



CREATE Working Paper 2018/08 (October 2018)

# Digitising the Edwin Morgan Scrapbooks: A Research Compendium

---

Author	Ronan Deazley Queen's University, Belfast R.Deazley@qub.ac.uk	Kerry Patterson University of Glasgow Kerry.Patterson@glasgow.ac.uk
	Victoria Stobo University of Liverpool Victoria.Stobo@liverpool.ac.uk	

---

CREATE Working Paper Series DOI: 10.5281/zenodo.1442583

This Working Paper is a compilation of the research sections from the web resource:  
[www.digitisingmorgan.org](http://www.digitisingmorgan.org).

It was created as part of *Digitising the Edwin Morgan Scrapbooks*, through support by the RCUK funded Centre for Copyright and New Business Models in the Creative Economy (CREATE), AHRC Grant Number AH/K000179/1.

<b>1. INTRODUCTION</b> by Ronan Deazley	3
<b>2. ABOUT THE PROJECT</b> by Ronan Deazley	4
<b>3. ORPHAN WORKS: LAW</b> by Ronan Deazley & Kerry Patterson	5
3.1 INTRODUCTION	5
3.2 ORPHAN WORKS IN CONTEXT	6
3.3 ORPHAN WORKS AND THE 2039 RULE	6
3.4 POSSIBLE SOLUTIONS FOR ORPHAN WORKS	7
3.5. THE ORPHAN WORKS DIRECTIVE: AN EXCEPTION FOR ORPHAN WORKS	8
3.5.1. THE ORPHAN WORKS DIRECTIVE: SOME PRACTICAL GUIDANCE	10
3.5.2. THE ORPHAN WORKS DIRECTIVE: ANALYSIS AND CRITIQUE	10
3.5.2.1. SCOPE	10
3.5.2.2. UNPUBLISHED WORKS	11
3.5.2.3. COMPENSATION	11
3.6. THE UK ORPHAN WORKS LICENSING SCHEME (OWLS)	12
3.6.1. OWLS: SOME PRACTICAL GUIDANCE	13
3.6.2. OWLS: ANALYSIS AND CRITIQUE	15
3.6.2.1. STATE-SPONSORED LICENSING	15
3.6.2.2. SEVEN-YEAR LICENCES	15
3.6.2.3. COMMODIFYING ORPHAN WORKS	15
3.7. UNDERSTANDING DILIGENT SEARCH	16
3.7.1. HOW DILIGENT IS DILIGENT?	16
3.7.2. WHAT ARE THE RELEVANT SOURCES?	18
3.7.3. IPO GUIDANCE ON DILIGENT SEARCH	20
3.7.4. DOES DILIGENT MEAN EXHAUSTIVE?	22
3.7.5. DILIGENT SEARCH IN PRACTICE	24
<b>4. DILIGENT SEARCH IN CONTEXT AND PRACTICE</b> by Kerry Patterson, Ronan Deazley and Victoria Stobo	26
4.1. INTRODUCTION	26
4.2. DILIGENT SEARCH IN CONTEXT	26
4.3. DEFINING DILIGENT SEARCH	28
4.4. DATA EXTRACTION AND CATEGORISING WORKS	29
4.5. PREPARING FOR DILIGENT SEARCH: CUTTINGS IN CONTEXT	30

4.6. DILIGENT SEARCH: SEARCHING FOR TEXT ONLINE	31
4.7. DILIGENT SEARCH: USING CONTENT BASED IMAGE RETRIEVAL (CBIR)	32
4.8. DILIGENT SEARCH EXAMPLE: PAGE FROM SCRAPBOOK 12	37
4.9. SELECTING WORKS FOR AN OWLS APPLICATION	42
4.10. MAKING AN OWLS APPLICATION	42
4.11. SELECTING WORKS FOR AN EUIPO REGISTER APPLICATION	47
4.12. MAKING AN APPLICATION TO THE EUIPO REGISTER	47
4.13. CLEARING THE SCRAPBOOKS	49
<b>5. DIGITISATION AND RISK</b> by Victoria Stobo, Kerry Patterson and Ronan Deazley	52
5.1. INTRODUCTION	52
5.2. COPYRIGHT AND DIGITISATION	52
5.3. RIGHTS MANAGEMENT IS RISK MANAGEMENT	54
5.4. CASE STUDIES: TAKING AND AVOIDING RISKS	57
5.4.1. COLLECTIONS BACKGROUND	57
5.4.2. RIGHTS AUDITING THE COLLECTIONS	58
5.4.3. THE RIGHTS CLEARANCE PROCESS	59
5.4.4. RESULTS OF THE RIGHTS CLEARANCE PROCESS	59
5.4.5. RISK MANAGEMENT DECISIONS	61
5.4.6. ANALYSIS	62
5.5. MANAGING RISK IN THE SCRAPBOOKS	63
5.5.1. RISK ASSESSMENT CATEGORIES	64
<b>6. CONCLUSION</b> by Ronan Deazley	66
<b>7. BIBLIOGRAPHY</b>	71

## **1. INTRODUCTION**

By Ronan Deazley

How do archives, museums and libraries enable digital access to works in their collections when it is difficult to identify or locate the copyright owners of that material? If permissions cannot be secured, institutions may simply avoid socially beneficial uses of these works, preferring instead to digitise material for which copyright has expired or is easy to clear. But, when decisions about the digitisation of heritage collections are influenced by the copyright status of the material itself, this skews the digital cultural record.

This problem has been addressed in part by the EU Orphan Works Directive 2012 and the UK Orphan Works Licensing Scheme. But, are these solutions fit for purpose? At the heart of both regimes lies the concept of diligent search. Drawing on unpublished scrapbooks created by the Edwin Morgan, the first Scots Makar, this project presents the first UK study addressing the legal and practical realities of diligent search since the Directive and OWLS came into effect. In addition, we include a set of resources providing clear, authoritative and practical guidance on a range of issues relevant for institutions engaging in similar digitisation initiatives (available at <https://www.digitisingmorgan.org/Resources>).

## **2. ABOUT THE PROJECT**

By Ronan Deazley

Copyright inhibits the digitisation of cultural heritage collections. The problem that orphan works pose for heritage memory institutions who want to enable the widest possible engagement with our shared cultural heritage is widely acknowledged. Orphans are works that are in copyright but for which the owner(s) cannot be identified or located. If permissions cannot be secured because rightholders cannot be contacted, institutions may simply avoid socially beneficial uses of these orphans, preferring instead to deal with material for which copyright has expired or is easy to clear. But, when decisions about the digitisation of heritage collections are shaped by the copyright status of the material itself, this skews the digital cultural record.

In the UK, there has been a twin-track approach to the problem of orphan works in the form of a new exception implementing the EU Orphan Works Directive 2012 operating in tandem with the Orphan Works Licencing Scheme (OWLS). The operation of both the exception and the licensing scheme is contingent on conducting a diligent search for every single orphan work to be digitised, regardless of the nature and scale of the project. Various commentators have suggested that the demands of diligent search inhibit the ambitions that institutions may have to engage in mass digitisation of their collections. This project is the first major UK study concerning the concept of diligent search since the Directive and OWLS came into effect.

In this resource, we report findings from a rights clearance simulation on a culturally significant set of unpublished scrapbooks created by Edwin Morgan, the first Scots Makar. Although unpublished, the scrapbooks are replete with orphan material, such as cuttings from newspapers, magazines and books. We conclude that mass digitisation and diligent search are fundamentally incompatible, however light-touch the diligent search obligation might be. In addition, we interrogate the legal and practical requirements of diligent search under both the Directive and OWLS, and explore the necessary role that a risk-managed approach to copyright compliance will continue to play in this domain, regardless of the introduction of the orphan works regime. Finally, we include a suite of resources providing clear, authoritative and practical guidance on a range of issues relevant for institutions engaging in cognate digitisation initiatives.

### 3. ORPHAN WORKS: LAW

Ronan Deazley and Kerry Patterson <sup>1</sup>

#### 3.1. INTRODUCTION

In recent years, successive governments have encouraged the cultural heritage sector to make their collections available online, enhancing access to meet user expectations in a digital, networked world. But, enabling online access to material that is still in copyright requires rights clearance, an activity widely perceived to be a barrier to the digitisation and dissemination of library and archive collections. There are a number of reasons for this:

- determining whether a work is in copyright or is public domain can be difficult
- the time and effort involved in identifying, locating and contacting known copyright owners can be prohibitive, especially when dealing with large numbers of works
- identifying and locating rights owners provides no guarantee of a response from that owner
- copyright owners cannot always be identified or located, leading to the so-called orphan works problem

The problem that orphan works pose for cultural heritage institutions in Europe and elsewhere is substantial, and it breeds uncertainty. The US Copyright Office recently summed up the consequences of the orphan works phenomenon in the following way:

[C]autious libraries, archives and museums may forgo socially beneficial use of orphan works, thereby excluding potentially important works from the public discourse and threatening to impoverish our national cultural heritage. Other types of socially beneficial uses of orphan works may be forestalled ... Filmmakers may avoid projects using orphan works as documentary source materials, businesses may not elect to commercially reissue lost works, and researchers may avoid potentially socially beneficial activities. According to one scholarly commentator, the orphan works problem “manifest[s] the greatest obstacle to copyright social utility in the developed world”. Hence, eliminating barriers to the use of orphan works would yield considerable societal benefits that would reverberate throughout the copyright system.<sup>2</sup>

Within this part of the resource we provide an overview of the legislative framework concerning the lawful use of orphan works in the UK, and critically assess the scope of both the European Orphan Works Directive (the Directive) and the UK Orphan Works Licensing Scheme (OWLS). We conclude by considering the concept of diligent search as it is currently defined within the UK copyright regime.

---

<sup>1</sup> We would like to thank Margaret Haig and Simone Schroff for taking the time to share their views on earlier drafts of this work; their comments and criticisms proved extremely helpful in shaping our final analysis of the demands of the diligent search requirement. The opinions stated herein are our own. The usual conditions apply.

<sup>2</sup> Pallante, M.A., *Orphan Works and Mass Digitization: A Report of the Register of Copyrights* (US Copyright Office, 2015), 38-39 (references omitted), available at: <http://copyright.gov/orphan/reports/orphan-works2015.pdf> (accessed: 12 December 2016).

### 3.2. ORPHAN WORKS IN CONTEXT

Essentially, a copyright-protected work is an orphan work if the owner of the rights in the work cannot be identified or located by someone seeking permission to perform one of the exclusive rights provided for under the copyright regime (for example, publication or making the work available online).

There are two factors intrinsic to the copyright regime that contribute to work becoming orphaned.<sup>3</sup> First, copyright does not require registration before work is protected, nor does a copyright notice have to be attached to a work. As soon as a qualifying work is created, it is automatically copyright-protected. This differs from other areas of intellectual property, such as patent law, where the acquisition of rights depends upon mandatory registration requirements. As a result, reliable information about the rights owner cannot always be easily or readily acquired for copyright works, particularly in the case of unpublished works.

The second factor concerns the very long duration of the copyright term. In the absence of a compulsory registration system, tracking a chain of title over a long period of time can become incredibly complicated. For example, the different economic rights provided by copyright might be separately assigned or licensed to third parties, or inherited by one or more heirs who may be unaware of their rights; they may have been assigned, licensed or inherited numerous times throughout the course of the copyright term. Alternatively, the corporate interests that own the rights might become bankrupt or simply go out of business. All of this can make identifying and locating the current copyright owner(s) extremely problematic.

### 3.3. ORPHAN WORKS AND THE 2039 RULE

The problem that the long duration of the copyright terms poses for orphan works within the UK is compounded by the '2039 rule'. The Copyright Designs and Patent Act 1988 (the CDPA) states that the duration of copyright in certain unpublished literary, dramatic, musical and artistic works, films and sound recordings, continues until 50 years from the end of the calendar year in which the 1988 Act came into force (that is, until 31 December 2039). Moreover, these provisions – collectively referred to as the 2039 rule – apply regardless of when the works in question were first created or when their authors died. Further details about the type of works that fall within the scope of the 2039 rule are set out in Schedule 1 of the CDPA.<sup>4</sup>

So: imagine that someone discovers a previously unknown manuscript by William Shakespeare. Shakespeare died in 1616, a time before the existence of copyright law as we know it. Yet, because the 2039 rule applies, this newly discovered manuscript remains in copyright. Anyone wanting to publish or perform the play would have to seek

---

<sup>3</sup> For relevant commentary, see van Gompel, S., 'Unlocking the potential of pre-existing content: how to address the issue of orphan works in Europe?' (2007) *International Review of Intellectual Property and Competition Law* 669-702.

<sup>4</sup> The 2039 rule affects various categories of unpublished work, including: anonymous or pseudonymous literary, dramatic, musical or artistic works (other than photographs) (Sch.1, paragraph 12(3)(b)); literary, dramatic and musical works, engravings as well as photographs taken on or after 1 June 1957 (paragraph 12(4)); sound recordings made on or after 1 June 1957, and films which have not been registered under the Cinematograph Films Act 1938 (and subsequent legislation) (paragraph 12(5)); Crown copyright works (paragraph 41(3)(b), (4) and (5)).

permission from the copyright owners, and seeking permission would involve tracking the chain of title across a period of 400 years through numerous generations and testamentary bequests, a time-consuming and costly undertaking. That effort may be worth it if the document in question was indeed a newly discovered manuscript by Shakespeare, but for the unpublished records typically held in archive institutions who would want to take on such a task, and who has the time?

Nor is this hypothetical example so ridiculous. The National Archives has estimated that there are over 100 million unpublished archive items subject to 2039 rule in the UK.<sup>5</sup> And while The National Archives may not house any unknown Shakespeare manuscripts, it does hold numerous unpublished private charters which are copyright-protected despite that they precede the introduction of copyright legislation to the UK by around 700 years. They even predate the introduction of the printing press into Britain by approximately 400 years.<sup>6</sup> We think it is absurd that any work created before printing was a viable technology is protected by copyright; we do not think many others would disagree.

The government has acknowledged that the 2039 presents a significant problem for the use and reuse of historic materials. In October 2014, the government consulted on reforming the 2039 rule but ultimately decided not to take any action.<sup>7</sup>

### 3.4. POSSIBLE SOLUTIONS FOR ORPHAN WORKS

Different solutions have been proposed for dealing with the phenomenon of orphan works.<sup>8</sup> Broadly speaking, they either require the prospective user to secure a licence to make use of the orphan work, or they do not. In relation to the latter, two main legislative solutions have been proposed: a statutory exception that allows for the use of an orphan work, and a limitation on liability rule. As for licensing-based solutions, again two basic models have been proposed: licensing by a public authority, or that collecting societies licence the use of orphan works in accordance with the principles of extended collective

---

<sup>5</sup> Intellectual Property Office, *Consultation on reducing the duration of copyright in unpublished ("2039") works in accordance with section 170(2) of the Copyright, Designs and Patents Act 1988* (October 2014), Annex A, available at: <https://www.gov.uk/government/consultations/reducing-the-duration-of-copyright-in-certain-unpublished-works> (accessed: 01 March 2016).

<sup>6</sup> BIS, *Enterprise and Regulatory Reform Act 2013: Policy Paper* (June 2013), 34, available at: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/209896/bis-13-905-enterprise-and-regulatory-reform-act-2013-policy.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/209896/bis-13-905-enterprise-and-regulatory-reform-act-2013-policy.pdf) (accessed: 01 March 2016).

<sup>7</sup> Intellectual Property Office, *Copyright Works: Seeking the Lost* (2014), available at: [www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/308372/consult-2014-lost.pdf](http://www.gov.uk/government/uploads/system/uploads/attachment_data/file/308372/consult-2014-lost.pdf) (accessed: 01 March 2016).

<sup>8</sup> See: Pallante, *Orphan Works and Mass Digitization*, 18-32; Favale, M., Kretschmer, M, and Mendis, D., 'The Treatment of Orphan Works under Copyright Law in Seven Jurisdictions: A comparative review of the underlying principles' (2013) CREATE Working Paper 2013/7 (July 2013), available here: <https://zenodo.org/record/8377/files/CREATE-Working-Paper-2013-07.pdf> (accessed: 01 March 2016); Vetulani, A., *The Problem of Orphan Works in the EU: An overview of legislative solutions and main actions in this field* (February 2008), 8-14, available at: [http://ec.europa.eu/information\\_society/activities/digital\\_libraries/doc/reports\\_orphan/report\\_orphan\\_v2.pdf](http://ec.europa.eu/information_society/activities/digital_libraries/doc/reports_orphan/report_orphan_v2.pdf) (accessed: 15 March 2013); de la Durantaye, K., 'Finding a Home for Orphans: Google Book Search and Orphan Works Law in the United States and Europe' (2010-2011) *Fordham Intellectual Property, Media and Entertainment Law Journal* 229-91, 247-58.



licensing. You can find out more about these different approaches here<sup>9</sup> (see pp.18-32 and 39-72) and here<sup>10</sup> (see pp.5-55).<sup>11</sup>

In Europe and the UK, a statutory exception and licensing by a public authority have been chosen as the preferred models. Specifically, an approach based on a statutory exception has been adopted in Europe in the guise of the *Orphan Works Directive*,<sup>12</sup> one of the key initiatives of the European Commission's *Digital Agenda for Europe*.<sup>13</sup> Within the UK, additional provision has been made by the introduction of an orphan works licensing scheme (hereafter: OWLS), following a recommendation of the *Hargreaves Review*.<sup>14</sup> OWLS is the world's first online licensing scheme for orphan works, and is much broader in scope than the European Directive. And so, Britain has two options for enabling the use of orphan works: the orphan works exception that applies across the EU (implementing the Directive) and the UK-based orphan works licensing scheme. Both are discussed in further detail below.

### 3.5. THE ORPHAN WORKS DIRECTIVE: AN EXCEPTION FOR ORPHAN WORKS

The Orphan Works Directive obliges Member States to introduce a new exception to copyright that enables specific types of use of some categories of orphan works. In the UK the Directive was implemented by the *Copyright and Rights in Performances (Certain Permitted Uses of Orphan Works) Regulations 2014* (hereafter: the *Orphan Works Regulations*) which introduced s.44B to the CDPA (Permitted uses of orphan works) along with the accompanying Schedule ZA1. As the Directive has been implemented almost verbatim by the 2014 Regulations, in the commentary that follows we make reference primarily to the provisions of the Directive.

The relevant institutions ('beneficiary organisations') entitled to rely on this new exception are cultural heritage institutions, in particular: publicly accessible libraries, educational establishments and museums, archives, film or audio heritage institutions as well as public-service broadcasting organisations (A.1(1)). Institutions wishing to rely on the exception must maintain records of any searches it has carried out to try and identify or locate the copyright owner. They then report the results of their diligent searches, providing information about the work and any relevant rightholders, to the European Intellectual Property Office (EUIPO).<sup>15</sup> EUIPO maintains a database of items registered as orphan works (available here<sup>16</sup>). In the first year of the Directive, only 10 organisations

---

<sup>9</sup> <http://copyright.gov/orphan/reports/orphan-works2015.pdf> (accessed: 24 February 2017)

<sup>10</sup> <https://zenodo.org/record/8377/files/CREATe-Working-Paper-2013-07.pdf> (accessed: 24 February 2017)

<sup>11</sup> See also van Gompel, 678-99.

<sup>12</sup> *Directive 2012/28/EU of the European Parliament and of the Council of 25 October 2012 on certain permitted uses of orphan works*, available here: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:299:0005:0012:EN:PDF> (accessed: 12 December 2016).

<sup>13</sup> *A Digital Agenda for Europe*, COM(2010) 245, available here: <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52010DC0245&from=EN> (accessed: 12 December 2016).

<sup>14</sup> Hargreaves, I., *Digital Opportunity: A Review of Intellectual Property and Growth* (May 2011), 8, 38-40, available here: <http://www.ipo.gov.uk/ipreview-finalreport.pdf> (accessed: 12 December 2016).

<sup>15</sup> The EUIPO was formerly known as the Office for Harmonization in the Internal Market (OHIM). With the entry into force of Regulation No 2015/2424, on 23 March 2016 OHIM became the European Union Intellectual Property Office (EUIPO). As such, throughout the remainder of the commentary we will make reference to EUIPO.

<sup>16</sup> <https://oami.europa.eu/ohimportal/en/web/observatory/orphan-works-database> (accessed: 24 February 2017)

in the UK registered with EUIPO as a beneficiary organisation (and only 61 organisations across the EU).<sup>17</sup> At the time of writing, EUIPO database contained just under 2000 entries (although there are many more embedded works on the database that this number would suggest).<sup>18</sup>

The type of material that falls within the scope of the exception includes: books, journals, newspapers, magazines and other writings, as well as phonograms, cinematographic and audiovisual works (A.1(2)). Under the Directive a work will be considered orphan if ‘none of the rightholders in [the work] is identified or, even if one or more of them is identified, none is located despite a *diligent search* for the rightholders having been carried out’ (A.2(1)) (emphasis added). However, the Directive does allow for *partially* orphan works. That is, where only some of the rightholders have been identified and located, the work may still be used under the Directive provided those known rightholders have granted permission for the use in question, and a diligent search has been carried out for the remaining owners.

The concept of diligent search is fundamental to the operation of both the Directive and OWLS. The Directive stipulates that a diligent search must be carried out for *each work prior to the use of that work*, and provides a list of sources for each category of work to be consulted as part of the diligent search process. These sources are set out in the Annex to the Directive. We consider the concept of diligent search in greater detail in section 7 below.

It is important to note that the diligent search need only be carried out in the Member State where the work was first published or broadcast (A.3).<sup>19</sup> This is because the Directive operates a principle of the mutual recognition of orphan works throughout the single market. That is, a work that is considered an orphan work in one Member State is deemed to be an orphan work in all Member States (A.4). This principle of mutual recognition has the benefit of identifying one relevant jurisdiction in which the diligent search is carried out;<sup>20</sup> it also removes the need to duplicate the diligent search in another Member State.

Once a work has been designated orphan, a beneficiary organisation can take advantage of the uses permitted under Article 6(1). That is, they are permitted to use orphan works contained within their respective collections in the following ways:

- copying the work for the purposes of digitisation, indexing, cataloguing, preservation, restoration and making the work available (A.6(1)(b))

---

<sup>17</sup> Intellectual Property Office, *Orphan Works: Review of the first twelve months* (2015), 16, available here: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/487209/orphan-works-annual-report.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/487209/orphan-works-annual-report.pdf) (accessed: 16 December 2016).

<sup>18</sup> On 27 November 2016 the database contained 1998 entries; see: [euipo.europa.eu/orphanworks/#search/basic/all](http://euipo.europa.eu/orphanworks/#search/basic/all) (accessed: 27 November 2016).

<sup>19</sup> This is subject to an exception concerning cinematographic or audiovisual works when the producer of the work in question has his headquarters or habitual residence in a Member State: in this case, the diligent search must be carried out in the Member State of the producer’s headquarters or habitual residence; A.3(3). Moreover, if there is evidence to suggest that relevant information on rightholders is to be found in other countries, then relevant sources of information available in those other countries should also be consulted; A.3(4).

<sup>20</sup> Although this is subject to the proviso concerning joint authors from different Member States set out in A.3(3).

- communicating the work to the public, including making it available online (A.6(1)(a))

Article 6(2) makes clear that beneficiary organisations are only permitted to make use of orphan works in order to achieve aims related to their public interest missions, ‘in particular the preservation of, the restoration of, and the provision of cultural and educational access to works and phonograms contained in their collection’. However, the Directive does state that relevant organisations may generate revenue in their use of orphan works ‘for the exclusive purpose of covering their costs of digitising orphan works and making them available to the public’.

The fact that a work has been designated an orphan does not mean that it will always remain an orphan. Member States must ensure that the owner of a work considered to be orphan can put an end to the work’s orphan status (A.5). Within the UK, the legislation states that the rightholder may put an end to the orphan work status of a relevant work by providing evidence of her ownership of the rights in question to EUIPO or to the relevant body (the beneficiary organisation) which carried out the relevant diligent search in the first place.<sup>21</sup>

### **3.5.1. THE ORPHAN WORKS DIRECTIVE: SOME PRACTICAL GUIDANCE**

The UK Intellectual Property Office has developed an *Eligibility Questionnaire* to help individuals and organisations determine whether they, and their proposed use of the orphan works, fall within the scope of the Directive; the questionnaire is available here<sup>22</sup>.

