



February 4, 2013

Office of Policy and International Affairs  
United States Copyright Office  
101 Independence Ave. S.E.  
Washington DC 20559

Re: Notice of Inquiry: Orphan Works and Mass Digitization, Docket 2012-12

Google Inc. welcomes this opportunity to submit comments in connection with the above-referenced Notice of Inquiry [NOI]. We share the Office's concern that "the uncertainty surrounding the ownership status of orphan works does not serve the objectives of the copyright system," and agree that "[f]or good faith users, orphan works are a frustration, a liability risk, and a major cause of gridlock in the digital marketplace."<sup>1</sup> We are therefore pleased to see the Copyright Office initiating a new round of comments on this topic. In these comments, we will discuss efforts that Google has made to reduce the orphan works problem in recent years, the experience of RightsFlow with orphan works, the scope of this Inquiry, and some possible solutions to the problem of orphan works.

### **Google Projects that Help Alleviate the Orphan Works Problem**

Over the past few years, Google has introduced new projects that help reduce the orphan works problem. First, we digitized and indexed many volumes of the U.S. Copyright Office's Catalog of Copyright Entries [CCE] and made them searchable in Google Books.<sup>2</sup> The CCE contains information about copyright registrations and renewals, which can be starting points for tracking down the current owner of works that are in copyright. (The records are also useful for determining whether a work is in the public domain, and thus not an orphan.) Formerly, a potential user of a work who wanted to consult these records would have to travel to a library that held these volumes, or pay someone to go there on his or her behalf; now anyone with internet access can search these volumes quickly and for free.

Second, Google Images introduced search-by-image in 2011. The Images home page (<http://images.google.com/>) now allows users to drag an image file into the search box. Google Images will then return results that are identical, or similar, to the image that was dragged into

---

<sup>1</sup> NOI at 64,555.

<sup>2</sup> Using CCE volumes provided by Stanford University, we also created a special page for searching CCE records related to books in particular, from the period between September 1922 and December 1977. See <http://books.google.com/googlebooks/copyrightsearch.html>.



the box. Users can also paste a url into the search box; in that case, Images will fetch the image located at the url and conduct a search to find matches.<sup>3</sup> This feature allows a user to search the web for copies of an image, and if the results include a copy that was posted by the owner, to find the owner. It also enables an owner to ensure that his or her image is not orphaned by posting a thumbnail version of it on the owner's website. Previously, potential users of images were limited to searching by text descriptions. As the Office's *Report on Orphan Works* noted, text descriptions are often inadequate as queries when the user is trying to find a specific photograph.<sup>4</sup> Search-by-image, however, does not have the limitations of text. Our understanding is that by 2006, several vendors were offering search-by-image as part of paid rights-enforcement services packages (and that they continue to offer these services today). What has changed since 2006 is that Google Images now offers powerful search-by-image for free, to everyone. Also, Google Images will return results that are not 100% bit-for-bit matches (for example if the image has been cropped or color-shifted) and the "Similar Images" feature in Images will cast an even wider net, pulling in images that have similar overall shapes and textures but different colors and significant modifications.

These new tools help people who are trying to find copyright owners in order to get permission to make uses of works. In addition, of course, Google Search itself continues to be an invaluable tool for finding owners. It should be noted that Google also helps people access orphan works when permission is *not* needed for their intended use: for example when a person wants to find a hard-copy book to read. Google Books is one of the most important ways that people discover and locate orphan books in libraries and bookstores today. As discussed in the NOI, Google Books provides a 21st-century "card catalog" for millions of books, many of them orphans. Anyone in the world who has internet access can, for free, use this index to discover the existence of books relevant to their research, and once they have identified a book of interest, use Google Books to locate the hard-copy book in a library.

### **The RightsFlow Experience with Orphan Works**

In 2011, Google acquired RightsFlow ([www.rightsflow.com](http://www.rightsflow.com)). Among other things, RightsFlow provides a mechanical rights licensing and royalty payment platform for licensees. RightsFlow licenses millions of musical works on behalf of thousands of artists, record labels, music distributors, and digital music services. In the course of doing so, they confront the orphan works problem on a regular basis. They are often unable to locate the owner of a musical work: for example, compositions by obscure artists. RightsFlow has therefore become quite familiar with Section 115(b)(1). That section was described in the *Report on Orphan Works* as one of the already-existing orphan works provisions in the Copyright Act.<sup>5</sup> It provides, with respect to the statutory license found in Section 115: "If the registration or other public records of the Copyright Office do not identify the copyright owner and include an address at which notice can

---

<sup>3</sup> For more details, go to <http://www.google.com/insidesearch/features/images/searchbyimage.html>.

<sup>4</sup> U.S. Copyright Office, *Report on Orphan Works* (2006) [*Report on Orphan Works*] at 75.

