

February 4, 2013

Via Electronic Submission

Maria Pallante
Register of Copyrights
United States Copyright Office
Library of Congress
101 Independence Avenue, S.E.
Washington, D.C. 20559-6003

Re: *Comments of Microsoft Corporation in Response to Copyright Office's Notice of Inquiry re Orphan Works and Mass Digitization*,
Docket No. 2012-12, 77 Fed. Reg. 64555 (Oct. 22, 2012)

Dear Ms. Pallante:

Microsoft Corporation ("Microsoft") respectfully submits these comments in response to the Copyright Office's Notice of Inquiry dated October 17, 2012, published in the *Federal Register* on October 22, 2012, 77 Fed. Reg. 64555 (the "Notice of Inquiry"). For nearly a decade, Microsoft has been actively engaged in and has supported efforts by the Copyright Office, Congress and others to devise a solution to the orphan works problem. Accordingly, we welcome and encourage the latest inquiry.¹

Since the Copyright Office last considered orphan works, the problem has not gone away. It remains an issue with significant implications for the development of technology, the fostering of creative activity and the preservation of our cultural heritage. The problem of orphan works is in need of a solution that both protects rights holders *and* avoids works lying fallow, thereby maximizing the public benefit of the copyright system.

¹ Microsoft is a global technology leader that develops, produces and distributes devices, software and services around the world. As both a creator and consumer of content, Microsoft has built its business around copyright and has a strong interest in a well-functioning, balanced copyright system. Microsoft has long been an active participant in the public discussion regarding orphan works, including the company's response to the 2005 Notice of Inquiry by the Copyright Office; the 2006 testimony of Associate General Counsel Thomas C. Rubin before the Senate Subcommittee on Intellectual Property; and through filing objections to the proposed settlement of the *Authors Guild, Inc. v. Google Inc.* litigation.

Developments over the past several years have brought into focus specific themes that should be reflected in any orphan works inquiry. Our comments will focus on four of those themes:

1. The promise and benefits of mass digitization are enormous.
2. The orphan works problem remains an impediment to mass digitization.
3. Additional solutions to enhance the 2008 legislative proposal may be needed.
4. Any solution must be structured to further the *public* interest and not merely benefit private parties motivated by self interest.

A solution to the orphan works problem that addresses these issues will be an important step toward better defining the relationship between copyright and creativity in the digital age. To serve the public interest, the ideal solution should come in the form of legislation, as several countries have recently enacted.²

1. The promise and benefits of mass digitization are enormous.

Since the Copyright Office's Report in 2006, mass digitization has come to the forefront of the orphan works discussion, bringing with it tremendous new opportunities. Efforts to collect, digitize and make searchable large collections of physical or analog works – such as books, photographs, films and sound recordings – have the potential to offer tremendous public benefit. Through mass digitization, works that have long been lost or inaccessible to all but a handful of archivists and researchers can be made available to the public in a central repository. Mass digitization can play an important role in helping to preserve works for future generations by greatly reducing the risks of decay and destruction. But that is just the beginning. Providing the public with the ability to search and analyze large volumes of digital works enables scholars, scientists, journalists, educators, students and others to gain new insight and knowledge not possible when forced to track down and view each work individually.

For a technology company like Microsoft, reducing the legal uncertainty around mass digitization of orphan works has the power to increase the speed and creativity of

² In the European Union, for instance, Directive 2012/28/EU permits nonprofit uses of audiovisual and printed works after conducting a diligent search, with limited damages if the copyright owner should reclaim his or her work. In Canada, the Copyright Board has discretion to issue licenses for orphan works after a reasonable search. Under Korean law, a diligent search triggers compulsory licensing, and Japan has taken a similar approach. *See also* Notice of Inquiry at 64560 (describing orphan works legislation in France, Hungary, Denmark and Finland). While each of these systems differs in certain respects and may not be the right solution in the United States, these countries have all recognized and responded to the problem of orphan works. The United States must do the same or else risk being left behind.

innovation in ways that are beneficial not just to the company but to society at large. Our experience in creating an online English/Haitian Creole translator following the devastating Haiti earthquake of 2010, which is detailed in a Microsoft Research blog post excerpted below, underscores the issue in concrete terms:³

When aid efforts began after the recent Haiti earthquake, a request came to the Machine Translation team within Microsoft Research’s Natural Language Processing (NLP) group from Microsoft volunteers involved in the community supporting assistance in Haiti: Was there a quick way to deliver an online English/Haitian Creole translator? . . .

