rights would not have been licensed.<sup>54</sup>

A major feature of research libraries' collections is that they hold many publications of an ephemeral nature, not produced by commercial publishers, which would otherwise disappear. These kinds of publications are commonly known as 'grey literature'<sup>55</sup> and they are reflected in the sample, though unevenly distributed over the decades. Examples of the types of publication and publishers are illustrated in Fig. 3 below.

**Fig. 3: Examples of publications by different types of publisher**

<b>Associations / Voluntary groups</b>	<ul style="list-style-type: none"> <li>– <i>Don't vote: organise</i>. [York]: Organisation of Revolutionary Anarchists, 1971</li> <li>– <i>The life and work of Stuart Morris / Sybil Morrison</i>. [London]: Stuart Morris Memorial Fund, Peace Pledge Union, 1969</li> </ul>
<b>Private companies</b>	<ul style="list-style-type: none"> <li>– <i>A history of Richard Lloyd Limited: to commemorate one hundred and fifty years of service / Edward G. D.</i> [Liveing. Birmingham]: Richard Lloyd Ltd., 1953</li> <li>– <i>Widdicombe Fair: illustrations, words &amp; music: they will be singing this old folk song with extra enthusiasm down in 'glorious Devon' this summer because it is a vintage year for Whiteway's cyder</i>. [Whimble, Devon]: H. Whiteway &amp; Co., 1934</li> </ul>
<b>Institutions</b>	<ul style="list-style-type: none"> <li>– <i>Whittlesey Straw Bear: with notes on the celebration of Plough Monday in the area / George</i> [Frampton. Peterborough]: Cambridgeshire Libraries Publications, 1989.</li> <li>– <i>Ammerdown: the House and the Centre / Bryan Little and Andrew Aldrich</i>. [Radstock, Bath]: published for the governors of the Ammerdown Study Centre, 1977</li> </ul>
<b>Commercial publishers</b>	<ul style="list-style-type: none"> <li>– <i>Practical nursing: including hygiene and dietetics / W.T. Gordon Pugh; assisted by Alice M. Pugh</i>. 15th ed. [Edinburgh]: William Blackwood, 1945</li> <li>– <i>Homespun: a study of a simple folk / Annie S. Swan</i>. [London and Aylesbury]: Hutchinson &amp; Co., 1893</li> </ul>
<b>Newspapers</b>	<ul style="list-style-type: none"> <li>– <i>Capital punishment and the Divine Law: a conversation in a village rectory / T. Gasquoine</i>. Reprinted from the "Bedfordshire Mercury" of Friday, August 16, 1907. [Bedford]: Published by the "Mercury" Press, High Street, [Bedford], 1907</li> <li>– <i>Practical politics: an object lesson: / Nunquam</i> [Robert Blatchford]. Reprinted from the "Clarion" of October 29, 1892. [Manchester]: printed by the Clarion Newspaper Co., 1892</li> </ul>
<b>Printer / Self published</b>	<ul style="list-style-type: none"> <li>– <i>Looking at West Cumbria: Silloth, Allonby, Maryport, Workington, Harrington, Whitehaven, St Bees / Herbert and Mary Jackson</i>. Maryport, [Cumbria]: H. Jackson, 1972</li> <li>– <i>History of St Peter's Church, Tiverton / by Edward J. Musgrove</i>. [Tiverton]: E.J. Musgrove, 1938</li> </ul>

<sup>54</sup> The reversion of rights may occur on expiry of a contract or is sometimes related to provisions within copyright law. Digital rights were not included in older agreements as they simply did not exist at the time. So called 'Martian clauses' that sought to cover all possible future uses only became standard practice following the development of digital technologies in the late 20th century.

<sup>55</sup> [http://en.wikipedia.org/wiki/Gray\\_literature#The\\_Definition\\_of\\_Grey\\_Literature](http://en.wikipedia.org/wiki/Gray_literature#The_Definition_of_Grey_Literature)

Commercial publishers are more likely to have records and rights information pertaining to the publication, author and any other rightsholders in the work as they are by definition rights-orientated businesses. However not all publishers are commercial concerns who perceive themselves as part of the 'publishing industry'. Publishers include organisations, institutions, government, commercial companies, voluntary groups ranging from short term single issue groups to more permanently established societies, political parties, charities and other non-governmental organisations, all of whom might issue reports or works of an ephemeral nature to the public from time to time. While bigger publishers may have good records, where an entity's function is not primarily as a "market" publisher, their record keeping may be idiosyncratic to the organisation's culture and in some cases non-existent. This makes it much harder to clear rights, since if they are unsure of ownership (e.g. where the work was done by volunteers without any employment contract or assignment of rights) they may not be in a position to give permission.

It used to be common for both publishers and printers to put their imprint on publications they printed for customers. Thus a number of the older catalogue records in the sample have the printer's details entered in the publisher field for lack of any other publishing information given on the physical work. It can at times be unclear whether a work was published by a publisher, or printed by a printer for the author or a corporate entity.

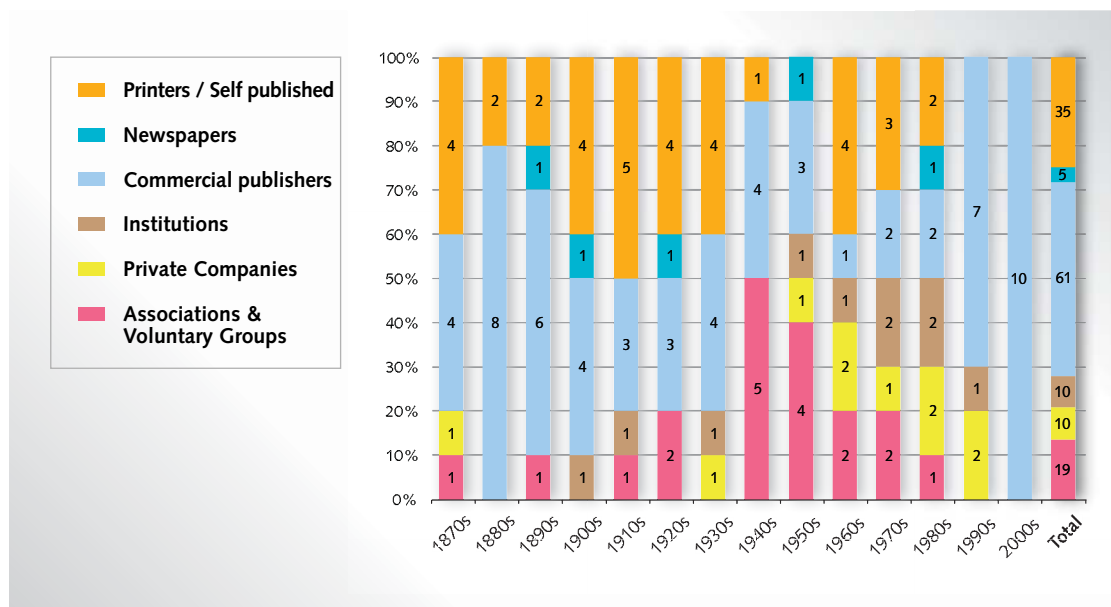
Some works both today and in the past have been self published by an author buying a print run to distribute either personally or to the booktrade. Today many authors make books available via websites or Internet bookstores either as a digital download or as print-on-demand item, where customer orders are fulfilled by a printer on an individual per order basis. According to *Publishers Weekly*, which offers its PW Select supplement to self-published authors as a means to present their titles to the publishing trade, contemporary self-publishing is a "phenomenon" that is "upending the publishing world."<sup>56</sup> While print on demand theoretically allows any book to remain in print forever, without a publishing business behind it and the record keeping that goes with that business, the risk of a work's rightsholder(s) becoming quickly untraceable increases as sales falter and websites lapse.

The types of publishers represented in the sample broke down into the following groups:

- Commercial publishers (61)
- Printers and the self-published (35)
- Associations and voluntary groups including political parties and movements (19)
- Institutions (10)
- Private companies whose main business was other than commercial publishing (10)
- Newspapers as ad hoc publishers of monographs (5)

Within the sample, commercial publishers were the largest single publisher group at 43.5% (61), but the majority of the works over the 140 years are published by non-mainstream publishers, who together represent 56.5% of the publishers in the sample.

Fig. 4: Publishers by category by decade



<sup>56</sup> <http://www.publishersweekly.com/pw/diy/index.html>

## 6 The search process

As previously stated the study followed four phases – identification of titles; searching for copyright status of the works; contacting rightsholders; and collating information received. The first three of these phases were comprised of the following stages:

- The Library created a spreadsheet containing selected MARC data in order to facilitate management of the results of the search.
- These records were verified for correctness and any missing data (such as additional authors and illustrators and place of publication) that had not been transposed from the MARC data correctly were entered into the spreadsheet.
- Searches were made using various reference sources available in the Library reading rooms to identify the copyright status of the works by finding any authors' death dates not in the MARC records.
- Searches were made to locate authors (or their estates) and other original rightsholders. Where the original rightsholders could not be located, publishers and third parties such as libraries or archives were identified and contacted for assistance in finding the rightsholder.
- The initially identified contacts were sent permission requests from the Library to digitise their work and make it available for public access free of charge online (see Appendix 4 – model letters sent to rightsholders). Initial requests were made by email, publishers' web forms, post and telephone, as appropriate, mostly between 22 October and 5 November 2010. Follow up letters were sent in the week of 13 – 22 December 2010 to those who had not yet substantively responded granting or withholding permission or opening up discussions, and a number of referrals were first contacted at the same time. Any still outstanding were followed up for a third and final time in early January 2011 and the period for responses kept open until the study was closed on 11 February 2011.

### 6.1 Supply of the MARC records

The sample records for the study were supplied by the Library in MARC 21 format in an ASCII text file. The Library also supplied an Excel spreadsheet with selected basic data from the MARC records in order to facilitate the searches and to record findings. As previously stated the physical works were not consulted directly, only the MARC record which provides a full description of the work, just as would be the case in a mass digitisation project where it would not be feasible to inspect the books themselves. Thus the manual search used the same data that was run against the ARROW database for comparison with the results of the manual search (see section 7).

The full MARC record provides a lot of information including the format and size of the work, death and birth dates, whether it contained illustrations or third party works and whether the authorship of any of these was ascribed to third parties who may or may not be rightsholders. Sometimes additional people will be named in the catalogue record who may have had an association with the work or have been its subject or previous owner, but were not themselves rightsholders: in such cases the reason for the entry of those names would be noted in the record.

A typical British Library MARC 21 record from the sample is as follows (this particular record has a year of death for everyone involved including a former owner, apart from the two libretto authors, Aubrey Hopwood and Harry Greenbank, who both had to be researched separately):

```
=LDR 01524cam a2200349 a 4500
=001 015441831
=003 Uk
=005 20100111101834.0
=008 091208s1898\\en\\k\\\\000\\1\\eng\\
=040 \\$aUk$beng$cUk
=042 \\$aukblsr
=100 1\\$aHopwood, Aubrey,$d1863-
=240 12$aA Runaway girl
=245 10$aMr. George Edwardes' new production, A
Runaway girl,$ba new musical play in two acts,$cby
Seymour Hicks and Harry Nicholls. Lyrics by Aubrey
Hopwood and Harry Greenbank. Music by Ivan Caryll and
Lionel Monckton ...
=246 14$aRunaway girl. A musical play in two acts ... Lyrics.
=260 \\$a[London] :$bWaterlow & Sons Limited, Printers,
London Wall, London , $c[1898] .
=300 \\$a36 p. ;$c19 cm.
=500 \\$alrprint from foot of title-page.
=500 \\$aFrom front cover: Price sixpence.
=500 \\$aBack cover contains adverts. for Chappell & Co.
=500 \\$aFrom head of title-page: Gaiety Theatre. Lessee
and manager, George Edwardes.
=500 \\$aIncludes lyrics for the songs, the majority written
by Aubrey Hopwood.
=500 \\$aFirst performed on the 21st May 1898.
=650 \0$aOperas$vLibrettos$y19th century
=700 1\\$aHicks, Seymour,$cSir,$d1871-1949
=700 1\\$aNicholls, Harry,$d1852-1926
=700 1\\$aGreenbank, Harry
=700 1\\$aCaryll, Ivan,$d1861-1921
=700 1\\$aMonckton, Lionel,$d1861-1924
=700 1\\$aNorthcott, Richard A.,$d1871-1931,$eFormer
owner.
=710 2\\$aGaiety Theatre (London, England)
=852 41$aBritish Library$bHMNTS$jNorthcott 38 (1)
```

## 6.2 Verification of the MARC records

The MARC records were transferred from the ASCII file to the more accessible Microsoft Word format. They were then checked, first against the originating ASCII files and then against the British Library Integrated Catalogue to ensure a clean transfer of data from one format to another and to verify whether the works actually had a shelfmark denoting their physical presence in the Library book stock. This was a time consuming and useful process to get the data correct as a few errors were discovered in the transfer process. These were both the result of human error – a duplicate record was discovered in the 1890s list and a record from the 1970s appeared in the MARC list for the 1900s for a book on the same subject by a different author which had been mixed up with the correct record in the spreadsheet.

### 6.3 Determining the copyright status of the works – searching for authors' death dates

As with any real mass digitisation project it was felt that for the purposes of this study a level of practicality around due diligence had to be applied. For this reason wills were not consulted however a wide range of sources were consulted from the extensive resources available within the British Library reading rooms or free of charge at the point of use on the Internet. These are listed below in Fig. 5.

Fig. 5: Sources used to find authors' death dates

Sources interrogated	Comments on usefulness to this study
British Library MARC Records supplied and British Library Integrated Catalogue (BLIC) including British Library Name Authority Files as necessary (Internet)	Most of the MARC records (apart from 9 of the 10 from the 2000s decade) were also in BLIC. MARC records are as good as the cataloguing conventions of the time. Verification of supplied records against the catalogue corrected a number of errors and omissions in supplied starter spreadsheet. Other than being the fundamental starting point the records were very useful for death dates and distinguishing one author from another with the same or similar names.
World Biographical Information System (WBIS) (includes British Biographical Archive)	Many matches – very good up to 1970s – strong on hits for earlier authors.
Who's Who/Who Was Who and International Who's Who (if author not British)	Who Was Who quite useful for this sample to 1970s.
Credo Reference (formerly Xreferplus) mainly Chambers Biographical Dictionary	Few matches from this sample.
The WATCH File (Writers and Artists and Their Copyright Holders) <a href="http://tyler.hrc.utexas.edu">http://tyler.hrc.utexas.edu</a>	Useful for contacts for some authors' literary estates, agents or heirs though some of the hits proved not contactable (information may be out-of-date in places).
Oxford Dictionary of National Biography (DNB)	Useful for people famous in their time or subsequently.
Literature Online	Not useful for this sample.
Literature Resource Center: includes: (1) LRC, Scribner Writers Series Online, Twayne's Authors Online, and (2) Biography and Genealogy Master Index – BGMI	LRC etc. only covers contemporary literary authors. BGMI provided some birth dates – not many death dates. It was not always possible to tell what was the subject field of the author so could not always distinguish between names. One contact e-mail was found from this source.
Library of Congress Authorities / LC catalog / Gutenberg Project	Useful for one or two names and author contacts for those not found earlier. Gutenberg Project's scanned US Copyright Office registration records very useful for identifying copyright registration status of older works.
WorldCat and WorldCat Copyright Evidence Registry beta website (Internet)	WorldCat, a cooperative catalogue but not that useful for this sample as most of the matching records were contributed by the British Library so held no extra information. WorldCat Copyright Evidence Registry only in beta form with little data, little copyright info and no link to it from main WorldCat pages.
COPAC (Internet)	University libraries cooperative catalogue used to verify cataloguing data and authors' names in a few cases.



Fig. 5: Sources used to find authors' death dates continued

Sources interrogated	Comments on usefulness to this study
National Library of Australia Catalogue & Name Authority File (Internet)	Catalogue used to verify cataloguing data and authors' names and dates in certain cases.
Internet search engines leading frequently to Google Books, Wikipedia, Amazon, authors' and fan websites and a range of online information sources including websites for special collections (in archives, libraries and museums) on relevant authors or movements with which they were associated	Search engines used especially for finding 1980s-2000s authors, many of whom still living and if literary authors, academics or researchers often a web presence could be established. A number of hits on Google Books and Wikipedia. Also the open Internet was the best source for verifying titles, publisher information and ISBNs for the post 2000 records since most of the supplied MARC records were incomplete pre-cataloguing records not in BLIC and with no shelfmark.
Internet Archive	Useful for disambiguating the authors Mary Bennett and Mary E. Bennett (1880s publication)
Genealogical websites and Ancestry UK index	Genealogical websites used to try to find death dates or identify heirs in a few cases mostly where author was from the peerage (useful). Ancestry UK is a charged service not available in the Library's London Reading Rooms but its free index was consulted online for any likely leads for establishing death dates of some authors strongly linked to a particular locality. As it can produce thousands of results, data such as full name, dates of birth or death and locality of either event (information that is usually not available in the MARC record or that may not have been gleaned from other sources) are needed to filter the results to a manageable number of outputs.
Society of Authors website	Society of Authors administers a limited number of literary estates directly. Website only useful for some contemporary authors. According to its website, only 1/3 of its membership agree to be listed online, <a href="http://www.societyofauthors.net/WritersAZ">http://www.societyofauthors.net/WritersAZ</a> . The Society of Authors was contacted directly re 74 authors not located or for whom no death date had been found, but, although very helpful, had no or insufficient information about them to be useful for this project
Virtual International Authority File (VIAF) beta website	VIAF started in 2009 – still in beta form. Interrogated to look for any outstanding authors' name authority information not already found in catalogues elsewhere. Resource identified later on in the project so no additional author death dates found.
Authors Licensing and Collecting Society (ALCS) contacted directly	Website only has a by-title public query interface to database of works for which ALCS holds unclaimed royalties. No public query interface to membership database. ALCS directly contacted re 74 authors not located or for whom no death date had been found but were unable to help with further information.
Design and Artists Copyright Society (DACS) website	DACS membership database queried via web interface to check if DACS represented (and so had details of) 9 named illustrators – no matches. DACS also contacted direct – very helpful but still no matches.

Not every source was used for every record since when a death year was identified no further search for that author needed to be made. However, works with joint or several authors or embedded third party rights in forewords and translations, as well as illustrations etc, where the MARC record did not already show year of death, required searches for each person named. At times the searches required to identify all third parties and their rights were substantial. After checking the British Library Integrated Catalogue and other national library authority files such as Library of Congress and the National Library of Australia, the two single most useful information sources for death dates were WBIS - the World Biographical Information Service and the open Internet. VIAF – the Virtual International Authority File, which was launched in 2009 and is still in beta form, was identified as another source of information at a later stage in the project. Authors from the sample for whom death dates had still not yet been found were checked against it but no successful matches were made.

The full British Library MARC record was an essential starting point since in describing the publication it provides indicators that help identify the right person when looking through directories and on the open Internet.<sup>57</sup> Researchers carrying out this kind of rights clearance activity need a good general knowledge in order to be able to recognise clues and make links and leaps of faith that would identify people with often very little information. As so many people actually share the same or very similar names and are contemporaneous, the additional clues that the catalogue record may provide from the title, subject headings and any other additional information all become important in the search for information about an author.

The death year of an author, or confirmation he or she is still living is crucial to determining the copyright status of a work. Both birth and death years are additionally essential for the purposes of identifying one author of the same name from another and in the event of disambiguating mix-ups between contemporaries with the same or similar names. A particularly confusing case requiring extensive research to disambiguate two 19th century children's authors of the same name, concerns 'Mary Bennett' and 'Mary E. Bennett b.1841' (see Appendix 5 – Case study: Inextricably intertwined in the world's catalogues: the two Mary Bennetts). It illustrates how important it is for rights clearances that national and cooperative library catalogue records are full and accurate as the two authors could only be disambiguated by physically looking at books.

