UNITED STATES COPYRIGHT OFFICE

Report and Recommendations of the Technical Upgrades Special Project Team

OFFICE OF THE CHIEF INFORMATION OFFICER

FEBRUARY 2015

UNITED STATES COPYRIGHT OFFICE

Report and Recommendations of the Technical Upgrades Special Project Team

OFFICE OF THE CHIEF INFORMATION OFFICER

FEBRUARY 2015

United States Copyright Office
Library of Congress · 101 Independence Avenue SE · Washington, DC 20559-6000 · www.copyright.gov

February 18, 2015

Dear Register Pallante:

I am pleased to deliver a report and final recommendations from the Technical Upgrades Special Project Team. The project team was charged with assessing technological functionality and business strategies related to Copyright Office services. I chaired this project in my capacity as the first Chief Information Officer of the Copyright Office.

The project team performed significant research and analysis on important concepts and technologies that would optimize key services for customers and learned a great deal from our interviews with stakeholders throughout the copyright community. We also solicited written comments from the public.

While the report contains many recommendations, our current IT environment presents a substantial limitation and we will need to work carefully with the Library of Congress to reconcile enterprise solutions with existing infrastructure and architecture. It seems unlikely, if not impossible, that the Copyright Office can administer the law effectively into the future without technological changes that reflect the importance and growth of its mission.

Thank you for entrusting me with this important project. I hope that the report will serve as a valuable resource for the Copyright Office as we consider future strategies.

Yours sincerely,

Douglas P. Ament

Douglas P. Ament
Chief Information Officer

Table of Contents

Executive Summary **4**

1. Background **12**

2. Project Objectives **12**

3. Project Approach and Methodology **13**

3.1 Intelligence Gathering and Analysis *13*

3.2 Resource Utilization *14*

4. Participating Parties **14**

5. Findings: Additional Service Recommendations **17**

5.1 Challenges with Current User Experience *17*
5.1.1 Feedback from External Users 17
5.1.2 Feedback from Internal Users 22
5.1.3 Additional Ideas and Features 24

5.2 Challenges with Existing Public Record *27*

5.3 Inadequacies of Current Copyright Data *29*
5.3.1 Data Exchange 29
5.3.2 Data Model and Data Standards 31
5.3.3 Data Repository 34

5.4 Outdated Architecture & Infrastructure *35*

6. Enabling Technologies **37**

6.1 Technologies *37*
6.1.1 Application Programming Interface 37
6.1.2 Cloud Computing 39
6.1.3 Integrated Solutions / Business-Driven Ecosystems 44
6.1.4 Service Oriented Architecture 45
6.1.5 Mobile Computing 47
6.1.6 Big Data 49
6.1.7 Data Analytics 50

6.2 Deployment Challenges *51*

6.3 Adoption of New Technology *52*

6.4 Cost/Benefit Considerations *54*

7. Modernizing the Copyright Office **54**

7.1 Enhance User Experience *55*
Enhance User Experience Recommendations 55

7.2 Enhance Public Record *57*
7.2.1 Public Record Database 58
7.2.2 Data Sources 58
7.2.3 Credibility of Data 59
Enhance Public Record Recommendations 60

7.3 Improve Data and Information *61*
7.3.1 Data Strategy, Data Management & Data Governance 63

7.3.2 Data Analytics 66
7.3.3 Data Exchange 66
Improve Data and Information Recommendations 71

7.4 Architecture 72
7.4.1 Enterprise Architecture (EA) 73
7.4.2 Technology Infrastructure: Requirements and Alternatives 74
7.4.3 Secure Repository for Works of Authorship 77
7.4.4 Remote Workforce 78
7.4.5 Mobile Technologies 79
7.4.6 Systems Development Strategy 79
7.4.7 Enterprise Solution 80
Architecture Recommendations 81

Closing 82

Appendices

Appendix A: Federal Register Notices

Executive Summary

On October 25, 2011, the Register of Copyrights released a two-year work plan entitled *Priorities and Special Projects of the United States Copyright Office*. The plan established ten special projects that were designed to "improve the quality and efficiency of services in the twenty-first century" and, more generally, to prepare the Office for future challenges.[1] This document is a report from the project team that was responsible for examining "Technical Upgrades," an assignment that commenced in November 2011 and included stakeholder input throughout 2013. The project team spent 2014 analyzing its findings and conducting additional research. Although the initial focus of the project was technology as it affects electronic registration, the scope expanded over time into a broader assessment of technological functionality and related business strategies.

The Technical Upgrades project should be viewed as an intelligence gathering initiative, not only to evaluate the technological environment in which the services of the Copyright Office are currently offered and used, but also to ensure the long-term flexibility that will be necessary as the Office plans for and implements changes to its administrative and regulatory functions. The project has highlighted important concepts and technologies that would enhance key services for customers – including copyright registration, the recordation of copyright-related documents, and the searchability of public records – and would facilitate the exchange of legal and business data with the global marketplace. This report documents various issues related to the Office's technology and services in a set of findings, along with recommendations that will inform the strategic direction of the Copyright Office.