### **3.5.2. THE ORPHAN WORKS DIRECTIVE: ANALYSIS AND CRITIQUE**

While many aspects of the Directive are to be welcomed, the scheme is unsatisfactory when considered from the perspective of cultural heritage institutions and particularly for the archive sector. In the commentary that follows we highlight three key issues: (i) the scope of works covered by the Directive; (ii) the Directive’s applicability to unpublished work; and, (iii) the requirement of fair compensation for reappearing rightholders.

#### **3.5.2.1. SCOPE**

As set out in A.1(2), the Directive applies *only* to the use of books, journals, newspapers, magazines and other writings, as well as phonograms, cinematographic and audiovisual works. Free-standing artistic works such as maps, drawings, plans, and photographs do not fall within remit of the Directive.<sup>23</sup> While estimates concerning the scale of the

---

<sup>21</sup> See CDPA, Sch. ZA1 paragraph 7(2).

<sup>22</sup> <https://www.orphanworkslicensing.service.gov.uk/eu-eligibility/answer/1> (accessed: 24 February 2017)

<sup>23</sup> The Directive does however provide that it extends to ‘works and other protected subject-matter that are embedded or incorporated in, or constitute and integral part of, the [works listed in A.1(2) and (3)]’; A.1(4).

problem of orphaned visual and artistic materials can vary quite dramatically,<sup>24</sup> the omission of these types of works is an obvious shortcoming.

This shortcoming is balanced to some extent by A.1(4) which states that the Directive also applies to works and other protected subject-matter embedded or incorporated in, or which constitute an integral part of, books, journals, newspapers, magazines and so on. That is, artistic works, such as photographs, embedded in works that fall within the scope of the Directive *also* fall within the scope of the Directive.

### 3.5.2.2. UNPUBLISHED WORKS

The extent to which the Directive applies to *unpublished works* is qualified in two ways that are very limiting. The Directive does apply to works and phonograms which have never been published or broadcast, so long as they 'have been made publicly accessible by [a relevant organisation] with the consent of the rightholders', and provided it is 'reasonable to assume that the rightholders would not oppose the uses referred to in Article 6' (A.1(3)). Taking the latter point first, a beneficiary organisation might rely on the Directive to digitise and make available online various unpublished works but only if it is reasonable to assume the unknown or untraceable owner would not object. In practice, though, when will it be reasonable to make such an assumption?

More significant is the prescription that unpublished material can only be digitised '*if it has been made publicly accessible with the consent of the rightholder*'. This poses a problem in relation to collections of records that include copyright material from third parties. Consider a collection of correspondence written by and to a noteworthy author. The unpublished letters written by the author may well have been originally deposited with the appropriate consent such that the material falls within the scope of the Directive. However, the letters written *to* the author will almost certainly not have been deposited with the kind of consent regarding access which the Directive requires. This severely limits the usefulness of the Directive when dealing with unpublished archive material.

### 3.5.2.3. COMPENSATION

With respect to owners who put an end to the orphan work status of their work under A.5, the Directive requires that Member States ensure they are provided with 'fair compensation' for the use that has been made of their work (A.6(5)). Within the UK,

---

<sup>24</sup> For example, a survey conducted by CEPIC of the commercial picture archives specialising in historic material reported that 50% of archives answered that orphan works comprise less than 10% of their stock, whereas 20% of respondents estimated that orphan works comprise between 20-30% of their stock; Fodor, S., *Results of CEPIC Survey on Orphan Works in Historical Archives* (September 2011), 4, available here:

[http://ceplic.org/news/ceplic\\_news/2011/10/results\\_ceplic\\_survey\\_orphan\\_works\\_historical\\_archives](http://ceplic.org/news/ceplic_news/2011/10/results_ceplic_survey_orphan_works_historical_archives). By contrast, in a study conducted by The National Archives in 2009 it was found that for works registered for copyright protection between 1883 and 1912, only 5% of the rights owners were traceable for 80,000 images still in copyright; cited in Vuopala, 30. For other estimates concerning orphaned visual material, see IPO, *Orphan Works, Impact Assessment No. BIS1063* (June 2012), 10, available here: <http://webarchive.nationalarchives.gov.uk/20140603093549/http://www.ipo.gov.uk/consult-ia-bis1063-20120702.pdf> (which, for example, reports that 25% of 500,000 pieces of artwork held by the National History Museum in London is orphan, as are 25% of the prints and drawings held in the London Metropolitan Archive).

Schedule ZA1 paragraph 7(3) states that, following the (re)appearance of the owner, the organisation that is using or has used the work must, within a reasonable period of time, provide the owner with fair compensation for that use together with information on how the fair compensation has been calculated (para. 7(3)). So, the level of compensation is, initially at least, left to the beneficiary organisation to determine. In the case of disagreement between the organisation and the owner, either party has the right to apply to the Copyright Tribunal to determine the amount to be paid (para. 7(4)).

But what will constitute fair compensation? Could this be interpreted to mean ‘no compensation’? This seems unlikely. Future claims to fair compensation might be calculated in line with the licence fees which the IPO currently levy under OWLS (discussed below), at least in relation to the types of non-commercial use allowed for under the Directive. At present, those fees are set at 10 pence (£0.10) per work for non-commercial use for a period of seven years. However, reappearing owners may well dispute the fairness of a settlement on those terms, and the Copyright Tribunal is certainly not bound to follow the tariffs suggested by the IPO.

In this respect, reliance on the exception scheme involves something of a gamble: it provides an opportunity to make use of the work today, set against the uncertain financial commitment that may be triggered by a reappearing owner(s) in the future. This inherent uncertainty about unspecified future costs may be one of the reasons why only 10 UK-based institutions registered as a beneficiary organisation to avail of the Directive in its first year of operation.

### 3.6. THE UK ORPHAN WORKS LICENSING SCHEME (OWLS)

When the *Enterprise and Regulatory Reform Act 2013* was passed, it included provisions to enable the government to establish an independent authorising body to grant licences to make use of orphan works within the UK (s.77). The 2013 Act was followed by the *Copyright and Rights in Performances (Licensing of Orphan Works) Regulations 2014* (hereafter: the *OWLS Regulations*) setting out the scope and detail of the Orphan Works Licensing Scheme (OWLS), and appointing the Intellectual Property Office (IPO) as the relevant authorising body. In turn, the IPO has established an *Orphan Works Register*, which can be accessed here<sup>25</sup>.

OWLS is much broader in scope than the European Directive. For example:

- while the scheme adopts the same definition of an orphan work as the Directive, it applies to *all types of copyright work* including free-standing artistic works (photographs, drawings, maps, and so on)
- *anyone* can apply for a licence under OWLS, not just libraries, educational establishments, museums and archives<sup>26</sup>
- OWLS enables both commercial and non-commercial uses of orphan works
- under OWLS the IPO will grant licences for all types of commercial or non-commercial activity, and not just those activities permitted under the Directive

---

<sup>25</sup> <https://www.orphanworkslicensing.service.gov.uk/view-register> (accessed: 24 February 2017)

<sup>26</sup> This is, however, subject to one exception: the 2014 Regulations set out that ‘[a]n orphan works licence may not be granted to a person authorised to grant licences’ (reg.6(4)).

In short, whereas the Directive enables the use of certain orphan works by certain organisations for certain purposes (across the EU), OWLS enables the use of *all orphan works by anyone for any purpose* (but only within the UK).

We can summarise the different nature and scope of the orphan works schemes provided for under the Directive and OWLS as follows:

	<b>DIRECTIVE (EU)</b>	<b>OWLS (UK)</b>
<b>NATURE OF THE SCHEME</b>	Copyright exception	Non-exclusive licence
<b>WHO CAN USE</b>	Cultural and educational institutions	Anyone
<b>WHAT WORK CAN BE USED</b>	Books, journals, newspapers, magazine and other writings, phonograms, cinematographic and audiovisual work	Everything
<b>HOW CAN THE WORK BE USED</b>	Copying to digitise, index, catalogue and preserve Communicate to the public, including making available online	Anything
<b>CONDITIONS</b>	Diligent search Non-commercial use only	Diligent search Application fee (upfront) & licence fee

**Table 1:** Nature and scope of the orphan works schemes

It should be remembered, of course, that the schemes are not mutually exclusive; it is perfectly acceptable to use both the Directive and OWLS within the same digitisation initiative, choosing the more appropriate route for each type of work to be digitised. Consider, for example, a collection of letters sent to and from soldiers on the frontline during the First World War, one of which contains a photograph. While the letters may be eligible to make available online under the Directive, the photograph will not: digitising the collection may well involve engaging with both regimes.

**3.6.1. OWLS: SOME PRACTICAL GUIDANCE**

As with the Directive, applicants under OWLS must conduct a diligent search in relation to each work for which they are seeking a licence. Crucially, the search undertaken should be properly documented. As part of the licensing process, applicants need to be able to clearly demonstrate that their search was diligent. Indeed, the IPO has produced a diligent search check list that must be completed (for each relevant rightholder that cannot be identified or located) and submitted with the licence application.

In addition to diligent search, applicants must also pay an application fee and a licence fee. The application fee is paid when the application is submitted, and is determined by the number of works for which a licence is sought (for example, to make an application to licence the use of one orphan work, the fee is £20; for 10 works, the fee is £40; for 20 works, it is £60). The maximum number of works that can be covered in a single application is 30 (for a fee of £80).

The licence fee is payable once an application has been successful, and the fee will vary depending on the type of orphan work and the proposed use. For example, while licence fees for commercial use are calculated in accordance with current market rates (wherever possible), the licence fee for non-commercial use is significantly lower. Indeed, the IPO

have set a licence fee of 10 pence (£0.10) per work for all non-commercial uses.<sup>27</sup> Moreover, it is possible to check how much a licence fee will cost *before* making a formal application, which helps with securing appropriate funding and/or changing the scope of the application to ensure affordability. VAT is payable on licence fees, but not on the application fee.

Licence fees paid under the scheme are held by the IPO on behalf of owners that are unknown or cannot be located. Should the owner reappear within eight years following the grant of the licence, the IPO will pay over any licence fee that was collected in respect of that work. After eight years, the owner no longer has a right to reclaim the relevant fee although the IPO can, at its discretion, make payment if there are good reasons for so doing. Any unclaimed licence fees will be used to offset the costs of administering OWLS; should there be a surplus it will be applied to fund social, cultural and educational activities. At present, the IPO have indicated that the decision-making process concerning the use of surplus funds will be established 'closer to the time' (that is, at some point within the first eight years of launching OWLS).<sup>28</sup>

All licences granted under OWLS are non-exclusive (meaning that the orphan work in question remains available for others to license and use) and are only valid for a term not exceeding seven years (reg.6(1)(b)). Thereafter, a new application must be submitted to enable the continued use of the work.

The IPO may refuse to grant a licence if they consider that a proposed use or adaptation of the work is not appropriate in the circumstances of the case, including whether the proposed adaptation constitutes a derogatory treatment of the work (or on any other reasonable ground). Should a licence be refused, the applicant can appeal the IPO's decision to the Copyright Tribunal.

During the first year of the scheme, 48 applications were made in relation to nearly 300 individual orphan works. Of these works, 79% concerned still images (typically, photographs). Application fees collected over this period totalled £1492, and 27 licences were granted to make use of 247 works, generating revenue of £8,001.97 (excluding VAT). Seven of the licences were granted for commercial use (in relation to 35 works), with the remaining 20 concerned with non-commercial use (in relation to 212 works). The revenue generated by non-commercial licences was £21.20 (excluding VAT). It is also worth noting that during this period no licences were refused by the IPO, and no rightholders reappeared in relation to licensed orphan works. However, the IPO have reported that rightholders *have been identified* by applicants while carrying out the required diligent search making use of the sources detailed in the IPO's guidelines.

General guidance as to the operation of the licensing scheme is available on the Government website.<sup>29</sup>

---

<sup>27</sup> Intellectual Property Office, *Orphan Works Licensing Scheme: Overview for Applicants* (October 2014), 2.32, available here:

[https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/450649/Orphan\\_Works\\_Licensing\\_Scheme\\_Overview\\_for\\_Applicants.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/450649/Orphan_Works_Licensing_Scheme_Overview_for_Applicants.pdf) (accessed: 01 March 2016).

<sup>28</sup> Intellectual Property Office, *Orphan Works: Review of the first twelve months*, 4.

<sup>29</sup> [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/450649/Orphan\\_Works\\_Licensing\\_Scheme\\_Overview\\_for\\_Applicants.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/450649/Orphan_Works_Licensing_Scheme_Overview_for_Applicants.pdf) (accessed: 24 February 2017)



### 3.6.2. OWLS: ANALYSIS AND CRITIQUE

In many respects, OWLS represents a much more progressive and enabling regime than the exception provided under the Directive. That said, there are issues specific to OWLS that are worth comment; they concern: (i) the appropriateness of state-sponsored licensing; (ii) time-limited licences; and, (iii) the implications of commodifying all archive and heritage collections.

#### 3.6.2.1. STATE-SPONSORED LICENSING

Some have questioned whether state-sponsored licensing is an appropriate mechanism at all, with the US Copyright Office rejecting this model as ‘highly inefficient’.<sup>30</sup> Nevertheless, licensing regimes have been implemented in Canada, South Korea, Japan, India and Hungary.<sup>31</sup> Of these, the Canadian system is longest established, dating to 1988. However, since that time the Canadian Copyright Board has issued relatively few licences regarding orphan works, which has led a number of commentators to criticise the Canadian scheme for being administratively burdensome and for delivering little in the way of actual public benefit.<sup>32</sup> It may be that OWLS will prove to be more successful. For one thing, it enables the use of a much greater variety of orphan works than the Canadian scheme,<sup>33</sup> which may encourage greater engagement with the scheme. However, initial figures regarding the uptake of OWLS (quoted above) are not particularly impressive or encouraging.

#### 3.6.2.2. SEVEN-YEAR LICENCES

The licences that can be granted under OWLS are capped at seven years, after which a new application must be made along with a new diligent search if the licensee wants to continue making use of the work. For cultural heritage institutions seeking to digitise their collections to enable free public access, the prospect of having to renew the orphan works licence every seven years may well deter applications to the scheme, especially for medium- and large-scale digitisation initiatives.

#### 3.6.2.3. COMMODYING ORPHAN WORKS

Whereas the transaction costs of engaging in diligent search (discussed below) and the need to renew licences every seven years may prove to be deterrents, the actual cost of the licence fee itself for non-commercial use is unlikely to be off-putting. In setting the

---

<sup>30</sup> Pallante, *Orphan Works and Mass Digitization*, 48.

<sup>31</sup> See, for example, Favale, M. et al, *Copyright and the Regulation of Orphan Works*.

<sup>32</sup> See, for example, British Screen Advisory Council, *Copyright and Orphan Works: A Paper Prepared for the Gowers Review by the British Screen Advisory Council* (2006), 11, available here:

[http://www.bsac.uk.com/files/copyright\\_orphan\\_works\\_paper\\_prepared\\_for\\_gowers\\_2006.pdf](http://www.bsac.uk.com/files/copyright_orphan_works_paper_prepared_for_gowers_2006.pdf)

(accessed: 12 March 2016), as well as the various criticisms considered in de Beer and Bouchard, 9-31.

For further detail on the Canadian regime, see de Beer, J., and Bouchard, M., *Canada’s “Orphan Works”*

*Regime: Unlocatable Copyright Owners and the Copyright Board* (December 2009), available here:

<http://www.cb-cda.gc.ca/about-apropos/2010-11-19-newstudy.pdf> (accessed: 12 December 2016).

<sup>33</sup> For example, in Canada, licences are only available in relation to *published* works and sound recordings, as well as fixed communication signals and performances.



licence fee at 10 pence (£0.10) per work for all non-commercial use, the IPO have strived to make licensing affordable while maintaining their commitment to ensuring that reappearing owners would receive some financial recompense for the use of their work. But, is it appropriate to commercialise orphan works in this way at all?, and particularly when considering orphans in archive collections? The records typically held in archives, have social, cultural, academic and historic significance although they were rarely created with the intention of commercial exploitation, and only a small proportion of these works have any *intrinsic* commercial value. Indeed, it is the organic nature of the records – the fact that they *evidence* individual and business actions and transactions – that makes them reliable, authentic and trustworthy, and so appropriate for inclusion within an archive collection. Should this material be subject to any form of commercialisation, however small the fee? And should institutions tasked with the cost of preserving these collections in the public good in perpetuity be subject to such fees?

### 3.7. UNDERSTANDING DILIGENT SEARCH

As mentioned above, the concept of diligent search is fundamental to the operation of both the Directive and OWLS, and must be carried out for each work prior to the use of that work. But what exactly does diligent search require?, and are the demands of diligent search the same when relying on the Directive or applying for a licence through OWLS? In this section, we consider how the concept of diligent search has been defined in UK copyright law, as well as the prescribed and recommended sources that one should consider when conducting a diligent search.

#### 3.7.1. HOW DILIGENT IS DILIGENT?

Within the UK, the concept of diligent search has been defined differently in the *Orphan Works Regulations* and the *OWLS Regulations*, a matter that has attracted no academic or practitioner attention to date of which we are aware.

The *Orphan Works Regulations* provide that for the purposes of establishing whether a relevant work is an orphan work ‘a relevant body must ensure that a diligent search is carried out in good faith in respect of the work by consulting the appropriate sources for the category of the work in question’ (Schedule ZA1, 5(1)). Compare the wording of the OWLS Regulations: it states that before applying for a licence, a licensee shall carry out a diligent search appropriate to the orphan work in question, and that ‘[a] diligent search must comprise a reasonable search of the relevant sources to identify and locate the right holders of the relevant work’.

So, whereas the search to be conducted out in accordance with the exception must be ‘carried out in *good faith*’, the search required to satisfy the licensing scheme must constitute a ‘*reasonable* search of the relevant sources’. Should we read these different standards – reasonableness and good faith – as synonyms within this regulatory landscape? It is not entirely clear. In relation to the exception, the requirement of a good faith diligent search is clearly articulated within the Directive.<sup>34</sup> But why not simply transpose the same good faith standard into the OWLS regime? Put another way, why

---

<sup>34</sup> A.3(1), and paragraph 13.

deploy different language if a different standard was not intended? And yet, both are tethered to the same 'minimum' list of sources to be consulted.

One explanation might lie in the fact that while 'good faith' as a legal term of art is familiar to continental lawyers it has tended to cause interpretive difficulties within common law systems such as the UK.<sup>35</sup> That said, 'good faith' is not entirely alien to British law or jurisprudence: it is a concept employed in consumer rights legislation, insurance law, and in relation to the actions of fiduciaries such as an agent, a solicitor, or the director of a company.<sup>36</sup> With consumer rights, for example, the law provides that any unfair term of a consumer contract will not bind the consumer, and defines a term to be unfair if 'contrary to the requirement of good faith, it causes a significant imbalance in the parties' rights and obligations' under the contract.<sup>37</sup> And in this context, Lord Bingham has described acting in good faith as engaging in 'fair and open dealing'.<sup>38</sup> Still, for lawyers trained in the common law system the concept has a vagueness which, from a legal perspective, can seem unsettling.<sup>39</sup> This may explain why the legislature has opted for the more familiar concept of reasonableness when implementing the UK-specific licensing scheme, while remaining faithful (literally) to the good faith standard set out in the Directive. But that still begs the question: when considering diligent search, are these standards one and the same?

Consider the concept of reasonableness within the context of copyright law: it can be found littered throughout the CDPA. The Act refers to making a 'reasonable inquiry' concerning the identity of the author of a work (s.9(5)), having 'reasonable grounds' for belief (s.25(1)), the 'reasonable terms' of a contract (s.31A), giving 'reasonable notice' (s.31BB), a 'reasonable proportion' of a work (s.42A), making a 'reasonable' assumption (s.57), reading a 'reasonable extract' (s.59), paying a 'reasonable royalty' (s.66), making a determination that is 'reasonable in the circumstances' (s.73A), providing a 'reasonable form of identification' (s.77(8)), exercising 'reasonable diligence' (ss.99 and 113),<sup>40</sup> using such 'reasonable force as is necessary' (s.109), waiting a 'reasonable time' (s.121) or a 'reasonable period' (s.135B), making 'reasonable payments' (s.133), the concept of a 'reasonable condition' (s.135C), and so on. What constitutes a reasonable condition, amount, time or payment will depend on the facts at hand. For our purposes, making a 'reasonable inquiry' or exercising 'reasonable diligence' seem particularly pertinent, although neither concept has attracted much consideration or attention either in the standard practitioner treatises or before the courts. In relation to the concept of 'fair dealing', Lord Justice Aldous has observed that 'the court must judge the fairness [of the

---

<sup>35</sup> Stannard, J.E., *Delay in the Performance of Contractual Obligations* (Oxford: OUP, 2007), 106.

<sup>36</sup> Virgo writes: 'A 'fiduciary' is essentially somebody, person A, who is in a relationship with another person, B, in which B is entitled to expect that A will act either in B's best interests or in their joint interests, to the exclusion of A's own interest'. Virgo, G., *The Principles of Equity and Trusts*, 2<sup>nd</sup> ed (Oxford: OUP, 2016), 498.

<sup>37</sup> *Consumer Rights Act 2015*, s.62(1)(4). In effect, this replicates the previous law as set out in the 1999 *Unfair Terms in Consumer Contracts Regulations* and, before that, the 1994 *Unfair Terms in Consumer Contracts Regulations*.

<sup>38</sup> *Director-General of Fair Trading v. First National Bank plc* [2001] UKHL 52 (Lord Bingham).

<sup>39</sup> In general, see Clarke, M., 'The Common Law of Contract in 1993: Is There a General Doctrine of Good Faith?' (1993) *Hong Kong Law Journal* 318-41.

<sup>40</sup> The CDPA provides that in certain circumstances someone can make an application for the delivery up of infringing materials after the end of the normal period of six years from the date on which the infringing materials were made; this is permissible when, for example, if during the relevant six year period the copyright owner was prevented by fraud or concealment from discovering the facts entitling him to apply for an order, and he could not with reasonable diligence have discovered those facts (s.113(2)(b)).

use] by the objective standard of whether a fair minded and honest person would have dealt with the copyright work [in the manner that the defendant did]'.<sup>41</sup> If pressed, when dealing with reasonable inquiry or diligence, the courts would almost certainly default to a similar objective standard: what steps would the honest and fair minded person think it is reasonable to take in the same circumstances?

With good faith, however, there are different ways to frame the appropriate standard. It might be determined objectively, as with the concept of reasonableness; here, the terms may indeed be synonyms. On the other hand, one might interpret it to mean no more than *an absence of bad faith or improper conduct* on the part of the person carrying out the diligent search, rather than imposing a more positive, substantive obligation (what would the reasonable person have done?). That is: so long as the researcher genuinely believes they have conducted a sufficient search and have not knowingly or intentionally omitted to check some relevant or obvious sources then the duty is discharged. Much will depend, however, on what constitutes a relevant source and when it should be consulted. We turn to these issues next.

### 3.7.2. WHAT ARE THE RELEVANT SOURCES?

With respect to the sources to be consulted when conducting a diligent search, in relation to the exception, Article 3(2) of the Directive lays down a clear direction to each Member State as follows: 'The sources that are appropriate for each category of works or phonogram in question shall be determined by the Member State, in consultation with rightholders and users'; that is, each Member State has a responsibility to outline which sources are appropriate to consult for each category of work.

The Directive continues that the list of sources shall include 'at least the relevant sources listed in the Annex'. In the interests of clarity, we reproduce the text of the Annex to the Directive below.