## 6.4 The importance of library catalogues and authority files to clearing rights

In general terms, a high standard of cataloguing and full description of the work makes the rights clearance process easier. Many national library catalogues provide birth and death year information through their Name Authority File and, in a number of instances, so did the British Library Integrated Catalogue (which is a simplified catalogue viewable on the web) and VIAF, the Virtual International Authority File. However, collection of authors' dates is patchy since cataloguers generally take such information from the book itself where available, but most publications are published during authors' lifetimes. Cataloguing standards have varied over the years and some records have fuller descriptions of works than others. The actual dates of birth or death, which might be regarded as sensitive personal data, are not usually needed, just the year. Since the name authority files did not provide comprehensive information, it was necessary to interrogate various biographical databases and the Internet to find a number of authors' death dates with regard to 79 (56.4%) of the works in the sample. This was because only 61 (43.6%) of the 140 catalogue records in the sample did not require further searching to establish their copyright status. Of these 61 catalogue records, 23 had no information recorded at all as to the author's identity, so the works in question were treated for copyright duration purposes as anonymously authored works whose authors could not be identified,<sup>58</sup> and 38 records included the year of death for all the named authors of the work.

---

<sup>57</sup> The 'open Internet' refers to the free-to-use resources available in what Troll Covey had called 'surface web' in her report since the content held in the 'deep web' she referred to is generally inaccessible to search engines and is kept behind subscription of pay-per-view paywalls, Troll Covey, Denise (2005). Op. cit. pp5-6.

<sup>58</sup> CDPA 1988 c. 48 ss.12(3)-(4), 57 <http://www.legislation.gov.uk/ukpga/1988/48/section/12> and <http://www.legislation.gov.uk/ukpga/1988/48/section/57>

## 6.5 Publication dates and authors' death ranges

The verification of the copyright status of the works in the sample revealed some interesting facts particularly with respect to the differences between UK and US rules.

- The earliest publication date for a work in the sample of **unknown copyright status** is 1877 – 134 years ago. At this age, the book may be out of copyright but as the author's death date could not be ascertained, the Library could not categorically state that the book is out of copyright.
- The earliest **in-copyright work** in the sample was published in 1892 (the author died in 1943).
- The most recent **public domain** title in the sample was published in 1967. This is because US treatment rules apply and this title fell under the transitional provisions of the 1976 US Copyright Act and therefore registration requirements rather than author death date form the basis of the calculation. The most recent **public domain** title under UK rules was published in 1935 – 32 years earlier than the American title.
- The latest death date for an author of a **public domain** work in the sample is 1996. This was because this author was American and his work was published in the 1940s. A shorter US copyright term applies to a work of that date (28 years after original registration unless renewed) hence the example from the study now being in the public domain. By contrast, the latest known death of a UK author of a public domain work in this sample was in 1939, 57 years earlier.
- As time goes on, however, these sorts of discrepancies will even out since the US now has a copyright term for literary works of life plus 70 years in parity with Europe.

## 6.6 Locating the rightsholders

After the establishment of the copyright status, searches were then made to try to locate any authors that were believed still to be alive, or their estates and representatives where it was unclear whether they were alive or not. Often the publisher is the most direct route to authors and artists, as not only will a commercial publisher be likely to be aware of where certain rights currently lie (i.e. if any rights assigned or licensed have since reverted to the author or if they have been transferred elsewhere), but they may know the author's current or most recent whereabouts, especially if they are commercially published and retain a literary agent. Also since author societies and author and publisher collecting societies may know their whereabouts if the person, publisher or agent is a member, it was decided to contact these organisations as well (see Section 6.9 below). However, it should be noted that it tends to be commercially published authors who join collecting societies so not all authors would be traceable in this way.

A wide range of sources were used to identify possible contact addresses for authors, their heirs, or their publishers. In some cases third party contacts were identified, such as the curators of special collections relating to certain authors held in libraries and archives, who were likely to have relevant information (see Section 6.7 below). It was much easier to find commercial publishers and evidence leading to those authors who were or are well known or active. Interestingly in some cases the best contact was the author's fan club.

Fig. 6: Sources used to find rightsholders' location

Sources interrogated	Comments on usefulness to this study
Nielsen's BookData Online	Useful for commercial publishers, but web searches proved more useful for finding current contact information for publishers. Was less useful for information on older publishers in sample which are no longer in business.
Publishers International ISBN Directory (CD-ROM)	Gives full names and addresses for publishers listed in the Directory but was less useful for the works in the study's sample as many pre-dated the use of ISBNs.
Internet search engines leading to publishers websites, authors' and fan websites, research archives, library and museum collections, genealogical websites (for heirs of peers) and a range of information sources.	Generally a good source for information on authors, current publishers and for a number of those out of business.
Wikipedia	Very useful for authors, publishers, newspapers, company takeovers and firms out of business.
Companies House	Very useful to find small publisher and non-standard publisher postal addresses not otherwise findable and to identify if firm still trading.
Google Books	Useful for exact imprint in order to distinguish between publishers with similar names. e.g. a common name followed by "Books", 'Press', 'Publishers', or 'Publications' etc.
Firms Out of Business – FOB <a href="http://tyler.hrc.utexas.edu/fob.cfm">http://tyler.hrc.utexas.edu/fob.cfm</a>	Useful for a few publishers in the sample.
Trade directories: Experian B2B Researcher, Publishers Directory UK and Irish Book Publishers, ISBN Directory	Used for contact information for publishers not found using search methods above, but no additional information found.
Writers and Artists Yearbook	Used for contact information for publishers not found using search methods above but no additional information found through this title.
Writers and Artists and Their Copyright Holders – The WATCH File <a href="http://tyler.hrc.utexas.edu">http://tyler.hrc.utexas.edu</a>	Useful for contacts for some authors' literary estates, agents or heirs though some of the matches proved not to be contactable.
Genealogical websites on the web	Useful in trying to identify heirs in a few cases, mostly where the author was from the peerage. On occasion used to try to identify death dates where the author was linked to specific locality although these searches were less successful if author was not well known and easily identifiable.

**Fig. 6: Sources used to find rightsholders' location continued**

Sources interrogated	Comments on usefulness to this study
Society of Authors website and contacted directly	Society of Authors contacted directly regarding 74 authors not located or for whom no death date had been found, but, although very helpful, either had no or insufficient information to allow the study to approach rights holders for permission to digitise. Its website says it only administers a limited number of literary estates directly. Website only useful for some contemporary authors and only a third of the membership agree to be listed publicly on the website.
Design and Artists Copyright Society (DACS) website	Used membership database query interface on website to check if DACS represented nine named illustrators' but there were no matches. Also contacted DACS directly who were very helpful and searched their records but still no matches were found.
Authors Licensing and Collecting Society (ALCS) contacted directly	Website only has a by-title public query interface to database of works for which ALCS holds unclaimed royalties. No public query interface to membership database. ALCS directly contacted regarding 74 authors not located but they were unable to help with further information.
Publishers Licensing Society (PLS) contacted directly	No relevant information on website. Contacted concerning 20 publishers not located. They were helpful but unable to furnish any productive leads in this case.
Writers Guild contacted directly English PEN contacted directly	Both contacted regarding 74 authors not located. No response before the study closed on 11/2/11.



Searching the web was quite productive as the commercial publishers who are still in business or who had taken over the imprints of former publishers were easy to find. Some of the big publishers even mentioned the imprints they covered and some usefully provided potted company histories in their 'About Us' pages that confirmed the takeovers of older publishing firms. However, once within publishers' websites, it was often quite difficult to track down what looked to be the right contact information for rights clearances.

Apart from business to business sales of rights such as foreign distribution and translation rights, publishers' rights clearance pages tended to focus on rights clearances to use extracts or parts of their publications and it was sometimes unclear where the boundaries lay when seeking to digitise a whole work for non-commercial purposes, e.g. whether the publisher's permissions department should be approached or the rights department. Since publishers generally wish to exploit the digital rights they hold in whole titles themselves, it is unsurprising that digital rights clearance is not mentioned on their websites.

Wikipedia proved a very useful source for information on some authors, on company takeovers and some firms out of business. The Companies House database was particularly useful for finding the postal addresses of less well-known firms and identifying if they were still trading. As far as the sample was concerned, the Firms Out of Business (FOB) database was of some limited help but less useful than Companies House, Wikipedia or the open Internet. Trade directories such as the *Experian B2B Researcher* database, *The Publishers Directory*, *The ISBN Directory* and *the Writers' and Artists' Yearbook* were also checked after the open Internet sources but did not produce any new contact information.

The search for author contacts led on a few occasions to genealogy websites, searching for the names and whereabouts of current descendants of (usually titled) families, and also to fan or enthusiast sites. Special collections on the web from museums, archives, libraries, and research centres were also helpful. In particular, these would at the very least provide contacts to the staff who possibly would have relevant rightsholder contacts on file for their special collections or be able to refer us on.

On the other hand, finding the heirs of one illustrator named in a MARC record, proved extremely simple since one of his daughters managed his estate and maintained a website about him with clear contact details for rights clearances. Other authors' heirs were found by following up mention of children in the biographical sources consulted about the authors, and then searching for and finding them on the Internet and in online telephone directories using whatever information was indicated about them.

## 6.7 Contacting the rightsholders

At the end of the search for rightsholders, the Library had managed to establish contact addresses of some kind (or at least details of third party intermediaries such as libraries, archives and museums as well as publisher and author collecting societies to approach with further enquiries to find rightsholders) for 80 of the 102 works for which it was trying to clear the rights. The remaining 22 had already been identified as orphan works because it could not locate anyone to whom to direct any enquiries.

The Library took note from Troll Covey's study that the Carnegie Mellon Library had initially taken a title-by-title approach to its clearance requests and had not bundled together all the requests relating to a specific rightsholder and that this had led to a large number of letters being sent. It was therefore decided that it would be more efficient for the Library and more helpful to the rightsholders if the Library took a by rightsholder approach and sent just one letter which covered all the works from the sample relating to the rightsholder in question.<sup>59</sup> This equated to 74 different lines of enquiry, or people and organisations contacted, as in a few cases a contact was to be approached about more than one work.

All identified contacts, in some cases very tentative ones, were sent letters from the Library seeking permission as part of a feasibility study to digitise their whole work(s) for placing on the Internet for open and free public access including downloading and printing (see Appendix 4). The letters explained why the Library thought the recipient might be the rightsholder or know about him or her, and asked for information as to the location of the rightsholder if they were not the right people to grant the required permission. Standard letters were sent, merged with the appropriate data, but letters were also adapted as necessary e.g. where third parties such as libraries, museums or archives were being approached in the quest for location or death date information about rightsholders.

Contact by the Library was by email where possible, or, if specified, via standardised web forms on publishers' websites. Where electronic contact was not an option, letters were sent by post or fax. Standardised web forms often proved constraining as their design was often based on the assumption that permission was being sought to use part or parts of the work in a new work, and not that permission was being sought to digitise the whole work, or even to digitise at all.

Altogether there were at least three attempts to contact rightsholders, although in some cases there were more due to referrals to other parties or because repeated phone calls were made in attempt to make a first contact. The first round of contact took place mainly between 22 October and 5 November 2010, though some first contacts and referral contacts were made later on in November due to constraints on available Library staff time. The 41 contacts that did not respond substantively within four weeks of first contact by granting permission or withholding it, were followed up by a chaser letter in the week of 13 – 22 December 2010 and a number of referrals were first contacted at the same time. Another month was then allowed before following up again for a third and final time, chasing the remaining 28 who had not yet responded. These final contacts took place between the beginning of January 2011 and the close of the study on 11 February 2011. This time the methods used were telephone, email and recorded delivery letter where a postal address was found and no other means of firm contact had been established.

Due to the time constraints of this study, a total of 11–15 weeks since first contact, depending on the date of the first letter sent, was allowed for responses before the contact phase was brought to a close. It was felt that this was entirely reasonable and was comparable with the Troll Covey study's general approach of one month first contact phase followed by 60 days for their second and final contact.<sup>60</sup> We were aware, however, that major publishers have to handle large volumes of requests but expected to at least receive an acknowledgement from them fairly quickly.

---

<sup>59</sup> Troll Covey, Denise (2005) Op cit. p.13

<sup>60</sup> Troll Covey, Denise (2005) Op cit. p.12

This generally was the case, although some of the larger publishers failed to respond at all. Also, in some cases, a secondary contact (a referral) was identified by the primary contact, which meant that the request went to the new contact straight away but that they would not have had the full four month period in which to respond. However, even those secondary contacts approached late in the contacting phase of the project would have had at least four weeks in which to make some kind of response if they wished.

As far as individuals and non-mainstream publishers such as organisations and associations were concerned, responses were quite patchy and many did not respond at all. Often when they did respond they were unsure how to handle the matter or whether or not they had the rights. They generally needed additional time to search for papers and make up their minds.

Third parties such as archives, libraries and museums with relevant special collections to whom the Library sent enquiries for information about rightsholders' death dates or whereabouts, were with one exception (an American museum library) generally quick to respond and were very helpful in supplying contact information for rightsholders and researching death dates from information they held.

## 6.8 Dealing with the rightsholders

This stage of the project involved a lot of human contact which sometimes bore quite mixed results. It is worth recording some of the experiences of the Library staff who had approached the rightsholders.

- **A national forest policy, by Bulmer Hobson. Dublin: B. Hobson, 1931:** The Library had only been able to trace a possible telephone number for the author's daughter and son-in-law after finding a reference on the web which mentioned their location and indicated that in the 1960s the author had moved to live with them until his death in 1969. It seemed unlikely that they would still be in the same location after 41 years, but after a few attempts the Library succeeded in speaking to the author's now elderly daughter. Her husband called back to say that he had checked the author's papers but he and his wife were unsure about who holds the copyright. The author's will made no mention concerning the specific disposal of his literary estate and copyright, but he had left his whole estate to his daughter and son. The son is now deceased without issue and his will also makes no mention of the rights in his father's literary estate. The daughter's husband confirmed that most of the author's works were self-published and said that he and a partner had set up their own publishing house. However, apart from finding a reference to this publishing house on the web, the Library was unable to establish any contacts for it. The Library was given the author's grandson's email address for further contact and he was sent a copy of the rights clearance request letter and book details, but he had not replied by the time the study was closed.
- **Yorkshire farming memories, by Stephen Harrison. York: Castle Museum, 1981:** The Castle Museum was contacted by the Library via its web enquiry form. The Museum's response was helpful. They acted as intermediary to obtain the author's clearance but they revealed that there were third party rights for dialect poetry in the work that the Library needed to clear separately with the Yorkshire Dialect Society (YDS). The Library had been unaware of these additional rights as there was nothing to indicate these in the MARC record for the book. The Library contacted the YDS, which granted permission subject to acknowledgement. However, the YDS said in response "What has stumped us is which poems and, specifically, photographs have been used! You don't have handy lists, do you?" This information could only have been gained by looking at the book itself as the MARC record does not have such a deep level of detail.

- **A Tyneside parish, by R. E. Holmes [Rev. Richard Ellis Holmes]. South Shields: E. Sword, Printer, 1911:** The author's birth date had been established as 20 March 1863 (using the World Bibliographic Information System resource) but no death date could be found. As he was a local writer, the Library decided to try South Shields Local Studies Library for information to try to establish a death date and if possible the location of any heirs. This library responded quickly and established from a press clipping that Holmes' age at death was 95 years, so the Library was able to calculate that he probably died in 1958. The book is thus in copyright until the end of 2018.
- **Scottish poet, publication in 1980s (anonymised as the author was contacted directly):** No trace had been found for the publisher's whereabouts but the author's email had been found in the Literature Resource Center database when checking to see if she was still living or not. The author responded to the Library's first chaser email granting permission. The publisher still had rights in the typographical layout in that edition but the author was unable to help find them "I've no idea how to contact [the publisher] I'm afraid, but as far as I'm concerned, go ahead."
- **The story of Winchelsea Church, retold and revised by Malcolm Saville; new photographs by A. Vaughan Kimber. 18th ed. St. Leonards: Kings Bros & Potts, 1978:** The Library had only been able to find intermediaries for the author; two societies formed in admiration of the author's work who were contacted. One society supplied the email contact for the author's grandson who is the administrator of his estate. The grandson responded to the Library's email saying he was unaware that his grandfather had written the book but that to the best of his knowledge the author's literary estate (of which he was an executor) owned the copyright. The Library had meanwhile also telephoned the Canon of Winchelsea Church, as the church publishes the current edition of the guidebook,

to seek any assistance he could offer. He revealed that the 18th edition was a "revised and retold" version of a book originally written by an author called Gertrude Leigh. This information was not in the MARC record for the book. The Library informed the author's grandson of this and he answered that in the light of this information he thought he could not give permission without permission also being secured from Gertrude Leigh's estate. This information came to light late in the course of the study, so the Library was unable to trace any information about Gertrude Leigh's death date or find her heirs in the time available. It remains unclear to what extent she held any rights in the 18th edition of the work. The Library then received a letter from the Canon granting permission with the statement "Copyright held by the ecclesiastical parish of Winchelsea" which suggests that the Parish thinks it is the copyright owner in the 18th edition but again this required further confirmation, as normally any assigned rights in a work published 33 years ago would by now most likely have reverted to the author(s). This confusion could not be resolved by the time the study was closed, so as the Library was still negotiating with the purported rightsholders the work was classed as 'permission not granted'.

# Results from the study

## 6.9 Contacting rightsholder societies

The Library also sent letters by email to four author organisations, the Authors Licensing and Collecting Society (ALCS), English PEN, Society of Authors and the Writers Guild, for help with 74 authors, or their heirs or literary executors for whom no point of contact could be found or year of death identified. The Publisher Licensing Society was also approached for help to find 20 publishers for whom the Library could find no contact details. DACS, the Design and Artists Copyright Society was also telephoned after having interrogated their membership database via their website for nine illustrators or photographers without finding any matches. These bodies were all told that “A library consortium is currently evaluating a digitisation project of a selection of 19th and 20th century books with a view to putting them on a public facing website for view and print purposes only. The project is not-for-profit and no charge will be made to members of the public to view or print out”. They were asked if they could help locate the authors or provide any further information to add to what the Library by running the names in the spreadsheet sent to them by the Library against any records they may have. The letters sent can be read in Appendix 4.

The Publishers Licensing Society, DACS and the Society of Authors were especially helpful but were unable to provide many new leads to those that had already been established. PLS was asked to help identify direct contact information for 20 publishers, or their subsequent owners where company takeovers had occurred, but could only offer three suggestions, none of which proved fruitful as the Library was unable to make contact with them. DACS had no records of the nine named illustrators the Library was seeking. The Society of Authors was able to return results for six authors: only one was a member, three were lapsed or resigned members and they had records for two who died in the 1940s but they had no contact details so could not forward anything for the Library. Of the three other authors' organisations that were approached, two did not respond before the study was closed and another said they were unable to assist.

## 7 Copyright status of works in the sample

Before contacting any rightsholders or intermediaries, the study had ascertained that 80 works were in copyright (57%), 38 works (27%) were in the public domain and 22 (16%) were of unknown copyright status because we could not by that stage discover when or if the author(s) had died. This meant that it was necessary to investigate clearance of the rights in 102 works (73% of the sample), comprising both the in-copyright works and those of unknown status. (See Fig. 7 for a graphical representation of this).

After having contacted the rightsholders, the information they supplied adjusted the copyright status of a few works resulting in the number of status unknown works dropping by 2% as two of them (from the 1910s and the 1940s) were confirmed by the publishers still to be in copyright – see Fig. 8. Two books became public domain works: one entered the public domain during the course of the study at the end of 2010, and another work from the 1890s where the editor's copyright might still have been extant, was confirmed by the publisher to be in the public domain. There was no change in the overall number of works in copyright (80 works, or 57% of the total). A comparison of Figs. 7 and 8 below shows that there was very little change in the copyright status of the works as a result of contacting the rightsholders.