In order to enable the suggestions recommended in this Report, and to manage the IT resources of the Copyright Office in a manner that supports its core mission and statutory responsibilities, the Office will need significant improvements. In the view of the project team, these improvements should be premised upon a new architecture and infrastructure that is specific to the Copyright Office. This infrastructure would support such initiatives as an enterprise application solution and a mobile technology strategy. In addition, the Copyright Office will need staffing resources that are more appropriate to its growing technology and data needs. It will need experts to support a systems development strategy and to handle implementation of new initiatives. This staff would have the benefit of being fully integrated into the daily business of the Copyright Office, and would also be familiar with the businesses that depend upon the Office.

The review process that served as a basis for the recommendations in this report commenced in November 2011 and included significant outreach to stakeholders and other Office constituents. As part of our work, the team interviewed stakeholders from

[1] *See also* Maria A. Pallante, *The Next Generation Copyright Office: What it Means and Why it Matters*, 61 J. COPYRIGHT SOC'Y 213 (2014) (summarizing the scope and preliminary findings of the Office's special projects).

throughout the copyright community and solicited written comments from the public.[2] Midway through the project, we published a set of questions in the Federal Register. These focused on (1) the capabilities of the Copyright Office's public portals; (2) the nature of the data captured during the registration and recordation processes; (3) data and metadata standards used within the copyright industries; (4) data storage practices and policies; (5) search technologies; and (6) the possibility of integrating third-party data or databases.

We received feedback from a variety of entities, including large and small businesses, individuals, trade and professional associations, rights management groups, legal practitioners, publishers, and content creators, as well as from internal Copyright Office staff. This input was in addition to significant research and analysis from the project team itself, which included business and technology experts.

Modernization Themes

The project yielded a number of common themes and identified areas in greatest need of improvement, falling into four broad categories: (1) challenges with the current user experience; (2) challenges with access to and the usability of copyright records; (3) inefficiencies with current copyright data; and (4) poor performance of outdated IT architecture and infrastructure.

Among other things, this report summarizes the 2007 implementation of electronic registration, known as "eCO," a successful effort that transitioned several paper-based processes to an online environment that may be considered a "first generation" system. In the words of one organization, the eCO system should be "*re-engineered so as to create a more intuitive user interface that is easier for new users to navigate, that allows users to print and save data in a way that can be easily printed, viewed, and forwarded outside of the system, and that allows clients to sign applications prepared by attorneys*."[3]

Likewise, the community would like to see richer, more accurate and timelier copyright data. The American Bar Association noted "*significant delays in the appearance of registration records*" in the online database,[4] and more than one stakeholder expressed frustration with the ability to find information when searching for it.[5] It would seem that confidence in Copyright Office records is directly related to confidence in copyright protection. In the view of the American Association of Independent Music (A2IM), the

[2] Notice of Inquiry: Technological Upgrades to Registration and Recordation Functions, 78 Fed. Reg. 17,722 (March 22, 2013). All comments sent to the Office in response to this public notice are available on the U.S. Copyright Office's website at http://copyright.gov/docs/technical_upgrades/comments/. Citations in this document will refer to the comment number as listed on this webpage.
[3] Comment 3, American Intellectual Property Law Association, at 2.
[4] Comment 2, American Bar Association, at 2.
[5] Comment 3, American Intellectual Property Law Association, at 6.

Copyright Office database can determine whether "*creators' works are easily identifiable and do not become orphan works.*"[6]

Looking ahead, many stakeholders would like the Copyright Office to develop data exchange capabilities between external and Office systems.[7] In addition, there was keen interest in the Office adopting and promoting data standards, especially recognized and accepted national and international standards. The project team recognized early on that data considerations would play a vital role in technical upgrades to the Copyright Office, and incorporated such concerns into its overall study.

Infrastructure is, of course, a major part of the analysis of this report. The project team met with a number of experts in the copyright business community with this in mind. Parties shared technological solutions that have proven successful for their businesses, as well as information about the challenges they face. It should be noted that the Copyright Office commenced the Technical Upgrades Project with prior knowledge that its copyright systems were outdated and overdue for upgrades. This is true in the area of the Office's underlying IT architecture as well as the supporting infrastructure.

General Conclusions

In order for the U.S. Copyright Office to deliver highly reliable services to stakeholders and provide a high degree of system availability, it needs to operate like a 24-7 business. To do so, it needs modern system architecture and supporting IT infrastructure. In the view of the project team, this ultimately should include an enterprise solution that is dedicated to the Copyright Office and deploys a highly reliable, high performing, highly redundant, and cost-effective server environment designed to meet the current and future IT needs of the Office.

In the current system, the Copyright Office does not have a fully-resourced technology shop, but employs 20 or so staff to perform planning, analysis, development, implementation and maintenance for the Copyright Office's line of business systems, which are primarily focused on the eCO system. All administrative control over the infrastructure, operating systems, database systems, storage systems, telecommunications systems, legacy systems, and other common IT resources are controlled by the Library of Congress.