Not surprisingly, different jurisdictions have taken different approaches when discharging their obligation under A.3(2). The Netherlands, for example, has simply transposed the list of sources set out in the Annex to the Directive into the Dutch copyright regime, an approach that has been described as 'consciously hands-off'.<sup>42</sup>

Other jurisdictions have chosen to incorporate a more complete list of sources within their enabling legislation. Article 61a of the German Copyright Act outlines the diligent search obligation in Germany stating that 'at the very least the sources set out in the Annex [to this Act] must be consulted'. The Annex to the German Copyright Act is based upon the Annex to the Directive, but provides additional detail and direction. For example, in relation to audio-visual works whereas the Directive identifies 'databases of film or audio heritage institutions and national libraries' as relevant, the German Annex refers to the following:

---

<sup>41</sup> *Hyde Park Residence Ltd v. Yelland* [2001] Ch 143, 158.

<sup>42</sup> Schroff, S., Favale, M., and Bertoni, A., 'The Impossible Quest: Problems with Diligent Search for Orphan Works' (*forthcoming*; manuscript copy on file with the author). About the Dutch law, the authors note as follows: 'The letter of the Directive has been transposed into national law. In fact, the Government Decree with which Diligent Search sources were transposed into Dutch legislation is essentially a translation of the OWD Annex.'

The databases of institutions and national libraries active in the field of cinematographic and audio heritage, in particular the Association of Film Archives, the Federal Archive, the Foundation of German Film Archives, the German Film Institute ([www.filmportal.de](http://www.filmportal.de) database and catalogue), the DEFA Foundation and the Friedrich Wilhelm Murnau Foundation, and the catalogues of the State Libraries in Berlin and Munich.

Similarly, whereas the Directive suggests ‘databases of other relevant associations representing a specific category of rightholders’, the German Annex clarifies as follows: ‘such as associations of film directors, screenwriters, film music composers, composers, theatre publishing houses, theatre and opera associations’.<sup>43</sup> The German Annex also includes a series of directions regarding unpublished works, which are not formally addressed within the Directive guidelines. Diligent search for unpublished material in Germany requires consulting with the current and original owners of the work, the National Registers of Estates, finding aids in the national archives, museum inventory lists, as well as credit agencies and telephone books.<sup>44</sup>

The approach adopted in the UK follows neither the Dutch nor the German model. We discuss it in detail in the next section.

<b>RELEVANT WORK</b>	<b>SOURCES TO BE SEARCHED</b>
1. Published books	Legal deposit, library catalogues and authority files maintained by libraries and other institutions The publishers’ and authors’ associations in the country in question Existing databases and registries, WATCH (Writers, Artists and their Copyright Holders), the ISBN (International Standard Book Number) and databases listing books in print The databases of the relevant collecting societies, including reproduction rights organisations Sources that integrate multiple databases and registries, including VIAF (Virtual International Authority Files) and ARROW (Accessible Registries of Rights information and Orphan Works)
2. Newspapers, magazines, journals and periodicals	The ISSN (International Standard Serial Number) for periodical publications Indexes and catalogues from library holdings and collections Legal deposit The publishers’ associations and the authors’ and journalists’ associations in the country in question The databases of relevant collecting societies including reproduction rights organisations
3. Visual works, including fine art, photography, illustration ... and other such works contained in books, journals, newspapers and magazines or other works	The sources referred to in paragraphs 1 and 2 The databases of the relevant collecting societies, in particular for visual arts, and including reproduction rights organisations The databases of picture agencies, where applicable
4. Audiovisual works and sound recordings	Legal deposit The producers’ associations in the country in question Databases of film or audio heritage institutions and national libraries

<sup>43</sup> Annex (to Article 61a): Sources for a diligent search. These provisions, in translation, are taken from an English language version of the German Copyright Act available here: [https://www.gesetze-im-internet.de/englisch\\_urhg/index.html#gl\\_p0982](https://www.gesetze-im-internet.de/englisch_urhg/index.html#gl_p0982) (accessed: 27 November 2016).

<sup>44</sup> *Ibid.*

	<p>Databases with relevant standards and identifiers such as ISAN (International Standard Audiovisual Number) for audiovisual material, ISWC (International Standard Music Work Code) for musical works and IRSC (International Standard Recording Code) for sound recordings</p> <p>The databases of the relevant collecting societies, in particular for authors, performers, sound recording producers and audiovisual producers</p> <p>Credits and other information appearing on the work's packaging</p> <p>Databases of other relevant associations representing a specific category of rightholders</p>
--	---

**Table 2:** Annex to the Directive with list of relevant sources

### 3.7.3. IPO GUIDANCE ON DILIGENT SEARCH

Within the UK, the sources listed in the Annex to the Directive are replicated verbatim in the *Orphan Works Regulations* (see Schedule ZA1, Part 2) with some additional guidance regarding unpublished works.<sup>45</sup> As with the Directive, these set out the 'minimum' 'appropriate sources' to be consulted for each category of work when carrying out a good faith search. Moreover, the same is true for the *OWLS Regulations*.<sup>46</sup>

In addition, however, the UK Intellectual Property Office has produced three sets of guidelines relating to (i) film, music and sound-related orphan works (September 2016), (ii) literary orphan works (November 2015), and (iii) still visual art orphans (November 2015).<sup>47</sup> These guidelines are 'primarily intended' for those wanting to make an application through OWLS, although each does state that they 'may also be of help to those conducting a diligent search in relation to the EU Directive'.<sup>48</sup> Each set of guidelines is accompanied by a Diligent Search checklist: these checklists set out, simply and unambiguously, the key organisations where inquiries regarding orphan works might be made, although they are not intended to be exhaustive. Taken together, the guidelines and the checklists provide very useful information and signposting when undertaking a diligent search. However, at no point do these guidance documents address either the concept of a 'reasonable' diligent search or a 'good faith' diligent search, or indeed the relationship between the two. The discussion is framed solely in terms of diligence: that is, 'applicants will need to show that their search was indeed diligent,'<sup>49</sup> but not that the diligent search must constitute a *reasonable* search.

Nevertheless, the recommended sources are structured in a way that map onto the Directive Annex, while also providing additional suggestions to help with diligent search.

<sup>45</sup> The Regulations recommend consulting 'Those sources that are listed in paragraphs 1 to 4 above which are appropriate to a relevant work which is unpublished'.

<sup>46</sup> The *OWLS Regulations* set out that the relevant sources to consult as part of a reasonable search for the relevant work must 'as a minimum' include 'any relevant sources listed for that category of work in Part 2 of Schedule ZA1': that is, the same sources listed in the *Orphan Works Regulations*.

<sup>47</sup> Under both the *Orphan Works Regulations* and the *OWLS Regulations* the Intellectual Property Office are empowered to produce on the appropriate sources to be consulted when conducting a diligent search. These guidelines are currently available [here](#). The IPO is committed to reviewing and revising these guidelines as appropriate, and indeed they have already been updated since OWLS was launched.

<sup>48</sup> See, for example, Intellectual Property Office, *Orphan works diligent search guidance for applicants: Literary Works* (November 2015), 1, available here: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/474401/orphan-works-literary-works.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/474401/orphan-works-literary-works.pdf) (accessed: 12 December 2016).

<sup>49</sup> *Ibid.*, 2.

Consider, for example, the list of sources set out in the Annex to the Directive regarding published books:

1. Published books	<p>Legal deposit, library catalogues and authority files maintained by libraries and other institutions</p> <p>The publishers' and authors' associations in the country in question</p> <p>Existing databases and registries, watch (writers, artists and their copyright holders), the ISBN (International Standard Book Number) and databases listing books in print</p> <p>The databases of the relevant collecting societies, including reproduction rights organisations</p> <p>Sources that integrate multiple databases and registries, including VIAF (Virtual International Authority Files) and ARROW (Accessible Registries of Rights information and Orphan Works)</p>
--------------------	--

**Table 3:** List of sources for published books from the Annex to the Directive

In relation to the five different categories of sources listed in the Annex, the IPO guidelines elaborate as follows:

Legal deposit, library catalogues and authority files maintained by libraries and other institutions	<p>Check the British Library catalogue. In addition, applicants might also check the following libraries:</p> <ul style="list-style-type: none"> <li>▪ The Bodleian Library, University of Oxford</li> <li>▪ Cambridge University</li> <li>▪ National Library of Scotland</li> <li>▪ The National Library of Wales</li> </ul>
The publishers' and authors' associations in the country in question	<p>Information on a right holder might be found with:</p> <ul style="list-style-type: none"> <li>▪ The Society of Authors</li> <li>▪ The Writers' Guild of Great Britain</li> <li>▪ The Association of Author's Agents</li> <li>▪ The Publishers' Association</li> </ul>
Existing databases and registries, watch, the ISBN (International Standard Book Number) and databases listing books in print	<p>The watch database might provide contact details for the author, the author's agent or literary executor</p> <p>The ISBN applies to books published after 1967. It identifies the title to which a work is assigned and provides details of the publisher</p>
The databases of the relevant collecting societies, including reproduction rights organisations	<p>Check the databases of all the relevant collecting societies, as an author might only be registered with one. Recommended databases for consultation include:</p> <ul style="list-style-type: none"> <li>▪ Authors' Licensing and Collecting Society</li> <li>▪ The Publishers' Licensing Society</li> <li>▪ The Copyright Licensing Agency</li> <li>▪ The Public Lending Right Register</li> </ul>
Sources that integrate multiple databases and registries, including VIAF and ARROW (Accessible Registries of Rights information and Orphan Works)	<p>There are sources that will hold information from multiple sources. These include VIAF (Virtual International Authority Files) which combines multiple name authority files into a single online computer centre.<sup>50</sup></p>

**Table 4:** Sources from the IPO Guidelines

That is, when conducting a search against the categories of sources listed in the Annex to the Directive, the IPO recommend 16 relevant sources for consideration; for ease of reference, let's call these the *core list of recommended sources* (or the *core sources*). As

<sup>50</sup> Note, although ARROW is expressly referred to within the Annex to the Database, the IPO guidelines make no reference to it.

noted above, the guidelines carry the general caveat that they are ‘primarily intended’ for those planning to make an application through OWLS. But, a common-sense interpretation would be to read these core sources as a *de facto* means by which the government is providing additional guidance on relevant sources under Article 3(2) of the Directive. Put simply, within the UK, these would appear to be the *appropriate* sources to consult when conducting a good faith diligent search for a work that falls within the scope of the Directive. Certainly, to ignore the potential relevance of these sources when conducting a search under the Directive would run the risk of compromising the good faith nature of that search.

Thereafter, the guidelines for published books set out an additional 22 sources that might be helpful in locating the owner of the rights in the work; let’s refer to these as the *additional sources*. These include a mixture of very specific suggestions (for example, the Firms Out of Business database and the Companies House Register) as well as more generic advice (for example, ‘general internet searching’).<sup>51</sup> Again, these are not presented as an authoritative or exhaustive list of sources. But, in presenting a set of additional sources that appear to lie outside the scope of the minimum requirements prescribed by the Annex to the Directive, perhaps we can detect an attempt to differentiate between the standard required of a good faith diligent search (under the Directive) and a reasonable search (under OWLS). Intuitively, logically, if a different standard of care is implied, then the way the IPO have structured and presented the range of sources within their guidelines may represent one way in which this different standard is realised and articulated, albeit tacitly. That is, when conducting a reasonable diligent search to make use of a work in accordance with a state-sponsored licensing scheme greater efforts are required (perhaps, *should* be required) than when conducting a good faith diligent search to enable access to orphan material within the context of a cultural heritage institution delivering on its non-commercial public interest mission. Or, more simply: the standard of reasonableness appears to be set at a higher threshold than that of good faith.

#### **3.7.4. DOES DILIGENT MEAN EXHAUSTIVE?**

The Directive requires that a diligent search be carried out in relation to each work ‘by consulting the appropriate sources for the category of works ... in question’ (Article 3(1)). The sources that are appropriate are to be determined by each Member State, but shall include at least the relevant sources listed in the Annex (Article 3(2)). But must every source listed as a relevant source be consulted in relation to each type of orphan work?, or might an archivist or librarian decide that certain of the relevant sources are in fact irrelevant in the circumstances and still be acting in good faith?

Consider, for example, an archivist conducting a search for the copyright owner of an extract from a newspaper article written by an unnamed author in the mid-20<sup>th</sup> century

---

<sup>51</sup> The various additional categories of sources are listed as follows: Credits and other information appearing on the work; FOB (Firms Out of Business) database; Companies House; The provenance of a works (i.e. where it was found); General internet searching; Records of literary agents; International Standard Text Code (ICTS); Copac; Author and book info database; Poetry library; International Standard Name Identifier (ISNI) database; Books in Print database; Copyright Hub; Academic and scientific databases; Online databases and catalogues; Digitised newspaper archives; Genealogy websites; Wills – search for family members or connections of the author; Archives; Treasury solicitors; Biographical directories online; Other sources identified.

and published in an unknown newspaper. There exists no contextual metadata that might assist in the diligent search. The IPO's Guidelines, as they relate to newspapers, magazines, journals and periodicals, list 40 organisations, registers and other sources that might be of assistance; but should an archivist really consult them all? Should they consult the UK National Centre of the International Standard Serial Number (that is, the British Library at Boston Spa), relevant library indexes and catalogues, legal deposit data, records held by all Publishers' and Authors' Associations, as well as the databases of all relevant collecting societies? If the newspaper article has nothing to do with food, travel or science writing, is there any point in consulting the Association of British Science Writers, the Garden Media Guild, the Guild of Food Writers, Travel Writers UK, or the British Guild of Travel Writers? Is the National Union of Journalists, the Chartered Institute of Journalists or the British Association of Journalists likely to be able to assist in identifying the relevant copyright owner?

For one thing, practically speaking, consulting all relevant sources simply might not be possible. In a study of diligent search requirements in the United Kingdom, the Netherlands and Italy, Favale et al found that, across all three jurisdictions, just over half of the sources that might fall within the scope of the Annex were freely accessible online. The remaining sources could only be consulted offline (that is, on site), only granted partial access to the relevant records, or could be accessed online but only on payment of a fee. On this basis, the authors recommend that a diligent search must be considered as carried out in good faith if 'all relevant freely and easily accessible sources have been consulted'.<sup>52</sup>

We endorse that recommendation but it still begs the question: when is a listed source *genuinely relevant* such that it *ought* to be consulted? It has been suggested that if a Member State has provided guidance on diligent search incorporating a list of relevant sources then *all sources* should be consulted; otherwise, the search cannot be deemed to be diligent. In other words, under the Directive the concept of diligence is simply a proxy for an exhaustive search.<sup>53</sup> We disagree. For example, if presented with a photograph of a building or public work of art, the British Society of Underwater Photographers is unlikely to provide any useful information about who the photographer might be.<sup>54</sup> Or take another example: conducting an ISBN search is recommended in both the Annex to the Directive and the IPO Guidelines for published books. But, the ISBN scheme was only created in 1967 and first introduced in 1970. So, should an ISBN search really be conducted for works that were published before 1970? Would a good faith diligent search be rendered invalid if ISBN was not consulted? That would surely be absurd. Indeed, carrying out futile and irrelevant search activity of this kind would seem to be the very antithesis of diligence, which requires care and conscientiousness in thought and action. It may be that some Member States have imposed a strict obligation to consult all relevant listed sources in their implementation of the Directive. Germany perhaps provides an example. The German Act states that '*at the very least* the sources set out in the Annex [to this Act] *must be consulted*' (emphasis added). That is, consulting each source listed is *necessary* to ensure a good faith diligent search; and yet, this may not be *sufficient* to

---

<sup>52</sup> Favale, M., Schroff, S., and Bertoni, A., *EnDOW Report 1: Requirements for Diligent Search in the United Kingdom, the Netherlands, and Italy* (February 2016), 38, available here: [http://diligentsearch.eu/wp-content/uploads/2016/05/EnDOW\\_Report-1.pdf](http://diligentsearch.eu/wp-content/uploads/2016/05/EnDOW_Report-1.pdf) (accessed: 12 December 2016).

<sup>53</sup> See Schroff, Favale and Bertoni, 'The Impossible Quest'.

<sup>54</sup> Of course, the photograph in question might have been taken by someone who ordinarily specialises in underwater photography. Our thanks to Margaret Haig for this example.



provide legal certainty. Depending on the context, consulting other sources not listed may also be required. We suggest, however, that the language adopted in the German regime goes beyond what it required by the Directive. After all, the Directive sets out sources that are *appropriate* to be consulted when conducting a diligent search, rather than *required* to be consulted *in every instance*.

If the Directive is not to be rendered a complete irrelevance whether for mass digitisation initiatives or otherwise, a more purposive or pragmatic approach to good faith diligent search must be adopted. What constitutes a relevant source should not be unthinkingly predetermined by the category of work, regardless of the nature of the work and the context in which it occurs. Consider again, our example of an extract from a newspaper article by an unnamed author published in an unknown newspaper. It would be unreasonable, indeed irrational, to interpret the concept of a good faith search as requiring an exhaustive engagement with all the relevant sources set out in the IPO Guidelines. Sending email information requests to all listed organisations and professional bodies would amount to little more than an exercise in generating spam, even if they are neither sent nor received in that spirit. And often, recipient organisations will simply not bother to respond. In truth, the realpolitik of a good faith search in relation to such a work is likely to start and end with no more than an internet-based search for the title of the article or for key passages of text. If the search returns nothing meaningful to go on, should more than this be required? We would suggest not. Moreover, our practical experience of conducting reasonable diligent searches under the OWLS regime confirms this view.

### 3.7.5. DILIGENT SEARCH IN PRACTICE

In the next main section of this online resource we present the results of a rights-clearance simulation carried out within the context of the UK orphan works regime (see here). It will be useful, however, to draw upon various examples from that work at this point to illustrate and contextualise some of the preceding analysis on the concept of diligent search.

To explore the nature and demands of the OWLS scheme, an application was made to OWLS featuring five different types of work including a published poem and an original black and white photograph. The author of the poem is Peter Appleton, as recorded by Edwin Morgan in Scrapbook No.12. When carrying out the diligent search for the copyright owner of this work, the project researcher addressed 15 of the recommended sources on the IPO's extended checklist. This included seven of the 16 relevant Annex-related sources, and eight of the 22 additional sources listed in the IPO guidelines. The Society of Authors, the Writers' Guild, and the Authors' Licensing and Collecting Society were unable to help. *Who's Who*, the *Dictionary of National Biography* and the Poetry Library led nowhere. VIAF (the Virtual International Authority File) provided information for other authors named Peter Appleton but not this specific poet.<sup>55</sup> Google also proved to be a dead end. After 90 minutes searching for the owner of the work, over a period of nearly four months, we determined the work to be orphan. The IPO agreed.

---

<sup>55</sup> For details on VIAF, see: <https://viaf.org/> (accessed: 12 December 2016).

For the original photograph (an image of a topless man with his hands behind his head: Image 9) only six sources were considered: the UK Orphan Works Register; the likely provenance of the work;<sup>56</sup> credits and other information appearing on the work [there were no credits or similar information]; web-based search tools for images (Google, PicScout and Tineye); the British Institute of Photographers;<sup>57</sup> the Association of Photographers;<sup>58</sup> and the British Association of Picture Libraries (BAPLA).<sup>59</sup> The total time spent on the search was 25 minutes, over a period of one day. The work was deemed to be an orphan. Again, the IPO agreed.

In both cases the IPO considered the diligent search to be a *reasonable search* of the *relevant sources* for the works in question, and this was despite that each of the recommended Annex-related sources had not been consulted. One assumes that, in the opinion of the IPO, these same efforts would also constitute a good faith search;<sup>60</sup> indeed, arguably even less effort might be required. The key point, however, is that, while the Annex to the Directive sets out a list of ‘minimum’ sources to be consulted when conducting a diligent search, all sources need not be consulted in every case, at least not in the UK. Just because a source is potentially relevant to a category of work, does not mean it will be a relevant source in relation to a specific example of that work. Much will depend on the content of the work and the context in which it is found, as well as the expertise and knowledge-base of the person conducting the search. Diligence does not demand a perfunctory or dogmatic adherence to a boilerplate check-list of sources, however useful and well-crafted. Moreover, the IPO guidelines indicate as much in noting that ‘there is no minimum requirement to be followed in every case’.<sup>61</sup>

---

<sup>56</sup> From the application submitted to the UK IPO: ‘There is similar material throughout the Scrapbooks (all uncredited). I think these are contemporary 1950s body building images, as appeared in Physique magazines. Likely purchased from an unknown magazine. Possibly US based?’

<sup>57</sup> The BIPP responded as follows: ‘Unfortunately, we’ve drawn a blank with this. I will keep the details on file, in case we do discover any information on the image, but at the moment, we’re unable to accredit it to a photographer’.

<sup>58</sup> The AOP replied as follows: ‘I have forwarded your information onto our board of directors and members. I will let you know if it gets any returns.’ No further correspondence was received.

<sup>59</sup> No positive responses were received from any image suppliers notified by BAPLA.

<sup>60</sup> Of course, as a free-standing artistic work the original photograph does not fall within the scope of the Directive.

<sup>61</sup> Intellectual Property Office, *Orphan works diligent search guidance for applicants: Literary Works*, 2.

## 4. DILIGENT SEARCH IN CONTEXT AND PRACTICE

Kerry Patterson, Ronan Deazley and Victoria Stobo

### 4.1. INTRODUCTION

Carrying out a diligent search is a requirement of both the EU Orphan Works Directive<sup>62</sup> and the UK Intellectual Property Office's Orphan Works Licensing Scheme (OWLS). Diligent search is a time-consuming exercise for any digitisation project and the task is even more challenging with the Edwin Morgan scrapbooks due to the number of de-contextualised and partial cuttings and the variety of sources used. In this part of the resource we explore the practical implications of diligent search in relation to both the Directive and OWLS, illustrated by a rights clearance exercise performed on a 10% sample of Scrapbook 12 from Edwin Morgan's series of scrapbooks.

### 4.2. DILIGENT SEARCH IN CONTEXT

This project is the first major UK study concerning the concept of diligent search since the Directive and OWLS came into effect. The costs and challenges of rights clearance activity and of dealing with orphan works have been identified as significant barriers to the digitisation of cultural heritage collections by various studies.<sup>63</sup> We begin by discussing a selection of those studies.

Denise Troll Covey's 2005 research reports the results of Carnegie Mellon University Library's 'feasibility study to determine the likelihood of publishers granting nonexclusive permission to digitize and provide surface Web access to their copyrighted books.'<sup>64</sup> An initial sample of 368 books selected at random was reduced to 277 titles; the reduction was attributable to mis-cataloguing and other reasons, including that a substantial number of works were deemed too complicated to include in the study because of third-party copyright issues (11% of the initial sample).<sup>65</sup> From the sample of 277 items from 209 publishers, a rightholder could not be located for 19% of the books. Of the remaining works, 27% of rightholders did not respond to the requests for

---

<sup>62</sup> DIRECTIVE 2012/28/EU OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 25 October 2012 on certain permitted uses of orphan works; see

[http://ec.europa.eu/internal\\_market/copyright/orphan\\_works/index\\_en.htm#maincontentSec1](http://ec.europa.eu/internal_market/copyright/orphan_works/index_en.htm#maincontentSec1) [accessed 13 December 2016]

<sup>63</sup> See for example: Deazley, R & Stobo V (2013) *Archives and Copyright: Risk and Reform*, available at <https://zenodo.org/record/8373/files/CREATe-Working-Paper-2013-03.pdf> (accessed 22 November 2016); Korn, N (2009) *In from the Cold: An assessment of the scope of 'Orphan Works' and its impact on the delivery of services to the public*, available at

[https://sca.jiscinvolve.org/wp/files/2009/06/sca\\_colltrust\\_orphan\\_works\\_v1-final.pdf](https://sca.jiscinvolve.org/wp/files/2009/06/sca_colltrust_orphan_works_v1-final.pdf) (accessed 22 November 2016); Vuopala, A (2010) *Assessment of the Orphan works issue and Costs for Rights Clearance*, available at [http://www.ace-film.eu/wp-content/uploads/2010/09/Copyright\\_anna\\_report-1.pdf](http://www.ace-film.eu/wp-content/uploads/2010/09/Copyright_anna_report-1.pdf) (accessed 22 November 2016); Akmon, D (2010) *Only with Your Permission: How Rights Holders Respond (or Don't Respond) to Requests to Display Archival Material Online*, available at

<http://link.springer.com/article/10.1007%2Fs10502-010-9116-z> (accessed 14 December 2016)

<sup>64</sup> Covey, D.T. (2005), *Acquiring copyright permission to digitise and provide open access to books*, DLF, Council on Library and Information Resources, Washington DC, available at

<http://www.clir.org/pubs/reports/pub134/reports/pub134/pub134col.pdf> (accessed 13 December 2016), 11

<sup>65</sup> Ibid.

permission to digitise while 30% expressly denied permission. Less than a quarter of rightholders (24%) eventually agreed to allow digitisation.