Fig. 7: Copyright status of works in the sample – prior to contacting rightsholders

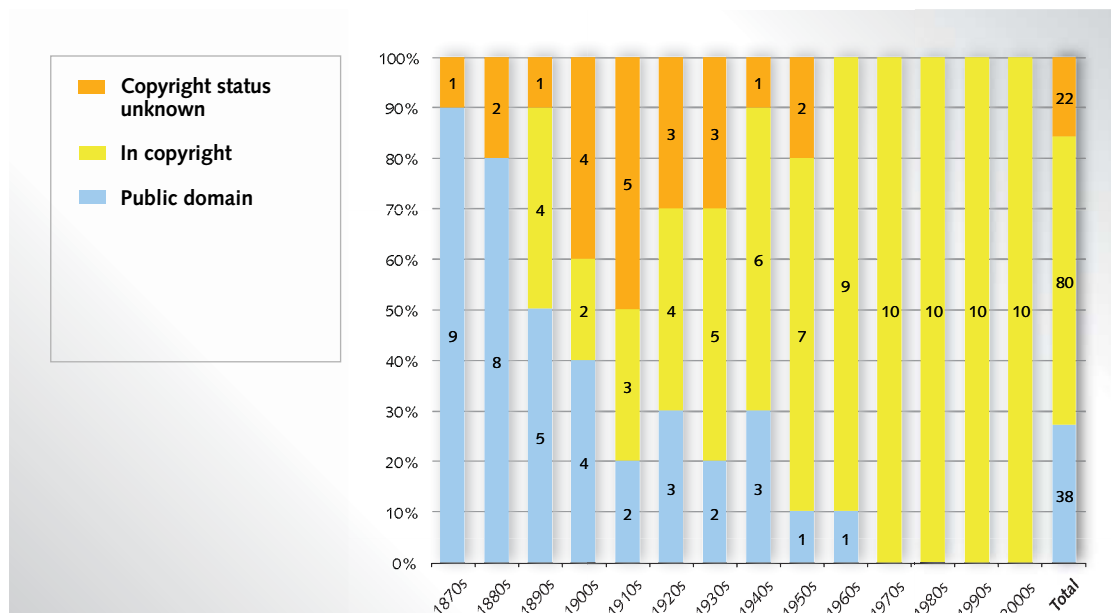
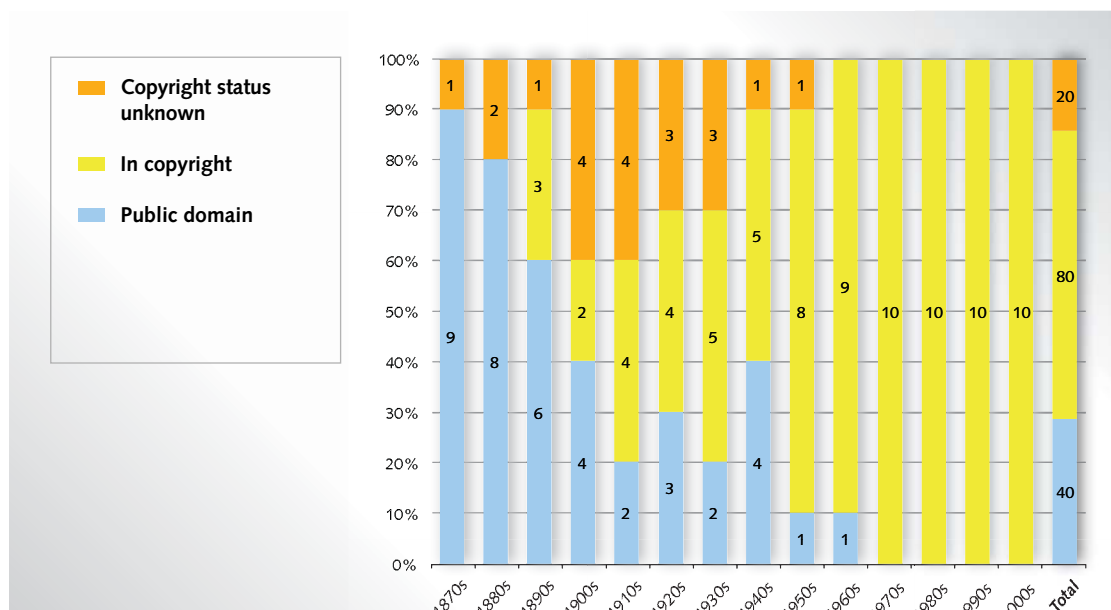


Fig. 8: Copyright status of works in the sample – at the end of rightsholder contact phase





## 7.1 Validation of copyright status and grant of permission through ARROW

A parallel track to this study was a comparison exercise carried out with the ARROW project team running the 140 MARC records in the sample against the ARROW system and comparing the results against the manual search findings.

The final report on the ARROW project<sup>61</sup> provides greater detail than this report on the system architecture and workflows involved, as does an accompanying paper produced by UKRRO - the group of UK rightsholder organisations involved in the project (see Appendix 6).<sup>62</sup> However in summary the ARROW system works by querying the databases of a number of partner organisations which are grouped into the following areas:

- European Libraries (TEL Central Index)
- Virtual International Authority File (VIAF)
- Books in Print (BIP)
- Reproduction Rights Organisations (RROs)

The ARROW system first matches the work title information submitted by a user against the European libraries' catalogues and then interrogates the VIAF file to determine authors' death dates and ultimately copyright status. The BIP database identifies whether the item is still in print and retrieves publisher information and finally the RROs' databases are interrogated to determine whether permission to digitise could be granted.

The comparison between the British Library manual search and the ARROW automated search 'pilot' highlighted some differences in the outcomes of using the two approaches.

Firstly, only 101 out of the 140 books could be uploaded to the ARROW system and run through the validation process. This was because either the country of publication was not handled by ARROW (for example where the

material was published outside Europe<sup>63</sup>) or because the place in the MARC cataloguing record in which the country of publication information was held was not compatible with the configuration of the ARROW system at the time of the pilot.<sup>64</sup>

A comparison of the copyright status results of the 101 books which could be loaded between the manual and the automated ARROW processes showed that in 51% of the cases the same results were returned, although this figure was subsequently improved upon after the study was completed. The differences in the remaining 49% of titles were as a result of missing death dates in the MARC record and the inability of the system to determine the copyright status of corporately owned or anonymous works at the time of the pilot.<sup>65</sup>

However despite the differences between the outputs of the manual search and the ARROW pilot, only eight records out of the 101 put through both processes differed in the ultimate rights clearance outcome (i.e. whether permission could be given to digitise).<sup>66</sup> In five of these cases the ARROW system returned a denial of permission to digitise when the manual search had determined the book was in the public domain. In two of the cases the ARROW system had granted permission on the basis that the work was public domain when the manual search had determined that the copyright status was unknown (and therefore the book could not be digitised). In only one case did the ARROW system identify a work as public domain and available for digitisation when the manual search had determined it was still in copyright and no permission could be found.<sup>67</sup>

The fact that the differences in the ultimate rights clearance responses were only marginal are extremely significant when considering the huge time savings that the ARROW system provides in identifying the copyright status of the works (see section 11 Time Spent).

---

<sup>61</sup> [http://www.arrow-net.eu/sites/default/files/D1\\_14\\_ARROW\\_public\\_report\\_final\\_def.pdf](http://www.arrow-net.eu/sites/default/files/D1_14_ARROW_public_report_final_def.pdf)

<sup>62</sup> Bide, Mark and West, Tom (on behalf of PLS/ALCS) (2011) ARROW – A service in development to assist “diligent search.”

<sup>63</sup> At the time of the validation the ARROW system was only capable of searching for books published in the UK, France, Germany and Spain. Expanding this out to other countries within Europe is the focus of the next phase of ARROW's development - ARROW Plus.

<sup>64</sup> In order to ensure optimum compatibility across all national libraries contributing to the ARROW project, the pilot system was configured to accept as many catalogue records as possible from the entire sample. As a result of differing standards in cataloguing across contributing libraries, the system could not accept some British Library records in which the country of publication was held in the '008' field. Further system improvements could address this issue should it be required in future.

<sup>65</sup> After the pilot the ARROW project team reconfigured the system to determine the rights status of anonymous or corporately owned works. This significantly improved the results with a 69% match between the manual and the automated search.

<sup>66</sup> Details of the rights clearance work undertaken by UKRRO is described in Bide, Mark and West, Tom (2011) (as detailed in Appendix 6). Op. cit. It was this work that ensured the ultimate rights clearance outcome was comparable for the 101 titles run through the ARROW system.

<sup>67</sup> This particular example was corrected in the subsequent post-pilot reconfiguration of the ARROW system so that both the automated and the manual searches returned the same result.

## 7.2 Examples of differences between manual and automated clearance results

- **Jones, Gascoigne Hastings Fowler 1850–1911: Sketches in York:** This work has been in the public domain since 1981, 70 years after the death of the author in 1911. However it was only possible for the researcher to determine this following a search using a range of Internet resources. The ARROW system returned a status of 'in copyright' as only the birth date was as in above example included in the MARC record. In cases where the death date is missing, ARROW applies a rule of 170 years from the date of the author's birth to ensure an estimate of the work's copyright status that would cover all but the most extreme examples.
- **Clarke, Mary Cowden, 1809–1898: The Complete concordance to Shakespeare: being a verbal index to all the passages in the dramatic works of the poet, New and rev. ed. London: W. Kent & Co., 1874:** This work has been in the public domain for many decades. The copy in the sample once belonged to Lord Alfred Douglas (Alfred Bruce Douglas 1870-1945) who was both an author in his own right and also famous for his relationship with the playwright Oscar Wilde. Therefore his name, birth and death dates had been entered in the '700' added names field of the MARC record stating that he was a former owner of the copy in question. Additionally, the entry from another of the fields in the MARC record states 'From the library of Lord Alfred Douglas, with his name in MS on fly leaf verso.' The '700' field is used for a range of added names of people known to be associated with the content or the specific copy of the work in varying roles, including additional authors or contributors. The various uses for the MARC '700' field by the cataloguing librarian at the time the record was created in this case have led to the ARROW systems classing Alfred Douglas as a co-author of the *Complete concordance to Shakespeare*, which he is not. As a result the book has been given an 'in copyright' status by ARROW even though it is in the public domain.

## 7.3 Summary of ARROW validation

Whilst there were differences between the results gained from the ARROW enquiry and the manual search, the validation was regarded as a success given the current level of development of the ARROW system. The 51% concurrence of copyright status results between the two methods was increased to 69% after the pilot took place showing that improvements to ARROW have the potential to enhance the system's capabilities even further and make it a central tool in the digitisation of library collections by cutting down time and resource needed to establish copyright status.

## 8 Orphan status of works in the sample

As previously stated an 'orphan work' is a copyright work for which the rightsholder(s) cannot be located or identified. Public domain works are by definition not orphan works since there is no need to locate their former rightsholder(s). Although some titles were evidently orphaned from the start as soon as it was established in the search phase of the project that as no trace could be found of the author or publisher, a final decision on which works could be considered orphan works and which not, was not possible to make until the contact phase of the study was completed.

In order to determine whether a work was an orphan, the following criteria were taken into account after carrying out a diligent search:

- If the Library was confident that the correct contact details for the rightsholder had definitely been found either directly or via a referral by the first contact approached, regardless of whether the rightsholder had responded, then the work was deemed to be **not orphaned**.
- If a response had been received and the matter was still in hand, the work was treated as **not orphaned**, although its status could have changed later should there be a response that the contact was not the rightsholder.
- Where there was uncertainty that the true rightsholder had been found *and* no response had been received, it was assumed that the work was **orphaned**.
- In cases where there was no trace of the rightsholder's location, or enquiries had reached a dead end, it was assumed that the work was **orphaned**.

### 8.1 Proportion of Orphan Works

At the end of the search phase, 22 of the 102 works (16%) were identified as orphan works as no contact point could be located (see Fig. 9). Following completion of the contact phase the figures were re-examined applying the criteria set out above and it was found that dead ends had been reached with regard to an additional 21 works, making a total of 43 orphans, i.e. 31% of the works in the sample (see Fig. 10).

15% (21) of the titles were definitely in-copyright orphan works and 16% (22) were orphan works of unknown copyright status (see Fig. 11).

By comparison, Troll Covey (2005 p.14) had identified 21% of their much larger final random sample of 209 publishers and 277 titles as non-locatable, i.e. orphaned. This should be considered in the context of the shorter term provision applied in the USA prior to the 1976 Act under which copyright in works had to be registered and could be protected up to a maximum of 56 years. The 1976 US Copyright Act brought in a life plus 50 years copyright term but this was not applied retrospectively although some provision was made by extending the renewal period by 19 years so that the maximum term became 75 years (see Appendix 2 for further details). Furthermore, since the copyright status of the works relied on registration, the Carnegie Mellon Library were able to refer to copyright registration files.

Figs. 9 and 11 illustrate the changes in the presumed numbers of orphan works in the sample at the beginning of the search phase and after seeking to locate the rightsholders.

Fig. 9: Orphan works in the sample before contacting rightsholders

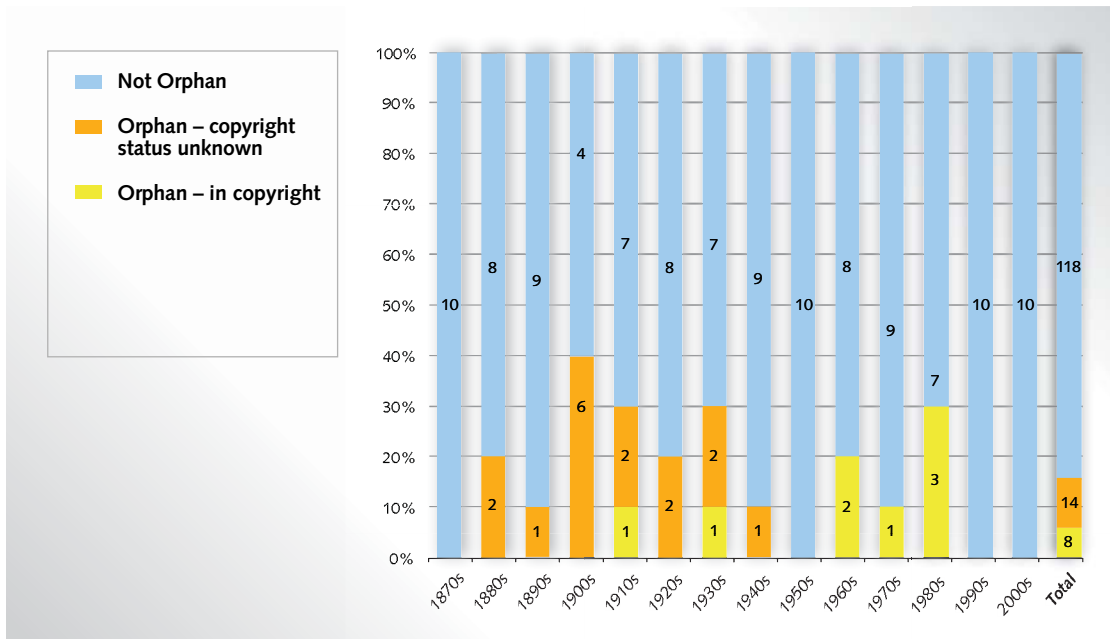


Fig. 10: Number of orphan works, public domain and in-copyright non orphan works after the contact phase

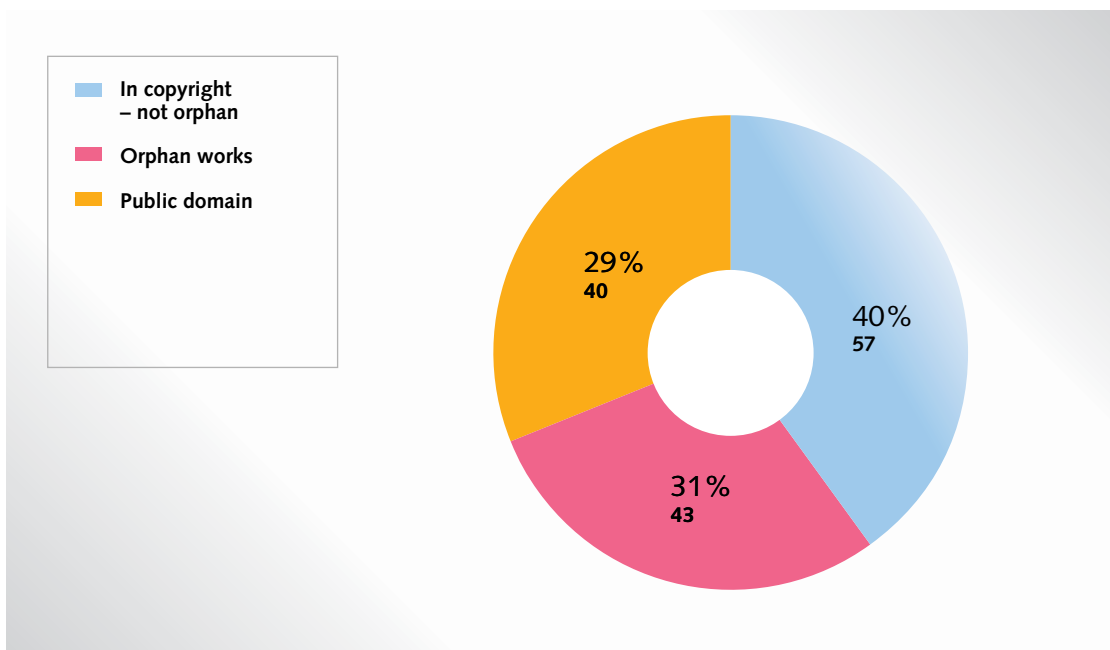
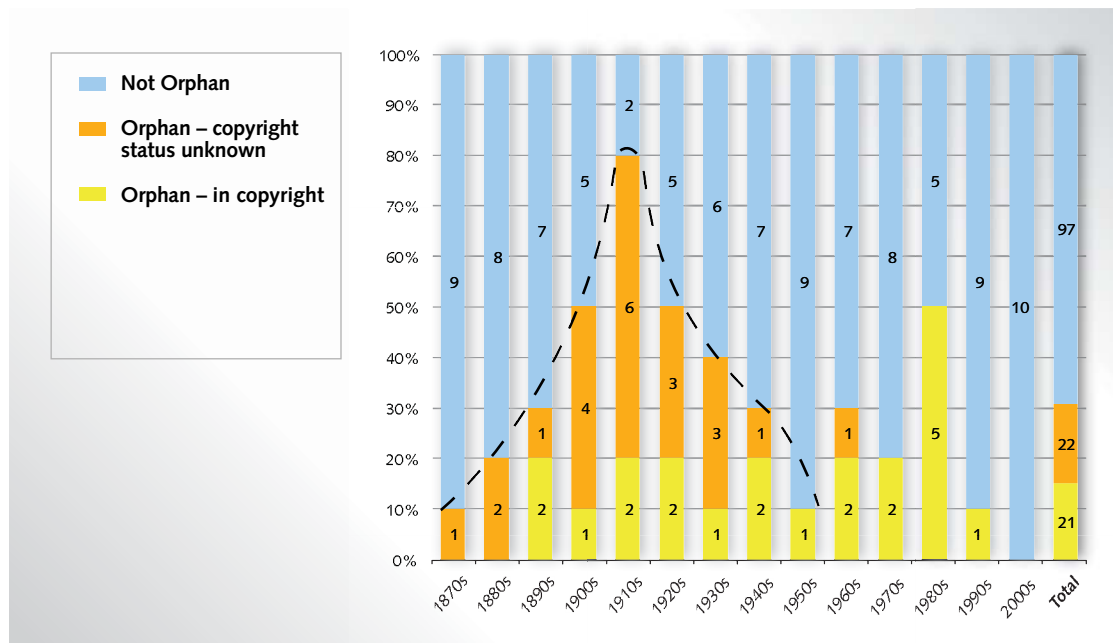


Fig. 11: Orphan works in the sample after contacting rightsholders



By the time the contact phase drew to a close, the number of orphans nearly doubled, rising from 16% of the sample to 31% (43 works). These 43 orphan works were published during the 13 decades between the 1870s and the 1990s. They were most prevalent in the 1910s (80% of works from that decade were orphans) followed by the 1900s, 1920s and 1980s (all of which had 50% orphan works). The only decade without any orphans was the 2000s.