<sup>5</sup> *Id.* at 44.



be served, it shall be sufficient to file the notice of intention in the Copyright Office.”<sup>6</sup> As the *Report* also noted, “the owner of the musical work ‘must be identified in the registration or other public records of the Copyright Office’ in order to receive any royalties under the license. Thus, a would-be user who is not able to locate the owner may make and distribute phonorecords of the orphan work according to the terms of the statutory license, provided she satisfies the conditions of section 115(b)(1) – and may continue to do so royalty-free until the owner files a registration or other record with the Copyright Office, at which point the work would no longer be an orphan work.”<sup>7</sup>

RightsFlow has relied on this provision numerous times and has found it to be useful. Because RightsFlow often needs to submit numerous notices, we applaud the Office’s recent amendments to its Section 115 regulations -- specifically, the amendments that permit notices to be submitted electronically.<sup>8</sup> We understand the electronic intake tool is in beta testing, and look forward to its full launch. We also applaud the Library of Congress’s Cataloging Distribution Service (“CDS”) ([www.loc.gov/cds](http://www.loc.gov/cds)), a web-based subscription service that enables subscribers to download the Library’s entire cataloging records. This enables users to conduct queries on their own equipment, including mass queries. We applaud this initiative, and think it’s a good model for possible future database solutions. One improvement would be to support mass queries on the Office’s website itself, without the need to subscribe to the CDS system.

RightsFlow (and indeed all licensees seeking to file notices under the Section 115 statutory license) also has a persistent problem with public ownership databases displaying inaccurate data -- including on copyright owners’ own websites. RightsFlow will send the notice required under Section 115 to the rightholder listed in a public database, only to find that the rightholder ignores the notice for a time, then responds by saying that it no longer owns the composition cited in the notice. This is a systemic problem that has plagued the industry for years. We would applaud any new incentives for rightholders to make their contact information public and accurate.

### **Not All Mass Digitization Projects Have Orphan Works Problems**

The efforts discussed above, however, have not eliminated, or even come close to eliminating, the orphan works problem. We therefore continue to believe, as we did in 2005,<sup>9</sup> that orphan works are an important problem that Congress should address. But in order to have a constructive discussion about solutions, it is important to understand the scope of the problem. The NOI asks questions about (a) orphan works generally and (b) orphan works in the context of mass digitization specifically. It is important to be clear that while some mass digitization projects will confront orphan works problems, others will not. For example, the Google Books

---

<sup>6</sup> 17 U.S.C. § 115(b)(1).

<sup>7</sup> *Report on Orphan Works* at 47-48 (citing 17 U.S.C. § 115(c)(1)).

<sup>8</sup> See 77 Fed. Reg. 71,101 (Nov. 29, 2012); <http://www.copyright.gov/fedreg/2012/77fr71101.pdf>.

<sup>9</sup> Our 2005 initial comments are at <http://www.copyright.gov/orphan/comments/OW0681-Google.pdf>. Our 2005 reply comments are at <http://www.copyright.gov/orphan/comments/reply/OWR0134-Google.pdf>.



Library Project digitizes hard-copy library books for the purpose of indexing them, and displays only short snippets of the books that are in copyright (including orphan books).<sup>10</sup> Google has always maintained that this constitutes a fair use (or in some cases *de minimis* use) -- and that position does not depend on whether a particular book is an orphan work or not. Therefore, no changes to the law need be made for this particular mass digitization project to prosper: there is no orphan works problem in the Library Project.<sup>11</sup>

We do not understand the NOI to be calling for comments on mass digitization generally, i.e. beyond the scope of the orphan works problem.<sup>12</sup> So when the NOI asks “How should mass digitization be defined[?]” (at 64,561), our response is that mass digitization should be defined, for purposes of this Inquiry, to include only mass digitization projects that do not come within existing limitations or exceptions in the Copyright Act.<sup>13</sup> Unless that distinction is made, this Inquiry risks becoming an ill-defined referendum on fair use, intermediate copying, and *de minimis* use, rather than a useful examination of new solutions to the orphan works problem.

That said, we would welcome a separate Inquiry into legal issues in mass digitization generally, provided that it focused on the most pertinent issues. In our view, orphan works would not be the most important issue in such an Inquiry, at least with respect to the mass digitization and full display of books. The issue of disputed rights, for example, likely has greater significance. Authors and publishers often disagree on the interpretation of old book contracts<sup>14</sup> and there are few judicial opinions that set out definitive guidelines. Where the trail of ownership leads definitively to either the author or the publisher (although they dispute between themselves which one), the work does not qualify as an orphan and thus would not be affected by any legislative changes directed to orphan works. Moreover, the disputed rights problem affects more popular books than the orphan works problem does, because contract difficulties plague popular as well as unpopular books indiscriminately, whereas orphan works are often orphaned because they are of little interest to most people. Therefore, if the Office is interested in removing the legal obstacles standing in the way of mass book digitization projects that do not rely on existing

---

<sup>10</sup> For more information, see <http://www.google.com/googlebooks/library/index.html>.

<sup>11</sup> Our 2005 comments pointed out that *additional* uses of the orphan works in the Library Project corpus would be greatly beneficial to the public, but that should not cause confusion with respect to the *existing* uses in the product.

<sup>12</sup> For example, the second “Subject of Inquiry” in the NOI is titled, “Orphan Works in the Context of Mass Digitization.” NOI at 64,560-61.