It is worth noting the length of time and labour required to achieve these results. The average length of time to receive a response from a publisher ranged from an average 101 days for a response of 'Permission granted,' to an average 124 days for a response of 'Permission denied.' Moreover, more than 60% of publishers required a second or third letter before responding; in total, 524 letters were sent.<sup>66</sup> The labour and material costs of searching for rightholders were estimated at \$200 USD per cleared work, not including agreed fees.<sup>67</sup>

Barbara Stratton's 2011 study focussed on rights clearance for 140 books published between 1870 and 2010 and held by the British Library.<sup>68</sup> Directly inspired by Covey's work, one of the report's stated goals was to 'measure and quantify the level of diligent search currently required to undertake mass digitisation of material from the last 140 years,'<sup>69</sup> in addition to identifying copyright status and the proportion of orphan works.

Before any rightholders were contacted, initial information had determined that of the total sample, 57% (80 books) of the books were in copyright, 27% (38 books) were in the public domain and the remaining 16% (22 books) had an unknown copyright status. Therefore rights clearance was investigated for the 73% of books (102 books) not believed to be in the public domain. These figures changed very little following rights research: the number of unknown works dropped by 2% when two books were confirmed to be in copyright and a further two works were confirmed as in the public domain.<sup>70</sup>

Following the period of rights clearance and rights research, it was determined that 43 of the books were orphan works, equating to 31% of the total sample of 140 books. Of these, 15% (21 books) were definitely still in copyright, while 16% (22 books) were orphan works of unknown copyright status.<sup>71</sup> On average, it took four hours to complete a diligent search for each title and more than 450 hours (80% of the project time) were spent on research and identification tasks before a single rightholder was even contacted.<sup>72</sup>

Taken together, Covey and Stratton's studies set out issues common to many digitisation projects, beginning with the 'weeding out' of the most complicated items – from a copyright perspective – at the beginning of the study, as described by Covey. Both studies evidence a high proportion of orphan work material within each of the random samples (31% by Stratton and 19% by Covey), a result that is particularly striking given that both studies concerned published materials for which one might expect more reliable, locatable data concerning rightholders. Time and the other associated costs also are

---

<sup>66</sup> Comprising 278 initial request letters and 246 follow-up letters; Covey, *Acquiring copyright permission to digitise and provide open access to books*, 13.

<sup>67</sup> These are noted to be transaction costs and do not include permissions paid to publishers for the rights to digitise for use online. Permission fees were paid up to \$100 per item.

<sup>68</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, Project Report for British Library/ARROW*, available at: <http://www.arrow-net.eu/sites/default/files/Seeking%20New%20Landscapes.pdf> (accessed 13 December 2016).

<sup>69</sup> *Ibid.*, 4.

<sup>70</sup> *Ibid.*, 33.

<sup>71</sup> *Ibid.*, 37.

<sup>72</sup> *Ibid.*, 51.

significant in each study, with four hours per item for rights clearance activity in Stratton's study in addition to the 80% of project time spent on pre-clearance research. Covey's study produced similar results with an average \$200 cost per item. These figures provide little comfort for any institution contemplating a digitisation initiative of any scale.

The title of Maggie Dickson's case study of the digitisation of the Thomas E. Watson papers, 'Due Diligence, Futile Effort,' sums up the experience of digitising a collection of late nineteenth and early twentieth century manuscripts.<sup>73</sup> A pilot for a larger digitisation project, one of the principal lessons that emerged was that using the methods adopted in the pilot to clear rights on a larger scale would be 'needlessly expensive and futile.'<sup>74</sup> Rather, the study led Dickson to conclude that risk had a central role to play in future digitisation initiatives within the sector: 'If we are willing to calculate and assume some degree of risk and to document our decisions, archives and libraries can move forwards with large-scale digitization.'<sup>75</sup> Indeed, the issue of risk management is impossible to disentangle from any digitisation process, and can be a stumbling block for many institutions.

One further rights clearance study by Dharma Akmon highlights the familiar issue of the complexity of copyright law, but also mentions the lack of case law in this area<sup>76</sup>, making it challenging for cultural heritage institutions to measure risk without examples from other organisations.<sup>77</sup> In the cited case to digitise the Jon Cohen AIDS Research collection, 85% of staff time on the digitisation project was spent on copyright permissions, with the average time of 1 hour 10 minutes per item,<sup>78</sup> figures very similar to those described by Covey and Stratton. Moreover, 'collections with a higher document to copyright holder ratio will probably cost less to usher through the rights process than collections with a low document to copyright holder ratio.'<sup>79</sup> This is certainly of relevance to the Edwin Morgan scrapbooks, where a single page could contain up to 30 individual rightholders.

As with Dickson, Akmon highlights that risks must be taken, emphasising that the unwillingness to take risks in the Cohen project meant that 18% of copyright items could not be displayed due to non-responding rightholders and a further 12% were not displayed due to unidentified rightholders<sup>80</sup>. We discuss risk in more detail in the next section of this resource. For now, suffice to say that even with the orphan works regimes introduced by the Directive and OWLS, managing risk remains a necessary aspect of any diligent search and digitisation activity.

### 4.3. DEFINING DILIGENT SEARCH

In *Orphan Works: The Legal Landscape* we provide a comprehensive analysis and critique of the orphan works regime, including the concept of diligent search, under both the

---

<sup>73</sup> Dickson, M (2010) *Due Diligence, Futile Effort: Copyright and the Digitization of the Thomas E. Watson Papers*, available at <http://americanarchivist.org/doi/10.17723/aarc.73.2.16rh811120280434> (accessed 22 November 2016).

<sup>74</sup> Ibid., 636.

<sup>75</sup> Ibid.

<sup>76</sup> Akmon, 2.

<sup>77</sup> Ibid., 2.

<sup>78</sup> Ibid., 26.

<sup>79</sup> Ibid., 26-27.

<sup>80</sup> Ibid., 27

Directive and OWLS. We will not revisit that commentary here, other than to reiterate that diligent search is fundamental to both the Directive and OWLS. Moreover, within the UK, the implementing legislation defines diligent search differently in relation to the orphan works exception (which requires a diligent search to be carried out ‘in good faith’) and OWLS (which requires the diligent search to be ‘reasonable’). We interpret these standards to mean different things; that is, we consider the concept of a reasonable search conducted under OWLS to set a higher threshold than that of a good faith diligent search conducted under the Directive.

Importantly, the UK Intellectual Property Office (the IPO) has produced published guidelines on conducting diligent searches in relation to (i) film and sound-related orphan works, (ii) literary orphan works, and (iii) still visual art orphans.<sup>81</sup> These guidelines are ‘primarily intended’ for those wanting to make an application through OWLS, although each does state that they ‘may also be of help to those conducting a diligent search in relation to the EU Directive’.<sup>82</sup> In addition, the guidelines are accompanied by ‘Diligent Search checklists’; the checklists set out the key organisations to approach regarding possible orphan works. Taken together, the guidelines and the checklists provide very useful information and signposting when undertaking a diligent search.

Ultimately, however, the nature and demands of a diligent search will depend on the content and context of each project; there are no hard and fast rules that apply in every scenario. As the IPO makes clear in its guidance: there is ‘no set way to conduct a diligent search as this will depend on the information available on the work,’ and ‘there is no minimum requirement to be followed in every case.’<sup>83</sup>

#### 4.4. DATA EXTRACTION AND CATEGORISING WORKS

It was decided to carry out a rights clearance process for a 10% sample of Scrapbook 12, a 30-page long section, informed by the IPO’s official guidelines on diligent search. The first step was a process of data extraction, in which we recorded details for each of the 432 works present in the 380 individual cuttings in the sample.<sup>84</sup> The information recorded included type of material and its completeness (if possible to determine), as well as information on creator, source of publication, date and country of origin. Some of this information was included within the cutting or had been provided by Edwin Morgan as an annotation, such as the date of a newspaper cutting. Other information could be determined based on other contextual information (for example, recognising the font of a

---

<sup>81</sup> Under both the *Orphan Works Regulations* and the *OWLS Regulations* the Intellectual Property Office are empowered to produce on the appropriate sources to be consulted when conducting a diligent search. These guidelines are currently available [here](#). The IPO is committed to reviewing and revising these guidelines as appropriate, and indeed they have already been updated since OWLS was launched.

<sup>82</sup> See, for example, Intellectual Property Office, *Orphan works diligent search guidance for applicants: Literary Works* (November 2015), 1, available here: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/474401/orphan-works-literary-works.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/474401/orphan-works-literary-works.pdf) (accessed: 12 December 2016).

<sup>83</sup> The IPO provide a comprehensive list of resources, with guidelines for each category of orphan work – Film and Sound, Literary Works and Still Visual Art. Each document is 44 to 50 pages long and has been compiled by experts in the relevant sectors.

<sup>84</sup> Some items had multiple instances of copyright e.g. a cutting with text and a photograph or a photograph of an artwork in which the artwork and photographer both have rights.



magazine that was used and named elsewhere in the Scrapbooks) or was found following research by the Project Officer.

This process allowed the Project Officer to distinguish between possible orphan works and those works that did not require any permission to digitise and make available online. The items in the latter category could be described as material that was without, or was not likely to attract copyright protection: items in the public domain, ephemera and cuttings that were deemed to be an insubstantial part of the original work. For these items, no permission was needed to digitise and make the work available online, so it was not necessary to consider the orphan works regime. That said, it was not always easy to decide whether a given cutting is substantial or not without having access to the original work. For the rights clearance process, the Project Officer recorded ‘completeness’ of cuttings, a concept not completely synonymous with ‘substantiality’ but which was nevertheless helpful in making determinations about the need to rely on the orphan works regime or not. In the sample, 50% of items were complete, 35% were incomplete with 15% unknown in that it was not possible to make a definitive determination.<sup>85</sup>

#### 4.5. PREPARING FOR DILIGENT SEARCH: CUTTINGS IN CONTEXT

Prior to starting a diligent search, an applicant should check the UK IPO’s Orphan Works Register and the EUIPO’s Database to ensure that the work in question is not already registered. As the schemes have only been in operation since 2014, the number of works registered with each is still relatively small: 1,950 works of all types on the EUIPO register and 456 on the OWLS database.<sup>86</sup>

As with the digitisation of many collections, and as advised by the OWLS Guidelines, looking to the provenance of the scrapbook gives an idea of the most appropriate sources to check. As Morgan compiled Scrapbook 12 between 1954 and 1960, we began by limiting our date selection for searching to roughly the decade between 1950 and 1960. Morgan does occasionally include older material, such as antique photographs or news cuttings (Images 1 and 2), however, these tend to be evident by their appearance; plus, Morgan often noted the year of non-contemporary material.

Morgan was a voracious consumer of print media. He read English Language and Literature at Glasgow University, but was also familiar with French, Russian, Italian and German; items in these languages appear throughout the scrapbooks. His personal papers are housed in the University of Glasgow Library, however, they do not include a list (of any kind) of the sources Morgan used in creating his scrapbooks.

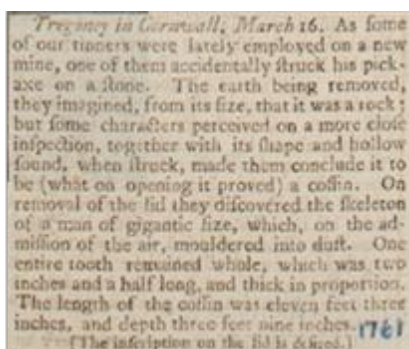


**Image 1:** Original antique photograph

---

<sup>85</sup> For further discussion, see Digitisation and Risk, Section 5: Managing Risk in the Scrapbooks.

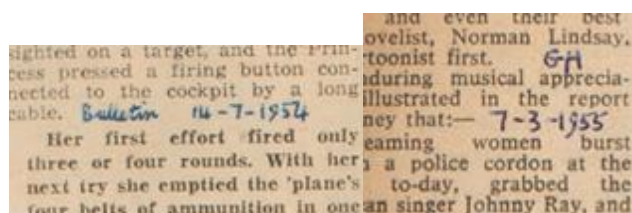
<sup>86</sup> OWLS database available at <https://www.orphanworkslicensing.service.gov.uk/view-register> (accessed 14 December 2016). Not all works registered on OWLS are granted a licence – some are pending and others are eventually withdrawn.



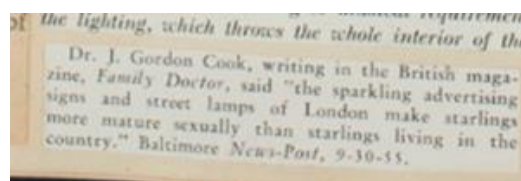
**Image 2:** Original newspaper cutting from 1761

That said, Morgan does provide some clues regarding some of the cuttings used. For longer length text cuttings, he wrote the initials of the newspaper and date on the cutting, or occasionally the full name of a published source (Image 3). Therefore, with knowledge of the types of sources Morgan favoured, gained through wider reading of the scrapbooks and using knowledge of newspapers that would have been available at the time, we were able to determine that 'GH' is *Glasgow Herald* and 'EN' is *Evening News*. Similarly, periodicals like *Life Magazine* or the *Illustrated London News* were often used by Morgan as sources for both images and stories.

The likely source of some uncredited text items can be found by looking at the typescript, the type of paper or even the subject matter, in order to match it up with known sources Morgan frequently used. For example, cuttings from *Doubt* magazine (Image 4), a favourite but niche periodical produced by the Fortean Society, can be identified by its distinctive font and glossy paper, combined with the subject matter of unexplained and paranormal events.



**Image 3:** Cuttings from the Glasgow Herald and Bulletin, annotated by Morgan



**Image 4:** Cutting from Doubt Magazine

#### 4.6. DILIGENT SEARCH: SEARCHING FOR TEXT ONLINE

Text-based searching using online search engines is a familiar and intuitive technique when trying to identify an unknown work, its creator or copyright owner.

For text-based cuttings included within the sample, Google yielded occasional results. Google News Archive has scans of some newspapers but this is not fully text searchable across all publications, although this capability is in development.<sup>87</sup> A list of all available newspapers available through the Google News Archive can be found [here](#).<sup>88</sup> The News Archive digitisation project has ended, with no plans to add any more publications.

Magazines are also available through Google, as part of the Google Books section (for a complete list see [here](#)<sup>89</sup>) but, like the newspapers, only some are text searchable. For example, a search of text from *Life Magazine* found the source article, while text from the *Glasgow Herald* newspaper did not. Another known source used by Morgan, the

<sup>87</sup> For further details, see: [https://en.wikipedia.org/wiki/Google\\_News\\_Archive](https://en.wikipedia.org/wiki/Google_News_Archive) (accessed: 21 December 2016).

<sup>88</sup> See: <https://news.google.com/newspapers> (accessed: 21 December 2016).

<sup>89</sup> <https://books.google.co.uk/books/magazines/language/en> (accessed 24 February 2017)



*Illustrated London News*, is fully text searchable for registered users of the Gale News Vault historical newspaper archive.<sup>90</sup>

#### 4.7. DILIGENT SEARCH: USING CONTENT BASED IMAGE RETRIEVAL (CBIR)<sup>91</sup>

While text-based searching using online search engines will be a familiar concept to researchers, searching for images by using images, is a more recent development. When faced with an image with no caption or clue to its context, Content Based Image Retrieval (CBIR) – also known as reverse image search and search by visual similarity<sup>92</sup> or image retrieval and image mining<sup>93</sup> – is an attractive and easy-to-use research option. It is one which would align well with the diligent search needs of the type of material (de-contextualised images) used in this project. As it is a more recently-developed and lesser used search technique it bears closer examination here.

When a user searches for images using a text-based query, search systems use metadata from the image, text surrounding the image and text in hyperlinks linking to the image,<sup>94</sup> to identify relevant images. CBIR works by indexing images based on their visual content ‘such as colors, textures, shapes and regions’<sup>95</sup> to match them with images available online. Typically, a user can upload their own image or use the link for an existing online image as the source image for a search. CBIR technology has traditionally been used by commercial organisations or photographers to identify unauthorised use of their images, but is now being used more widely. The IPO acknowledged the potential usefulness of these tools by including web-based search tools for images in its diligent search guidelines. The IPO guidelines only refer to TinEye ([www.tineye.com](http://www.tineye.com)) and PicScout ([www.picscout.com](http://www.picscout.com)) but other sites are available, such as Image Raider ([www.imageraider.com](http://www.imageraider.com)), and an image search function is also embedded in various web browsers. For example, Google has offered a reverse image search function since June 2011, allowing you to upload an image to be compared to visually similar images. Similarly, Bing also offers an Image Match function.

TinEye is free and can be used without registration while PicScout offers a free three-month trial period following registration, features that likely influenced their inclusion in

---

<sup>90</sup> See: <http://gale.cengage.co.uk/product-highlights/history/illustrated-london-news.aspx> (accessed: 21 December 2016).

<sup>91</sup> This section is based on an extract from Patterson, K (2016), *Can I Just Google That? Orphan Works and Image Recognition Tools*, in Andrea Wallace and Ronan Deazley, eds, *Display At Your Own Risk: An experimental exhibition of digital cultural heritage*, 2016, <http://displayatyourownrisk.org/patterson/> (accessed 22 November 2016)

<sup>92</sup> Marques, O (2016), *Visual Information Retrieval: The State of the Art*, IT Professional, vol. 18, no. , pp. 7-9, July-Aug. 2016, doi:10.1109/MITP.2016.70, available at <https://www.computer.org/csdl/mags/it/2016/04/mit2016040007-abs.html> (accessed 14 December 2016)

<sup>93</sup> Girija O K & M Sudheep Elayidom (2015), *Overview of Image Retrieval Techniques*, International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE), Vol.4, Special Issue 1, June 2015 , available at <http://www.ijarcce.com/upload/2015/si/icrtcc-15/IJARCCE%2019.pdf> (accessed 16 November 2016), Section 1

<sup>94</sup> Niuwenhuysen, P (2013), *Search by Image through the WWW: an Additional Tool for Information Retrieval*, published in proceedings of the international conference on Asia-Pacific Library and Information Education and Practices = A-LIEP 2013 <http://megaslides.es/doc/4053174/search-by-image-through-the-www--paul-nieuwenhuysen> (accessed 21 November 2016)

<sup>95</sup> Niuwenhuysen, P (2013), *Search by Image through the WWW: an Additional Tool for Information Retrieval*

the IPO's Guidelines.<sup>96</sup> For this reason, they are attractive to the user who is only searching a few images and doesn't wish to sign up to a site or have to pay. TinEye is free for non-commercial users and includes extensions that allow for easy searching in a web browser toolbar. The PicScout Platform is aimed at commercial users. Their search tool is designed to 'enable image buyers to identify and license the images they'd like to use,' and they have '200 million owner-contributed image fingerprints.'<sup>97</sup> As a subsidiary of Getty Images, PicScout would seem an obvious choice when searching for the type of commercial photography that features so heavily in the scrapbooks.

Image Raider relies on Google, Bing and Yanex to get results. It offers a long-term image monitoring service and allows the user to run multiple searches concurrently, features attractive to photographers who wish to monitor potential copyright violations of their work. It uses a credit model, where users can purchase credits or earn credits by tweeting about the site.

Cultural institutions engaging in diligent search activity will rightly be concerned about copyright and image security when uploading images from their collection, and this will no doubt influence their choice of search tool. On this issue, different CBIR systems adopt different approaches. For example, Google's Help Forum states: 'When you search using an image, any images or URLs that you upload will be stored by Google. Google only uses these images and URLs to make our products and services better.'<sup>98</sup> This vague statement will certainly be discouraging to some potential users of the service, who would not want their images retained by Google, particularly in the case of a mass digitisation project. However, as the world's most popular search engine with a global share of 75.2%,<sup>99</sup> it seems a problematic omission from the IPO's guidelines. Would a returning rightholder be satisfied that a diligent search had been carried out without the use of Google?

In contrast, TinEye have a clearer, more acceptable policy: 'Images uploaded to TinEye are not added to the search index, nor are they made accessible to other users. Copyright for all images submitted to TinEye remains with the original owner/author.'<sup>100</sup> Search images submitted to TinEye by unregistered users are automatically discarded after 72 hours, and links to these searches will stop working after 72 hours, unless a registered user happens to save the same image.

Bing's privacy statement does not specifically mention what happens to images,<sup>101</sup> and the Project Officer was unable to find information relating to this issue on PicScout or Image Raider.

---

<sup>96</sup> PicScout was able to be used for free and without registration at the time the UK IPO's Guidelines were last reviewed (September 2016) and also at the time the article on which this section is based, was published (June 2016)

<sup>97</sup> The PicScout FAQs from which this information was taken are now no longer available (checked 16 December 2016) but previously found at <http://www.picscout.com/about-us/faqs/> (accessed 8 April 2016)

<sup>98</sup> Google Search Help available at <https://support.google.com/websearch/answer/1325808?hl=en> (accessed 14 December 2016)

<sup>99</sup> Desktop search engine market share statistics available at <https://www.netmarketshare.com/search-engine-market-share.aspx?qprid=4&qpcustomd=0> (accessed 14 December 2016)

<sup>100</sup> TinEye FAQs on image uploading available at <http://www.tineye.com/faq#uploading> (accessed 14 December 2016)

<sup>101</sup> Bing details available as part of Microsoft's privacy statement available at <https://privacy.microsoft.com/en-gb/privacystatement/> (accessed 14 December 2016)

So how useful are these search tools? The two IPO recommended tools were tested alongside Google Images, as the leading search engine. The results, when searching for orphan images from the scrapbooks, were variable, especially when dealing with partial or cropped images. Within the scrapbooks, Morgan often cropped down images from their original state in newspapers, magazines and books. These irregular-shaped items tend to decrease the likelihood of an uploaded image search yielding beneficial results, although identification of partial images is still possible.



**Image 5:** Villa Doria Pamphili, Rome (Souvenir d'une Villa)

One example of a successful search is this image of an oil painting (Image 5), taken from Scrapbook 12. Despite the fact that Morgan had cropped the image, Google Images and TinEye were both able to point to sources to identify the cutting showing the centre third section of the oil painting *Villa Doria Pamphili, Rome (Souvenir d'une Villa)* 1838-39 by Alexandre Gabriel Decamps (1803-60). Naturally, the key to the success of the search tools is the fact that Decamps' painting can be found on multiple websites. The more ubiquitous the image is online, the greater the chance of identifying it. PicScout, however, was unable to identify the painting.

In general, the kind of hit rate you can expect to get from image search will vary. In a further test of these tools, two pages were selected at random from Scrapbook 12, incorporating a total of 14 viable images. From this small sample, Google provided the best results (identifying two images), followed by TinEye (one), with PicScout unable to provide anything at all. An example of one of the images found by both TinEye and Google is a crop of an animal that appears on a tapestry from the Middle Ages (Images 6 and 7).