## 8.2 How do works become orphaned?

While it was not the purpose of the study to establish the reasons that the works become orphaned, the results of the study enable us to draw some conclusions from the 43 orphan works that were identified with publication dates between 1877 and 1999.

The age of the work would be the obvious answer and this does appear to be a significant factor in a work becoming orphaned. There are more orphans in the sample in the period 1900-1959 than in the period 1960–2009. A crude arc can be seen in Fig. 11 – there are fewer orphans in the earliest decades of the sample where more of the titles are in the public domain but gradually the number increases to a peak in the 1910s and then decreases again as more recent books are more likely to be in-copyright and in print.

Significantly when thinking about age, the 1980s rather obviously bucks this trend with five orphan works. However this reflects that all 10 publications in the sample from that decade were the product of non-mainstream publishers and small presses. The orphan works from this decade included a slim volume of poetry published in 1986 by Plaistow Poets' Press in London, a local history booklet published possibly by a printer, and a volume of short stories published by Golden Cow Productions Ltd, a company of which now no trace could be found. As discussed in section 5, smaller publishers are harder to trace and less likely to keep records so the kind of publisher also appears to be a major factor.

The second largest group of orphan works (11) in the whole sample had been published by traditional commercial publishers. They had either disappeared, did not respond to enquiries so who now owned the rights remained unknowable, or they had no information as to the whereabouts of the authors or their heirs after the expiry of the publishers' rights in the works.

### A breakdown of these by decades is as follows:

**1880s:** both publishers, one based in Dublin and one in London had gone out of business.

**1890s & 1900s:** one publisher from the 1890s and another from the 1900s are now merged into the same large international publishing house which did not respond to enquiries before the study closed.

**1910s:** two publishers from the 1910s are still in existence – one merged into a new company which did not respond to the Library's enquiries before the study closed and the other is a remnant of the old company which had no information about the rightsholder.

**1920s:** the commercial publisher still exists but had no information about the rightsholder.

**1930s:** one publisher appeared to have disappeared and the other still exists as the Indian subsidiary of the original UK based international publisher but did not respond to enquiries before the study closed.

**1980s & 1990s:** both the commercial publishers were small presses that have since gone out of business. One in particular, published in 1999 by Paragon in Leeds was classed as 'active', i.e. in print, by Nielsen BookData but with the caveat added to Nielsen's database during the period of the study that the publisher's address is invalid with no forwarding address. The Library was unable to trace any contact details for the author or the publisher. Likewise the Publishers Licensing Society could not trace the publisher and the Society of Authors had no records for the author. In just 12 years since publication by a small commercial publisher, tracking down the rightsholders had become impossible.

Another likely cause in the orphan status is the kind of publication and the following examples from across the sample bear out that the more ephemeral the work the more likely it is to become orphaned:

**1890s:** Useful recipes for coachmen, grooms, carters, farmers etc - Recipes with advertisements for horse and cattle medicines, oils and drugs, to be supplied by J. Tully, veterinary chemist and druggist, St. George's Road, Hastings. The Library could not trace any contact details for the author or the publisher.

**1900s:** Lines on Wolds Hill and the surrounding landscape. The researcher was unable to trace contact details for either the author or publisher.

**1950s:** The architectural history of Lamport. The Library was unable to trace contact details for the author. The researcher contacted the Northamptonshire Archaeological Society which was founded in 1974. They came back to say that they had nothing to do with the original society.



Fig. 12: 43 orphan works by type of publisher after contacting rightsholders

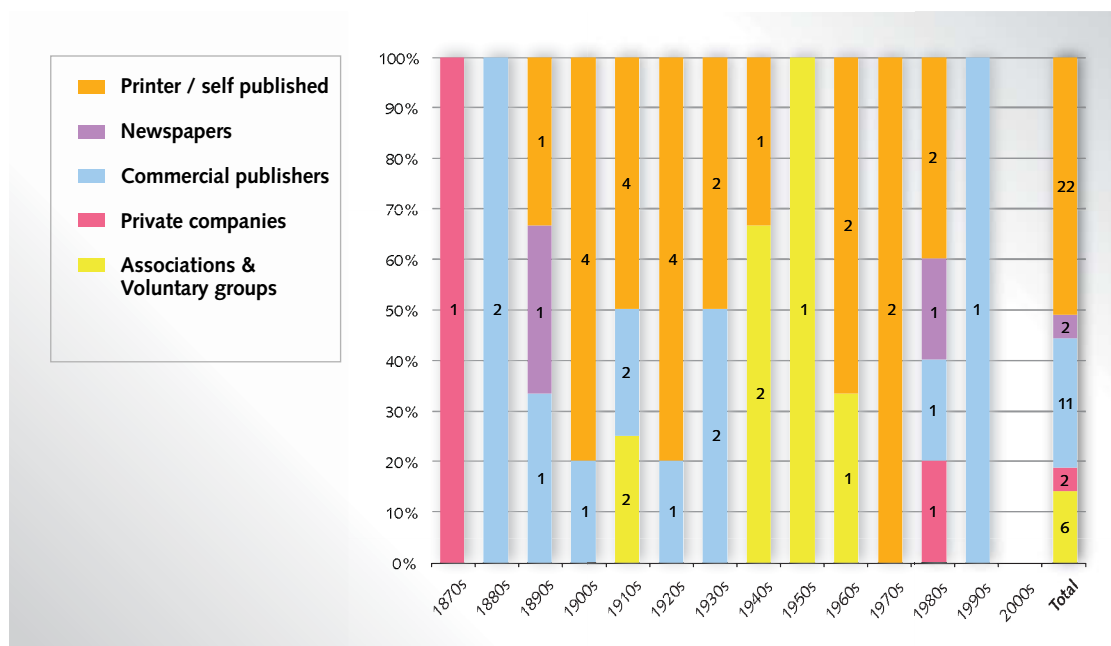


Fig. 12 shows the distribution of orphan works at the conclusion of the project by type of publisher across the decades. This shows that the largest group (51%) of the 43 orphan works had been self-published but the second largest group of orphans (11 works or 30%) had been published by what were believed to be commercial publishers. Of the 43 orphan works identified in the study (31% of the whole sample of 140 works), 22 (or 51%) of the orphans appeared to have been self-published or produced by small presses or printers.

The examples and patterns illustrated above suggest that a myriad of reasons can cause works to become orphaned, from people moving away to firms going out of business and as time passes the risk increases. It would seem, however, that the more personally produced and distributed or the less commercial the publication is, the greater that risk becomes a reality. This greatly affects the possibility of digitising collections in libraries archives and museums because as public interest institutions they hold a huge range of works which are not commercially published or, in the case of archives, not published at all.

## 9 In print status of works in the sample

The titles were checked against *Nielsen BookData* to determine whether they were in print<sup>68</sup> and whether there is any relationship between in print and orphan status or the likelihood of rights being cleared (see Figs. 13 and 14). This established that:

- 21% of the works in the sample were in print, i.e. 29 of the 140 works.
- Of the 29 works in print, 59% were in copyright, 38% were public domain and 3% were copyright status unknown. (See Fig. 14).
- 22 (76%) of the publishers of the in print books, were based in the UK, 4 (14%) were from the USA with the remaining 3 (10%) in India. 28 (97%) of these publishers were commercial publishers, including print on demand publishers, and the remaining publisher was an institution.
- 17 of the 80 in-copyright books in the sample were in print: 10 of these were published in the 2000s decade (in fact in 2008-09) and four in the 1990s (i.e. 60% of the 1990s books were out-of-print). This shows that much of the older material in library collections is no longer published.
- 11 of the 40 books in the public domain are listed as in print, all of which were offered on a print on demand basis through US based publishers.
- Of the 43 orphan works in the sample only three of them are still in print and two of these are available only via print on demand US publishers, where the titles are now in the public domain.

In Fig. 13 an arc is visible for the in print status with peaks in the earlier years of the sample (where much of the material is in the public domain, particularly in the US up until 1923) decreasing in the first half of the 20th century as material remains in copyright in the EU but is out of date and peaking again in the 1990s and 2000s where material is more recently published.

There are two variations in this curve: the 1940s where as discussed in section 8.2 much of the material is ephemeral in nature and in the 1960s where the sample included a very varied range of types of publisher (see Appendix 1).

Conversely, the figures also show that 111 of the 140 works in the sample are out-of-print.<sup>69</sup> Of these, 29 (26%) are public domain works but 82 (74%) are status-unknown or in-copyright works (63 works or 57% of the out of print sub-set were in copyright).

In the sample, 100% of the 1980s works were out-of-print, as were 60% of the 1990s titles which indicates that the commercial viability of many publications and thus authors' royalty earnings from sales of hard copy books is in many cases limited to about 20 years at best. As stated above, all of the sampled works from the 1980s were also ephemeral works or slim volumes which were self-published or published by small presses and 50% of them were found to be orphans. The nature of the publications probably made it more likely that they were one-offs and more likely to go out-of-print and not being purely commercial publications they are also less likely to get picked up by commercial publishing databases such as Nielsen BookData. With regard to the 1990s, seven publishers were commercial of which three were international. The remaining three were careers and employment organisations and a small press (Paragon of Leeds – see Section 8 above) which seems to have gone out of business. This demonstrates that it is only the very small minority of published material that has a long term commercial life and this would suggest that the vast majority of items in library and archive collections are unlikely to be digitised by rights holders themselves in their own commercial interests.

---

<sup>68</sup> 'In print'. Said of a book which is available from the publisher." Prytherch, Raymond: *Harrod's Librarians' Glossary and Reference Book*. / 10th ed. Aldershot, Hampshire and Burlington, Vermont: Ashgate, 2005. p.343.

<sup>69</sup> "Out of print. A book is out of print when the publisher has no more copies for sale and no intention to reprint. Abbreviation O.P." Prytherch, Raymond (2005) *Op cit*. p. 515.

Fig. 13: Analysis of print status by publication decade including copyright status of in print works

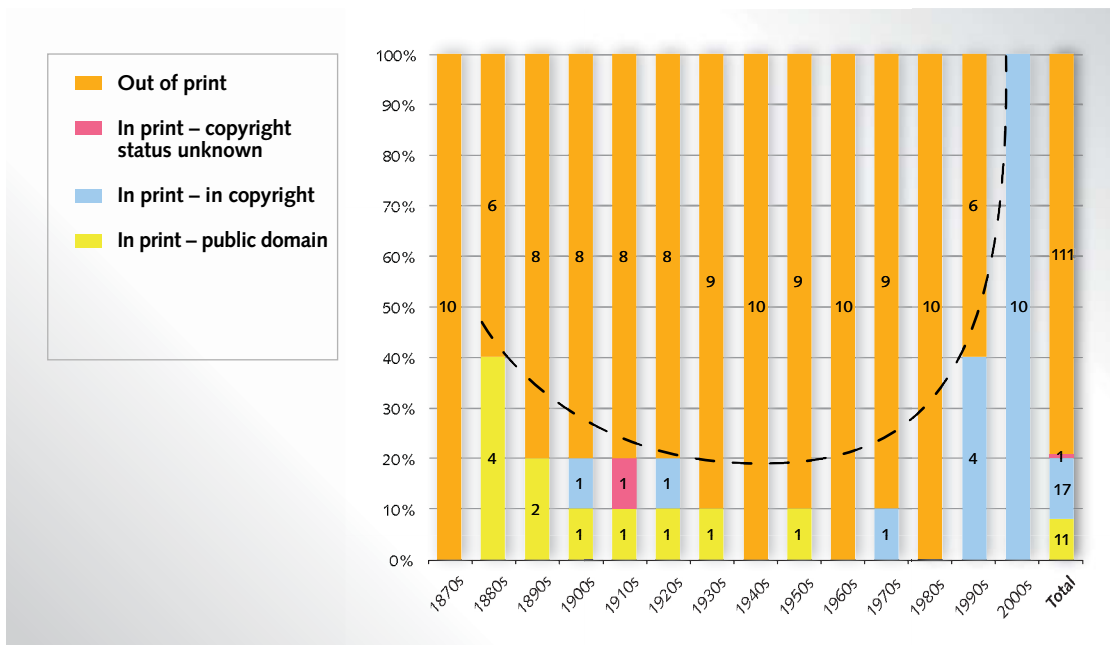
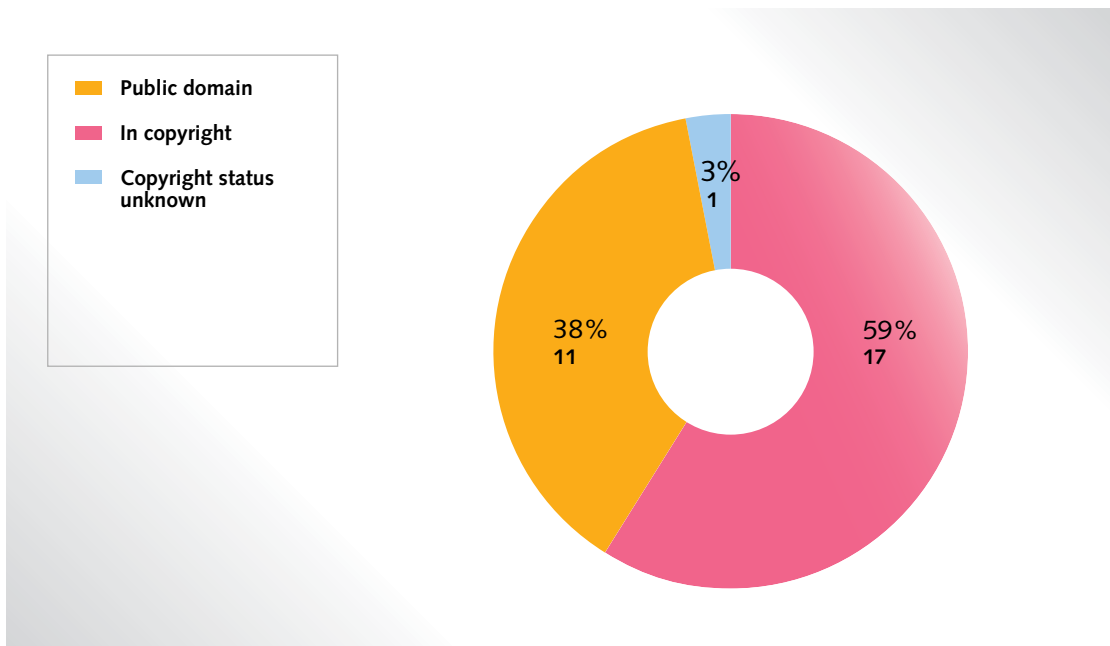
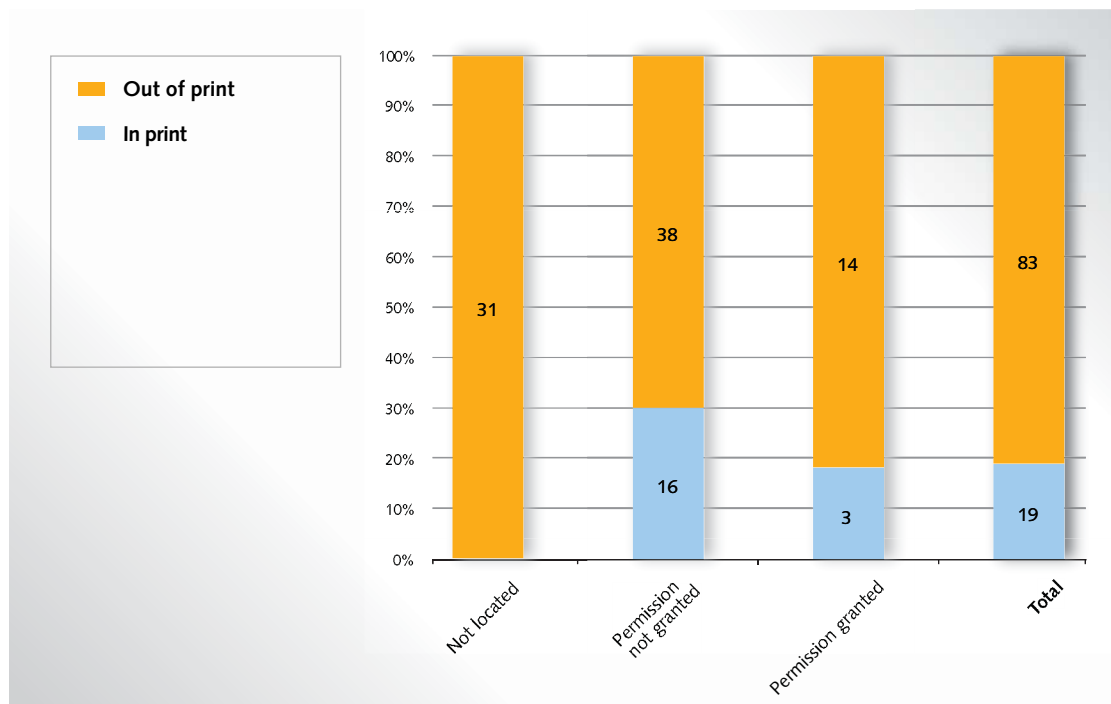


Fig. 14: Copyright status of the 29 In Print works



It seemed logical that whether or not a work is in print would probably affect a rightsholder's decision whether to allow digitisation for public access on the Internet. As far as this study was concerned, the status only had a noticeable impact on clearances for the works in the 2000s decade where all the works were actively in print (see Fig.13 above) and clearance to digitise was obtained only in respect of one work from that decade. Permission was not granted in respect of four titles and the publisher had not responded in respect of the remaining five before the study closed. Of the titles where permission to digitise was granted, 82% of the works were out of print and of those where permission was not granted (which includes instances where there had been no response, or where negotiations were still in train by the time the study closed), 70% were out of print (see Fig. 15 below). The group most difficult to locate were the self-published and printers, in fact not a single printer in the sample had been successfully located.

**Fig. 15: Analysis of rights clearance results of 102 potentially in copyright works by print status**



Although most of the out of print works tend to be older and are therefore more likely to be orphaned, as can be seen by comparison of Figs. 9, 11 and 12 with Fig. 13 it does not seem to be the print status of the work itself that necessarily causes the work to be orphaned, but above all its age and how it was published.

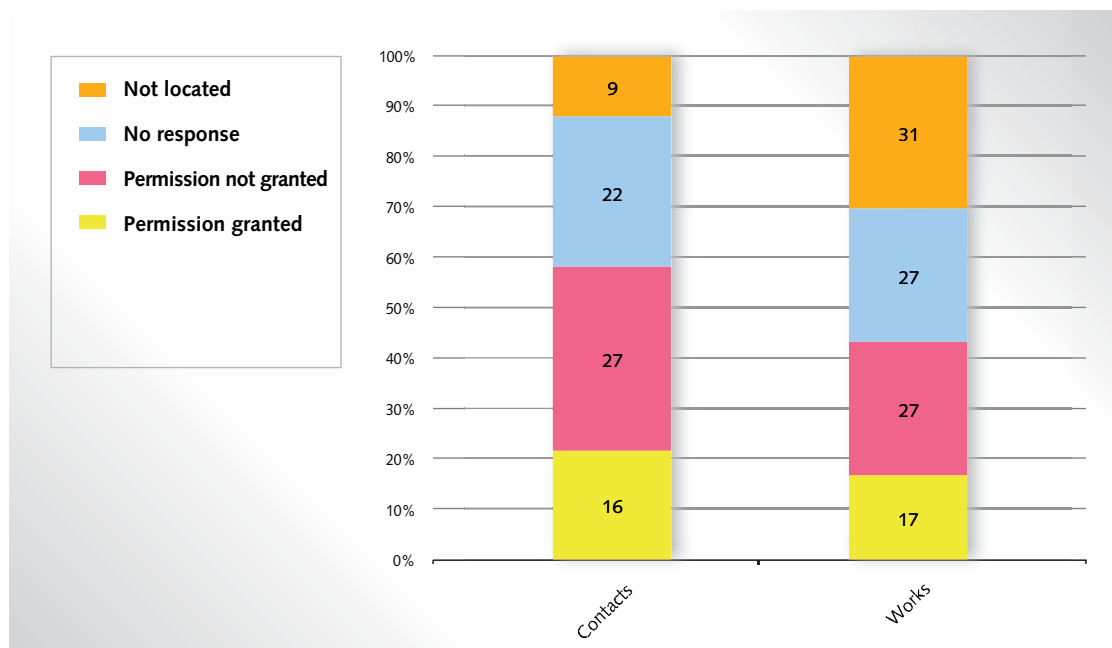
International Standard Book Numbers (ISBN)<sup>70</sup> were introduced in the 1970s and can be of assistance in tracing publishers, but they are still not used in many smaller publications, especially grey literature. In the sample only two 1970s works have ISBNs, three in the 1980s, nine in the 1990s and 10 in the 2000s (all works in the sample from the 1990s and 2000s were published commercially or by institutions).

