



Roylance, Abrams, Berdo & Goodman, L.L.P.  
Intellectual Property Attorneys

Megan E. Gray  
202-530-7365  
[mgray@roylance.com](mailto:mgray@roylance.com)  
February 28, 2007

## **The Case for an Image-Recognition Database at the Copyright Office**

### *Executive Summary*

The fundamental problem that “orphan work” legislation poses for visual-art industries is that, because there is currently no practical way to effectively search visual art, much of it will be inadvertently deemed orphaned, to the great economic detriment of these core American enterprises (textile, architecture, computer graphics, product packaging, greeting cards, ceramic tile, dishware, jewelry, etc.). The only way to avoid this result is if there is an image-recognition database. The question then is, should the Copyright Office undertake this responsibility or should the copyright owners themselves? The Copyright Office is, by far, the better candidate, for a number of reasons:

- The database must be comprehensive across visual-art industries. Visual art is not medium-specific (e.g., a gift-wrapping paper design could have been knocked off from a ceramic dishware piece).
- The database must be populated with previously created copyrighted works, which currently exist only in non-digital form. A company like Google can digitize the millions of copyrighted works in the Copyright Office’s possession, which are already indexed and aggregated, but such a company won’t endeavor to travel to each of the millions of individual textile mills, graphic art studios, and greeting card companies, scattered all over the United States, to digitize their disheveled collections.
- In the event that money is needed to develop the requisite database, the Copyright Office can obtain funds for it, unlike private enterprise.
- With a government agency working with a technology company, the database is likely to be completed faster and more efficiently than if hundreds of different industry groups pull competing technology companies in a variety of different directions.
- Visual-art companies rarely have technological or large-scale, white-collar project management experience, while the Copyright Office does.
- Only the Copyright Office can integrate the database with the Copyright Office’s online registration process, which is scheduled to debut this summer.
- A Copyright Office database is permanent, while a private database could go out of business or change its business model at any time.
- With an image-recognition database, the Copyright Office is transformed into a repository that provides a concrete public benefit – the registration system as it exists now does not.

The factual and political reality is that visual-art enterprises cannot – simply cannot – develop the necessary image-recognition technology, collect all their previously created works, digitize them, and submit them to the database, without massive and probably insurmountable financial and employee dislocations. But the Copyright Office can.

## *The Orphan Work – Visual Art Conundrum*

Orphan Works legislation is easily summarized -- if someone makes a good-faith effort to find the copyright owner but is unable to do so, then he can go ahead and use the copyrighted work without fear of a large money judgment later on if the copyright owner sues him for infringement. There is nothing objectionable about this scenario as a theoretical matter. The problem lies in the practical application of the legislation to visual art.

The core problem, one that is completely unavoidable, is that there has been no real way to search visual art. You cannot run a Google search on the design from your blouse, greeting card, wallpaper, or what-have-you. You cannot view visual art registered with the Copyright Office from its website. You cannot walk into the Library of Congress (which houses the Copyright Office) and ask to see all depictions of tigers registered in the last 30 days, or during 1990, or otherwise. Abstract art can't be accurately or uniformly described in text. There is no Dewey Decimal system for indexing visual art. The majority of published works of pictorial, graphic, and sculptural art are disseminated either without attribution (contrary to the owner's efforts) or with attribution information easily (and often) removed.

Because there is no practical way to search visual art, the end result is that the majority of visual artwork is likely to be deemed orphaned. In other words, as far as visual art is concerned, almost any search is likely to be deemed to be reasonably diligent, even if that search has essentially a zero chance of actually identifying the copyright owner. For all the designs currently or previously exploited by copyright owners, Orphan Works legislation makes them the baby thrown out with the bathwater.

A real-world example is the best illustration -- if you would like to copy the art from a ceramic dishware set onto wallpaper, you could physically examine the dishware for a copyright notice. However, there is a significant chance that the dishware doesn't carry a copyright notice because it is a knock-off from China, whose manufacturer originally stole the design from a graphic design company (unknown to you). Unable to identify the copyright owner from the dishware itself, you perhaps also review the websites of a few dishware retailers, but you don't see the at-issue design, which is unsurprising because that design is a few seasons old and is no longer advertised or cataloged, or because the graphic design company sold or licensed the art to a textile company and not a dishware company.

In short, the basic premise of the Copyright Office's Report of Orphan Works is flawed. The Report expressly, and erroneously, states that the copyright owner can almost always be found if he wants to be found. The Report also specifically, and incorrectly, states that, if a work has been registered at the Copyright Office, it will not likely be deemed an orphan work. For visual art, those statements are simply not true. In the example above, the at-issue design could be splashed across the homepage of the graphic design company, emblazoned on all its business cards, SEC filings, and catalogs, and lauded in a San Francisco art gallery, and registered for many years at the Copyright Office - yet that still would not increase the likelihood that you would be able to link the design to the copyright owner.