<sup>13</sup> The Office should therefore *not* adopt the following definition, which appears in the Office’s *Legal Issues in Mass Digitization [USCO Analysis]* (at 8-9):

“[M]ass digitization’ is not a scientific term. In the context of books, it has come to mean large-scale scanning. It may also refer to a systematic methodology or approach. There seems to be a consensus that the Google Books project, which has scanned and digitized more than 15 million books from research libraries and continues to scan more at a rapid pace, qualifies as a mass digitization.”

As we explain above, it is our view that the digitization efforts behind the existing Google Books Library Project do not require any changes in the law.

<sup>14</sup> See *USCO Analysis* at 18, 28.



exceptions and limitations to copyright,<sup>15</sup> it might consider conducting an Inquiry into ways to solve the disputed rights problem.

### **Possible Solutions for Mass Digitization Projects that Have Orphan Works Problems**

While there are mass digitization projects that have no orphan works problems because they come within existing limitations and exceptions to copyright, there certainly can be others that do not come within those limitations and exceptions, and thus may have orphan works concerns. Moreover, orphan works can pose difficulties for efforts that do not involve mass digitization (for example, restoration and duplication of old personal photographs by retail photo finishers<sup>16</sup>). For these reasons, a solution to the orphan works problem would be a welcome development, both for those considering large-scale digitization efforts, as well as those who encounter orphan works in smaller-scale projects.

As we said in our 2005 comments, we think that any orphan works solution suitable for mass digitization projects that do not come within existing limitations and exceptions in the Copyright Act should meet these minimum criteria:

- The criteria for a qualifying owner search would have to be objective: for example a query of a specific database or set of databases (which could be operated by the Copyright Office or by a private entity). In order to be able to control risk for large-scale projects, the would-be user needs to have a clear and definite stopping point for the searches.
- The owner information databases must support automated searching.
- Once a qualifying search has been performed, but no owner found, there should be limits on remedies for use of the orphan work.
- Both for-profit and not-for-profit uses should be encompassed.
- Use of US works as well as non-US works should be encompassed because it is often the case that the origin of the work cannot be ascertained from the work itself.

While there are a range of ways to satisfy these criteria, the Copyright Office's proposal in its *Report on Orphan Works* and the legislation introduced in the 110th Congress are not in that range. In particular, the conception of "reasonably diligent search" used in those proposals is incompatible with the needs of mass digitization efforts, as it does not contemplate automated searches that meet objective, pre-determined criteria. Accordingly, if the Office chooses to return to this issue, we hope it will revisit the criteria suggested above.

---

<sup>15</sup> See *id.* at 14 ("While scanning and electronic dissemination of books could undermine the benefit of [the Copyright Act's] term of protection, mass digitization and dissemination also may serve important public interest goals that justify restricting or limiting certain exclusive rights for works that are subject to a lengthy copyright term.").

<sup>16</sup> See *Report on Orphan Works* at 24-25 (citing comments of Wal-Mart Stores, Inc.).



## A More General Approach to the Orphan Works Problem

There may also be a more general approach to the orphan works problem that could resolve much of the orphan works problem for both mass digitization and smaller scale efforts:

- Amend Section 412 to say that statutory damages and attorney fees are not available in a case where the infringement occurred at a time when the plaintiff owner's contact information was not current in the registration records of the Copyright Office. Section 506 should also be amended to eliminate criminal liability when these conditions are met.
- Amend Chapter 7 to require the Office to make the contact information of owners accessible publicly, online, in a machine-searchable format.

The practical effect of this would be that potential users could locate owners of copyrighted works easily, or in cases where they could not locate them easily, they could use works with only the threat of actual damages and profits hanging over them. (Or perhaps the Act could be further amended to provide a remedy of "reasonable royalty" in this case.) For many projects, actual damages and proportional profits might be negligible, and in other projects, the project sponsor might be perfectly amenable to paying appropriate damages or profits to a later-appearing owner. This system would also comply with the Berne Convention because infringement damages would still be available -- and statutory damages are not required under the Convention. Indeed, as discussed above, the current Copyright Act already has a system very much like this for non-dramatic musical works: the records of the Office must "identify the copyright owner and include an address at which notice can be served" or the owner is not entitled to any royalties under the Section 115 license. Our proposal would not go as far as Section 115 does, because the owner would still be entitled to ordinary damages (rather than nothing).

The Office would of course have to shoulder the logistical burdens of digitizing the registration records for all in-copyright works and of building a database around them that could handle automated mass queries, but a big part of that work is already being done: the post-1978 records have long been digital and, as the NOI notes (at 64,558), the Office is in the process of digitizing the pre-1978 records already. Adding this benefit to registration will incentivize more registrations, which will not only be good for the copyright marketplace, but also good for the Office and the Library of Congress (because of the deposit requirements).

We are open to exploring other solutions as well. The core issues of orphan works -- owner information and damages reform -- are broad enough that many different approaches could bring about good results.

\* \* \*



We applaud the Office's efforts in this area, and look forward to continued engagement on this topic.

Respectfully submitted,

*/s/*

Oliver Metzger  
Senior Copyright Counsel  
Google Inc.