## 10. Responses from rightsholders

The timescale of the project meant that a cap had to be put on the time allowed for rightsholder responses to be included in the study and a total of 11–16 weeks was allowed between October 2010 and mid-February 2011. In that period the Library succeeded in clearing the rights and receiving permission to digitize just 17 of the in copyright books in the sample and 17% of the 102 works for which the Library sought clearances. The study had to close on 11 February 2011 so where, for whatever reason, responses had not been received in time or negotiations had not been concluded by that date, it was deemed that permission had not been granted to digitise. Similar conclusions are likely during a full-scale digitisation project as the permission seeking stage cannot be open-ended and there has to be an assessment of how many works can go forward to be digitised.

Fig. 16 shows the results from contacting the 74 rightsholders about 102 works<sup>71</sup> for which the Library was seeking clearance to digitise. Six rightsholders could not be located at all and a further three also came to dead ends despite referrals from intermediaries. 22 did not respond at all to any of the three contacts prior to the close of the study, making a total of 31 rightsholders not located.

Fig. 16: Analysis of rights clearance results of 102 works from 74 contacts



<sup>70</sup> <http://www.isbn-international.org/> and <http://www.isbn.nielsenbook.co.uk/>

<sup>71</sup> Note that some rightsholders were contacted regarding more than one title and some titles have more than one rightsholder, therefore the numbers of rightsholder contacts and works do not tally.

The Library was additionally unable to conclude discussions with 18 rightsholders primarily due to the time taken to clear any third-party rights embedded in the work, being referred elsewhere by the first contact, or because the rightsholders were unaware of their rights or had no knowledge of the work. In such cases, it was assumed that for the purposes of this study, permission had not been granted as without express permission the Library could not have proceeded to digitise those works. These 18, taken with the nine outright refusals of permission received, make up the total of 27 instances of permission not granted. When added to the 22 no-responses (where rightsholders were contacted but no response was sent) this means 49 of the 74 contacts actively or passively did not grant permission, equating to 54, (or 52%) of the 102 titles for which permission was sought.

Only one rightsholder asked for payment when granting permission to digitise. Otherwise, the only requirement was that their work be fully cited and acknowledged.

With regard to the no-responses, it is worth noting that (automated acknowledgement emails apart which were discounted), as can be seen from Fig. 17 below, 50% of the rightsholders in the works that were published in the 2000s had not made substantive responses either granting or withholding permission within the 11–16 week contact period between October 2010 and mid-February 2011. All these works were commercially published in 2008–09 so the Library did not expect to get clearance to digitise, although it did receive one permission.

The lack of responses from so many publishers in this timescale could seriously hold up a real digitisation project, increasing already substantial transaction costs and staff time.

Fig. 18 shows an analysis of the rights clearance results by publisher type. Proportionately, in relation to the number of books for which they were responsible, it was the associations and voluntary groups that proved the most likely to grant permission (for six works out of 18 enquired about or 33%) and institutions (such as museums, schools and professional bodies) the least likely – out of six works where enquiries were made they withheld permission for four works and did not respond regarding another. Taken together permission was not granted for five of the six works (83%).

The figures in the graph in Fig. 18 above also reveal that commercial publishers gave permission to digitise six works out of the 39 enquired about, but refused permission for 13 works (33%) and in the timeframe available did not respond with regard to a further 14 works (36%). Taken together permission was not granted for 27 (69%) of the 39 works.



Fig. 17: Analysis of rights clearance results of 102 works by publication decade

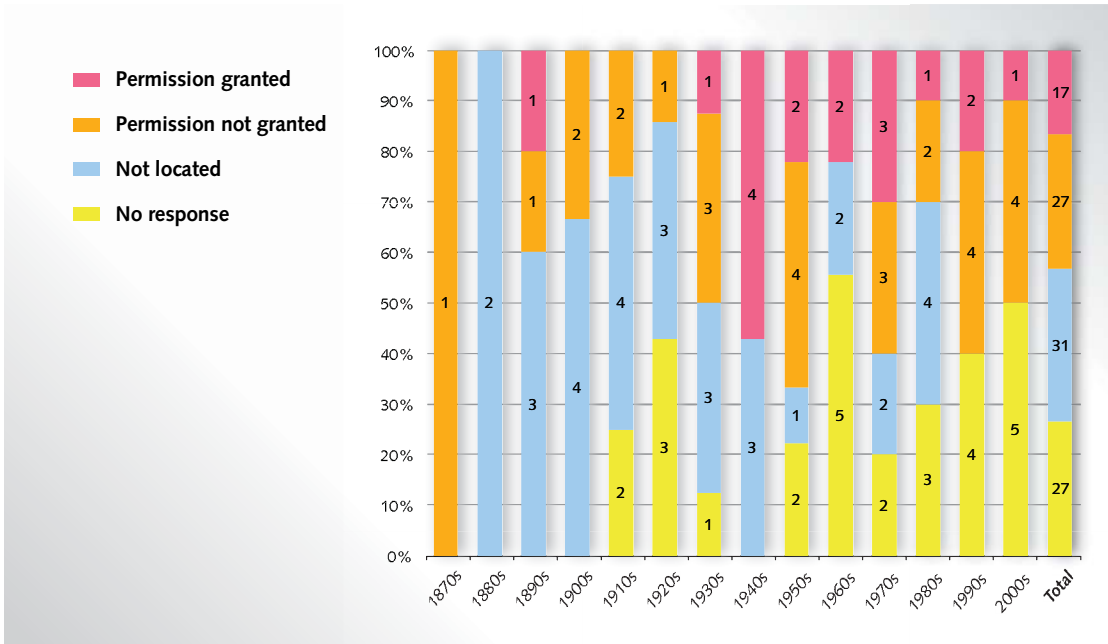
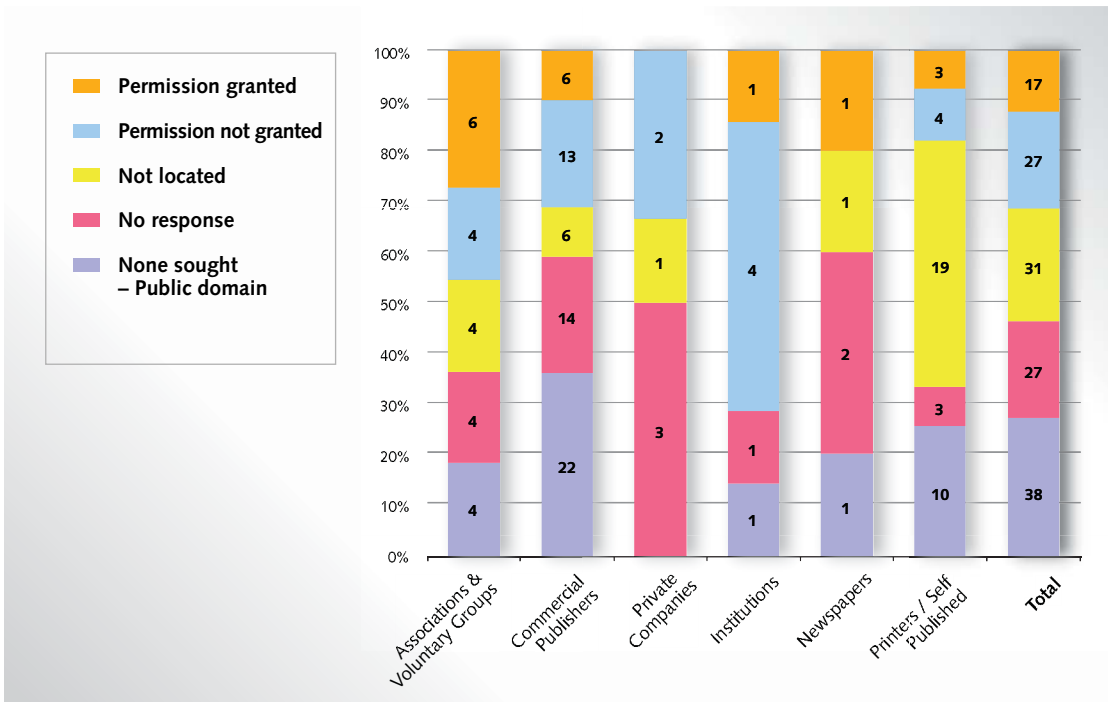


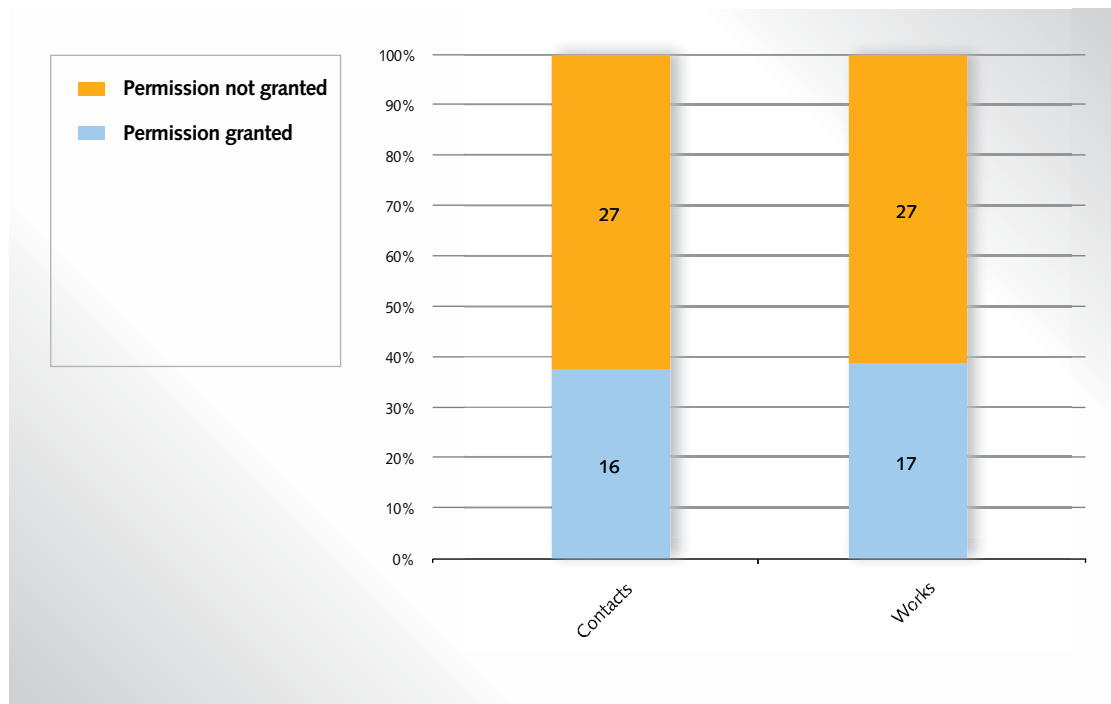
Fig. 18: Analysis of rights clearance results by publisher type (no. of works)



The overall clearance results shown in Fig. 19 were disappointing as the time and effort expended resulted in permission being granted to digitise just 17 of the works initially identified as in copyright at the end of the search phase (17% of the 102 works for which permission had been sought). These included a book confirmed by the rightsholder to be in the public domain plus another book that had entered the public domain during the final contact phase of the project.

If digitisation were to go ahead with the rights clearances received at the point the study closed, the Library would only be able to digitise just 55 works out of 140 (39%) – the 40 public domain works plus the 15 in-copyright works that had been cleared. This leaves 85 works which would have to be discounted from the project, 43 of which were orphan works.

**Fig. 19: Analysis of completed negotiations**



## 11 Time spent

Within the parameters of the study outlined in Section 4.1, the total time spent on all phases of the study averaged 4 hours per title, although some titles were very quick to deal with and others very slow. This was in excess of what would have been possible in a live mass digitisation situation – an extensive diligent search of this nature in such a situation would have been unfeasible and would have provided a major disincentive to digitise anything other than public domain works.

There were some time constraints in that the project had to be completed within the span of 11 months including waiting time for responses to come in from rightsholders and intermediaries, and because the personnel involved were not engaged full-time on the project. However, within these parameters, an effort was made to conduct as diligent a search as possible to identify and locate the rightsholders, following leads as far as possible using the resources available in the Library. The time spent to conduct the diligent manual search, broken down into the respective phases of the project, has been estimated in Fig. 20:

The average figure of 4 hours was calculated by adding the total time spent on rights clearance and contacting rightsholders and dividing by the total number of books in the sample.

Based on these figures, the shortest amount of time it took to establish the digitisation status of a book was on average 16 minutes 30 seconds which breaks down to: checking the MARC record for author death dates, checking when published for typographical rights and establishing third party rights in the work. This relates to works that had a death date in the MARC record, were found to be in the public domain and therefore had no rights to clear.

Where every step had to be applied clearance took an average of 5 hours and 56 minutes although some titles took much longer and at the end of the search of course some rights remained unidentified or uncleared.

On the other hand the ARROW search took just 4½ hours to establish the copyright status of 101 titles and use of this service would reduce the time taken by a researcher to establish the copyright status in the works to a matter of minutes per title. Not only was this shown to be quicker in the ARROW validation exercise undertaken alongside this study, the clearance results of those records that were compatible with the system largely concurred with the manual search results. However there are still some areas where the ARROW system can not currently locate the more difficult to find rightsholders and there remains a need for the system to reflect more accurately member state by member state differences in the application of copyright law.

**Fig. 20: Analysis of total time spent on rights clearance activities**

<b>Phases 1 and 2 – Establishing rights status and general admin work</b>			
	<b>Number of records / contacts</b>	<b>Average Time spent per record / contact</b>	<b>Total Time spent on activity</b>
MARC record checking:	140	5 minutes	11 hours 40 minutes
Author death date searching	79	1 hour 58 minutes	156 hours 20 minutes
General Admin	140	58 minutes	136 hours
Establishing 3rd party rights in MARC Record	140	2 minutes	4 hours 30 minutes
Searching DACS database	9	5 minutes	45 minutes
Establishing print status	140	6 minutes 30 seconds	16 hours
<b>Total amount of time spent on phases 1 and 2</b>			<b>325 hours 15 minutes</b>

<b>Phase 3 – Locating and contacting Rightsholders</b>			
	<b>Number of records / contacts</b>	<b>Average Time spent per record / contact</b>	<b>Total Time spent on activity</b>
Establishing contact details	102	1 hour 15 minutes	128 hours
Contacting Rightsholders	74	1 hour 28 minutes	109 hours
<b>Total amount of time spent on phase 3</b>			<b>237 hours</b>

<b>Total of time spent on all phases of study</b>			
	<b>Number of books</b>	<b>Average Time spent per book</b>	<b>Total Time spent on activity</b>
<b>Total</b>	140	4 hours	<b>562 hours 15 minutes</b>

# Conclusions

## 12 Conclusions

This study set out to examine the process and challenges of rights clearance for mass digitisation. In the course of the project it became clear that a diligent search is a multi-phase process for which the amount of time required cannot be underestimated. More than 450 hours (80% of the project time) were spent on research and identification tasks before a single rightsholder was contacted. The study therefore confirmed the impracticality of large scale clearance rights on a title by title basis. Scaling up the findings, at 4 hours per book it would take one researcher over 1,000 years to clear rights in half a million books – the sort of numbers that might be considered for real mass digitisation in the context of the rich collections of European cultural institutions. The study does not even begin to consider the more complex nature of rights clearance in unpublished or audio-visual materials.

The process itself sometimes requires negotiations with multiple parties. Some of the books have multiple rightsholders – some books were a mixture of public domain and in-copyright material – while others required onward referrals or persistence in making contact. Researchers need a varied portfolio of skills including an understanding of copyright and good general knowledge as well as the ability to interrogate and interpret a variety of sources.

The study established that, although age is a factor and there were considerably more orphans in the first half of the 20th century than the second, orphan works are not a phenomenon of a particular period but can occur at any point in time – the 2000s was the only decade for which all rightsholders could be traced for all the works. Furthermore, the study found that 43% of the in-copyright works or 31% of the sample was comprised of orphan works, a similar finding to the Troll Covey report, suggesting that similar collections in other libraries can also expect to find this level of orphans. It also became apparent that the vast majority of the works in the full sample – 79% – were out of print and therefore not creating any economic value for the authors or the publisher. Considering this in the context of library collections, it could be concluded that much of the material has been put on shelves and is essentially no longer available to anyone other than those who do research in the large libraries of Europe. Making material available digitally and accessible to new audiences and to new research techniques such as data and text mining can open up a new lease of life for such material as a cultural, economic or research resource.<sup>72</sup>

Given its relatively early stages of development, the results from the validation against the ARROW system were very encouraging, indicating the potential for huge efficiency gains in rights clearance projects, particularly around establishing copyright and in print status through an automated search point. Once the planned enhancements to the system, such as involvement of libraries and RROs in other countries, have been put in place its value as a resource will increase even further.

In order to make a success of programmes such as Europeana as well as facilitate the online availability of much of the past 140 years of European culture that sits in our great libraries and other cultural institutions, it is vital that action be taken to facilitate efficient rights clearance processes covering traceable rightsholders, as well as find a solution for the many orphan works that exist. A system such as ARROW must play an important role in this, as a cost effective mechanism for a single diligent search. It should support legal mechanisms that are established, and could also act as a central European register of orphan works. This could ease the search process by negating duplication of work and providing a “bulletin board” for rightsholders to claim their work.

In the meantime, libraries are disincentivised to enter into mass digitisation projects for in-copyright material. Identifying rightsholders and clearing rights are both complex and time consuming activities and in an era of constrained budgets and minimal public investment, libraries face the reality that because of expensive rights clearance – which in many cases will simply lead to a dead end – more ambitious digitisation projects become a financial impossibility. This means that libraries will only be able to offer users access to public domain works, thus limiting the cultural and research value of their online resources.

If public institutions are able to make the most comprehensive range of European materials available to researchers, citizens, innovators and businesses, then more efficient ways of clearing rights and giving legal certainty to institutions are required. Without digital access to our cultural heritage, we withhold from citizens the opportunity to increase their understanding of our history, our traditions and the world we live in, especially if rights clearance or inappropriate legislation causes the exclusion of harder to clear items or certain types of copyright works from that history. At a time of economic difficulty across the globe, the argument to open up access to more digital content is compelling and presents a unique opportunity to stimulate innovation and creativity as well as support new business or service developments for the benefit of all European citizens.

---

<sup>72</sup> The recent British Library Historical Collections iPad app created by Bibliolabs has given access to thousands of 19th century digitised works. Users have documented the gems they have unearthed at <http://britishlibrary19c.tumblr.com/> and their input will help the Library better understand this collection of forgotten works.