### ***Balkanized Private Databases Would Only Exacerbate the Orphan Works Problem***

The only possible manner in which we can avoid inadvertently casting most or all visual art into the orphan works net is if there is a centralized, searchable, image-recognition database. The Copyright Office, as the repository for visual art since the 1800's, is the natural location for such a database. However, the Copyright Office does not want this responsibility, and instead opines that private industry segments will spontaneously form and create suitable databases. Even if one were to assume that somehow all of the competitors in each visual-art industry (textile, wallpaper, jewelry, illustrators, photographers, cartoonists, stained glass, etc.) would actually form heretofore non-existent associations that would develop, populate, and launch image-searchable databases of their members' artwork (a coalescence that borders on the preposterous), you remain on the horns of the dilemma, because the ultimate problem is still not solved: the user who wants to identify the copyright owner of a visual work is still unable to do so.

The dilemma remains because the Copyright Office overlooks two fundamental characteristics of any effective visual database. First, visual art is not medium-specific. A greeting card design could have been stolen from a computer screensaver, for example. The user does not know what industry the designer occupies; the user only knows what industry his copy of the design (which very well may be a knockoff) occupies. In order to conduct a diligent search, the user would need to identify and search each of the hundreds or thousands of visual-art industry association databases, and no doubt pay a significant search fee to each database in order to conduct the search. This system would be bad for users legitimately trying to find owners and bad for copyright owners themselves.

Second, it is not enough to simply create the database (a difficult enough task in and of itself). That database has to be populated with millions of designs that only exist in a physical dimension (not electronically)(e.g., textile, jewelry, greeting cards), and many of which are difficult to locate, out of stock, or are stored only at the Copyright Office. (Many non-business people erroneously assume that every company has a catalog of all its creations, forgetting that is usually only done at national, retail levels). For example, paper mache sculpture and ice sculpture rot or disintegrate over time. This visual art is referred to as "fugitive" media, and once the copyright owner has memorialized the work pursuant to the Copyright Office's specifications (a photograph usually) and registered the work, it disappears, and the only record of its existence is held by the Copyright Office. Along similar lines, companies strive to sell all their product so that they are completely out of stock, such that a textile company does not retain samples of a fabric design after registering it with the Copyright Office. Insisting that private industry populate a database with what is currently in their possession only means that there will be many gaps in the database, and like a canoe with many holes, such a database isn't particularly useful to anyone.

### ***Copyright Office Database as a Panacea***

Orphan Works legislation can work in the context of visual art if, but only if, there is a centralized, image-searchable database populated with previously created designs and constantly updated with newly created designs. Who is best tasked with this database responsibility? The unavoidable answer is the Copyright Office. Consider the following:

- Cross-Industry Central Repository: The Copyright Office is the only place where a large amount of visual art, crossing all visual-art industry segments, is centrally collected.
- Commercially Exploited Central Repository: The Copyright Office is the only place where the most commercially valuable copyrighted works in recent times are located. In some instances, a company trying to re-assemble all the artwork that it created/registered over the last 30 years could only obtain them from the Copyright Office, because a commercial enterprise generally does not stock products that are not in its current season product line. (For example, a textile company may have saved the generic description, color palette numbers, substrate material, weaving screen, etc. for a particular design, but the depiction of the design itself may only be in the Copyright Office records).
- Organized Central Repository: The Copyright Office is the only place with a uniform and meticulous indexing and organizing system for all the items in its possession. Unlike most small businesses (which comprise the vast majority of all American enterprises), which often use cramped and disheveled closets, docks, and bins for storage, the Copyright Office has an extensive warehouse where it can generally find whatever it is searching for. Moreover, the Copyright Office stores these items in a secured location, where materials cannot be easily stolen, vandalized, or harmed by atmospheric conditions.
- Organized Central Repository Appeals to Free Digital Archiving Groups: As a result of all these characteristics, there are a variety of companies or associations that would be immensely interested and excited at the prospect of gaining entry to the Copyright Office's warehouse and making digital copies of all the copyrighted works (which incidentally are all subject to disclosure as public records) and collecting them into an electronic database. Indeed, Google would likely be willing to do this without charge, for free.

Companies are already deeply engaged in similar archiving projects. Google Books' Library Project, for instance, has teamed up with numerous major libraries across the country to digitize their collections, providing on-line access to 34 million volumes from over 100 libraries nationwide. The Open Content Alliance, with backing from Microsoft and Yahoo!, has currently digitized over 100,000 volumes, with literally millions more to follow. Contributing institutions include Smithsonian Institution Libraries, American Museum of Natural History, Boston Public Library, Columbia University, Johns Hopkins University Libraries, San Francisco Public Library, National Archives (United Kingdom), National Library of Australia, University of California, The University of Chicago Library, University of Virginia, HP Labs, MSN, the Xerox Corporation, among many others.