**Image 6:** Cropped image identified as a section from a tapestry by Google and TinEye



**Image 7:** Original tapestry from which cropped image taken  
([www.tchevalier.com/unicorn/tapestries/sight.html](http://www.tchevalier.com/unicorn/tapestries/sight.html))

These findings echo the results of other researchers. Kirton and Terras compared results from TinEye and Google Images in their Reverse Image Lookup study investigating re-use of images from The National Gallery, London.<sup>102</sup> They note that Google produced a 'significantly larger number of results,' giving the example of the painting *Whistlejacket* by George Stubbs which had 109 results on TinEye and 271 on Google.<sup>103</sup> Paul Nieuwenhuysen's study offers similar findings, taking as an example images a set of publicly-available photographs available on a central university server for several years.<sup>104</sup> Nieuwenhuysen submitted nine images to TinEye and Google: Google found six but TinEye found only three.<sup>105</sup> Further searches led to the conclusion that 'search by image for duplicate images functions with an efficiency that is highly variable from case to case.'<sup>106</sup>

Kirton and Terras note that in addition to Google having a much larger database than TinEye (over 10 billion images versus TinEye's just over two billion at the time of their study), they found that Google's database was more up to date.<sup>107</sup> Moreover, as there was little crossover between the results found by Google and TinEye, each system processed the search and presented results in a different way, with TinEye presenting results in a more straightforward and transparent manner and Google constantly readjusting search results as the user moves through pages.<sup>108</sup> These differences help to explain the disparity in search results, and Kirton and Terras argue that the vastly different results 'undermine the use of these tools for anything but a guide as to how to understand image reuse.'<sup>109</sup> Nieuwenhuysen's overview of online image searching reached similar conclusions about the lack of transparency of CBIR systems and the lack of knowledge about their functionalities.<sup>110</sup>

Reverse image search technology can certainly be beneficial to cultural heritage institutions, if used with care. Image recognition tools have a role to play in helping identify any very 'obvious' works which are still in copyright, i.e., those which are usually by well-known creators. Works that are in the public domain are perhaps more likely to be found as they present less risk for users to use and share online; in turn, they are more likely to feature on multiple websites increasing the likelihood of detection. Of course, simply finding an image may not answer the copyright questions you have about the work, but it is a starting point. One example of a useful outcome from the scrapbook sample came from the image search of an advert that originally featured in *The New Statesman*. One result identified the particular issue in which it originally featured as containing spoof publisher adverts and in-jokes, which was not evident when the advert was removed from that context. This demonstrates an additional use that image recognition

---

<sup>102</sup> Kirton, I and Terras, M (2013), *Where Do Images of Art Go Once They Go Online? A Reverse Image Lookup Study to Assess the Dissemination of Digitized Cultural Heritage*. In *Museums and the Web 2013*, N. Proctor & R. Cherry (eds). Silver Spring, MD: Museums and the Web. Published March 7, 2013. Available at <http://mw2013.museumsandtheweb.com/paper/where-do-images-of-art-go-once-they-go-online-a-reverse-image-lookup-study-to-assess-the-dissemination-of-digitized-cultural-heritage/> (accessed 14 December 2016)

<sup>103</sup> Ibid.

<sup>104</sup> Nieuwenhuysen (2013).

<sup>105</sup> Ibid.

<sup>106</sup> Ibid.

<sup>107</sup> Kirton and Terras (2013), section: 'TinEye versus Google Images Search'.

<sup>108</sup> Ibid.

<sup>109</sup> Ibid.

<sup>110</sup> Nieuwenhuysen (2013).

tools might offer beyond the identification of a possible rightholder: they can help cultural heritage organisations contextualise and research the material within their collections.

There are also practical considerations to bear in mind when using these tools. Preparing images to upload for search may involve considerable effort that is not scalable when engaging in a mass digitisation project. For example, we estimate that all 16 volumes of the Edwin Morgan scrapbooks contain 41,472 orphan works requiring diligent search and, from the sample, 56% of items are artworks and photographs from published sources, leading to a large amount of potential administration time for carrying out checks using CBIR tools. Moreover, institutions should also consider which tool is the most appropriate tool for its needs. Although Google was the most likely to provide results, individuals and organisations will have understandable reservations about uploading large amounts of images to Google due to their policy of storing images for their own use. Equally, other sites with unclear or unstated policies about image storage may also be an unattractive choice for users.

Ultimately, in the case of the scrapbooks, the nature of the de-contextualised works means that in some cases, CBIR systems form a significant way of conducting an-IPO approved diligent search. The technology is still relatively new and results are variable; it should be used with this in mind. However, as the technology is continually developing and improving, it seems likely that the usefulness of image recognition tools for cultural heritage institutions engaging in digitisation and rights clearance activities will only increase in the future.

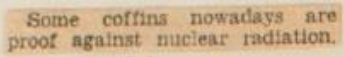
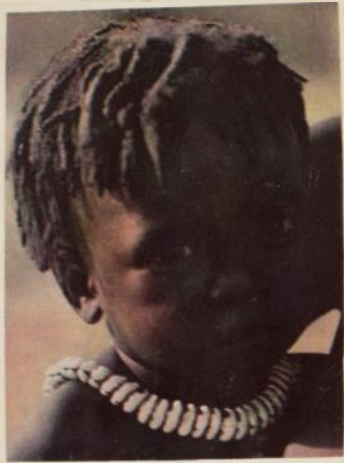


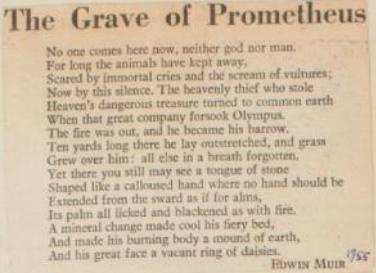
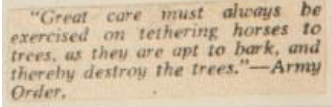
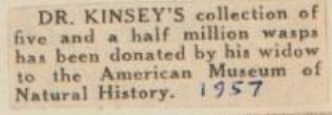


#### 4.8. DILIGENT SEARCH EXAMPLE: PAGE FROM SCRAPBOOK 12

In this section we present details of the search activity carried out in relation to one complete page from our sample. Table 5 provides narrative detail of the search history for each cutting, as well as the time spent on each search, as well as the total period of time over which each search was conducted. We consider each of these examples to constitute a good faith diligent search.








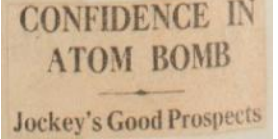
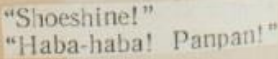
Image 8: Page 2245 from Scrapbook 12

IMAGE	SEARCH HISTORY	TIME SPENT	TIME PERIOD
	<p>Tried web searches for a match to sentence in cutting. No matches.</p>	<p>10 mins</p>	<p>1 day</p>
	<p>Cropped image from page so it could be checked by Image Recognition Tools. PicScout and TinEye did not recognise, Google Images recognised it as part of a cover from a 1954 French magazine called Réalités from May 1954 with title strapline 'le village le plus secret de l'afrique'. Magazine not available online or in the Glasgow University library. World Cat showed that it is available in the <a href="#">National Library of Scotland in Edinburgh</a>. Further online research showed that an English edition of the magazine was also produced. Searched for images from a couple of photographers who were named as frequent contributors - Jean-Philippe Charbonnier and Édouard Boubat. In an image search filtered by colour images, none similar came up. No results for image search of full cover of magazine. Time spent: 40 mins.</p> <p>Further web research and looking at names of photo agencies to approach – decided to try an orphan works search through BAPLA (25 mins). Emails with Getty and Camera Press. Reply from LAT Photo (negative) Camera Press (negative) PA Images Getty (negative) Lee Miller Archive (negative) (35 mins) Time spent: 1 hour.</p> <p>A copy of the magazine was located in the National Library of Scotland in Edinburgh. After applying online for a library card, a letter was delivered to my home address with the card number so the magazine could be ordered to the reading room for my arrival at the library (36 mins). Following a return train trip to Edinburgh (3 hours 40 mins for travel time and time spent at library), I did not find the photograph I was looking for in the magazine. The May 1954 issue of the magazine in the Library had a different cover. It was the English version so it is possible that Morgan used the French version and it may have had a different cover. Or, it came from a different issue of the magazine altogether (I checked January to June 1954). On the positive side, I did discover several other images from the magazines that were used in the Scrapbook.</p> <p>Train fare: £24.90</p>	<p>5 hours 56 mins</p>	<p>Dec.2015 – Sep.2016</p>

 <p><b>The Grave of Prometheus</b>  No one comes here now, neither god nor man.  For long the animals have kept away,  Scared by immortal cries and the scream of vultures;  Now by this silence. The heavenly thief who stole  Heaven's dangerous treasure turned to common earth  When that great company forsook Olympus.  The fire was out, and he became his barrow.  Ten yards long there he lay outstretched, and grass  Grew over him: all else in a breath forgotten.  Yet there you still may see a tongue of stone  Shaped like a calloused hand where no hand should be  Extended from the sward as if for alms,  Its palm all licked and blackened as with fire.  A mineral change made cool his fiery bed,  And made his burning body a mound of earth,  And his great face a vacant ring of daisies.  EDWIN MUIR 1957</p>	<p>After discovering dates for Edwin Muir were 1887-1959, I found him online on the <a href="#">Faber website</a>. Initial web research and then online application for permissions, jointly with another work (40 mins) No response after stated 6-8 week reply period. Follow up correspondence (10 mins) but further follow up was required several month later as some questions remained unanswered and issue took around seven months to be resolved. Finally, Faber concluded that, 'I'm afraid we don't have contact information for the Estate of Edwin Muir. We can, however, suggest that you look into Fair Dealing in case you feel that the usage falls under that particular remit.' (10 mins).</p>	<p>60 mins</p>	<p>Nov. 2015 – Aug. 2016</p>
 <p><i>"Great care must always be exercised on tethering horses to trees, as they are apt to bark, and thereby destroy the trees."—Army Order.</i></p>	<p>Tried web searches for a match to the sentence. No matches.</p>	<p>10 mins</p>	<p>1 day</p>
 <p>DR. KINSEY'S collection of five and a half million wasps has been donated by his widow to the American Museum of Natural History. 1957</p>	<p>Tried a web searches for a match to the sentence. No matches.</p>	<p>10 mins</p>	<p>1 day</p>
	<p>Decided item could be deemed not substantial enough to be covered by copyright law. It is cut from a larger image and it is hard to determine the subject of the photograph.</p>	<p>5 mins</p>	<p>1 day</p>
	<p>Cropped image from page but there were no image search matches on Google, TinEye or Picscout.</p>	<p>15 mins</p>	<p>1 day</p>



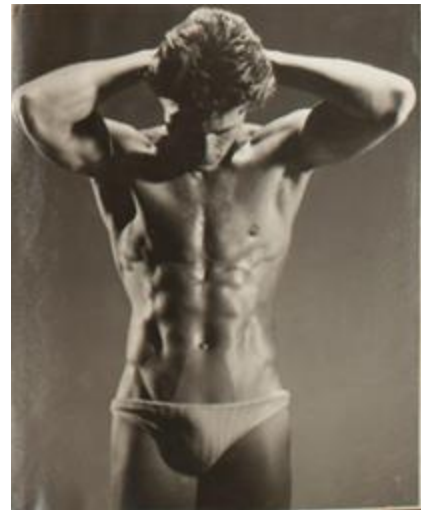
	<p>Cropped image from page but there were no image search matches on Google, TinEye or Picscout.</p>	<p>15 mins</p>	<p>1 day</p>
	<p>Contact with Newsweek over 3 month period via emails, web contact form and finally via Twitter, with difficulty determining who was able to answer the question on rights. Finally resolved, Newsweek granted rights for us to use the text, they are unable to grant image rights as images are credited to the US Airforce (60 mins) Works created by the US Government are in public domain. Not completely clear if this includes the forces and the US Government may assert copyright for US Gov works used outside of the US: <a href="https://www.usa.gov/government-works">https://www.usa.gov/government-works</a>. Spent time researching US Airforce copyright and general US copyright, emailed military contact details found online over a period of three months but no response. Decided to assume that images were in the public domain and use (70 mins).</p>	<p>2 hrs 10 mins</p>	<p>Nov.2015 – Jan. 2016</p>
	<p>As with the cutting above, decided that credit to US Airforce meant that item was in the public domain and could be used. The caption was deemed to be sufficiently factual so as to not to attract copyright protected.</p>	<p>10 mins</p>	<p>1 day</p>

	<p>Cropped image from page but there were no image search matches on Google, TinEye or Picscout.</p>	<p>15 mins</p>	<p>1 day</p>
	<p>Cropped image from page but there were no image search matches on Google, TinEye or Picscout.</p>	<p>15 mins</p>	<p>1 day</p>
	<p>Tried web searches for a match to words in cutting. No matches</p>	<p>10 mins</p>	<p>1 day</p>
	<p>Tried web searches for a match to words in cutting. No matches</p>	<p>10 mins</p>	<p>1 day</p>

**Table 5:** Details of diligent search activity for p.2245

#### 4.9. SELECTING WORKS FOR AN OWLS APPLICATION

To explore the parameters of OWLS, we made an application using different types of work: two cartoons, a poem, a text cutting from a magazine and an original black and white photograph. Predominantly, these were works where a significant amount of time had been spent in trying to find the rightholder, as the name of a potential rightholder was included in four of the five items (see Table 6). The names of the cartoonists and poet were provided on the cuttings and although the name of the author of the magazine text was unknown, the source publication was given. A diligent search was conducted for each of these items relying on the IPO guidelines.



**Image 9:** Original black and white photograph submitted for OWLS application

The remaining item was an original photographic work (Image 9), rather than one cut from a magazine or other source. We believe it to be a purchased black and white studio photograph, contemporary with the creation of the scrapbooks; however, Morgan provides absolutely no information about its origin. With this studio photograph, we were interested to explore the IPO's response to the use of image recognition tools as the primary means of diligent search in an application made to the OWLS scheme. Although they are an IPO-approved method for diligent search, the results can be variable, as previously discussed. Google, PicScout and TinEye were used to search for the photograph, with no results; an application was then submitted to the OWLS scheme on the basis of just those three searches. The IPO's response was that the requirements of the scheme would be satisfied by a further search of three additional sources: the Association of Photographers (AOP), British Association of Picture Libraries (BAPLA) and British Institute of Professional Photographers (BIPP). This involved sending an email to each contact and did not result in identification of the work: the AOP undertook to forward details on to their members and Board of Directors and notify us of identification was made but no further contact was received, BAPLA sent out details to their image suppliers but there was no positive response and BIPP could not identify but would keep the details on file. The total time spent on the search was 25 minutes, over a period of one day. The work was deemed to be an orphan. Again, the IPO agreed.

This result should be encouraging to cultural heritage institutions who intend to apply to OWLS, to know that a diligent search carried out using CBIR tools can form a significant part of their application.<sup>111</sup>

#### 4.10. MAKING AN OWLS APPLICATION <sup>112</sup>

This section explores the mechanics of the diligent search and application process for OWLS, focussing on the cartoon by Paton, looking at the diligent search undertaken and the process from making the application to the licence being granted. Following the

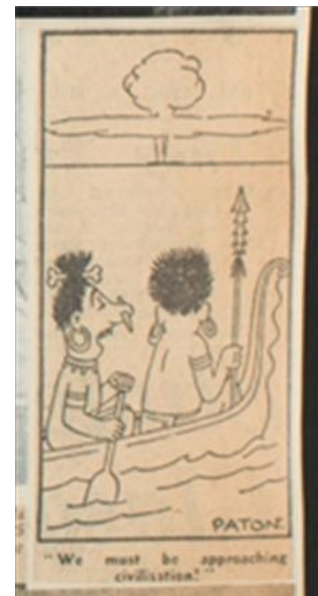
---

<sup>111</sup> This section is based on an extract from Patterson, K (2016), *Can I Just Google That? Orphan Works and Image Recognition Tools*, section IRTS and Diligent Search

<sup>112</sup> Section based on Stobo, V., Patterson, K., and Erickson, K. (2017) Forthcoming.

guidance provided by the IPO, the Project Officer undertook a diligent search for the rightholder of this cartoon by using official sources listed in the guidelines,<sup>113</sup> in addition to other sources as appropriate. While comprehensive, the guidelines do not claim to be complete, but provide ‘details on the relevant sources that applicants must consult and provides a non exhaustive list of additional sources.’<sup>114</sup> Thirty minutes of the diligent search consisted of general web research, including a search of the existing UK orphan works registry and contacting the British Cartoon Archive (BCA) for information on Paton. The BCA replied after a number of days with a web-link to a different cartoon by Paton that had appeared in the magazine Parade, but could provide no other information. The BCA are not listed on the guidelines but were identified during web searching as a possible source of relevant information.

The IPO guidelines for Still Visual Art list eight ‘Illustration Associations’ which were of particular note for this type of work. Of these, four were disregarded immediately as not relevant, as they related specifically to railways, sculpture, architecture and medical illustrations. The other four were the Professional Cartoonists Organisation (PCO), the Comic Creators Guild (CCG), the Cartoonists Club of Great Britain (CCGB) and the Association of Illustrators (AOI). The CCG is primarily concerned with strip cartoons so was not appropriate. Paton was not a listed member on the sites of either the PCO or the CCGB, and the PCO were unable able to help with research enquiries. An additional step taken beyond the boundaries of the guidelines was to post on the forum of the CCGB website, which is reasonably active. Twenty-five minutes spent registering and engaging with the CCGB forum resulted in a reply but it did not identify the artist.<sup>115</sup> The Project Officer spent fifteen minutes emailing both the AOI and Punch (not included in the guideline but a known source of cartoons of this type) but received a non-response and negative response, respectively. In total, one hour and 10 minutes was spent on diligent search for the item, with no positive results.



**Image 10:** Cartoon by Paton submitted for OWLS application

Additional time was spent on the administrative task of applying for a licence to make use of the orphan work. The OWLS application process takes place entirely online. It requires comprehensive information about the work but for the purposes of this section, and in the interest of brevity, we provide a condensed account of the process only.<sup>116</sup>

In completing an application, the first problematic issue encountered concerned providing a title for the work: where a work has no obvious title, how can applicants be assured that they are choosing a title which would enable the work to easily be found by other users? For example, a portrait photograph of a soldier could be titled as ‘Portrait of

<sup>113</sup> UK IPO Guidelines available at <https://www.gov.uk/government/publications/orphan-works-diligent-search-guidance-for-applicants> (accessed 22 November 2016)

<sup>114</sup> As found in the introductions to the Guidance relating to still visual art, film and sound, and literary works: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/474401/orphan-works-literary-works.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/474401/orphan-works-literary-works.pdf) (accessed 22 November 2016)

<sup>115</sup> The reply is available at: [http://www.ccg.org.uk/q\\_and\\_a\\_forum/simpleforum\\_pro.cgi?fid=01](http://www.ccg.org.uk/q_and_a_forum/simpleforum_pro.cgi?fid=01) (accessed 14 December 2016)

<sup>116</sup> See appendix in Stobo, V., Patterson, K., and Erickson, K. (2017) Forthcoming

a soldier’, or by someone with knowledge of regimental badges as ‘Portrait of a soldier from X regiment,’ a much more specific title. For most published sources this is unlikely to cause significant problems, although there could potentially be issues with items that are published under different titles in different jurisdictions. It is possible to note that the item has no title, and to provide a description instead. Licensees should be aware that rightholders and other users might search the register periodically, so any title given for a work should include information likely to be used as part of a keyword search. If titles given to works in archive or museum catalogues are unlikely to satisfy this requirement, some consideration should be given to the time and effort that devising appropriate titles for the application process might generate at this stage of the overall rights management process.<sup>117</sup>

During the application, it was also necessary to make a number of assumptions about the work in order to proceed. The most difficult assumption to make was whether to identify Paton, the cartoonist that created the work, as the rightholder in the work. We know nothing about the publication the work was taken from, and nothing about Paton. It could be the case that the publisher holds the rights to the work. Indeed, this was a recurring issue for the scrapbooks as whole given the huge amount of the material contained in the volumes taken from newspapers published in the 1950s. The work of freelance journalists is specifically mentioned in both the 1911 and the 1956 Copyright Act, with copyright being retained by the writer when working in a freelance capacity.<sup>118</sup> However, without access to employment records, it is difficult to gauge whether the journalist (or in this case, cartoonist) was working under a contract of employment or whether they were a freelance worker which would impact who owned the rights in question.

One also has to indicate whether the application relates to a commercial or non-commercial purpose. On this point, the definition of commercial use employed by the IPO is worth considering. According to the IPO, their definition ‘has been chosen to reflect the practice of licensors. For the avoidance of doubt, it is not intended as a definition in UK or European copyright legislation’. They continue:

Commercial use covers any uses (including by individuals as well as organisations) that make money from the work – such as selling copies of the work or charging directly for access to it. As well as activities that generate revenue, such as merchandising or selling copies of a publication, commercial use would also cover any other uses that are commercial in nature, such as any use in commercial advertising, marketing or promotion activities. This applies equally to not-for-profit organisations.<sup>119</sup>

---

<sup>117</sup> Archivists, librarians and curators are generally very adept at creating titles for works in their collections, through cataloguing, but it may often be that a title in a catalogue isn’t specific enough - they may be relying on an identifying number rather than a descriptive title. For example, to identify individual cuttings within the scrapbooks, the project officer had to devise a numbering scheme. This meant extra time had to be allotted to the creation of descriptive titles when the applications were submitted.

<sup>118</sup> See the Copyright 1956 Act Part, 1 s.4(2) available at [http://www.legislation.gov.uk/ukpga/1956/74/pdfs/ukpga\\_19560074\\_en.pdf](http://www.legislation.gov.uk/ukpga/1956/74/pdfs/ukpga_19560074_en.pdf) (accessed 14 December 2016), and also the 1911 Copyright Act available at [http://www.legislation.gov.uk/ukpga/1911/46/pdfs/ukpga\\_19110046\\_en.pdf](http://www.legislation.gov.uk/ukpga/1911/46/pdfs/ukpga_19110046_en.pdf) (accessed 14 December 2016).

<sup>119</sup> Intellectual Property Office (2015) *Orphan Works Licensing Scheme overview for applicants*, 2.34 - 2.38, available at

The IPO's definition of commercial use could affect even very small scale endeavours. Take the example of a local history society wishing to use four or five orphan images in a booklet with a small print-run. Even if they don't intend to profit from the publication, but only want to cover printing costs by charging a small fee, the society would still be charged the same commercial rate as a much larger publisher to make use of the orphan works. Indeed, the cost of the licence could make the planned publication unfeasible, unless they are willing to raise the price per copy.<sup>120</sup>

A final consideration for institutional users of the scheme is the fact that debit and credit cards are the only accepted method of payment, which is due in two stages: an application fee to start the process and a licence fee once the IPO accepts the application and grants a licence. For smaller applications containing only a few works, it may be difficult for local authorities and large institutions to pay a fee of as little as ten pence (£0.10) for the non-commercial use of a single work, without the option of raising an invoice or paying an additional card-handling charge.

The application process, in its entirety, took one hour and 10 minutes. Diligent search took the same amount of time, providing a total of two hours and 20 minutes spent on the single Paton cartoon. Table 6 presents these results alongside the other works for which an application was made to the OWLS scheme. A licence was granted for the non-commercial use of all five works.

---

[https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/518251/Orphan\\_Works\\_Licensing\\_Scheme\\_Overview\\_for\\_Applicants.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/518251/Orphan_Works_Licensing_Scheme_Overview_for_Applicants.pdf) (accessed 17 November 2016).

<sup>120</sup> This mirrors the EU Directive on Re-use of Public Sector Information, which requires that all users be charged the same amount for re-using public sector information.



**Table 6:** Results of diligent search for works on UK IPO application

NAME OF WORK	TIME SPENT ON DILIGENT SEARCH (DS)	TIME COMPLETING DS CHECKLIST	TIME ON APPLICATION	TIME ON DS REQUESTED BY IPO	TOTAL TIME	APPLICATION FEE (AVERAGE)	COMMERCIAL COST	TERMS OF USE	NON-COMMERCIAL COST
Paton Cartoon	70 mins	20 mins	50 mins	10 mins	140 min	£2.67	£75.12 +VAT	Reproduction in a published book (not a textbook), Inside the book, 1/16 page or less, one year or less, 1000 copies or less	£0.10
Appleton Poem	85 mins	15 mins	10 mins	5 mins	110 min	£2.67	£96.43 + VAT	Reproduced as a book, up to 1,000 copies, sale price of £10	£0.10
Giovannetti Cartoon	100 mins	25 mins	20 mins	15 mins	160 min	£2.67	£89.15 + VAT	Reproduction in a published book (not a textbook), Inside the book, 1/4 page or less, one year or less, 1000 copies or less	£0.10
Doubt Magazine cutting	180 mins (60 online/120 library)	15 mins	20 mins	15 mins	230 min	£2.67	£56.25 + VAT	Reproduction in a published book (not a textbook), Inside the book, 100 words or less, one year or less, 1000 copies or less	£0.10
B&W photo of man	10 mins	15 mins	10 mins	15 mins	50 min	£2.67	£77.13 + VAT	Reproduction in a published book (not a textbook), Inside the book, 1/8 page or less, one year or less, 1000 copies or less	£0.10
<b>Total:</b>								<b>690 min</b>	
<b>Average:</b>								<b>138 min</b>	

#### 4.11. SELECTING WORKS FOR AN EUIPO REGISTER APPLICATION

Selecting works for the EUIPO register, there is a more limited choice as artistic works are not eligible. That is, works such as the original photograph registered under OWLS are ineligible. However, embedded artistic works do fall within the scope of the Directive,<sup>121</sup> and these form the majority of the artistic works contained in the scrapbooks. And yet, within the context of the scrapbooks these embedded artistic works pose their own problems. In their original context – the newspaper or magazine from which they were taken – these cuttings are indeed embedded. But do they cease to qualify as an embedded work when neatly cut from those newspapers and magazines to be pasted into a scrapbook alongside other freestanding artistic works such as the black and white photograph previously discussed? Would a returning rightholder regard the cutting of what was an embedded work to be no longer embedded? Possibly not. Indeed, often with the images included in the scrapbooks it can be difficult to distinguish between an original photograph and a reproduction taken from a magazine or a book. In any event, for the purpose of this project we interpreted the Directive to apply to embedded artistic works even when those works had been removed from their original publication.