# Appendix 1

Author	Title	Year	Publisher
<b>1870s</b>			
	A Welcome guest from robin's nest.	1879	George Routledge and Sons, London
Religious Tract Society	Heartfelt prayers.	1878	Religious Tract Society, London
	A Book of joy for mamma's boy.	1879	George Routledge and Sons, London
Jones, G. H. Fowler	Sketches in York.	1878	York
Foster, Clarence	Venice: a dream of the Adriatic.	1877	Thomas Cook & Son, Ludgate Circus; London
Prichard, J.	An Account of the ladies of Llangollen.	1876	H. Jones, at the Atmospheric Gas Printing Works, "Advertiser" Office, Llangollen,
Clarke, Mary Cowden,	The Complete concordance to Shakespeare:	1874	W. Kent & Co.
Burke, Bernard,	The Sovereigns of England from the Norman Conquest, in rhyme.	1872	Hodges, Foster, and Co.
Stanley, Arthur Penrhyn,	As dying, and behold we live.	1872	Spottiswoode & Co, Printers.
March, James,	March's magical moving figures	1871	March, publisher,
<b>1880s</b>			
Zola, Émile,	Man metamorphosed	1886	Lewis & Co,
Hartwig, G.	The sea and its living wonders.	1882	Longmans, Green, and Co.
Hatton, Frank,	North Borneo.	1885	S. Low, Marston, Searle, and Rivington, London
	Sleeping beauty.	1881	Dean & Son, London
Fenn, Alice Manville & Louisa Manville Fenn	A summer's day.	1888	Griffith, Farran, Okeden & Welsh
Ed. George Charles haite			
Kelly, William,	Fifty years' Masonic reminiscences briefly told.	1888	Leicester
Blake, William,	There is no natural religion	1886	Edmonton
Bennett, Mary E.	The Canadian girl; or, the pirate's daughter.	1885	Milner and Company, Limited, London
McClay, Jas. Lynn	Emigration,	1880	Hodges, Foster, and Co.
Homer	The Iliad and Odyssey of Homer,	1880	Frederick Warne & Co, London
<b>1890s</b>			
Tully, J	Useful recipes for coachmen, grooms, carters, farmers.	1890	Hastings,
	My darling's story book	1898	Griffith Farran Browne & Co.
Liberal Central Association	Liberal songs.	1890	Published by the Liberal Publication Department
Swan, Annie S	Homespun: a study of a simple folk	1893	Hutchinson & Co.
Hopwood, Aubrey,	Mr. George Edward's' new production, A Runaway girl	1898	Waterlow & Sons Limited, Printers
Weir, Harrison,	Uncle Charlie's favourite fables	1898	Griffith Farran Browne & Co. Limited
Browning, Elizabeth Barrett	The poetical works of Elizabeth Barrett Browning	1897	Smith, Elder, & Co., London;
Blatchford, Robert	Practical politics: an object lesson	1892	Clarion Newspaper Co.
Meredith, George,	Diana of the Crossways	1891	Chapman and Hall, London
Froude, James Anthony	English Seamen in the sixteenth century.	1895	Longmans & Co.



Author	Title	Year	Publisher
<b>1900s</b>			
Fitchett, W. H.	The tale of the great mutiny	1904	Smith, Elder, & Co.
Roberts, Harry,	The book of old-fashioned flowers and other plants which thrive in the open-air of England	1904	John Lane
Gasquoine, T	Capital punishment and the Divine Law.	1907	Published by the "Mercury" Press
Shackleton, T. S.	Lines on Wolds Hill and the surrounding landscape,	1900	Borough Printing Works, Keighley
Crabb-Boucher	The violet-leaf treatment: its value and efficacy.	1903	John Wright & Co. Printers
Manchester School of Embroidery.	The Third 'book of Hows' or how to knit and crochet wools.	1900	Manchester School of Embroidery
Marshall, Percival	Telephones and microphones.	1900	Published "The Model Engineer and Amateur Electrician" by Dawbarn and Ward, Limited
Gregory, Edward W	Canterbury weavers past and present:	1905	London The Press Printers Ltd.
McGonagall, Sir William Topaz	The death of the Queen. A new poem	1901	Edinburgh
Lal, Shanker.	Study of legal literature	1902	Agra, Moon Press India

## 1910s

Brazil Foreign Ministry Eng. version / introduction and notes by Andrew. Boyle.	The Brazilian Green Book :	1918	Allen and Unwin
Goodenough, Anna,	A corner in the Garden of the Hesperides	1910	S. Sidders & Co. Printers
Margaret Llewelyn Davies, A. Honora Enfield and Lilian Harris	Co-operation and labour unrest.	1919	Cooperative Union, Manchester
Holmes, R. E.	A Tyneside parish.	1911	E. Sword, Printer,
Merrick, Leonard,	While Paris laughed.	1918	Hodder & Stoughton
Elliott, Michael James. Rev	Homicidal anarchy	1911	Peace Society,
Brandon, Jocelyn.	The time problem of a great city	1911	Cockayne & Co., Ltd
Gerber, Maria A.	Passed experiences, present conditions, hope for the future	1917	Ramsey-Burns Printing Co.
Parker, B.	Frolic farm	1915	W & R Chambers
Shaw, Pran. Krishna	The menace of plague and how to meet it	1917	Barnagore Home Press

## 1920s

Davies, R. J.	Japanese rule in Siberia	1922	Caledonian, London
Shastid, Thomas Hall	How to stop internation war.	1928	Duluth Min Pax Club
Lewis, Sinclair,	Arrowsmith	1925	Harcourt, Brace and Company
Snowden, Philip	The individual under socialism	1924	Independent Labour Party, London
Douglas, Alfred Bruce,	The devil's carnival	1922	Robert Dawson & Son
Groves, Ernest W. Hey	Text-book for nurses	1921	Henry Frowde and Hodder & Stoughton
Sankunni Menon, P	Some revelations	1921	Trichur the author
Hornby, M. L. + Allen, F.E.	Christmas carols and poems	1920	Liverpool printed for private circulation
Griffin, H. J.,	Historical articles on old Tottenham and Edmonton	1926	Crusha & Son, Limited
Merrick, Leonard,	The actor-manager.	1920	Hodder & Stoughton,

Author	Title	Year	Publisher
<b>1930s</b>			
	Widdicombe Fair:	1934	H. Whiteway & Co.,
Hobson, Bulmer.	A national forest policy /	1931	[B. Hobson], DUBLIN "Privately printed.
	Notes on – the churches of St. Helen, St. Nicholas, St. Michael, Abingdon	1932	W.H. Hooke, Abingdon
Musgrove, Edward J.	History of St Peter's Church, Tiverton /	1938	[E.J. Musgrove], Tiverton
Rust, William.	Socialism and peace	1936	Communist Party, Scottish Committee, Glasgow
Kipling, Rudyard,	The jungle book /	1935	Macmillan, London
Kincaid, Charles Augustus	A book of golden tales	1932	Oxford University Press, London
Singh, M. G	Epic tales of the East /	1932	Oxford University Press, Bombay
Hollis, A. H.	Sir Francis Drake and the Drakes of Ashe.	1931	H.G. Cockerton,
Anderson, Tom	The class state /	1930	Liberty Press,

## 1940s

	Birthplace of Elgar.	1940	Novello and Company Ltd.
Hobhouse, Stephen	Can the dictators hold Europe?	1940	Peace Pledge Union
Bramley, Ted.	We are many!	1940	London District Communist Party?
White, William Lindsay	Lost boundaries.	1948	Harcourt, Brace and Company,
Coles, Howard W.	The cradle of freedom:	1943	Oxford Press
Mowat, Jack & Power, Albert	Our struggle for socialism!!!:	1949	N.G Atkinson
Canada.Royal Canadian Airforce	The R.C.A.F. Overseas	1944	Oxford Univer. Press
Eldrid, Trevor.	The smugglers of Lee /	1948	T. Eldrid,
Pugh, W. T. Gordon	Practical nursing:	1945	William Blackwood,
Dewar, Hugo.	The mask of democracy /	1946	Independent Labour Party, London

## 1950s

	The Royal Masonic Institution for Boys	1958	Royal Masonic Institution for Boys, London
Performing Right Society.	The Performing Right Society Ltd.	1951	Performing Right Society, London
	Orchestral works	1955	Oxford University Press, London
Coke, Barbara.	The power of an idea: the Red Cross, June 24th 1859, June 24th 1959	1959	British Red Cross Society,
Isham, Gyles,	The architectural history of Lamport /	1953	Northamptonshire Architectural and Archaeological Society
Liveing, Edward G. D.	A history of Richard Lloyd Limited:	1953	Richard Lloyd Ltd
Owen, F Mrs., & Bailey, M.E. Miss	Electricity and home appliances	1953	The Electrical Association for Women
Morris, Stuart & Reynolds, Reginald	That which essentially belongs to man	1952	Peace News
Wynne Mason, Walter	Prisoners of War	1954	Oxford University Press
Campbell, John Wood,	Islands of space	1956	Fantasy Press

Author	Title	Year	Publisher
<b>1960s</b>			
McQuade, John.	Man of The people	1968	Belfast
Dixon, Arnold.	Ship master	1967	Market Arcade
Curzon, Robert	Parham conversion list.	1961	London
	Bash Street Kids go to the sit-in.	1968	England
Jacobs, Arthur	Choral music	1969	Penguin Books,
Morrison, Sybil,	The life and work of Stuart Morris	1969	Stuart Morris Memorial Fund, Peace Pledge Union,
Associated-Rediffusion	People and holidays.	1962	Associated-Rediffusion, London
Twemlow, Charles Fletcher.	The Betley story, 1086–1961	1961	Betley and District Village Hall Committee, Betley
Communist Party of Great Britain.	Meet the Communists.	1961	Communist Party of Great Britain, London
	Full ahead!	1964	Austin & Pickersgill Ltd. Sunderland

## 1970s

Saville, Malcolm.	The story of Winchelsea Church.	1978	St. Leonards King Bros & Potts
Nations, Opal L.	Creation of a nation, death of a mind	1972	Transgravity Press
Thomas, J. D.	Fifty years of Whitleyism in the Inland Revenue, 1920–1970	1970	Somerset House, London
Little, Bryan D. G.	Ammerdown the House and the Centre.	1977	Radstock, Bath for the Ammerdown Society & Aldrich, Andrew
Gibson, John C. L.	Textbook of Syrian Semitic inscriptions	1971	Clarendon Press, Oxford
Bearman, Robert.	The Gregorys of Stivichall in the sixteenth century	1972	Coventry Branch of the Historical Association
Jackson, Herbert.	Looking at West Cumbria:	1972	H. Jackson, Maryport
Organisation of Revolutionary Anarchists.	Don't vote: organise	1971	Organisation of Revolutionary Anarchists, York
United Kingdom Atomic Energy Authority.	Dounreay: United Kingdom Atomic Energy Authority	1970	Richmond Press, Wilmslow
Howells, John G.	Nosology of psychiatry	1970	Ipswich, England Calver Press

## 1980s

Watson, Maureen.	Liverpool: lessons for the left	1985	Socialist Worker, London
Gorman, John	Lively as a linnet.	1986	Plaistow Poets Press, London
Bacon, Jean	Dunwich Suffolk.	1988	
Segment, Harrison, Stephen,	Yorkshire farming memories.	1981	Castle Museum, York
Jackson, Herbert. & Jackson, Mary	Looking at Workington.	1980	H. and J. Jackson, Maryport
Blackhall, Sheena	The spik o' the lan'.	1986	Rainbow Books
Brandon, David,	Just leave your luggage at the door.	1986	Channel Four Television,
Golden Cow Productions.	Cow tales and under atrocities.	1986	Golden Cow Productions Ltd
Frampton, George	Whittlesey Straw Bear: with notes on the celebration of Plough Monday in the area.	1989	Cambridgeshire Libraries
Press Club (London, England)	100 years of Fleet Street:	1982	Press Club

Author	Title	Year	Publisher
<b>1990s</b>			
Dube, Hope.	Love me forever	1999	Paragon Leeds
Das, S. K.,	Civil service reform & structural adjustment	1998	Oxford University Press, Delhi and Oxford
Walker, Liz.	Oxford/Cambridge: applicant's guide to Oxford and Cambridge	1998	Lifetime Careers, Harrow
Falla, Tim.	Watch this! Teacher's book	1999	Oxford University Press, Oxford
Lloyd, Chris.	Discover Oxford:	1999	Bravura, Abingdon
Gloucestershire Records Office.	Abolition & emancipation	1999	Adam Matthew, Marlborough Wilts.
Voet, Donald & Voet, Judith G.	Biochemistry.	1996	Wiley, New York & Chichester
Jones, Nick,	Victorian shorts: ten pre-20th century short stories by major writers & what to do with them	1997	Devon Curriculum Advice,
Gould, Mark.	Parental leave is coming:	1999	New Ways to Work,
Warner, Alan.	Financial statements	1999	MTP Aylesbury

## 2000s

Moore, Wendy,	Wedlock: how Georgian Britain's worst husband met his match	2009	Phoenix, London
Erik Borgman, Maria Clara Bingemer, and Andrés Torres Queiruga	Monotheism – divinity and unity reconsidered.	2009	SCM Press, London
Hart, Ellen.	Faint praise	2009	Bella; Ferndale,
Ken Follett	Hornet flight.	2009	Pan,
Shah, Janat	Supply chain management: text and cases	2009	Dorling Kindersley India
Matthew L. Nelson, Michael J. Shaw, Troy J. Strader	Value creation in e-business management:	2009	Springer, New York
Pradip N. Khandwalla	Management of corporate greatness: blending goodness with greed.	2008	Dorling Kindersley India
A., Joel	Mutual arousal: self-help encouragement words: for healthy living, happiness relationships.	2009	Strategic Insight, Horsham
Anjali Gera Roy and Nandi Bhatia	Partitioned lives: narratives of home, displacement, and resettlement.	2008	Pearson Longman,
Ecclestone, Kathryn	Transforming formative assessment in lifelong learning.	2009	Open University Press

# Appendix 2

**Duration of copyright in literary, dramatic, musical and artistic works in the UK** (excluding Crown Copyright, Parliamentary Copyright and international organisations copyright)<sup>73</sup>

## 1 General Notes

- i) Literary works may be written or performed and the term 'writing' is broadly defined in CDPA s.178 to 'include any form of notation or code, whether by hand or otherwise and regardless of the method by which, or medium in or on which, it is recorded'.<sup>74</sup> They include works such as books, journals, magazines, newspapers, articles, forewords, notes, tracts, poems, knitting patterns, letters, reports, accounts, databases and computer programs. CDPA s.3(2) requires that the work be recorded in some form, such as writing, for copyright to subsist.<sup>75</sup>
- ii) Dramatic works are intended for performance – for example, plays, film screenplays, librettos and dance and mime choreography.
- iii) Musical works are designed to be performed. They do not include accompanying words (libretto).
- iv) Artistic works include engravings, lithographs, collages, drawings, paintings, photographs, sculptures and artistic installations, architecture, plans, charts and maps.
- v) To qualify for copyright protection in the UK, prior to the European Court of Justice's (ECJ) 2009 decision in the Infopaq case, a work had only to satisfy a low standard of originality, i.e. be the product of labour, skill and effort or 'sweat of the brow' and it did not have to be original in its usual meaning of 'uniqueness' or 'creativity', or have any merit or quality. The ECJ's Infopaq ruling emphasised that the EU standard for works to qualify for copyright protection must be that the work is the author's own intellectual creation.<sup>76</sup>
- vi) An 'author' is always a natural person. The term 'author' in relation to copyright generally includes all creators who are natural persons, for example artists, illustrators, photographers, composers and so forth.
- vii) With some exceptions (e.g. for work done in the course of employment where the employer is the first owner of copyright) the author is the first owner of copyright in a work.<sup>77</sup> The ownership will pass to his or her legal heirs unless he or she bequeaths or assigns it elsewhere.<sup>78</sup> Normally, if not assigned or bequeathed elsewhere, the author's copyright would pass to whoever is the residuary beneficiary of an author's will. If there is no will, complex general intestacy rules would apply.
- viii) Reversionary provisions apply concerning assignment of copyright or grant of licences: where the author had assigned or granted a licence of his or her rights in a literary, dramatic, musical or artistic work 'otherwise than by will', CDPA provides that if there was no assignment of the author's reversionary interest, the rights would revert to the author's estate 25 years after death. This provision applies to assignments made between the passing of the 1911 Copyright Act and the commencement of the 1956 Copyright Act.<sup>79</sup>
- ix) The UK Copyright Act of 1956 introduced the concept of measuring the length of the postmortem period of UK copyright term to run from the end of the calendar year in which the author died. This remains the case under the CDPA.

---

<sup>73</sup> Information mostly sourced from Padfield, Tim: Copyright for Archivists and Records Managers. 4th ed. London: Facet Publishing, 2010.

<sup>74</sup> CDPA s.178 "writing" includes any form of notation or code, whether by hand or otherwise and regardless of the method by which, or medium in or on which, it is recorded, and "written" shall be construed accordingly.'

<sup>75</sup> CDPA s.3(2) 'Copyright does not subsist in a literary, dramatic or musical work unless and until it is recorded, in writing or otherwise; and references in this Part to the time at which such a work is made are to the time at which it is so recorded.'

<sup>76</sup> Case C-5/08. Infopaq International A/S v Danske Dagblades Forening, <http://curia.europa.eu/jurisp/cgi-bin/gettext.pl?lang=en&num=79909283C19080005&doc=T&ouvert=T&seance=ARRET>

<sup>77</sup> CDPA ss.9(4) and (5)

<sup>78</sup> CDPA s.90(1)

<sup>79</sup> CDPA Sch.1 para 27

## 2 UK Copyright term<sup>80</sup>

NB. See note at 1(vi) above concerning the use of the term 'author' with regard to copyright.

- i) Where the author's **identity is known**, copyright endures for life + 70 years from the end of the calendar year in which he or she dies. However, there are different provisions with regard to known authors of some photographs created between 1 June 1957 and 31 July 1989 and some engravings created before 1 August 1989. (For more detail see para 2(vi) below with regard to photographs and the Copyright Duration Chart in Fig. 21 below).
- ii) Where the author's **identity is unknown**,<sup>81</sup> including where the creator is a corporate body and not a natural person, the author is regarded as anonymous or, where relevant, pseudonymous (i.e. when the true identity of the author is not known). In cases of true anonymity of the author, it is assumed that copyright belongs to the publisher. Copyright in such works lasts for 70 years from the end of the calendar year in which the work was made, or was first legally made available to the public. This applies to all works of unknown authorship apart from some photographs. In the case of pseudonymous works, where the author's identity is known or becomes known, the work is then treated as being of known authorship. Copyright in anonymous or pseudonymous works is not infringed where 'it is not possible by reasonable inquiry to ascertain the identity of the author, and (b) it is reasonable to assume (i) that copyright has expired, or (ii) that the author died 70 years or more before the beginning of the calendar year in which the act is done or the arrangements are made.'<sup>82</sup>
- iii) In the case of works of **joint authorship**, copyright expires 70 years after the death of the last known author.
- iv) The Berne Convention and the Universal Copyright Convention provide for '**national treatment**' i.e. treating and protecting foreign works in the same way as national works, but the provision is not mandatory.<sup>83</sup> The UK did not introduce national treatment until required to do so on implementation of the Term Directive.<sup>84</sup> However, national treatment does not apply where the copyright term in the home nation (where the work is to be copied) is greater than that in the foreign nation where the work was published, in which case the lower foreign term should be applied. For example, a literary work published in Canada is protected in the UK for life + 50 years, not the EEA copyright term of life + 70 years; however, a literary work published in Mexico, where the copyright term is life + 100 years, is protected in the UK for life + 70 years. The lower term always applies.
- v) In general, upon implementation of the Term Directive in the UK, the copyright term for **unpublished works** by authors who died after 1969 is now life + 70 years whereas in the case of authors who died before 1969 it remains in copyright to the end of 2039.<sup>85</sup> However there are some variations (see also Copyright Duration chart Fig. 21):
  - Copyright was extended or revived in works where the author was an EEA national who died after 31 December 1924 (where UK copyright in works by non-EEA nationals had expired by 1 July 1995, there was no extension or revival).
  - Copyright was extended in anonymous works created or first made available to the public in the UK after 31 December 1944, and revived in anonymous works that had once qualified for UK copyright protection *and* been created or made available to the public in another EEA state before 1 January 1995, *and* still qualified for copyright protection in that state on 1 July 1995.