Normally, adding a new language to the machine-translation engine can take weeks, if not months. . . . The NLP team knew that its biggest challenge would be identifying parallel data between English and Haitian Creole for training the engine. . . . But team members quickly replaced skepticism with dogged determination and reached out for help. That was when they discovered other groups who had made language resources available.

“For instance,” [Microsoft researcher Chris] Quirk says, “Carnegie Mellon University had a repository for parallel Haitian Creole and English spoken and text data. Government agencies released parallel documents and glossaries, and Web sites such as CrisisCommons and haitisurf.com were happy to share glossaries and translation resources.”

Such assistance was invaluable.

“If not for the efforts of the community, who made data and dictionaries available with minimal license restrictions,” [Microsoft senior product manager Vikram] Dendi says, “this Haitian Creole machine translator would not be available.”

Yet access to a broader range of digitized works, including orphan works, could have resulted in the development of a more refined translation engine even faster. Indeed, after the release of the first version of the translator Microsoft put out a call for more data so it could continue to improve the technology. The creation of the English/Haitian Creole translator underscores both the transformative benefits of mass digitization and the way in which the orphan works problem inhibits that potential.

³ Janie Chang, *Translator Fast-Tracks Haitian Creole*, MICROSOFT RESEARCH, Feb. 4, 2010, <http://research.microsoft.com/en-us/news/features/haitiancreole-020410.aspx>.

2. The orphan works problem remains an impediment to mass digitization.

One of the lessons learned in the Copyright Office's 2006 Report was that the orphan works problem places a strain on the copyright system, as ordinary citizens who would benefit from the use of orphan works are often frustrated by copyright law's lack of a sensible solution for their uncertainty about the status of a work. The copyright system remains similarly under strain today, as information-sharing technology has become even more convenient and prevalent, leading to an even greater disconnect between what people reasonably expect to do with works and what copyright law allows them to do.

A significant part of that strain could be resolved by addressing orphan works, particularly in the context of mass digitization. What makes mass digitization beneficial – the broad scope of works being brought together – is also what makes it so difficult to implement under existing law. There is no clear way to identify who owns the copyright in a given work. Further, even if a prospective user conducts an exhaustive search and finds nothing, the copyright owner might still appear sometime in the future to object to and demand compensation for (and/or cessation of) the use. Where mass digitization is involved, the cost of such searching and uncertainty is multiplied by a factor of thousands, if not millions. As the Notice of Inquiry correctly observes, “the works may in fact have copyright owners, but it may be too labor-intensive and too expensive to search for them.” Notice of Inquiry at 64557.

Put simply, it undermines the credibility of the entire copyright system for the law to continue to impose on responsible libraries and archives these risks and costs where so many works are at issue. Most importantly, not addressing the problem inevitably means foregoing many of mass digitization's benefits to the public.

3. Additional solutions to enhance the 2008 legislative proposal may be needed.

The Copyright Office's 2006 Report focused primarily on situations where someone wishes to use only a single orphan work or a small number of orphan works, such as making reprints of an old family photograph or incorporating archive footage into a documentary film. See United States Copyright Office, REPORT ON ORPHAN WORKS 24, 36-40 (2006). As the Notice of Inquiry observes, the issue of orphan works raises very different policy considerations in the context of mass digitization than in the context of such occasional or isolated uses. See Notice of Inquiry at 64557. To reflect this reality, orphan works legislation could take a dual-option approach, where the user's potential liability depends on the type of use. At the start of this Inquiry, we pass along these preliminary ideas for further study and exploration.

Diligent Search Standard for Isolated Uses: In 2008, in response to the Copyright Office's Report on Orphan Works, Congress considered (but did not pass) legislation that would have limited remedies for copyright infringement, on a case-by-case basis, where the user of the orphan work was able to show that he or she first conducted a good-faith, reasonably diligent search to locate the copyright owner. *Id.* at 64556. This approach might

continue to be a viable solution for isolated uses, such as those involving only one or a small number of orphan works (thus minimizing the burden of searching) and which often involve significant investment by the user (for example, in creating a new work).

Conditional Safe Harbor for Public-Minded Mass Digitization: As described above, requiring a diligent search for each work included in a mass digitization project would, in effect, make that project impractical, if not impossible. It simply doesn't scale. Accordingly, any orphan works legislation must specifically address mass digitization so as to encourage such projects where socially valuable.

One possible solution is a conditional safe harbor for mass digitization. The touchstone of the safe harbor could be that the institution creating digital copies must make them broadly available, in the manner of a public library or museum, to all users: individuals, companies, scholars, competitors, everyone. This would ensure that truly orphan works – which belong to no one – don't become "privatized" and that such projects provide maximum benefit to the public at large. The Copyright Office could explore whether public availability could be furthered by requiring deposit of digital copies of orphan works with the Library of Congress, to ensure fulfillment of obligations regarding public access and to enrich that national treasure.