Critically, Google or the Open Content Alliance would be willing to do this because the material is all organized and centrally located – these groups will go to one place to digitize the material but will not go to 9,000,000 individual textile, wallpaper, greeting card, wrapping paper, jewelry, architecture, carpet, ceramic tile, dishware, and handbag companies scattered all over the country. Both the standard “private industry” model and the “domain name” model of a copyright database completely ignore the problem of how to populate the database with previously created designs without crippling small businesses and copyright owners with massive employee and technology diversions.

By partnering with these digital archiving groups, the Copyright Office could obtain a broadly populated, searchable, image-recognition database with probably very little effort or expense on its own part, and without unfairly penalizing or burdening copyright owners or copyright users – and it could be done quicker than if left up to a million different industry segments and competing software providers, all working on unharmonized versions of the same wheel.

- Public Funding Available: The Copyright Office has access to government funds, if needed. The Copyright Modernization Act of 2006 (from the 109<sup>th</sup> Congress, introduced by Rep. Lamar Smith) appropriated \$12 million dollars for law enforcement of intellectual property crimes. The Copyright Office database expense would be small potatoes compared to this. Moreover, it is possible that, with the cooperation of the digital archiving groups, the Copyright Office will not incur any significant expense whatsoever. In contrast, in a private database system, it is likely that either copyright owners will face exorbitant fees in order to be included in the database or that users will face steep prices in order to search the database. Either users or copyright owners will have to bear the financial burden. With the Copyright Office acting as the ringleader, it will eliminate the need for an undue burden to be borne by either users or owners.
- Accelerated Software Development: Image-recognition software is not yet fully developed but it will be in the near future, especially if the Copyright Office pushes for its development and/or digital archive companies like Google or the Open Content Alliance have an even bigger incentive to fast-track the development. Within a decade, it is likely that the Copyright Office could even buy “off-the-shelf” software, not requiring any customization, in order to have a searchable, image-recognition database.
- Sophisticated Project Managers Overseeing Development: The Copyright Office employs many white-collar professionals with advanced degrees and training, who are more than experienced enough to oversee the outsourcing of a potentially unwieldy project. This is in marked contrast to the majority of visual artists and their employers, most of whom are enterprises with only a few employees. It is preposterous for the Copyright Office to assert, in essence, “We don’t have the ability” to undertake this project, but then point a finger at a 24-person textile mill with no technology expertise as being more capable of creating a massive electronic database.
- Integration with Online Electronic Copyright Registration: The Copyright Office would not likely need significant computer programming in order to link up the contemplated database with the Copyright Office’s online electronic registration system, set to debut this summer. The system would be primed for automatic dual public benefits – a copyrighted work could be simultaneously registered as well as included in the image-searchable database.
- Modeled on Internet Superhighway Development: A collection of hundreds or thousands of private databases is the wrong solution to most “collective action” problems like those posed by intellectual property protections, which have both public and private benefits. The “collective action” analysis holds that the market fails when individual consumer rationality and firms’ profit-seeking do not lead to efficient provision of public goods (i.e., where another level of provision would provide a higher utility at a lower cost). (See [http://en.wikipedia.org/wiki/Collective\\_action](http://en.wikipedia.org/wiki/Collective_action)). This is why roads and highways –

much like the Internet superhighway – must be developed, at least initially, by the government.

- Fulfills Core Copyright Office Mission: Even the Copyright Office will acknowledge that the key question is "how does the registration system benefit the public?" Currently the registration system is of zero benefit to the public. The Copyright Office recognizes that the registration system is meaningless and of no public benefit without image-search functionality. Congress itself, in passing The E-Government Act of 2002 (see 44 U.S.C. Section 101 et. seq.), recognized that modern government means that agencies need to make their records Internet-accessible.
- Permanent Record of Visual Creations: The Copyright Office and its associated records and database are more permanent than any collection that could be created by a private entity, which could become bankrupt or completely change its revenue model, dependent as it is on market conditions.

### ***The Bottom Line – Factual and Political Reality***

The unavoidable factual and political reality is that visual-art companies cannot collect previously created visual art without severe business disruption (to the point of complete business cessation), cannot digitize scattered and unorganized tangible products without massive financial investment, and will not pay huge sums of money and divert a large percentage of their workforce to re-submit these copyrighted works to an as-yet-to-be-formed database when all these works were already submitted to the Copyright Office at considerable expense. Any Orphan Works legislation must tie the effective date to an image-searchable database, which in turn must incorporate images previously registered with the Copyright Office.

If the Copyright Office is handed this database responsibility, we can get a centralized, searchable, image-recognition database, one that includes previously registered art as well as art that is created in the future. And we can get it for a fraction of the cost and time that would be entailed in the numerous Balkanized databases, each riddled with gaps in submissions, which would arise (assuming they arise at all) if this project is left to private industry. And this contemplated database fulfills the core mission of the Copyright Office itself. In order to accomplish the database goal, all that the Copyright Office would need to do is open its doors to an electronic archiving enterprise (e.g., Google), contract out the continued development of image-recognition software, and slightly modify its online registration system.