---

<sup>80</sup> CDPA s.12

<sup>81</sup> CDPA s.9(5) provides that 'For the purposes of this Part the identity of an author shall be regarded as unknown if it is not possible for a person to ascertain his identity by reasonable inquiry; but if his identity is once known it shall not subsequently be regarded as unknown.'

CDPA s.102(2) 'Where a name purporting to be that of the author appeared on copies of the work as published or on the work when it was made, the person whose name appeared shall be presumed, until the contrary is proved – (a) to be the author of the work;

(b) to have made it in circumstances not falling within section 11(2), 163, 165 or 168 (works produced in course of employment, Crown copyright, Parliamentary copyright or copyright of certain international organisations).'

CDPA s.102(4) 'Where no name purporting to be that of the author appeared as mentioned in subsection (2) but – (a) the work qualifies for copyright protection by virtue of section 155 (qualification by reference to country of first publication), and (b) a name purporting to be that of the publisher appeared on copies of the work as first published, the person whose name appeared shall be presumed, until the contrary is proved, to have been the owner of the copyright at the time of publication.'

<sup>82</sup> CDPA s.57(1)

<sup>83</sup> Berne Convention Art. 7(8)

<sup>84</sup> Directive 2006/116/EC Art. 7 (codified version of the Term Directive 93/98/EEC)

<sup>85</sup> CDPA Sch.1 S.12(4)



- vi) **Photographs** are classed as artistic works and are subject to the same copyright term as other artistic works, but CDPA makes labyrinthine transitional provisions for **older photographs and some engravings**, summarised below (as amended by implementation of the Term Directive 1993) (for more detail see the Copyright Duration chart Fig. 21)
- Author is known *and* photograph was taken before 1 June 1957: copyright term is life + 70 years.
  - Author is known *and* died before 1 January 1969 *and* photograph was not published before 1st August 1989 *and* was created on or after 1 June 1957 but before 1 August 1989: copyright endures until 31 December 2039.
  - Author is known *and* photograph or engraving was published before 1 August 1989 *and* the author died before 1 January 1969, where the photograph was taken on or after 1 June 1957 but before 1 August 1989, or the engraving was made before 1 August 1989: copyright endures for 50 years from the end of the calendar year in which the photograph or engraving was first published.
  - Author is known *and* photograph was created on or after 1 June 1957 *and* published before 1 August 1989 *and* author died less than 20 years before publication: copyright is for life + 70 years.
  - Author is unknown: length of copyright term depends on when the photograph was first made or made available to the public. If the year is known, copyright is for 70 years.
  - Author is unknown *and* photograph was created on or after 1 June 1957 but before 1 January 1969 *and* was not published before 1 August 1989 *and* was not made available to the public before 1st January 2040: copyright endures until 31 December 2039.
- vii) **Typographical arrangements** in published editions are protected for 25 years after publication. This includes the typographical layout of a new edition of a work (also of a public domain work) but not of a straight reprint.<sup>86</sup>

---

86 CDPA s.15



## Appendix 3

**Duration of copyright in literary, dramatic, musical and artistic works in Canada, India, Ireland and the USA.** Since no government, parliamentary or international organisation publications appear in the sample, the rules for these are omitted below. The information below is drawn from Padfield.<sup>87</sup>

NB. See note at Appendix 2 para 1(vi) above concerning the use of the term 'author' with regard to copyright.

### i) Canada<sup>88</sup>

- Literary, dramatic or musical works, or engravings by a known author
  - Published before 25 July 1997, where the author died before 25 July 1947: **publication + 50 years** (s.7(1))
  - Unpublished on 25 July 1997, where the author died before 25 July 1947: **copyright expired** - entered the public domain (s.7(4))
  - Unpublished on 25 July 1997, where the author died between 25 July 1947 and 24 July 1997: **in copyright to 2047** (s.7(3))
  - Unpublished where the author died after 24 July 1997: **life + 50 years** (s.6)
  - Published literary, dramatic or musical work by a known author created after 24 July 1997, or created before that date and published during the author's lifetime: **life + 50 years** (s.6)
- Artistic works by a known author: **life + 50 years** (s.6)
- Corporate authored photographs: **creation + 50 years** (s.10(1))
- Anonymous works: **the shorter of publication + 50 years or creation + 75 years** (s.6(1))
- Typographical arrangement in published editions: **no copyright**
- Moral rights: **same term as the copyright** (s.14(2))

### ii) Ireland<sup>89</sup>

- Literary, dramatic, musical and artistic works (or an original database): **life + 70 years**.
- Works created before the Irish Copyright Act of 2000 came into force continue to be covered by the 1995 regulations that implemented the **life + 70 year** term introduced by the European Term Directive 1993. These are pretty much identical to those set out in the Act itself.

### iii) India<sup>90</sup>

The copyright term of works created before 4 June 1957 are governed by the provisions of the UK Copyright Act 1911, implemented in Indian law by the Indian Copyright Act 1914.

- Literary, dramatic, musical or artistic works, photographs excepted, and published during author's lifetime
  - Where created on or after 4 June 1957: **life + 60 years** (1957 s.22)
  - Where created before 4 June 1957: **life + 50 years** (1914 Sch. 1)
  - If anonymous
    - When published: **publication (or publication of adaptation) + 60 years**
    - Where unpublished: **copyright is perpetual until published**
- Literary, dramatic, musical or artistic works, or engravings created on or after 4 June 1957 and published after the author's death (1957 s.24)
  - When published: **publication (or publication of adaptation) + 60 years**
  - Where unpublished: **copyright is perpetual until published**
- Literary, dramatic, musical or artistic works, or engravings created before 4 June 1957 and published after the author's death (1914 Sch.1 para 17(1)):
  - When published: **publication + 50 years**
  - Where unpublished: **copyright is perpetual until published**
- Artistic works other than photographs or engravings unpublished during author's lifetime: **copyright is perpetual until published**
- Photographs
  - Created on or after 4 June 1957 (1957 s.25)
    - When published: **publication (or publication of adaptation) + 60 years**
    - Where unpublished: **copyright is perpetual until published**
  - Created before 4 June 1957: **creation + 50 years** (1914 Sch.1 para 21)
- Typographical arrangement in published editions: **no copyright**
- Moral rights: **none**

<sup>87</sup> Padfield, Tim (2010). Op cit.

<sup>88</sup> CANADA: Copyright Act (1985 c.42). There is currently a Copyright Reform Bill (Bill C-32) before the Canadian Parliament (tabled 2 June 2010).

<sup>89</sup> IRELAND: Copyright and Related Rights Act 2000 s.33, Sch 1 s.9, 30; SI 1995/158 Reg. 5

<sup>90</sup> INDIA: Copyright Act 1914; Copyright Act 1957 s.79(5). The Copyright (Amendment) Bill 2010 is, at the time of writing, before the Indian Parliament.

iv) **USA**<sup>91</sup>

Prior to 1978, US copyright was only available to works registered with the Library of Congress Copyright Office. The 28-year copyright term had to be renewed on expiry for the work to remain in copyright for a second 28 year term, which by 1976 had been extended to 47 years meaning that a work could be in copyright for a maximum of 75 years from the date of first registration.

However, the USA did not accede to the Berne Convention until 1988 and its membership took effect from 1 March 1989. In order to comply with Berne's requirement for no formalities and a copyright term based on the author's life + 50 or more years, the US abandoned compulsory registration for works published outside the US and adopted a copyright term based on the author's life + 50 years. This was later extended to life + 70 years by the Copyright Term Extension Act.<sup>92</sup> Despite this, registration of copyright in the US Copyright Register remains compulsory for works first published in the US. Registration provides proof of copyright ownership which is a prerequisite for entitlement to statutory damages in respect of any infringement proceedings before the US courts.

- Work created on or after 1 January 1978 by a known author: **life + 70 years**<sup>93</sup>
- Anonymous work or work made for hire, created on or after 1 January 1978: **earlier of publication + 95 years or creation + 120 years**<sup>94</sup>
- Work created but not published or copyrighted before 1 January 1978 and not published before 1 January 2003<sup>95</sup>
  - If author known: **life + 70 years**
  - If anonymous or made for hire: **creation + 120 years**
- Work created but not published or copyrighted before 1 January 1978 but published between 1 January 1978 and 31 December 2002: **later of 2047 or relevant standard term applicable where author is known or anonymous or if the work is made for hire**<sup>96</sup>
- Work copyrighted on or after 1 January 1950 and still in its first registration term (28 years) on 1 January 1978: **expired if copyrighted before 1 January 1964 and not renewed as appropriate between 1978 and 1991. Otherwise original registration + 95 years**<sup>97</sup>

- Work copyrighted or first published on or after 1 January 1923 where the copyright was renewed and still in force on 27 October 1998: **original registration + 95 years**<sup>98</sup>
- Work published before 1 January 1923: **copyright expired**
- Typographical arrangement in published editions: **no copyright**
- Moral rights: **none** except with regard to paintings, drawings, prints, still photographs and sculptures created for exhibition only. The work must be a single copy or in a limited edition of up to 200 signed and numbered copies.<sup>99</sup>

---

<sup>91</sup> UNITED STATES: Copyright Act 1976 ss.104-105, 401, 411

<sup>92</sup> Copyright Term Extension Act 1998 (CTEA) known also as the "Sonny Bono Act" or the Mickey Mouse Protection Act". See Wikipedia [http://en.wikipedia.org/wiki/Copyright\\_Term\\_Extension\\_Act](http://en.wikipedia.org/wiki/Copyright_Term_Extension_Act)

<sup>93</sup> 1976 s.302(a), 1992

<sup>94</sup> 1976 s.302(b), 1992

<sup>95</sup> 1976 s.303

<sup>96</sup> 1976 s.303

<sup>97</sup> Automatic renewal applies to works first copyrighted on or after 1 January 1964. Any work first copyrighted before then must be renewed by the rightsowner. (1976 s.304(a), 1992)

<sup>98</sup> 1976 s.304(b)

<sup>99</sup> Visual Artists Rights Act 1990 (VARA) See Wikipedia [http://en.wikipedia.org/wiki/Visual\\_Artists\\_Rights\\_Act#Covered\\_works](http://en.wikipedia.org/wiki/Visual_Artists_Rights_Act#Covered_works)

# Appendix 4

## Model letters sent to rightsholders

### 1 Letter to authors' heirs, literary executors or similar

Dear

Re [XXXXXXX] [dates where known eg. '1850–1920' or 'born 1850' or 'died 1960'] **Request for Permission to use works in attached list.**

A library consortium is currently evaluating a digitisation project of a selection of 19th and 20th century books with a view to putting them on a public facing website for view and print purposes only. The project is not-for-profit and no charge will be made to members of the public to view or print out.

We have been trying to locate the current copyright holder for the above author and our researches indicate that you might be in a position to help us either because you are a relative or descendant or because of your specialist interest in the author. We are therefore writing to you in the hope that you might be able to help us locate the current copyright holder so we can seek any necessary permissions to use the works listed as attached.

We would be grateful if you could confirm whether you yourself are the copyright holder of the title(s) in the attached list. If you are not, or do not hold all the relevant rights, it would be most helpful to us if you could assist us with any information you might have to enable us to identify and locate the current rightholder(s), including advising us of the year of the author's death where applicable (if no year of death is listed above we have been unable to establish this information which is very necessary to confirm the duration of the copyright).

To make the process as easy as possible please find attached a form we would ask you to fill out and return to us together with this letter countersigned as below. **Please return to me by e-mail to [insert email] where possible.**

Yours faithfully

*(signature and date)*

(Name and Position)

(Countersignature from Addressee and date)

(Name [BLOCK LETTERS] and Position)

(Contact e-mail and phone no.)

## Letter to authors' heirs, literary executors or similar continued

[Pre-insert name of Addressee]

### Please return this Form as a Word document

1. I can confirm that we are / do represent the copyright holder in the following titles:
2. I can confirm that we **do not** represent the copyright holder in the following titles:
3. I /My organisation has the following information / contact details regarding the copyright holder (please also advise of year of author's death if possible where not listed above):
4. Fees:

If you are, or represent the rightholder, without prejudice to further negotiations, please indicate how you would intend to allow access to these titles:

- a) No fee payable.
- b) A one-off fee of:
- c) Subject to further discussion.
- d) Permission not granted. **(Please indicate why if possible)**
- e) Other:

Please be aware that in the event of digitisation of these titles a notice and takedown policy will be implemented so the books are removed from view within 24 hours in the event of any claim in regards to the books.

Please note full acknowledgement will appear next to the text. If the required acknowledgement has altered since the publication of the book please indicate below.



## 2 Letter to publishers

Dear

**Request for Permission to use works in attached list.**

A library consortium is currently evaluating a digitisation project of a selection of 19th and 20th century books with a view to putting them on a public facing website for view and print purposes only. The project is not-for-profit and no charge will be made to members of the public to view or print out.

Please kindly confirm whether you are the rightholder of the title(s) in the attached list. If you are not, or do not hold all the relevant rights, we would be grateful if you could help us with any information your organisation might have to enable us to identify and locate the current rightholder(s), including year of death where applicable.

To make the process as easy as possible please find attached a form we would ask you to fill out and return to us together with this letter countersigned as below. **Please return by e-mail where possible.**

Yours faithfully

*(signature and date)*

(Name and Position)

(Countersignature from Publisher and date)

(Name [BLOCK LETTERS] and Position)

(Contact e-mail and phone no.)

## Letter to publishers continued

*[Pre-insert name of Publisher]*

### **Please return this Form as a Word document**

1. I can confirm that we are / do represent the copyright holder in the following titles:
2. I can confirm that we are / do not represent the copyright holder in the following titles:
3. My organisation has the following information / contact details regarding the copyright holder:
4. Fees:

If you are, or represent the rightholder, without prejudice to further negotiations, please indicate how you would intend to allow access to these titles:

- a) No fee payable.
- b) A one-off fee of:
- c) Subject to further discussion.
- d) Permission not granted. **(Please indicate why if possible)**
- e) Other:

Please be aware that in the event of digitisation of these titles a notice and takedown policy will be implemented so the books are removed from view within 24 hours in the event of any claim in regards to the books.

Please note full acknowledgement will appear next to the text. If the required acknowledgement has altered since the publication of the book please indicate below:

### 3. Letter to authors' organisations

Dear

A library consortium is currently evaluating a digitisation project of a selection of 19th and 20th century books with a view to putting them on a public facing website for view and print purposes only. The project is not-for-profit and no charge will be made to members of the public to view or print out. We have been trying to locate the authors listed in the attached spreadsheet so we can seek any necessary permissions to use their works, all of which are either definitely in copyright or it is unclear whether they are yet in the public domain.

Unfortunately although in many cases we have been able to find out some information about them (included in the spreadsheet) we have not been able to identify current direct contact information for them and, in some cases a death date where applicable, and we wonder whether you can possibly help on both counts?

We would be grateful if you could run the names against any records you may have and fill in any contact information or additional information you may hold in the spreadsheet column for your organisation and return it to me by email to *[insert email]*. If there is a charge for your services or if you cannot help please let me know.

Yours faithfully

### 4. Letter to Publishers' Collecting Society

Dear

A library consortium is currently evaluating a digitisation project of a selection of 19th and 20th century books with a view to putting them on a public facing website for view and print purposes only. The project is not-for-profit and no charge will be made to members of the public to view or print out. We have been trying to locate the publishers listed in the attached spreadsheet so we can seek any necessary permissions either from them or from their authors as appropriate.

Unfortunately although in many cases we have been able to find out some information about them (included in the spreadsheet) we have not been able to identify direct contact information for them, or their subsequent owners where company takeovers have occurred, and we wonder whether you can possibly help?

We would be grateful if you could run the names against any records you may have and fill in any contact information or additional information you may hold in the spreadsheet column headed PLS and return it to me by email to *[insert email]*.

If there is a charge for your services or if you cannot help please let me know.

Yours faithfully

## Appendix 5

### Case study: Inextricably intertwined in the world's catalogues: the two Mary Bennetts

In the course of the study it was found that a cataloguing error had mixed up 'Mary E. Bennett b. 1841' with 'Mary Bennett'. The British Library Integrated Catalogue (BLIC) described the author of the work in the sample, *The Canadian girl; or, the pirate's daughter* (1885), as 'Mary E. Bennett b. 1841' but this did not tally with the fact that BLIC records also showed that 'Mary E. Bennett' was apparently also the author of *The Canadian girl, or the pirate of the lakes*, thought to be first published in 1838 before she was supposed to be born. The catalogue also revealed that there was a 'Mary Bennett' who was a 19th century writer of novels with a similar style of titles for and about girls and young women all published in the UK, but there were no dates for her in BLIC. 'Mary Bennett's' earliest work listed in BLIC was apparently *The gipsey bride: or, the miser's daughter* published in 1841.

The birth and death dates for 'Mary Bennett' (1813-1899) were listed within the article for her brother, John Saunders, in the Oxford Dictionary of National Biography (ODNB)<sup>100</sup> so her works were in the public domain. The article mentioned that she was author of *The Jew's Daughter* (1837), *The Cottage Girl* (1838), and *The Gipsy Bride* (1840), but did not mention the work in the sample, so it was still necessary to establish who was its author. Additionally the scanned 1909 edition of *Herringshaw's Library of American Biography* in the Internet Archive, revealed 'Mary E. Bennett' to be an American children's author born in 1841 but without a death date. The titles cited for her were quite different from the ones ascribed to 'Mary Bennett' in BLIC and the ODNB.<sup>101</sup>

The Internet Archive holds five scanned editions of *The Canadian girl, or the pirate of the lakes*, four of them ascribed to 'Mary E. Bennett b.1841'<sup>102</sup> and the fifth to "the Authoress of *The Jew's daughter*" and all published in London. On inspection, they were revealed to be written by 'Mary Bennett' mentioning some of her other works including those mentioned in the ODNB. Finally, the Archive's scanned facsimile of an undated edition of *The Jew's daughter; or the witch of the water-side* (also ascribed in its catalogue to 'Mary E. Bennett'), published in London by Milner & Co., was, on inspection online, found to be by 'Mrs Bennett' with a list of many of Mary Bennett's other known works on the title page including *The Canadian girl*. This edition also had an extensive publisher catalogue in its rear pages which confirmed "the Author of 'Jane Shore'" (identified elsewhere in the same publisher catalogue as 'Mary Bennett') to be the author of the book in the sample, *The Canadian girl' or the pirate's daughter*, so it was clear which Mary Bennett was its author and that the work in the sample was in the public domain.<sup>103</sup>

---

<sup>100</sup> Law, Graham (2004): 'Saunders, John (1811–1895)', Oxford Dictionary of National Biography, Oxford University Press, 2004

<sup>101</sup> "Bennett, Mary E., litterateur, author, was born in 1841 in Connecticut. She is the author of *Cyril Rivers*; *Six Boys*; *Asaph's Ten Thousand*; *Jefferson Wildrider*; and other works." *Herringshaw's Library of American Biography* V.1 pn293 p306, American Publishers Association, 1909  
<http://www.archive.org/stream/cu31924088421593#page/n293/mode/2up>

<sup>102</sup> [http://www.archive.org/details/cihm\\_18128](http://www.archive.org/details/cihm_18128)

<sup>103</sup> Information from publisher's catalogue for 'Mary Bennett' in the rear pages of *The Jew's daughter; or the witch of the water-side* / Mrs Bennett. London, Milner, n.d., <http://www.archive.org/stream/jewsdaughterorwi01benn#page/n1/mode/2up>

## Appendix 6

# ARROW<sup>1</sup>

## A service in development to assist “diligent search”

Mark Bide and Tom West on behalf of the Publishers Licensing Society (PLS) and the Authors' Licensing & Collecting Society (ALCS)

### Introduction

A recent paper from the British Library<sup>2</sup>, written in the context of the ARROW project, eloquently describes one of the challenges for any library that sets out to undertake a mass digitisation program: finding the rightsholders for those works that are still in copyright in order to secure the necessary permissions.