To evaluate the time and resource costs involved in using the EUIPO orphan works database and exception, five works were chosen: three news cuttings and two embedded photographic images, one from a magazine and one from a newspaper. These items were chosen for their similarity in length and provenance to the works used in the OWLS exercise, but they were unique in order to simulate a first-time rights holder diligent search.

#### 4.12. MAKING AN APPLICATION TO THE EUIPO REGISTER<sup>122</sup>

We registered five works with the EUIPO register. As with OWLS, registering works on the EUIPO orphan works database takes place online. Before registering works, institutions must register as a 'beneficiary organisation' with the EUIPO. The application process is straightforward, but it took five working days for registration to be confirmed by the EUIPO permitting log-in and registration of works on the database.

Many of the features noted during the OWLS application apply equally here: instead of creating a title for the work, users can record that the work has no title, and provide a full description instead. Assumptions about rightholders still have to be made: in this case, the project team assumed that the newspaper publisher would hold copyright in the cutting, although there is certainly an



**Image 11:** News cutting submitted to EUIPO for Orphan Works Register

<sup>121</sup> Article 1(4) of the EU Directive states: 'This Directive shall also apply to works and other protected subject-matter that are embedded or incorporated in, or constitute an integral part of, the works or phonograms referred to in paragraphs 2 and 3'; that is, although not included in their own right, artistic works printed within the 'books, journals, newspapers, magazines or other writings' listed in Article 2(a), are considered to be embedded works.

<sup>122</sup> This section is based on Stobo, V., Patterson, K., and Erickson, K. (2017) Forthcoming.



argument that the journalist could (or perhaps should) be listed as an additional rightholder (depending on the specifics of the contractual agreement between journalist and employer).

That said, the process of applying to the EUIPO register is less onerous than the UK licensing scheme: the online form is much shorter and the database offers a bulk upload function that simplifies the registration process. EUIPO request that users submit spreadsheets to them before upload: this is simply to check that data can be processed; EUIPO does not audit individual diligent searches. This bulk upload function makes it significantly quicker to register works; although the same information fields required by the web form has to be completed in the spreadsheet, completing the information in a spreadsheet format will be quicker than individual online registration.

However many works one is registering, it is important to record the narrative of all diligent searches, the sources used, the results, as well as documenting how the work is used; moreover, keep those records for at least as long as the work is registered and in use.

Finally, when relying on the Directive, it is important to keep in mind some significant differences between the Directive and OWLS. We have already addressed the issue of free-standing artistic works. The other main differences are discussed in *Orphan Works: The Legal Landscape*; they include: (i) the limited application of the Directive to unpublished works (OWLS applies to all types of work whether published or not); (ii) the need to provide ‘fair compensation’ to a reappearing orphan work owner under the Directive (under OWLS, a reappearing owner is only entitled to claim the licence fee already been paid under the scheme); and (iii) that the Directive enables non-commercial use only (OWLS permits both commercial and non-commercial use).



**Image 12:** A magazine photograph submitted to EUIPO

Table 7 presents details of the time and other associated costs when registering our sample of orphan works with EUIPO.



**Image 13:** A magazine photograph submitted to EUIPO

NAME OF WORK	TIME SPENT ON DILIGENT SEARCH (DS)	TIME SPENT ON APPLICATION	TOTAL TIME SPENT
Newscutting 'Atomic Particle's Travels'	25 mins	17 mins	43 mins
Newscutting – Football Pools winner	20 mins	14 mins	34 mins
B/W newspaper photograph – mushroom cloud	10 mins	16 mins	26 mins
B/W magazine photograph – doorways	10 mins	11 mins	21 mins
Newscutting – BIS takes you there	15 mins	10 mins	25 mins
<b>Total</b>			158 mins
<b>Average</b>			31.6 mins

**Table 7:** Time and associated costs under the EUIPO scheme

#### 4.13. CLEARING THE SCRAPBOOKS <sup>123</sup>

Having carried out the data extraction exercise and made applications under OWLS and the EUIPO register, we used this information to estimate the amount of time it would take to clear the 16 volumes of scrapbooks under both orphan works regimes.<sup>124</sup>

We estimated the total number of works in the collection to be 51,480, based on an average of 15 cuttings per page across 3,600 pages in 16 volumes. Table 8 provides an overview of the different types of works included in the sample. Having identified that 69% of the sample were artistic works, the Project Officer conducted further analysis to identify the proportion of works which were (originally) embedded within another work. Only 13 works were found to be standalone artistic works; the remaining 285 works were deemed to be embedded works for the purpose of the Directive.

TYPE OF WORK	NO. OF WORKS	PERCENTAGE OF SAMPLE
<b>Artistic Works</b>		
Book Artwork	7	
Book Photograph	10	
Newspaper Artwork	8	
Newspaper Photograph	23	
Periodical Artwork	15	
Periodical Photograph	177	
Original Artwork	11	
Original Photograph	37	
Photograph of a TV broadcast	1	
Ephemera	9	
<b>TOTAL</b>	<b>298</b>	<b>69%</b>

<sup>123</sup> The results presented in this section are taken from Stobo, V., Patterson, K., and Erickson, K. (2017) Forthcoming.

<sup>124</sup> Both regimes would need to be used as the Directive does not extend to free-standing artistic works.

<b>Text-based works</b>		
Book Text	9	
Newspaper Text	74	
Periodical Text	44	
Original Handwriting	7	
<b>TOTAL</b>	<b>134</b>	<b>31%</b>
<b>Overall Total</b>	<b>432</b>	<b>100%</b>

**Table 8:** Types of works included in the sample

In Table 9 we summarise the rights status of all 432 works in the sample. 52% of the sample were revealed to be orphan works; scaling up for all 16 scrapbooks, this gives an estimate of 26,770 orphan works in total.<sup>125</sup>

<b>TYPE OF WORK</b>	<b>NO. OF WORKS</b>	<b>PERCENTAGE OF SAMPLE</b>
Works created by Edwin Morgan	37	9%
Works without or unlikely to attract copyright protection		
Works in Public Domain	22	5.1%
Insubstantial parts of works	84	19%
Ephemera (e.g. adverts and tickets)	9	2.1%
Permission Granted		
Works where permission granted by rightholders (15) for use free of charge (4 works subject to specific instructions)	29	7%
Works where permission granted by rightholders (7) for use dependent on fee (which project staff decided not to pay)	19	4.4%
No response		
No response from rightholders (2), with accurate contact details	2	0.5%
No response from rightholders (3), with uncertain contact details (these could be defined as orphan works)	3	0.7%
Works where publisher was unable to license for online use (advised staff to use an exception instead)	1	0.2%
Definite orphan works (excluding non-responders where address details are uncertain)	226	52%
<b>Total works in sample</b>	<b>432</b>	<b>100%</b>
<b>Total time spent on diligent search</b>	<b>1080 hours</b>	
<b>Equivalent salary cost</b>	<b>£11,653.20</b>	

**Table 9:** Rights status across the sample of 432 works from Scrapbook 12

<sup>125</sup> This number was extracted by calculating the percentages from the number of orphan works within the sample (226, or 52%) by the number of estimated works across the 16 volumes of scrapbooks (51,480). The three works where a response has not been received and the project staff are unsure they have the correct contact details for the identified rightholders, could be considered orphan works according to the definition provided in the CDPA 1988. However, we do not count them as such for the purposes of this analysis.

The total number of estimated orphan works, and the indicative status of works within the scrapbooks, allows us to estimate the likely costs of undertaking diligent search activity and clearing rights for all 16 scrapbooks. If we assume the total number of orphan works to be 26,770, then 94% (or 25,164) can be uploaded to the EUIPO orphan works database and 6% (or 1,606) can be licensed through OWLS.

In relation to the 1,606 works that would have to be licensed through OWLS, assuming a standard salary cost of £10.79 per hour,<sup>126</sup> the total estimated costs are as follows: the application and licence fees would amount to £4,448.62; the salary costs and time spent on the rights clearance process would be £39,860.92 and 1.8 years respectively (based on 52 weeks at 40 hours per week).<sup>127</sup> For those works falling within the scope of the Directive (25,164 in total) the average time per work complying with the diligent search requirements and interacting with the online system was 31.6 minutes. This means the salary costs and time spent on the process would be £142,931.52 and 6.4 years, respectively. It is likely, however, that this total could be further reduced by making use of the bulk upload mechanism and grouping works together.

The total cost of using both OWLS and the EUIPO Orphan Works database in tandem to make all orphan works contained in the scrapbooks available online would be £187,241.06 (including application and licence fees, and salary costs) and would take 8.2 years.

---

<sup>126</sup> We report using the average hourly rate of £10.79, calculated from the most conservative archivist annual salary estimate of £22,443 as reported by the Archives and Records Association (ARA). 'The ARA recommends that the minimum starting salary for recently qualified archivists, archive conservators and records managers is between £22,443 and £38,000'; see: <http://www.archives.org.uk/careers/careers-in-archives.html> (accessed 22 December 2016).

<sup>127</sup> These figures are based on the application and licence fees required by the UK IPO, which we have calculated at £2.67 per work and £0.10 per work respectively. This is based on the assumption that applications will be made in batches of 30 works at a cost of £80 in application fees (this is the maximum amount allowed in a single application). The non-commercial licence fee of £0.10 is the standard amount charged under the scheme. The average time to complete a diligent search in line with the expectations of the IPO, and engage with the application process, was 138 minutes; at a standard salary rate of £10.79 per hour, this generates salary costs of £24.82 per work.

## 5. DIGITISATION AND RISK

Victoria Stobo, Kerry Patterson and Ronan Deazley

### 5.1. INTRODUCTION

Decisions about copyright clearance, when to do it, how to do it and how much to do, are always considerations based in the end on a vision of risk, and of risk tolerance in a particular institution ...

Professor Peter Jaszi

Curators, archivists and librarians have to balance a series of demands, expectations and risks when they make digitised collections available online. Before the introduction of the EU Directive and the development of the UK's Orphan Works Licensing Scheme (OWLS), cultural heritage institutions (CHIs) had to determine their own approach to using orphan works. Sectoral guidance included a recommendation that diligent search be performed and recorded but, without clear legal guidelines, risk assessment had a key role to play<sup>128</sup>. A solely risk-assessed approach to the digitisation of orphan works is still preferred by some organisations, while others use it in combination with the orphan works regime. In this section we consider some of the ways in which CHIs avoid, accept, mitigate and manage the risks associated with digitising copyright-protected material and making it available online. We consider risk assessment as an alternative to strict compliance with copyright law, and outline how we engaged in the process of risk assessing the Edwin Morgan Scrapbooks for this project.

### 5.2. COPYRIGHT AND DIGITISATION

The opportunities presented by widespread digital access to our shared cultural heritage are transformative. Recent surveys of the archive and library sector have found that users expect all library and archive materials to be digitised and available.<sup>129</sup> Of course, professional codes of conduct also state that archivists and librarians must respect the rights of individuals: they are committed to complying with the requirements of the Copyright Designs and Patents Act 1988, among other legislation. But, because of the restrictions of the legislation and the costs of rights clearance,<sup>130</sup> it is clear that copyright

---

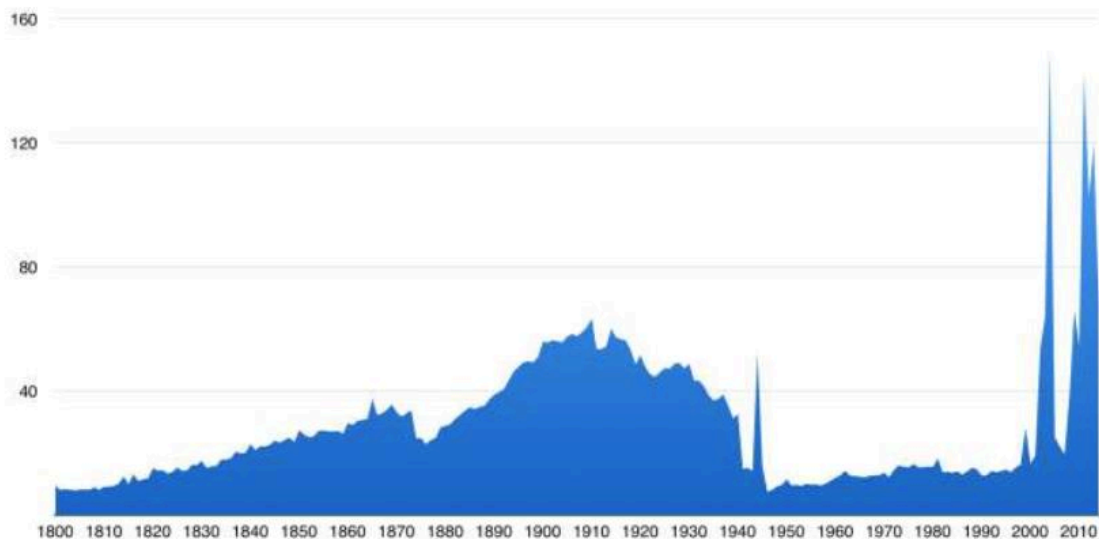
<sup>128</sup> See, for example, the module on Orphan Works and Risk Management available at: [www.web2rights.com/SCAIPRModule/rlo3.html](http://www.web2rights.com/SCAIPRModule/rlo3.html) (accessed 10 January 2017), or in Pedley, P., *Copyright Compliance: Practical steps to stay within the law* (London: Facet Publishing, 2008).

<sup>129</sup> For example, see Dooley, Jackie M., Rachel Beckett, Alison Cullingford, Katie Sambrook, Chris Shepard, and Sue Worrall (2013), *Survey of Special Collections and Archives in the United Kingdom and Ireland*. Dublin, Ohio: OCLC Research. Available at: [www.oclc.org/resources/research/publications/library/2013/2013-01.pdf](http://www.oclc.org/resources/research/publications/library/2013/2013-01.pdf) (accessed 16 November 2016).

<sup>130</sup> Research has shown that the cost of rights clearance usually outstrips both the cost of digitisation and the monetary value of the work itself: Vuopala, A., "Assessment of the Orphan works issue and Costs for Rights Clearance" (May 2010), available at: [www.ace-film.eu/wp-content/uploads/2010/09/Copyright\\_anna\\_report-1.pdf](http://www.ace-film.eu/wp-content/uploads/2010/09/Copyright_anna_report-1.pdf) (accessed 16 November 2016), p. 5; Korn, N., *In from the Cold: An assessment of the scope of 'Orphan Works' and its impact on the delivery of services to the public* (April 2009), available at: [www.jisc.ac.uk/publications/reports/2009/infromthecold.aspx](http://www.jisc.ac.uk/publications/reports/2009/infromthecold.aspx) (accessed 16 November 2016), p.21; Stratton, B., *Seeking New Landscapes: A rights clearance study in the context of mass digitisation of 140 books published between 1870 and 2010* (2011) London: British

plays a significant role in determining what material is selected for digitisation. Archivists either tend to select material that is in the public domain, or material for which they can be reasonably certain that their organisation owns the rights (or has been assigned them by a depositor with the authority to do so).<sup>131</sup> One consequence of this copyright-driven selection process is illustrated,<sup>132</sup> in part, by the 20<sup>th</sup> century black hole in the Europeana data set pictured below (Table 10): note the significant drop in the amount of material available through Europeana that was created from the 1950s onwards, as this material is more likely to be protected by copyright than earlier material.

**Chronological distribution of *dcterms:issued* values in Europeana dataset (1800-today)**



**Table 10:** Date range of materials in the Europeana dataset <sup>133</sup>

There are two problematic issues with using rights status as a method to select works for digitisation. First, the proportion of public domain material in UK collections is significantly reduced compared to most other jurisdictions because of the 2039 rule.<sup>134</sup> Second, asking depositors or donors to assign rights to institutions is not always possible: the depositor may not be the rightholder, and collections often contain third-party materials. Certainly, asking depositors to assign copyright where they can is advisable but staff may find that the agreements for older collections do not include relevant assignments, or that documentation for older collections simply does not exist. As a result, many institutions choose to digitise material where they can be certain that they hold the copyright in the works selected, rather than digitising the material that best fits a particular research theme, user request or strategic priority for the institution.

Library/ARROW. Available at: [www.arrow-net.eu/sites/default/files/Seeking%20New%20Landscapes.pdf](http://www.arrow-net.eu/sites/default/files/Seeking%20New%20Landscapes.pdf) (accessed 16 November 2016), p.5.

<sup>131</sup> Dryden, "Copyright issues in the selection of archival material for internet access" (2008) *Archival Science*, p.123.

<sup>132</sup> Other factors will influence selection processes: staff skills and equipment, budget constraints, and designing the project to meet the needs of a funder, parent institution, or specific group of users.

<sup>133</sup> Europeana Factsheet (2015) The 20<sup>th</sup> Century Black Hole: How does this show up on Europeana? Available at [pro.europeana.eu/files/Europeana\\_Professional/Advocacy/Twentieth%20Century%20Black%20Hole/copy-of-europeana-policy-illustrating-the-20th-century-black-hole-in-the-europeana-dataset.pdf](http://pro.europeana.eu/files/Europeana_Professional/Advocacy/Twentieth%20Century%20Black%20Hole/copy-of-europeana-policy-illustrating-the-20th-century-black-hole-in-the-europeana-dataset.pdf) (accessed 17 November 2016).

<sup>134</sup> For more information, see Legal Landscape Section 3, and [www.create.ac.uk/blog/2014/06/02/will-uk-unpublished-works-finally-make-their-public-domain-debut/](http://www.create.ac.uk/blog/2014/06/02/will-uk-unpublished-works-finally-make-their-public-domain-debut/) (accessed 16 November 2016).

Ultimately, the result of digitising material based on rights status is that users who access collections online are only ever seeing material that has been filtered through this selection process, rather than material which has been deliberately chosen to illustrate the full breadth and depth of the institution's complete holdings, and by extension our shared cultural heritage: from the oldest manuscripts through to born-digital records. In short, the digital historical record becomes skewed towards material that presents no or minimal rights clearance issues. This is a concern for at least three reasons. First, if digital is now the principal method of access to records for many CHI users, those users may not be aware of or attentive to the records that are absent from the digital collection. Second, the skew towards older public domain works means these tend to be the materials that shape research opportunities and activity; this is a particular problem in disciplines such as the digital humanities, which rely on large datasets to conduct research and where researchers are not always able to travel to relevant institutions in person. Third, it creates a fundamental barrier to digital access and therefore to the digital preservation of more recently-created works.<sup>135</sup>

### 5.3. RIGHTS MANAGEMENT IS RISK MANAGEMENT

Risk is typically expressed as the severity of an outcome (or the extent of a benefit resulting from an outcome) occurring, multiplied by the likelihood or probability of it occurring.<sup>136</sup> Risk normally occurs as the result of interaction with uncertainty, for example: we may be uncertain about the rights status of material in our collections; about the likelihood of a rightholder making a complaint about the use of material; and about the likelihood of consequences, such as financial obligations or reputational damage, arising from a complaint. We may be willing to tolerate the risk of making material available despite uncertainty, on the basis that the benefits realised by digitisation outweigh the potential severity of any negative outcomes.

This formulation can often be difficult to apply to the outcomes of CHIs digitising copyright-protected collections as clear data on the rights clearance efforts from previous digitisation projects are not widely available, and very few case studies have been published. Additionally, there is no case law where UK CHIs have been sued for copyright infringement; allied with a lack of data on near-misses and complaints,<sup>137</sup> this makes it difficult to predict the probability of litigation against a CHI or reputational damage

---

<sup>135</sup> Digital preservation is typically carried out when work is in such poor condition that it is close to being lost or suffering further permanent damage. But, digital preservation is also often triggered by a user request to make use of the work in some way, whether in the searchroom, for display in an exhibition, or as part of a research project. That is, delivering digital access is often the prompt for preservation activity. In this way, if copyright status presents a barrier to digital access, similarly it can also impede preservation activity.

<sup>136</sup> For example, the Institute of Risk Management defines risk as "the combination of the probability of an event and its consequences. In all types of undertaking, there is the potential for events and consequences that constitute opportunities for benefit (upside) or threats to success (downside)." See Institute for Risk Management (2002) *A Risk Management Standard*, p.1, available at: [www.theirm.org/media/886059/ARMS\\_2002\\_IRM.pdf](http://www.theirm.org/media/886059/ARMS_2002_IRM.pdf) (accessed 22 November 2016).

<sup>137</sup> The authors define 'near-miss' in this context as a complaint about copyright infringement which could result in litigation, or where litigation is threatened, but which is resolved, either by negotiation or by capitulation, before proceedings are issued, or where proceedings are abandoned.



occurring as the result of a complaint. That said, while the lack of litigation is a revealing metric in itself, in that it underlines the seeming unlikelihood of litigation arising within the heritage sector, we should be cautious of reading too much into this given the fact that reliable data on near-misses and complaints is unavailable.<sup>138</sup>

Furthermore, the CHI sector could be more vocal and proactive in articulating the impact and value of digitised collections, which would make it easier to calculate the benefits of digitisation as against the risk of infringement. One way of doing this would be to use the Balanced Value Impact Model to articulate the different kinds of values, benefits and impacts generated by digitisation.<sup>139</sup> For example, by clearly articulating the social value of digitising local film collections, a local history museum could balance the benefits (improved user experiences, new outreach activities, an increased sense of place and belonging for participants, donations of film materials, increased knowledge about collections, and so on) against the risks (copyright infringement, sensitivity, complaints from rights holders, potential loss of good reputation). By doing this, they could then put in place strategies to minimise those risks and maximise the benefits; for example: by creating a local film history group; publicising the search for rights holders and the people who appear in the films; run screenings where viewers can provide feedback, information and memories; work with local social care providers to run memory sessions; and, contribute to local schools' learning resources.

Risk	Probability (1-5)	Severity (1-5)	Score (PxS)	Action to Prevent or Manage Risk
Legal	3	5	15	The EThOS and EThOSnet projects are addressing the legal aspects of collecting, digitising and making this type of material available.
More theses to be digitised than expected	5 (realised)	2	10	The project has delivered 4 times the number of theses to be digitised than originally expected. This means greater logistical involvement for the British Library, but the additional resource can be made available.
Institutions attempting to clear rights with authors	3 (small number realised)	3	9	A small number of institutions are contacting authors for clearance to make their theses available. 3 or 4 of the bigger institutions are doing this impacting on the logistics of the project. This is containable by applying a time limit of late May for decisions and the addition of further theses to replace those withdrawn.

**Table 11:** An example of a completed balanced scorecard for risk assessment.

There are three common ways of expressing and calculating these risks. One example of a traditional method is the balanced scorecard approach, which can be used as part of the project management process. A completed example, taken from a JISC digitisation project, is given above (Table 11).<sup>140</sup> Users of this method are expected to assign a numerical

<sup>138</sup> For example, the authors know of at least one action initiated against a UK archive institution, which was dropped before reaching court.