In both scenarios – isolated use and mass digitization – Congress should focus on ways of limiting the remedies available against qualifying users of orphan works. This is likely the most flexible and least disruptive solution to the orphan works problem, as it minimizes the effects of legislation on the existing structure and substance of the Copyright Act. The safe harbor provisions of the Digital Millennium Copyright Act (DMCA) offer one example of how such an approach might be implemented. See 17 U.S.C. § 512.

To be sure, mass digitization can present substantial challenges to authors and other copyright owners that are different from the effects of individualized uses. While the diligent search standard goes a long way toward protecting authors and creators with respect to those individualized uses, other mechanisms may need to be developed and adopted for mass digitization projects. For example, a robust notice and takedown requirement to address a complaint from a resurfacing author could be a condition of the safe harbor, and the safe harbor might apply only to older copyrighted works.

Similarly, it is important to recognize that books are not the only works capable of mass digitization. Many other categories of orphan works, including photographs, film and sound recordings, can be digitized and made available for public access. Yet these different categories may raise somewhat distinct practical considerations, such as the ease of determining the author of the work or when the work was created. Accordingly, orphan works legislation should be flexible enough to accommodate these differences when applying the safe harbor requirements. Given these challenges, which must be weighed against the huge societal benefit to solving the orphan works problem, it may be necessary to consider legislation that addresses only certain categories of works.

Finally, as Microsoft has stated previously, the development and availability of registries of works would be highly beneficial and should play a role in any solution. Along these lines, the parties to the Google Books litigation gathered an enormous amount of data in building what would have become the Book Rights Registry – data that could prove helpful in the effort to craft orphan works legislation. For example, data on the number of authors and publishers who did or did not claim works could lead to more accurate estimates of the scope of the problem that currently exists. Unfortunately, as far as we are aware, this data has not been made available to anyone other than the parties to the litigation. Microsoft encourages those parties to find a way to make that data available to the Copyright Office and others as part of this Inquiry.

4. Any solution must be structured to further the *public* interest and not merely benefit private parties motivated by self interest.

Since Congress last considered orphan works legislation in 2008, the issue has received significant attention outside the legislative branch, particularly in the consideration of the proposed settlement of the Google Books litigation, the decision in the Authors Guild’s lawsuit against HathiTrust and the increased awareness that organizations engaged in mass digitization, such as the Internet Archive, have received as a result. This attention has provided further insight into potential solutions for the orphan works problem. While there may be disagreement about the best form of legislation to address orphan works, it has become clear that setting public policy through litigation and settlement agreements between private parties is not the answer, as it cannot properly balance the interests of – or provide solutions for – the many diverse stakeholders outside the litigation. As the Copyright Office, the U.S. Department of Justice and the District Court overseeing the case all found, the proposed Google Books settlement, which would have given a prospective license to orphan works to a single private party, provides a stark example of a solution that was *not* in the public interest.

Indeed, the Copyright Office acknowledged the need for legislation and echoed concerns raised by Microsoft and others that the proposed Google Books settlement “would encroach on responsibility for copyright policy that traditionally has been the domain of Congress.” *Hearing on Competition and Commerce in Digital Books: The Proposed Google Book Settlement*, 111th Cong. 2 (2009) (statement of Marybeth Peters, Register of Copyrights). In rejecting the settlement, the court concluded that these important questions – including “who should be entrusted with guardianship over orphan works, under what terms and with what safeguards” – are best left for Congress to decide. *See Authors Guild, Inc. v. Google Inc.*, 770 F. Supp. 2d 666 (S.D.N.Y. 2011). This Inquiry should keep those teachings in mind and ensure that the broad public interest is served by any solution to the issue.

Conclusion

Microsoft supports this Inquiry and the goal of finally resolving the issue of orphan works. We are committed to pursuing a solution for orphan works that takes into account the interests of *all* stakeholders, including copyright owners, users of orphan works and the public more broadly. Events over the past several years in the United States and

abroad have made clear that an orphan works solution has the potential to unleash huge benefits from a wide array of potential uses, ranging from individual remixes to mass digitization. Whatever form the solution ultimately takes, it should directly address mass digitization of orphan works as well as isolated uses, and be one of broad applicability that benefits the public interest at large. Inaction is no longer a viable solution to the problem.

Microsoft appreciates the opportunity to provide these comments in response to the Notice of Inquiry and would be happy to provide additional information or testimony if that might be useful to the Copyright Office.

Sincerely,

Thomas C. Rubin
Chief Intellectual Property Strategy Counsel