This is precisely the challenge that ARROW was established to overcome. Undertaking the necessary “diligent search” for a rightsholder starting without access to the appropriate data is clearly a daunting task. But there is no need for anyone to work “blind” in this way. ARROW has been designed specifically to create a network of data and expertise to support diligent search processes and has shown – even in its preliminary pilot phase – how effective this network can be.

This brief paper sets out to describe the functions and capabilities of ARROW, specifically from the perspective of the UK, and the lessons learned from the initial pilot as we move towards offering a “production” service.

### What is ARROW?

ARROW was established as a European Commission supported project, which ran between 1st September 2008 and 28th February 2011; it was led by the Italian Publishers Association (AIE)<sup>3</sup> and had an extremely broad group of partners from across Europe – see Appendix 1 at the end of this document for full details.

The purpose of the project was to explore the extent to which existing data sources could be used to support a “diligent search” process for libraries involved in projects for the mass digitisation of books and similar textual works<sup>4</sup> in their collections. Data providers to the project (directly or indirectly) included the cultural sector represented by national libraries (national bibliographies); the (typically commercial) organisations which create and maintain “books in print” databases across Europe; and finally the (not-for-profit) collective management organisations which maintain a network for the collective management of textual reproduction rights on behalf of authors and publishers (“reproduction rights organisations” or RROs).

Following the successful completion of the ARROW project, a successor project (ARROW Plus) commenced immediately. This project is designed to extend the implementation of the ARROW infrastructure beyond the four countries that ran live pilots in the initial project – France, Germany, Spain and the UK – to a wider circle of European countries (including some with less sophisticated data infrastructures than those of the four initial pilot countries) and to explore the possibility of including images, as well as other types of written works.

---

<sup>1</sup> Accessible Registries of Rights Information and Orphan Works towards Europeana

<sup>2</sup> Stratton B (2011) *Seeking New Landscapes – A rights clearance study in the context of mass digitisation of 140 books published between 1870 and 2010*

<sup>3</sup> Associazione Italiana Editori <http://www.aie.it/>

<sup>4</sup> Including monographs

ARROW developed a core technical infrastructure for the management and routing of requests from libraries to national RROs, through a process of data enrichment and aggregation. National RROs are the natural ultimate target for requests to digitise textual works on a large scale, since they already have a network of contacts with the relevant rightsholders (in the ARROW context, book authors and publishers) and typically have extensive “repertoire databases” – linking content items with their rightsholders.<sup>5</sup>

Following the end of the project phase, ARROW can therefore be seen as a technical infrastructure:

- a shared central data ingestion, validation, enrichment and routing mechanism; this European resource continues to be managed and made available by AIE
- distributed nationally-based “diligent search” systems, based on local capabilities and requirements; in the UK, this is now jointly supported and staffed by ALCS and PLS

## What is the research process

At a high level, the ARROW-enabled process can be described as following these steps:

1. A library requests permission to digitise a particular book or books in its collection, by providing its own catalogue records for the items concerned; this information is critical to the process of identifying the right book (the first step in tracking down the relevant rightsholders)
2. The information is checked against relevant data held by The European Library (TEL)<sup>6</sup> which is essentially a “union catalogue” of the holdings of European National Libraries; all books published in Europe (or in the holdings of a European National Library) should be represented, and the record provided by the requesting library can be enhanced by additional information from the data held by TEL. This potentially includes (for example) enhanced author information (including dates of death – something which is one of the key factors in determining whether or not a work is still in copyright).
3. TEL also provides a service of “work clustering” – looking for other editions of the same title and relating these to the book that the library wants to digitise; this can be very important in identifying (for example) whether an ebook version is commercially available.
4. Once this has been completed, the records are checked against “books in print” databases to find out whether the book (or another edition of the book) is still commercially available (since, if it is, there is a presumption that the rightsholders will be unwilling to allow a library to digitise).

---

<sup>5</sup> The structure of the RRO in the UK is unique, in that it has a number of constituent parts. The Publishers Licensing Society (PLS: <http://www.pls.org.uk>) is responsible for managing relationships with publishers (including seeking their “mandates” for CLA licences) and for distributing money received from collective licences to them; and the Authors’ Licensing and Collecting Society (ALCS: <http://www.alcs.co.uk>) undertakes a similar role on behalf of authors. The Copyright Licensing Agency (CLA: <http://www.cla.co.uk/>) is the joint agent of ALCS and PLS and manages the process of licensing copyright users and collecting money from them for the use of copyrights; in addition, the Design and Artists Copyright Society (DACS) is the rights management organisation which represents the interests of visual artists in the UK and through an agency agreement, enables CLA to license the copying of artistic works on their behalf. This group of organisations is sometimes known collectively as “UKRRO”, and that is how we will refer to them in this document.

<sup>6</sup> <http://search.theeuropeanlibrary.org/portal/en> “The European Library is a free service that offers access to the resources of the 48 national libraries of Europe in 35 languages. Resources can be both digital (books, posters, maps, sound recordings, videos, etc.) and bibliographical. Quality and reliability are guaranteed by the 48 collaborating national libraries of Europe.”

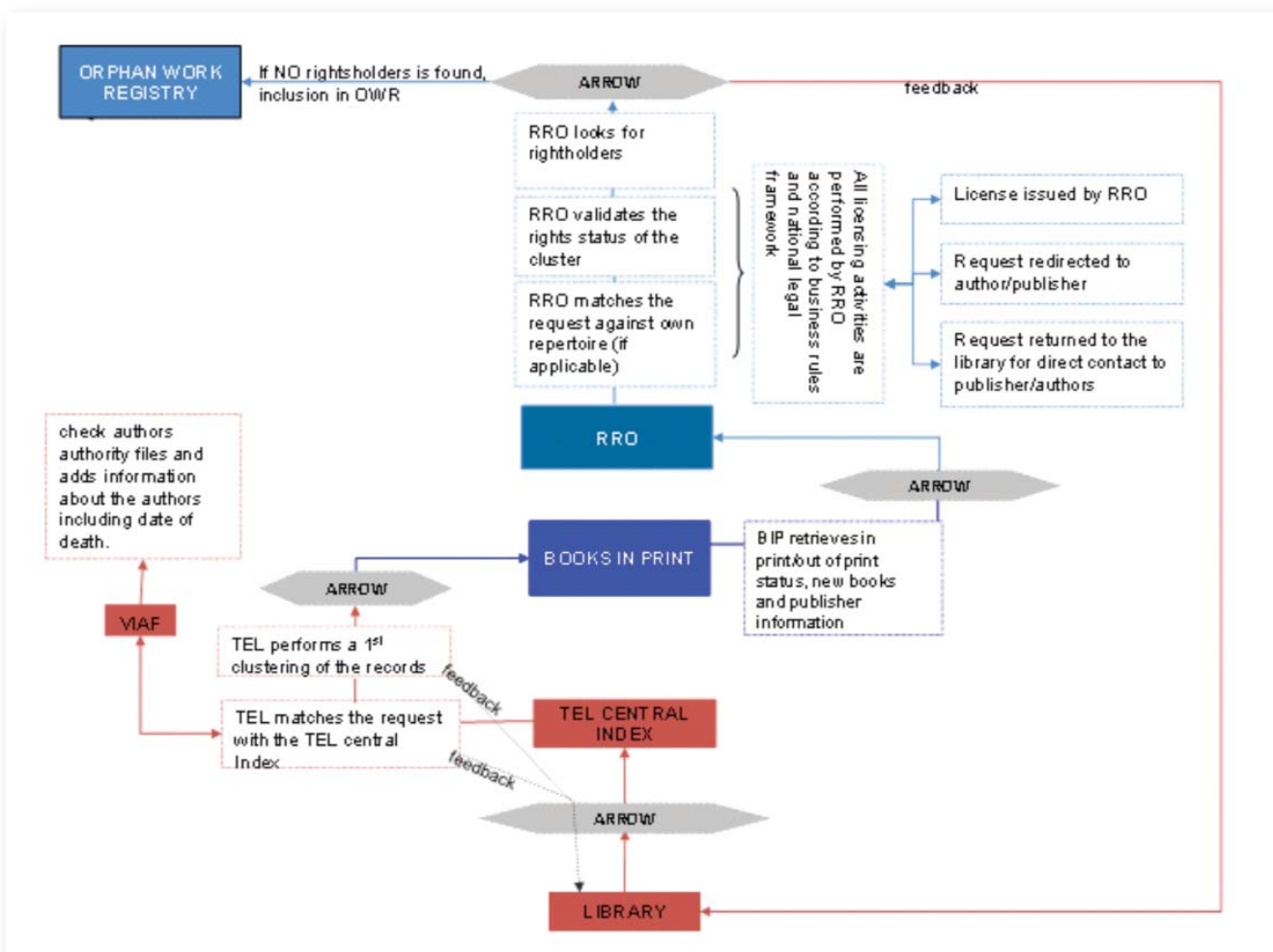


By this time, the ARROW process has removed from the need for further consideration:

- works that are out of copyright (where no permission is required to digitise) – in other words, where all the creators involved have been dead for 70 years or more
- works that are still in commerce (where digitisation permission is almost certain to be withheld)

From this point onwards the hard work of the RRO begins. Sometimes the identification of rightsholders is straightforward – because the work is already recorded in the relevant repertoire database. Where this is not the case, research processes to track down authors and publishers have been well honed by RROs, for whom tracing rightsholders is a daily task. RROs have extensive data resources and networks of contacts in the UK and overseas.

### Overview of the ARROW workflow

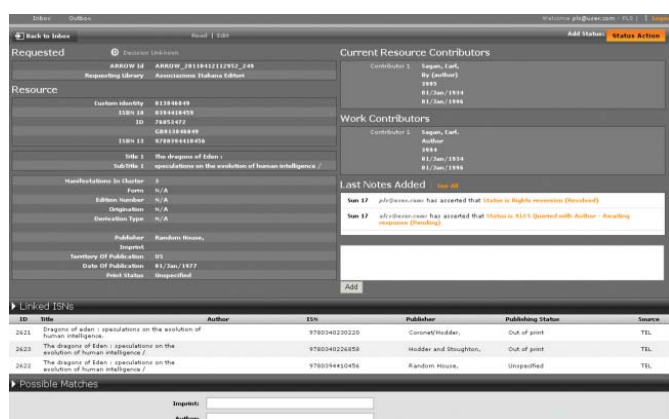


Nevertheless, not all rightsholders are likely to be found; for example, where authors are dead but their work is still in copyright, it can be extremely hard to track down literary heirs.

A proportion of works at the end of a diligent search process will be “orphan works” (works for whom the rightsholder cannot be identified or, if identified, located). In the case of the 101 records that could be validated by ARROW of the sample provided by the British Library<sup>7</sup> the proportion of works declared orphan was 21%.<sup>8</sup>

However, the advantage of implementing a system like ARROW to support diligent search is that (as a data-based system) it is capable of “learning”, of accumulating data and experience. Over time, it should be possible to identify a steady improvement in the efficiency of the system. Critically, from the point of view of the library requesting permissions, the enormous effort detailed in their paper is reduced to close to zero (simply the cost of uploading the catalogue records and performing whatever checks they choose on the responses that they receive). For the RRO, this is simply “business as usual” – facilitating the grant of rights between a rightsholder and a rights user.

### Identifying rightsholders through the UKRRO ARROW system



**7** The sample used for the project, which was taken from the British Library’s catalogue (see the BL report for a more detailed summary of the selection process), provided a particularly robust test of ARROW in that a significant percentage of the works in the sample consisted of grey literature published over the last 140 years (grey literature is typically produced on a non-commercial basis by organisations for which publishing is not their primary function, such as government departments, professional organisations, special interest groups etc., and so not subject to the same bibliographic controls, on the part of the publisher, as works intended for commercial exploitation), thus making the task of tracing the rights holders a much greater challenge.

**8** This figure differs from the 31% noted in Section 7 of the British Library’s paper, Seeking New Landscapes – A rights clearance study in the context of mass digitisation of 140 books published between 1870 and 2010. This is largely due to the fact that the British Library’s results refer to all 140 records in the original sample whereas this figure refers only to the 101 records that passed initial ARROW validation, of which a further 4 records failed validation at the RRO processing stage.

**9** It should be noted that a full like-for-like comparison with the BL results of the manual diligent search as reported in the Library paper is not possible as those results apply to 140 records while the UKRRO sample was a subset of 101. 39 records could not be validated by ARROW; some because they were published outside the UK (and were therefore out of the scope of the UK pilot), and others due to particularities in the catalogue data provided.

**10** Of this figure, 5% of works were still in-commerce, 39% in copyright (but not in-commerce), 22% out of copyright. For a full breakdown of the results from the UK pilot of ARROW see Appendix 2 overleaf – Results of the UK Pilot.

### The future

What lessons have been learned in the UK from the first “pilot” phase of implementation of ARROW?

The most important lesson of all is that ARROW works. Still somewhat imperfectly – as can reasonably be expected from a pilot – but nevertheless the tasks that ARROW was designed to complete were delivered. In a fraction of the time that the British Library took to achieve broadly similar results<sup>9</sup>, ARROW was able to identify (and to locate, in the case of works still in copyright) rightsholders for 66%<sup>10</sup> of the 101 British Library records that passed validation.

There are some requirements to improve the internal workflow and make it more flexible as ARROW moves towards “production” capability, and processes are required to work towards continuous improvement of data. However, the fundamental architecture proved robust and effective. UKRRO is already in advance talks about deploying ARROW for a digitisation project for the Wellcome Library, part of the Wellcome Trust. This, however, exposes what is perhaps the most significant gap in the current infrastructure – the ability to issue licences for digitisation, particularly for the digitisation of orphan works. At the time of publication it should be noted that the Hargreaves Report of 18 May, recommended legislation to enable potential licensing of orphan works by collective management organisations, which would allow the ARROW enabled UKRRO to make a “full service” offering to any library wishing to undertake a mass digitisation project of books or similar published textual works. The Government response to the Hargreaves Report was supportive of this approach, subject to the necessary consultation processes concerning its implementation.

# Appendix 1

## ARROW Partners

### Contract Partners

#### Libraries

Biblioteca Nacional de España (BNE)  
Bibliothèque Nationale de France (BNF)  
British Library (BL)  
Deutsche Nationalbibliothek (DNB)  
Koninklijke Bibliotheek (KB)  
Narodna in Univerzitetna Knjiznica (NUK)  
National Library of Norway (NLN)  
University of Innsbruck (UIBK)

#### Publishers Organisations

Associazione Italiana Editori (AIE)  
Federación de Gremios de Editores de España (FGEE)

#### Reproduction Rights Organisations

Copyright Licensing Agency (CLA)

#### International Organisations

Federation of European Publishers (FEP)  
International Federation of Reproduction Rights Organisations (IFRRO)

#### Technology Developers

Consorzio Interuniversitario per il Calcolo Automatico dell'Italia Nord Orientale (CINECA)  
NUMILOG

#### Bibliographic Agencies

Marketing und Verlagsservice des Buchhandels (MVB)

### Associate Partners

#### Libraries

The European Library (TEL)  
Finish National Library (FNL)

#### Authors Organisations

Authors Licensing and Collecting Society (ALCS)  
Stichting LIRA

#### Publishers Organisations

Syndicat National de l'Edition (SNE)  
Publishers Licensing Society (PLS)  
The Swedish Publishers Association

#### Reproduction Rights Organisations

Centro Español de Derechos Reprográficos (CEDRO)  
Centre Français d'exploitation du droit de Copie (CFC)  
Copy-Dan Writing Information (COPYDAN)  
KOPINOR  
KOPIOSTO  
SAZOR

# Appendix 2

## Results of the UK pilot of ARROW

The results in the table below show the outcome of processing the 101 validated records from the British Library's sample through ARROW's UK workflow.

Out of copyright	22
In Commerce	5
Orphan	21
In copyright (but not in-commerce): rights holder identified and located	39
In copyright: Unable to fulfil request <sup>11</sup>	10
Unable to process <sup>12</sup>	4
<b>Total</b>	<b>101</b>

---

<sup>11</sup> This category refers to records where insufficient data was available to identify the work or the author, or where, although the works in question were not orphan and were in copyright, the rights issues were considered to be too complex to resolve in the pilot e.g. books containing multiple works from multiple authors such as anthologies.  
<sup>12</sup> Unable to validate in UK system due to data particularity in incoming MARC21 record.

# Glossary

**Aleph** – the system software used by the British Library as an administrative interface for the Integrated Library System which delivers its internal and external services. The system's full name is Aleph 500 and it is provided by the company Ex Libris.

**ALCS** – Authors' Licensing and Collecting Society. Collecting society representing UK writers that collects royalties and campaigns on behalf of its members. Established in 1977.

**ARROW** – Accessible Registries of Rights Information and Orphan Works towards Europeana. A project of a consortium of European national libraries, European and national publishers and CMOs representing publishers and writers. The ARROW system is intended to support digitisation through identification of rightsholders and rights and clarifying whether a work is in print or is an orphan.

**ASCII** – (pronounced ass-kee) American Standard Code for Information Exchange. A character encoding scheme that enables text to be represented on computers. There are 128 characters in the ASCII scheme and it is the code upon which the majority of modern character encoding schemes are based.

**BLIC** – The British Library Integrated Catalogue. The catalogue of the British Library which contains the majority of items within the Library's collection.

**CDPA** – The Copyright, Designs and Patents Act 1988. UK legislation governing copyright and related intellectual property rights.

**CLA** – Copyright Licensing Agency. UK Collecting society and RRO, owned by PLS and ALCS which, in conjunction with DACS, provides licences for business and other organisations to make reproductions of textual and visual material. Established in 1983.

**CMOs** – Collective Management Organisations. see 'collecting societies'.

**Collecting societies** – also known as collective management organisations or CMOs. These are organisations that represent the rights of copyright owners in order to licence their works for use by businesses, educational/cultural institutions, community groups etc. Those collecting societies who administer licensing for reproduction of copyright material are called Reproduction Rights Organisations or RROs.

**DACS** – Design and Artists Copyright Society. UK collecting society representing artists that licences its members' works and campaigns on their behalf. Established in 1984.

**Dewey Decimal System** – library classification scheme developed by Melvil Dewey in 1876. It remains the most commonly used in the world, enabling stock to be arranged on shelves in a specific order to make it easier to locate particular items or to browse in specific subject areas.

**Diligent search** – the steps that must be undertaken in rights clearance to mitigate against risk of copyright infringement.

**Grey literature** – refers to published material which is distributed through non-commercial means and not via conventional publishers. Examples of grey literature include technical reports from government departments or scientific research groups and working papers from research groups and committees.

**ISBN** – International Standard Book Number. A unique numeric commercial book identification system developed by the International Organization for Standardisation (ISO) which came into widespread use during the 1970s.

**Monograph** – The term 'monograph' is used by librarians to mean a non-serial publication complete in one volume or in a finite number of volumes. In the context of this study it covers books, treatises, pamphlets and other single subject publications.

**MARC** – MACHine-Readable Cataloguing. A set of format standards which allow the communication of bibliographic information in a machine readable form. MARC 21 is a development of the MARC standard, so named as it redefined the format for the 21st century requirements of the international library community.

**Orphan Works** – copyright works for which the rightsholder cannot be identified or located.

**PLS** – Publishers' Licensing Society. UK collecting society established in 1981 representing book, journal and serial publishers.

**Public Domain** – term used to signify that the copyright subsisting in a creative work has expired and therefore the restrictions on its usage no longer apply.

**Shelfmark** – the code used by Libraries to locate a collection item on a shelf.

**RROs** – Reproduction Rights Organisations. See 'collecting societies'.

**UKRRO** – The term used to describe the collection of UK based RROs – CLA, PLS, ALCS and DACS – that contributed to the ARROW project.



[www.bl.uk](http://www.bl.uk)