<sup>139</sup> Tanner, S. (2012) Balanced Value Impact Model.

<sup>140</sup> This scorecard is taken from the UK Thesis Digitisation Project Project Plan, available at: [webarchive.nationalarchives.gov.uk/20140702233839/http://www.jisc.ac.uk/media/documents/programes/digitisation/ukthesesp.pdf](http://www.jisc.ac.uk/media/documents/programes/digitisation/ukthesesp.pdf) (accessed 17 November 2016) p.6.



value against the probability of an event occurring, with '1' meaning no to low probability and '5' meaning the event is highly likely to take place. A numerical value is then assigned to the severity of an event, with '1' meaning little to no effect, and '5' meaning severe consequences for project outcomes. The values for probability and severity are then multiplied to give a total score, and a section of the table is provided to record in detail how the risk identified will be mitigated or avoided. This allows project managers to see at a glance the project elements which carry the most risk and how they are being managed. This scorecard uses 5x5 scoring, but 3x3 and 'High, Medium and Low' scoring is also common.

The second approach adopts a similar methodology in that the scorecard was developed into a risk calculator, created by the Web2Rights project for an Open Educational Resources Toolkit. The risk calculator assigns numerical values to different types of material and the different ways in which they can be used, giving a high, medium or low ranking for different uses in addition to a numeric score. An example taken from the risk calculator, using an artistic photograph as the subject, can be seen below (Image 14).<sup>141</sup>

The image shows a screenshot of the Web2Rights OER Risk Management Calculator. At the top, there is a header 'Artistic photograph' next to a photograph of a camera and film. Below this is a dropdown menu with '- Artistic photograph' selected. The calculator consists of several colored sections, each with a question and a dropdown menu:

- Created with commercial intent?** (Yellow background) with 'Unknown' selected.
- Includes clinical content/images of identifiable individuals/children for which rights have not been cleared?** (Light blue background) with 'No' selected.
- Licence intended for your end product?** (Light green background) with 'CC BY NC ND' selected.
- Nature of creator** (Light blue background) with 'Known' selected.
- Contact with Rights holder?** (Pink background) with 'Known(low profile & traceable but don't respond)' selected.

Below these sections, the **Indicative Risk Level: 384** is displayed in an orange bar. At the bottom, a dark grey box contains the text: **Medium risk = 151 - 500** and a paragraph: *This Indicative Risk Value indicates that the level of risk associated with the content that you wish to use is likely to require serious consideration. If your organisation is not comfortable with this level of risk, you should consider lowering the potential risk by providing access under a more restrictive end use licence and/or selecting material which reduces the Indicative Risk*

**Image 14:** Example of an artistic photograph and use ranked using the Web2Rights OER Risk Management Calculator

As we can see, the user of the calculator has decided to explore the risk associated with making an artistic photograph available. It is not known whether the photograph was created with commercial intent, but we do know that the photograph does not include clinical content, or images of identifiable individuals or children. The user wants to make the image available under a Creative Commons Attribution NonCommercial NoDerivatives licence. The creator is known, with a low profile, and the user has found contact details and approached the rightsholder for permission, but they have not responded. The

<sup>141</sup> The Risk Management Calculator is available at: [www.web2rights.com/OERIPRSupport/risk-management-calculator/](http://www.web2rights.com/OERIPRSupport/risk-management-calculator/) (accessed 22 November 2016).

calculator gives a score of 384, which places it within the 'Medium' band (which includes scores of 151-500).

A third and final option is to define bespoke 'criteria' or 'categories' of risk for specific institutional digitisation projects; the Wellcome Library case study outlined below provides just such an example.

Whichever approach is adopted, it's important to bear in mind that rights clearance is cyclical and costs will build up throughout a project; that is, costs are not based purely on the amount of money that is paid to rightholders to obtain licences. Indeed, the vast majority of rightholders contacted during rights clearance for archive and library digitisation projects grant permission without requesting fees. Rather, the factors that can contribute to the potential costs range from the time taken to create item-level metadata at the beginning of a project, to maintaining records once the project is completed, and so on. For example, staff must: identify the copyright status of the works selected for digitisation, bearing in mind that one item can include multiple rights and therefore multiple rightholders; identify the rightholder(s) in the works (which could include estates, multiple heirs or successor companies); locate and contact the rightholder(s); negotiate permission for specific uses based on the way in which they intend the work to be use (this may or may not incur licence fees); and, create and maintain associated metadata and records of all of these actions, transactions and results.

#### **5.4. CASE STUDIES: TAKING AND AVOIDING RISKS**

In this section, we consider three published case studies concerning institutions that have taken different approaches to managing the risks associated with rights clearance. The first explores the Jon Cohen AIDS research collection at University of Michigan.<sup>142</sup> The second is taken from the Southern Historical Collection and Carolina Digital Library and Archives, based at the University of North Carolina at Chapel Hill, where the Thomas E Watson Papers were digitised.<sup>143</sup> The final example is the Wellcome Library (working with five partner institutions) and its pilot mass digitisation project, Codebreakers: Makers of Modern Genetics.<sup>144</sup> The commentary within this section is arranged thematically comparing and contrasting the actions of each institution at each stage in the digitisation process.

##### **5.4.1. COLLECTIONS BACKGROUND**

The collections include: the Jon Cohen AIDS research collection, comprising the personal papers of Jon Cohen, a journalist and science writer who covered the development of a vaccine for the condition during the 1980s; the personal papers of Thomas E Watson, a prominent US senator active in the late 19<sup>th</sup> and early 20<sup>th</sup> century; and 20 collections of

---

<sup>142</sup> Akmon, D. (2010), "Only with your permission: how rights holders respond (or don't respond) to requests to display archival materials online", *Archival Science*, 10(1), pp. 45-64.

<sup>143</sup> Dickson, M. (2010), "Due Diligence, Futile Effort", *The American Archivist*, 73(2), pp. 626-636.

<sup>144</sup> Stobo, V., Deazley, R., and Anderson, I.G. (2013) "Copyright and Risk: Scoping the Wellcome Digital Library", Working Paper 2013/10, CREATE, University of Glasgow, Glasgow, available at: [zenodo.org/record/8380/files/CREATE-Working-Paper-2013-10.pdf](http://zenodo.org/record/8380/files/CREATE-Working-Paper-2013-10.pdf) (accessed 25 November 2016).

personal papers created by geneticists of the 20<sup>th</sup> century, including James Watson, Francis Crick, and Rosalind Franklin, among others.

#### 5.4.2. RIGHTS AUDITING THE COLLECTIONS

Staff at the University of Michigan audited the Jon Cohen collection and found 13,381 items. 6,026 items (45%) were found to be either newspaper or journal articles, and the library staff decided not to digitise these as the majority were already available online elsewhere. 1,892 (14%) of the items were US Federal Government documents<sup>145</sup> which could be made available without seeking permission. For only 209 documents was Jon Cohen the copyright holder (2%), for which he granted permission. This left the staff with the task of securing permissions for the remaining 5,254 items (39%) where copyright was held by a third-party. The staff identified 1,376 unique rightholders in this material. The archivists at University of North Carolina sought and were granted permission to digitise the material created by Thomas E. Watson from his surviving heirs. They audited the collection and found 3,304 correspondents, with identifying information for 3,280 correspondents. They found birth and death dates for 1,709, leaving 1,571 with no dates. 1,101 of the correspondents died after 1939, which was the cut-off date the archivists had decided to use in determining whether a letter was still protected by copyright or not. 608 died before 1939, allowing the archivists to assume that these letters were in the public domain.<sup>146</sup> This process took four and a half months and cost almost \$6000. At the end of this process, the archivists determined that the collection could be categorised as follows: 14% were created by Watson family members and permission had been granted by the estate; 3% were created by freelance workers and excluded from the scope of the project because of time and expense; 4% of the letters were unsigned, illegible or used pseudonyms; 21% of the letters were in the public domain; 27% were still in copyright; and 31% of the letters had an undefined copyright status.

The Wellcome Library staff and their partners realised there would be many thousands of third-party rightholders across the 20 collections of personal papers selected for digitisation: altogether, almost 3M pages were digitised. To manage this process, they asked their partner archives to identify rightholders in their collections according to a set of risk criteria shown in the table below (Table 12). Using these criteria, the Wellcome and their partners were able to iteratively reduce the number of rightholders they intended to contact from thousands to just 160 across all the collections.

---

<sup>145</sup> US Federal government documents do not benefit from domestic copyright protection, and are therefore in the public domain (s.105 of the 1976 Copyright Act).

<sup>146</sup> This digitisation project took place in 2009. The assumption that material created pre-1939 would be in the public domain is based on the US copyright term of life plus 70 years: 1939 + 70 = 2009.

HIGH RISK	MEDIUM RISK
Author is a well-known literary figure, broadcaster, artist	Author has (or had) a high public profile
The author/estate/publisher is known to actively defend their copyright	Author is alive and known to have a literary estate as recorded in the WATCH file
The relationship between the institution and the author/estate/publisher is awkward	The material appears to have been published or broadcast and/or prepared for commercial gain rather than to advance academic knowledge or in a not-for-profit context

**Table 12:** Risk criteria used by the Wellcome Library during Codebreakers: Makers of Modern Genetics

### 5.4.3. THE RIGHTS CLEARANCE PROCESS

Where contact details were found, the staff at the University of Michigan sent the rightholder a letter that explained the project, described the material they wanted to digitise, and asked for a non-exclusive licence to digitise the item and make it available online, without the offer of a licence fee. The letter also included a statement of support from Jon Cohen and the funders for the project, along with a consent form.

The staff at the University of North Carolina attempted to find identifying information for rightholders using 'ancestry.com, the Congressional Biographical Directory, the Historical Marker Database online, the Library of Congress authority database, the New Georgia Encyclopedia, print references, the Social Security Death Index, the WATCH File, Wikipedia, and WWI draft registration forms.'<sup>147</sup>

Wellcome Library staff searched Who's Who, the WATCH File, Google, Wellcome Trust internal databases, third-party archives, the Dictionary of National Biography, obituaries and Wikipedia to find contact details for rightholders and their heirs.<sup>148</sup> The results of this process, including initial contacts, follow-up attempts and details of permissions granted and refused were carefully documented. Rightholders were sent a letter that explained the project, outlined the material to be digitised, and asked for permission to make the material available under a non-exclusive Creative Commons Attribution Non-Commercial 4.0 licence. The Wellcome did not offer a licence fee to rightholders.

Staff from both Michigan and the Wellcome sent follow-up letters and used email and phone where possible to secure permissions. Many rightholders requested copies of the material before making a decision, and these were sent by email, fax or post where required. It is worth noting that such requests are common during digitisation projects and time should be factored in at the project planning stage to accommodate them.

### 5.4.4. RESULTS OF THE RIGHTS CLEARANCE PROCESS

The following table presents the results of the rights clearance process at each of the case study institutions (Table 13).

<sup>147</sup> *ibid*, p.629. The article does not provide information on the content of the permission letters.

<sup>148</sup> Stobo, V. et al (2013) p.26.

	<b>Jon Cohen AIDS Collection</b>	<b>Thomas E Watson Papers</b>	<b>Codebreakers</b>
<b>No. of Copyright Owners identified</b>	1,377	3,280	160
<b>No. Copyright Owners traced</b>	1,023 (74% of those identified)	4 (0.12% of those identified)	134 (84% of those identified)
<b>Replied</b>	748 (68% of those contacted) <sup>149</sup>	3 (75% of those contacted)	103 (77% of those contacted)
<b>Permission granted for all items</b>	679 (91% of respondents)	3 (100% of respondents)	101 (98% of respondents)
<b>Permission granted for some items</b>	23 (3% of respondents)	n/a	n/a
<b>Permission denied</b>	46 (6% of respondents)	n/a	2 (2% of respondents)
<b>Non Response</b>	352 (32% of those contacted)	1 (25% of those contacted)	26 (19% of those contacted)
<b>Orphan Works</b>	354 (26% of those identified)	3,276 (99.88% of those identified)	22 (14% of those identified)
<b>In Progress</b>	n/a	n/a	4 (4% of those identified)

**Table 13:** Rights Clearance Results

We can see that the archivists at the University of Michigan had a success rate of 74% in finding contact details for rightholders. 68% of those contacted responded, with 94% of respondents granting permission for all or some of the material requested. 6% of respondents refused. 32% of those contacted did not respond, and contact details could not be found for 26% of the rightholders identified in the collection. The archivists decided not to make non-respondent or orphan material available online. If we add these items to those where permission was refused, the University of Michigan was unable to make 1,973 items available, or 36% of the total collection. This is despite receiving dedicated funding for the digitisation and rights clearance process, and spending, on average, 70 minutes per rightholder on securing express permissions. Moreover, the study revealed that commercial rightholders are more likely to refuse permission requests, and that after four months, subsequent permission requests deliver diminishing returns.

The archivists at the University of North Carolina managed to find contact details for the estates of only four correspondents out of the 3,280 identified, either by using the WATCH file or by contacting other repositories where manuscript collections were known to contain material created by identified correspondents. Of the four, three estates granted permission and one did not respond.

Staff at the Wellcome managed to find contact details for 84% of the selected rightholders. 77% of those contacted replied, with 98% of respondents granting

---

<sup>149</sup> The figures for responses do not tally completely because 83 rightholders were contacted on two occasions and were asked for permission to digitise two different sets of items: those permissions were recorded separately.

permission. 19% of those contacted did not respond, but after re-assessing the likelihood of those rightholders objecting to publication, most of the material was made available online, subject to takedown requests. Contact details could not be found for 14% of the selected rightholders, meaning that this material is orphaned. In contrast to the University of Michigan, the Wellcome decided to make orphan work and non-respondent material available online in batches, excluding material which was deemed to be very high risk. As a consequence, staff were able to make most of the collections they selected for digitisation available online: indeed, the material associated with 91% of the selected rightholders was made available, in addition to all of the other material where rightholders were not contacted at all.

#### 5.4.5. RISK MANAGEMENT DECISIONS

The University of Michigan decided to adhere to a policy of strict copyright compliance and only make third-party material available where they had been granted express permission. As a result, they made no orphan works or non-respondent material available. At the end of the project, almost 36% of the collection was not available to view online.

If the archivists at the University of North Carolina had followed the same path, they would only have been able to make available online the Watson family material, public domain material, and material for which they were able to secure permission: just 35% of the total collection. However, rather than digitise only when express permission was granted, the archivists turned to the US copyright doctrine of fair use.<sup>150</sup> Following the initial attempts to clear rights in the material, the archivists involved in the project presented their findings to the legal counsel for University of North Carolina University Libraries and explained that they wanted to discontinue any further copyright investigation for this collection. In turn, they were given permission by the University to make the material available online: the University had been convinced by the fair use argument.<sup>151</sup> A year later, staff reported that no takedown requests had been received in relation to the Thomas Watson material.

---

<sup>150</sup> Section 107 of the US Copyright Act 1976 provides that 'the fair use of a copyrighted work, including such use by reproduction in copies or phonorecords or by any other means specified by that section, for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright'. In making a determination about whether an act constitutes fair use In determining whether the use made of a work in any particular case is a fair use, s.107 requires that the following four factors be taken into account: 'the purpose and character of the use, including whether such use is of commercial nature or is for non-profit educational purposes; the nature of the work itself (whether it is a factual or creative work); the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and the effect of the use upon the potential market for or value of the copyrighted work'.

Fair Use justifications like those used by University of North Carolina have been used to facilitate many digitisation projects at US CHIs. Such justifications still rely on elements of risk management, given the need to interpret the four factors. See Aufderheide, P. and Jaszi, P. (2011) *Reclaiming Fair Use: How to put balance back in Copyright*, Chicago: University of Chicago Press for more details, and the *Code of Best Practices in Fair Use for Academic and Research Libraries* (available at: [www.arl.org/storage/documents/publications/code-of-best-practices-fair-use.pdf](http://www.arl.org/storage/documents/publications/code-of-best-practices-fair-use.pdf)) for an example of fair use guidance.

<sup>151</sup> Dickson, 636.

Given the huge size of the combined collections selected for digitisation, the Wellcome Library staff and their partners realised that contacting all third-party rightholders would be a significant undertaking with little guarantee of comprehensive success. Unlike the staff at the University of North Carolina, they could not avail of a generalist defence like the US fair use argument; they understood that, for the project to be successful, they would have to accept a much larger degree of risk. They managed this by developing the criteria for identifying rightholders likely to object to publication, which they combined with a 'best endeavours' search to trace and contact those rightholders. The Wellcome Library also has a takedown policy that applies to all of the material made available on their site. To date, the Wellcome Library has received only two takedown requests in relation to Codebreakers material. No reasons were given for the requests, no compensation was sought by the requestors, and no litigation has ensued.<sup>152</sup>

#### 5.4.6. ANALYSIS

The staff involved in each of these rights clearance exercises made decisions to set strategic boundaries for their respective projects. For example, the staff at both the Wellcome and the University of Michigan decided not to digitise newspapers and journals for two reasons: it was considered to be a waste of effort where the material was available elsewhere, and the rights clearance process for the material was perceived to be too onerous. The University of North Carolina decided to exclude material made on a freelance basis from their project, on the grounds that including the material would have created further time and expense for the archivists involved in rights clearance.<sup>153</sup> At the University of Michigan, non-response was taken as denial of permission, while the University of North Carolina decided to make non-responders' materials available on the basis that digitisation amounted to fair use. In contrast, the Wellcome decided to make most of their non-respondent and orphan material available, unless it was deemed to be very high-risk.

CHIs must be aware of the trade-offs in making such decisions, balancing resources and perceived risks against the benefits of making more collections available online. As previously mentioned, digitisation efforts should be focused on the most appropriate material for a particular project. In turn, staff at cultural institutions should consider weighing the risks of making copyright-protected material available online without express permission against clearly-articulated benefits of doing so; with this in mind, they should formulate strategies that minimise the risks and maximise the benefits. The case studies discussed in this section illustrate the potential value of risk-based strategies, strategies that we believe will become increasingly significant as more institutions digitise their collections in the absence of meaningful legal reform in this arena.

---

<sup>152</sup> Stobo et al (2013) Copyright and Risk: Scoping the Wellcome Digital Library Project, CREATE Working Paper 2013/10, CREATE, University of Glasgow, available at: (accessed 5 December 2016).

<sup>153</sup> Including material made on a freelance basis would have generated more work for the project archivists, as they would have had to assess whether the copyright was held by the freelance worker, or by the employer the freelancer was working for. Usually copyright vests with the freelance worker in such situations, but this can be subject to an agreement to the contrary, where the copyright will be retained by the employer. Assessing all the freelance works in the collection would have taken up considerable time with no guarantee of successful clearance.



## 5.5. MANAGING RISK IN THE SCRAPBOOKS

Created in the mid-20<sup>th</sup> century from predominantly published material, the Edwin Morgan Scrapbooks contain a huge number of works that are still in copyright. Taking a 30-page sample (10%) of Scrapbook 12, we carried out a full data extraction exercise that included assessing the risk level of each item. From this sample of 432 works, 64% were in copyright to both known and unknown rightholders.<sup>154</sup>

Throughout the assessment process, uncertainty was a constant undertone. Aside from the concerns of carrying out a search that would be sufficiently diligent, the incomplete nature of many of the works presented a challenge. From the sample, 50% of works were either incomplete or it wasn't possible to tell if they were a complete work. Of these, 35% were clearly incomplete, whereas for 15% it was not possible to determine definitively whether they were incomplete or not. With this in mind, the 'completeness' of each cutting was assessed as: Yes-No-Unknown.

Some of the works that were incomplete could be deemed to be an insubstantial part of the larger work, allowing for lawful use without permission. However, incompleteness and insubstantiality are not the same thing: a very small cutting might display the most significant or recognisable part of an artistic or literary work even though it is quantitatively insignificant. For example, the extract might contain the most famous passage from a novel; as such, it would almost certainly be regarded as a substantial part of the work itself. In some cases, however, it was relatively easy to determine that a cutting was not substantial, for example, if it contained a few words of non-specific text or an indistinct section of an image.

Naturally, substantiality needs to be decided on a case-by-case basis, with consideration given to the content of the work and its context; this can be a daunting task, with many works falling within a grey area. From our sample of 432 works, we categorised 217 as either incomplete or unknown; of these, 84 (19% of sample total) were deemed to be insubstantial in nature.

Thereafter, we addressed the risk of making these materials available online without permission, adopting three risk categories: low, medium and high.<sup>155</sup> As with the issue of substantiality, making this determination was not always clear-cut. We discuss our risk assessment categories in the next section.

---

<sup>154</sup> For these purposes we exclude works created by Morgan himself, works in the public domain, insubstantial parts of works that may or may not be in copyright, and ephemera. For further details, see *Diligent Search in Context and Practice*, Table 9.

<sup>155</sup> Sometimes, items classed as low risk could also be described as 'no risk,' as this category includes works which are not subject to copyright protection, as well as works for which rights were subsequently cleared as part of the digitisation project.

Images 15, 16 and 17 illustrate some of the challenges posed by the material contained in the Scrapbooks. All three were considered by the Project Officer to pose a low risk from the perspective of digitising for use online. Readers may or may not agree. The first cutting (Image 15) is small in size,<sup>156</sup> taking up approximately 1/16 of the scrapbook page. The image might be a part of a larger image. For example, the right side of the cutting looks misshapen, as if it had been cropped from a larger picture. But, this is speculation only. All that can be said is that the image composition does not give any meaningful indication as to whether the image was originally intended to be as shown. The image may



**Image 16:** Colour image of waves (perhaps)

or may not be complete, or it may be a substantial part of a larger work. It is not possible to tell.

The second example is a colour image of what appears to be waves (Image 16). The image is small, taking up less than 1/16 of the page. It is definitely cropped from a larger image. The subject matter is indistinct and difficult to make out; as such, we considered it to be insubstantial.



**Image 15:** Colour photograph of a man operating a control panel

The third (Image 17) consists of two images. One is the wooden frame of the television and the other is the black and white image of the figure inside. Together they take up less than 1/16 of the scrapbook page. Both are definitely incomplete, and it is likely that the television surround came from a magazine advert (such things are used elsewhere in the scrapbooks). By and large, cuttings from magazine adverts of this kind were considered to be ephemera and extremely low risk. The black and white photograph appears to have been greatly cropped. As such, we considered it likely to be an insubstantial part of the larger work. As with the other two cuttings, we determined that the risk of making use of the image without permission was low.



**Image 17:** A collage

### 5.5.1. RISK ASSESSMENT CATEGORIES

After an initial assessment of the scrapbooks by the Project Officer, a set of risk categories were developed to allow categorisation of the material during the data extraction process

<sup>156</sup> See, for example, the newspaper text above the image which provides a sense of scale.

on the 10% sample. The criteria were developed partly in response to the type of material found in the scrapbooks but also taking into account more general risk criteria informed, in part, by the Wellcome Codebreakers digitisation project.<sup>157</sup> In Table 14 we set out the guidelines underpinning the risk classification of each item from the sample.

<b>NO/LOW RISK</b>	Item was created by Edwin Morgan Item is no longer in copyright Item is a piece of ephemera <sup>158</sup> Item appears to be an insubstantial part of a larger work Published works authored by private individuals for a non-commercial purpose, e.g., a letter written to a newspaper or magazine
<b>MEDIUM</b>	Personal photographs or other similar items, not produced for commercial purposes, where the author is thought to be a friend or correspondent of Morgan Work thought to be in copyright where an author is named but no further information can be found Work thought to be in copyright with an identifiable publisher that no longer appears to exist
<b>HIGH</b>	Item is still in copyright and the rightholder is identifiable Author is known to have an active estate/publisher defending copyright Substantial extract from a book or article, and particularly when other extracts taken from the same work appear elsewhere in the Scrapbooks

**Table 14:** Risk Guidelines from Edwin Morgan Project

During the data extraction process, each item was initially classified as either low, medium or high risk. Following this initial classification, the Project Officer targeted the high risk material to see if permission could be secured for use. As we noted in *Diligent Search in Context and Practice*, a number of rightholders (15) granted permission to make use of their work free of charge; whenever permission was secured the relevant work was re-designated ‘no risk’. In addition, five orphan works were cleared for use through OWLS; they too can be considered ‘no risk’, at least until the seven-year licence expires. A further five orphan works were cleared for use through the Directive.

After more than 1000 hours spent on diligent search and rights clearance activity by the Project Officer, 61% of the sample was deemed to be no/low risk, 34% was medium risk, with 5% remaining in the high risk category. Naturally, the high risk category includes each of the 19 works for which the relevant rightholder was only prepared to grant permission conditional on payment of a fee. No licence fees were paid.

<sup>157</sup> For details, V. Stobo with R. Deazley and I.G. Anderson, *Copyright & Risk: Scoping the Wellcome Digital Library Project* (2013) CREATE Working Paper 2013/10, DOI: 10.5281/zenodo.8380.

<sup>158</sup> Ephemera are items without lasting significance, intended to be used for a short period of time e.g. tickets or advertisements. For this reason they can be classed as low risk, as long as they do not contain material from a source still in evident copyright, such as a photograph or artwork. Even if these ephemera are protected by copyright, the owner is highly unlikely to object to use (so long after the event) given that these works were originally created for functional/informational purposes at the time

## 6. CONCLUSION

Ronan Deazley

The problem that orphan works pose for cultural heritage institutions who want to enable the widest possible digital engagement with our shared cultural heritage is substantial. There has been a twin-track response to this problem in the UK, in the form of the exception implemented under the Orphan Works Directive operating in tandem with the Orphan Works Licensing Scheme (OWLS). This approach offers greater opportunities than many other European countries for developing digitisation initiatives that can be tailored to institutional collections, needs, budgets and ambition. Crucially, however, both schemes are tethered to the requirement of diligent search.

Within Europe, Member States are required to determine what sources should be consulted as part of the diligent search requirement; at the same time, their guidance must include 'the relevant sources listed in the Annex' to the Directive. Different jurisdictions have taken different approaches to the implementation of this obligation. Some countries have simply transposed the list of sources set out in the Directive into their national copyright regime, without providing any further details or guidance. Other jurisdictions have chosen to articulate a more complete list of sources to be consulted within their enabling legislation. In the UK, the sources listed in the Annex are replicated verbatim in the UK's *Orphan Works Regulations*. In addition, the Intellectual Property Office (IPO) has produced three sets of guidelines that are 'primarily intended' for those wanting to make an application through OWLS, although each does state that the guidelines may be helpful to those conducting a diligent search in relation to the Directive. Each of the guidelines is accompanied by a Diligent Search checklist. Taken together, the guidelines and the checklists provide very useful information and signposting when undertaking a diligent search.

These different approaches to implementing the Directive underscore the lack of clarity that envelopes diligent search, both conceptually and a matter of practical implementation across the EU. Academics disagree on the nature of the obligation diligent search imposes. It has been suggested that if a Member State has provided guidance on diligent search incorporating a list of potentially relevant sources then *all those sources* should be consulted; otherwise, the search could not be considered diligent. In other words, this concept of diligence requires an exhaustive search of identified sources. We disagree. We argue that a more purposive interpretation of the diligent search requirement should be adopted across the EU. Diligence should not be characterised by an unthinking adherence to a check-list of sources, however useful and well-crafted. Much will depend on the content of the work and the context in which it is found, as well as the expertise and the knowledge-base of the person conducting the search. Within the UK, the IPO have clearly signalled in its guidance and practice that diligence must be context-specific: that is, there is no minimum requirement to be followed in every case. We encourage other jurisdictions to adopt a similarly practical, pragmatic approach.

However, the benefit of this context-specific approach is complicated by the fact that the enacting legislation within the UK has introduced (or seems to have introduced) two different diligent search standards: good faith diligent search and reasonable diligent search, for the Directive and OWLS respectively. At present, it is unclear whether the

imposition of a different obligation was intended by the legislature, and the IPO do not address the issue in their guidance. That said, the IPO have structured their guidance in a way that suggests, albeit tacitly, a different standard of care is implied, and that the standard of reasonableness is set at a higher threshold than that of good faith. From our perspective, we do not think it desirable, or necessary, to complicate the UK orphan works regime with different standards regarding diligent search. We prefer an approach that treats good faith and reasonableness as synonymous in this context, rather than introducing a hierarchy of practice. This would provide a simplicity of process, while adequately safeguarding the interests of unknown and unlocatable rightholders. What is needed, however, is clear direction from the IPO on this issue.

The IPO might also influence the public understanding and practice of diligent search in another way. At present, across most Member States, details of diligent search remain with the institution availing of the exception under the Directive. Only the details of the work and relevant rightholders are formally reported to the EUIPO, although information on the search must, of course, be retained by the institution itself. As such, the Directive does not encourage or enable a *shared practice of diligent search*. In the UK, however, there is an opportunity to do just that: to make available documentation evidencing the practical reality of diligent search activity in relation to a variety of works, contexts and institutions. As helpful as the IPO's guidelines are in providing an overview of sources that may be useful when conducting a search, the sheer volume of potential sources is likely to overwhelm many archivists or institutions contemplating making use of OWLS or the Directive. But, supplementing those guidelines with examples of real-life search activity – search activity that has been considered and approved by the IPO as part of the licensing process – would prove invaluable in helping the heritage sector better understand the nature of the diligent search requirement in both theory and practice. It would allow a shared practice of diligent search to emerge, one that could prove instrumental in positively shaping the heritage sector's engagement with the orphan works regime in the UK and across Europe.

Despite these opportunities for potentially improving and enhancing engagement with the Directive and OWLS, the simple fact remains that for anything other than small-scale digitisation initiatives, the long-term efficacy and relevance of both schemes appears fatally compromised by the demands of diligent search. The costs and challenges of rights clearance activity are a significant barrier to the digitisation of cultural heritage collections. Existing literature evidences this reality, and our research – the first major UK study concerning the concept of diligent search since the introduction of the Directive and OWLS – confirms that diligent search and mass digitisation are fundamentally incompatible, however light-touch the nature of the diligent search obligation. With respect to the 16 scrapbooks created by Edwin Morgan, we estimate that it would take one researcher over 8 years to undertake the diligent search activity alone, at a cost of more than £185,000. And of course, this would not guarantee that the scrapbooks could be made available in their entirety. Known and locatable rightholders might make permission contingent on the payment of unaffordable licence fees, or might simply withhold permission on other grounds. No cultural heritage institution, however well-resourced, would ever take on such a speculative and costly venture.

As for the scrapbooks, from 1953 Morgan made sustained efforts to get them published without success. Similarly, the Edwin Morgan estate and the University of Glasgow are

keen to share them online with the rest of the world. However, they will remain accessible only within the physical confines of the University Library building for the foreseeable future, apart from a selection of the pages that we digitised for this project. In this respect, they evidence the very real phenomenon of the so-called 20<sup>th</sup> century black hole. Because copyright status shapes selection processes regarding the digitisation of heritage materials, collections containing work from the mid- to late 20<sup>th</sup> century are less likely to be digitised for public consumption. In short, our digital historic record is skewed towards material created in the nineteenth and early twentieth century, material that resides in the public domain.

Moreover, the orphan works regime may unintentionally introduce another variable likely to distort digitisation selection processes. We have argued that the diligent search requirement must be interpreted to require less than a non-exhaustive search of all possible relevant sources: that is, the nature of the obligation will be determined by the content and context of the work. But in addition, the experience of this project is that when presented with a work lacking any meaningful contextual information, diligence will often require little more than a google search (for text), or the use of reverse image search technology (for image). For one black and white photograph, the subject of an exploratory application to OWLS, reference to six sources only was deemed necessary to satisfy the diligent search requirement, with the time spent on search activity and completion of the application totalling less than an hour. In other words, *the less you know about the work in question* the easier and less costly it can be to conduct an effective diligent search (whether good faith or reasonable). Just as institutions tend to privilege public domain material when making selections for digitisation, so too institutions might be tempted to select the low-hanging fruit in the copyright garden: material that is in copyright but that lacks any contextual metadata and so minimising the search burden. Put simply, items may be selected for digitisation on the basis that little to nothing is known about them.

Whenever choosing to make copyright-protected material available online, risk also plays an inevitable role in shaping digitisation strategies. Indeed, the solution to the problem of orphan works adopted by the European Directive requires taking risks, in that reliance on the exception provides no guarantee of immunity from future litigation. A reappearing rightholder might challenge the robustness of the diligent search carried out in relation to their work; if successful, what was perceived to be a lawful use would be deemed to fall outside the scope of the Directive. OWLS on the other hand directly mitigates this risk: it provides users with a reassurance about regulatory compliance which the exception cannot, as well as a shield against future liability should any rightholders reappear. It would be entirely understandable if risk-averse institutions opted for the safety-net of the OWLS regime instead of the Directive, even if it comes with an additional financial cost attached, and even if use turns on having to renew the relevant licence every seven years.

For anything beyond a small-scale digitisation initiative, risk-management will not just shape a digitisation strategy, it will lie at its very core. The Wellcome Library's Codebreakers project provides a compelling illustration.<sup>159</sup> To enable online access to as

---

<sup>159</sup> For details, see V. Stobo with R. Deazley and I.G. Anderson, *Copyright & Risk: Scoping the Wellcome Digital Library Project* (2013) CREATE Working Paper 2013/10, DOI: 10.5281/zenodo.8380.

much work and correspondence of the pioneers of genetic research as possible, a risk-managed approach to copyright compliance was adopted, and necessarily so in the opinion of the Wellcome Library. In nearly four years since the project was first launched, the Wellcome have only received two requests to remove Codebreakers material from their online resource. For both, no reasons were given for the takedown request, no compensation was sought, and no litigation ensued. The initial risk taken on by the Wellcome in deploying their strategy was significant; but the benefits have far outstripped any potential negative consequences whether financial or reputational.

And within this project, concerning just 432 works meticulously arranged across 30 pages of a scrapbook in the mid-20<sup>th</sup> century, we too have assumed a certain quantum of risk. At heart, this project is concerned with an academic exploration of the practicalities – and implications – of diligent search when digitising unique cultural artefacts. Many of the 432 works from our initial sample are orphans. But, for only five of those orphans did we comply with the requirements of the Orphan Works Directive; for a further five, we secured a licence for non-commercial use through OWLS.<sup>160</sup> In short, our use of the overwhelming majority of orphan works that we have made available through this resource lies outside the scope of both OWLS and the Directive. In addition, many of the works were not orphans. We identified and contacted 32 rightholders, 15 of whom (47%) agreed to our proposed use without asking for payment; notably, those granting permission for free included the Herald and Times Group (in relation to three different publications within their stable), Trinity Mirror-Sunday Mail, Newsweek and DC Thomson. However, a further seven rightholders (22%) requested fees, from £15 to \$2000. No licence fees were paid.

So, on what basis do we make the majority of orphans available? And, on what basis do we make available those works for which we have been refused permission by the rightholders? We reproduce that material in accordance with sections 29 and 30 of the Copyright Designs and Patents Act 1988 permitting use for non-commercial research and for the purposes of criticism, review and quotation. Wherever reasonably practicable, we have acknowledged and attributed the original author of the work and the original source of publication. But, as with reliance on the orphan works exception, our claim to lawful use carries risk: indeed, risk and uncertainty are structurally embedded within all copyright exceptions. A rightholder might object that our use is not critical, or does not constitute quotation or non-commercial research. They might also argue that, in any event, our use is not fair. These are known unknowns. In good faith, we consider that our use of this material falls within the scope of sections 29 and 30, but *we do not know that is does*; nor can we. And these will remain known unknowns unless and until threatened litigation proceeds to court. Only with a judicial pronouncement might the fairness or lawfulness of our activity within this project be determined definitively; only then might these known unknowns become known knowns.

And yet, there were certain risks that we were not willing or able to take; or, to put it more accurately, there were risks that the project team were not able to agree upon. Attentive readers will have noticed our concession that only a *selection of pages* from this project have been made available online. Why not all? Because of extracts from two poems that featured in our sample: *The Vanishing Ballad* by Charles Madge and *The*

---

<sup>160</sup> Licence number OWLS000057, issued on 09 June 2016.



*Clearing* by Robert Graves. The publisher was approached for permission and requested only £15 per poem for use of the extracts online. The complicating factor was that the publisher in question – Carcanet Press – is also Edwin Morgan’s publisher. With events planned to commemorate Morgan’s centenary in 2020, the University Library had already established a positive relationship with Carcanet; as such, sensitivity was required. Different options were discussed. One option was to pay the licence fee to Carcanet, given the circumstances and strategic importance of the centenary plans. Alternatively, the extracts in question might be permanently redacted from our online version of the sample pages, either on their own or in combination with all material that had been designated high risk. From the library’s perspective, it was essential not to lose Carcanet’s cooperation and support for the events in 2020 and beyond.

From an academic perspective, neither payment nor redaction on a selective basis was considered appropriate. This is an academic research project specifically concerned with rights clearance, risk, and how the copyright regime addresses the challenges of enabling access to digital cultural heritage. To deal with one rightholder differently from any of the others that refused permission without payment – however modest the fee that had been requested – would compromise the intellectual integrity of this project. The extracts in question featured in the last 12 pages of our 30-page sample. The decision was taken to make only the first 18 pages available online, just as Morgan had originally assembled them, with every work visible in colocation to each other; in this way, the integrity of Morgan’s work is also preserved. And of course, the 12 pages from our sample that are not online remain available to view in the Level 12 reading room of the University Library, along with the other 3,570 pages that make up Morgan’s scrapbooks.

## 7. BIBLIOGRAPHY

Akmon, D., (2010) 'Only with your permission', *Archival Science*, 10: 45, doi:10.1007/s10502-010-9116-z. Available at <http://link.springer.com/article/10.1007%2Fs10502-010-9116-z>

de Beer, J. and Bouchard, M., (2009) 'Canada's "Orphan Works" Regime: Unlocatable Copyright Owners and the Copyright Board'. Available at: <http://www.cb-cda.gc.ca/about-afpropos/2010-11-19-newstudy.pdf>

Bielstein, S.M., (2006) 'Permissions: A Survival Guide', Chicago: University of Chicago Press

BIS, (2013) 'Enterprise and Regulatory Reform Act 2013: Policy Paper'. Available at: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/209896/bis-13-905-enterprise-and-regulatory-reform-act-2013-policy.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/209896/bis-13-905-enterprise-and-regulatory-reform-act-2013-policy.pdf)

British Screen Advisory Council, (2006) 'Copyright and Orphan Works: A Paper Prepared for the Gowers Review by the British Screen Advisory Council'. Available at: [http://www.bsac.uk.com/files/copyright\\_orphan\\_works\\_paper\\_prepared\\_for\\_gowers\\_2006.pdf](http://www.bsac.uk.com/files/copyright_orphan_works_paper_prepared_for_gowers_2006.pdf)

Clarke, M., (1993) 'The Common Law of Contract in 1993: Is There a General Doctrine of Good Faith?', *Hong Kong Law Journal*

Covey, D.T., (2005) 'Acquiring copyright permission to digitise and provide open access to books', DLF, Council on Library and Information Resources, Washington DC. Available at: <http://www.clir.org/pubs/reports/pub134/reports/pub134/pub134col.pdf>

Deazley, R. and Stobo V., (2013) 'Archives and Copyright: Risk and Reform'. Available at: <https://zenodo.org/record/8373/files/CREATE-Working-Paper-2013-03.pdf>

Dickson, M., (2010) 'Due Diligence, Futile Effort: Copyright and the Digitization of the

Thomas E. Watson Papers,' *The American Archivist*, Fall/Winter, Vol. 73, No. 2, pp. 626-636, doi: <http://dx.doi.org/10.17723/aarc.73.2.16rh811120280434>. Available at: <http://americanarchivist.org/doi/10.17723/aarc.73.2.16rh811120280434>

Dryden, J., (2008) 'Copyright issues in the selection of archival material for internet access', *Archival Science*, 8: 123 doi:10.1007/s10502-009-9084-3. Available at: <http://link.springer.com/article/10.1007/s10502-009-9084-3?no-access=true>

de la Durantaye, K., (2011) 'Finding a Home for Orphans: Google Book Search and Orphan Works Law in the United States and Europe,' *Fordham Intellectual Property, Media and Entertainment Law Journal*, 229. Available at: <http://ir.lawnet.fordham.edu/iplj/vol21/iss2/6>

European Commission, (2012) 'Directive 2012/28/EU of the European Parliament and of the Council of 25 October 2012 on certain permitted uses of orphan works.' Available

at:<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:299:0005:0012:EN:PDF>

European Commission, (2010) 'A Digital Agenda for Europe', COM. Available at: <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52010DC0245&from=EN>

Favale, M., Kretschmer, M. and Mendis, D., (2013) 'The Treatment of Orphan Works under Copyright Law in Seven Jurisdictions: A comparative review of the underlying principles', CREATE Working Paper 2013/7. Available at: <https://zenodo.org/record/8377/files/CREATE-Working-Paper-2013-07.pdf>

Fodor, S., (2011) 'Results of CEPIC Survey on Orphan Works in Historical Archives'. Available at: [http://web.archive.org/web/20120122162827/http://www.cepic.org/news/cepic\\_news/2011/10/results\\_cepic\\_survey\\_orphan\\_works\\_historical\\_archives](http://web.archive.org/web/20120122162827/http://www.cepic.org/news/cepic_news/2011/10/results_cepic_survey_orphan_works_historical_archives)

Girija O. K. and Sudheep Elayidom, M., (2015) 'Overview of Image Retrieval Techniques', *International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE)*, Vol.4, Special Issue 1, June 2015. Available at: <http://www.ijarcce.com/upload/2015/si/icrtcc-15/IJARCCE%2019.pdf>

van Gompel, S., (2007) 'Unlocking the potential of pre-existing content: how to address the issue of orphan works in Europe?', *International Review of Intellectual Property and Competition Law*. Draft pdf available at: <http://www.ivir.nl/publicaties/download/713>

Hargreaves, I., (2011), 'Digital Opportunity: A Review of Intellectual Property and Growth.' Available at: <http://www.ipo.gov.uk/ipreview-finalreport.pdf>

Hirtle, P.B., Hudson, E. and Kenyon, A.T., (2009) 'Risk Management: How to digitise safely', *Copyright and Cultural Institutions: Guidelines for Digitisation for US Libraries, Archives and Museums*, Cornell: New York. Available for download as PDF at: <http://hdl.handle.net/1813/14142>

Kirton, I. and Terras, M. (2013), 'Where Do Images of Art Go Once They Go Online? A Reverse Image Lookup Study to Assess the Dissemination of Digitized Cultural Heritage', *Museums and the Web 2013*, N. Proctor & R. Cherry (eds). Silver Spring, MD: Museums and the Web. Published March 7, 2013. Available at: <http://mw2013.museumsandtheweb.com/paper/where-do-images-of-art-go-once-they-go-online-a-reverse-image-lookup-study-to-assess-the-dissemination-of-digitized-cultural-heritage/>

Korn, N., (2009) 'In from the Cold: An assessment of the scope of 'Orphan Works' and its impact on the delivery of services to the public.' Available at: <http://www.jisc.ac.uk/publications/reports/2009/infromthecold.aspx>

Korn, N. and McKenna, G., (2009) 'Copyright: A practical guide,' London: Collections Trust

Marques, O., (2016) 'Visual Information Retrieval: The State of the Art', *IT Professional*, vol. 18, no. 04, pp. 7-9, July-Aug. 2016, doi:10.1109/MITP.2016.70. Available at: <https://www.computer.org/csdl/mags/it/2016/04/mit2016040007-abs.html>

Niuwenhuysen, P., (2013) 'Search by Image through the WWW: an Additional Tool for Information Retrieval', published in proceedings of the international conference on Asia-Pacific Library and Information Education and Practices = A-LIEP 2013. Available at: <http://megaslides.es/doc/4053174/search-by-image-through-the-www--paul-niuwenhuysen>

Padfield, T., (2015) 'Copyright for Archivists and Records Managers', London: Facet

Pallante, M.A., (2015) 'Orphan Works and Mass Digitization: A Report of the Register of Copyrights', US Copyright Office. Available at: <http://copyright.gov/orphan/reports/orphan-works2015.pdf>

Patterson, K., (2016) 'Can I Just Google That? Orphan Works and Image Recognition Tools', in Wallace, A. and Deazley, R., eds., *Display At Your Own Risk: An experimental exhibition of digital cultural heritage*. Available at: <http://displayatyourownrisk.org/patterson/>

Pedley, P., (2015) 'Practical Copyright for Library and Information Professionals', London: Facet

Stannard, J.E., (2007) 'Delay in the Performance of Contractual Obligations', Oxford: OUP

Stobo, V., Deazley, R. and Anderson, I.G., (2013) 'Copyright & Risk: Scoping the Wellcome Digital Library', CREATE Working Paper 2013/10. Available for download as PDF at: <http://www.create.ac.uk/wp-content/uploads/2013/12/CREATE-Working-Paper-No.10.pdf>

Strategic Content Alliance, 'IPR Toolkit n.d.' Available at: <http://www.webarchive.org.uk/wayback/archive/20101001101427/http://www.jisc.ac.uk/media/documents/publications/scaiprtoolkitoverview.pdf>

Stratton, B., (2011) 'Seeking New Landscapes: A rights clearance study in the context of mass digitisation of 140 books published between 1870 and 2010', London: British Library/ARROW, available at: <http://www.arrow-net.eu/sites/default/files/Seeking%20New%20Landscapes.pdf>

Tsiavos, P., (2009) 'Case studies mapping the flows of content, values and rights across the public sector', JISC: Strategic Content Alliance. Available at: [https://sca.jiscinvolve.org/wp/files/2009/04/sca\\_2009symp\\_ipr\\_casestudies-final.pdf](https://sca.jiscinvolve.org/wp/files/2009/04/sca_2009symp_ipr_casestudies-final.pdf)

UK Intellectual Property Office, (2014) 'Consultation on reducing the duration of copyright in unpublished ("2039") works in accordance with section 170(2) of the Copyright, Designs and Patents Act 1988', Annex A. Available at: <https://www.gov.uk/government/consultations/reducing-the-duration-of-copyright-in-certain-unpublished-works>

UK Intellectual Property Office, (2014), 'Copyright Works: Seeking the Lost'. Available at: [www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/308372/consult-2014-lost.pdf](http://www.gov.uk/government/uploads/system/uploads/attachment_data/file/308372/consult-2014-lost.pdf)

UK Intellectual Property Office, (2016) 'Orphan Works Diligent Search Guidance for Applicants'. Available at <https://www.gov.uk/government/publications/orphan-works-diligent-search-guidance-for-applicants>

UK Intellectual Property Office, (2012) 'Orphan Works, Impact Assessment No. BIS1063'. Available at: <http://webarchive.nationalarchives.gov.uk/20140603093549/http://www.ipo.gov.uk/consult-ia-bis1063-20120702.pdf>

UK Intellectual Property Office, (2014) 'Orphan Works Licensing Scheme: Overview for Applicants'. Available at: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/450649/Orphan Works Licensing Scheme Overview for Applicants.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/450649/Orphan_Works_Licensing_Scheme_Overview_for_Applicants.pdf)

UK Intellectual Property Office, (2015), 'Orphan Works: Review of the first twelve months'. Available at: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/487209/orphan-works-annual-report.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/487209/orphan-works-annual-report.pdf)

Vetulani, A., (2008) 'The Problem of Orphan Works in the EU: An overview of legislative solutions and main actions in this field'. Available at: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.178.8855&rep=rep1&type=pdf>

Vuopala, A., (2010), 'Assessment of the Orphan works issue and Costs for Rights Clearance'. Available at: [http://www.ace-film.eu/wp-content/uploads/2010/09/Copyright\\_anna\\_report-1.pdf](http://www.ace-film.eu/wp-content/uploads/2010/09/Copyright_anna_report-1.pdf)

Web2Rights n.d., [Information on intellectual property rights, licensing, risk management, etc]. Available at: <http://www.web2rights.com/>

(Bibliography created 12th January 2017)



RCUK Centre for Copyright and  
New Business Models in the  
Creative Economy

College of Social Sciences / School of Law  
University of Glasgow  
10 The Square  
Glasgow G12 8QQ  
Web: [www.create.ac.uk](http://www.create.ac.uk)