This arrangement is not optimal given the general IT challenges at the agency level, and perhaps difficult to rationalize given the specific importance of the Copyright Office to the overall national copyright system and global digital economy. Additionally, in some cases, the Library's needs in relation to copyrighted works – which revolve around acquisition, preservation, and access – may compete with those of copyright owners, who are most concerned with legal protection and security.

[6] Comment 1, A2IM, at 1.
[7] By "data exchange" we mean the sending and receiving of data, which may include the use of data standards and schemas, to support interoperability.

In its report, the project team identified several technologies that should be central to a modern Copyright Office:

Application Programming Interfaces (API): API technology is used to enable standards-based data exchange between two entities. APIs are the de-facto standard to support business-to-business (B2B) data exchange. It was clear in interviews with the copyright community and from the public's written comments that data exchange with the Office is a critical service that needs to be offered. However, the Office must expand the capabilities of its IT systems to provide data exchange via APIs.

Cloud Computing: Cloud computing has emerged as a viable alternative to the traditional approach of deploying applications within an organization. Cloud computing has characteristics that separate it from traditional computing environments, such as the provision of on-demand service that is billed like a utility service and based on usage. This characteristic provides users of cloud technologies with advantages in efficiency, agility, scalability, and innovation, advantages that should be a priority for an agency at the center of a booming marketplace of technology platforms, devices, and content.

Service Oriented Architecture (SOA): SOA refers to a collection of services (software) that have the ability to communicate with each other. It can range from simple data transfer between two services to multiple services working in tandem to perform an activity. SOA requires services that are well-defined, self-contained, and have the ability to operate independently without requiring the context or state of other services. Organizations that follow SOA integrate multiple services to fulfill a specific business function. In the case of the Copyright Office, this would promote better planning and cost savings.

Integrated Solutions/Business-Driven Ecosystems: Organizations have recognized that one-size-fits-all solutions often fail to meet the needs and expectations of all users. Because of this, many businesses are developing solutions that provide internal and external system developers with the flexibility to provide solutions that are targeted as well as integrated. This kind of approach is possible with Service Oriented Architecture, flexible web-services, and expanding cloud solutions. In order for the Copyright Office to serve its data driven customers, it will have to develop a new ecosystem model.

Mobile Computing: Mobile devices are being used by organizations for access to internal systems and by end-users to access business systems. End-users have a plethora of applications for everyday tasks (*e.g.*, email, mobile banking, access to weather, account information) available to them through mobile devices. Moreover, organizations are beginning to develop interfaces into their business systems that are designed specifically for mobile devices.

The Copyright Office should seek to deploy mobile capabilities to meet customer demand. This would result in new software and web interfaces optimized for mobile device utilization. Benefits of deploying mobile capabilities would include:

1. Improved Business Productivity – Provide immediate access to systems anytime, anywhere.
2. Reduced Operation Costs – Enhanced access to business systems (no need to be at the office/desk) to correct issues or logistics problems.
3. Improved Customer Relationships – Provide additional opportunities to connect with systems to obtain information anytime, anywhere.

Big Data: "Big Data" is a term used in information technology to refer to extremely large data sets that are stored in large database systems and rely on high-performance systems to capture, store, and analyze data to provide useful business intelligence. In order to take advantage of Big Data technology and realize the most benefit, organizations will need to deploy a Data Analytics solution (see below). A combination of both would provide the Copyright Office with the business intelligence to gain efficiencies while meeting its customer needs, thus maximizing its services. Specifically, the development of Big Data technology would help the Copyright Office determine customer needs and develop new business goals; increase efficiency in existing operations; refine existing services and products that better meet customers' needs; and improve the Office's decision-making regarding both short-term and long-term objectives.

Data Analytics (DA): DA is an essential technology for managing large amounts of data (see Big Data, above). DA provides organizations with information to help make them more competitive and target products and services that provide the most return on investment. Organizations use DA as one method for predictive analysis based on data captured by their systems. Both governmental and commercial organizations continue to deploy DA as a tool to increase profitability and efficiency and it is difficult to see the Copyright Office moving forward without investing in and developing these analytics tools.

Specific Recommendations

The project team's conclusions and recommendations speak to the most pressing deficiencies in existing environment, but are nonetheless a step towards positioning the Copyright Office for future success. As the agency adopts a more sophisticated IT direction, it will of course require the flexibility to perform ongoing assessments and adjustments simply to keep pace with its customers. These customers have been patient with the Copyright Office, but as they have politely but firmly stated in recent years, they need the Office to be significantly more innovative and accountable than it is now. With this in mind, the project team is pleased to offer these initial recommendations. These recommendations address key aspects of the four modernization themes, and are supported by a number of more detailed recommendations that provide a comprehensive strategy for the future of the Office's technical upgrades.

Enterprise Solution: An architecture and infrastructure that directly and capably supports the many complex duties of the Copyright Office is fundamental to all other

goals discussed in this report. Ultimately, the Office should have a new and dedicated enterprise copyright system, which might be perceived as "eCO 2.0." By this, the team means a large-scale application package that includes the capabilities for registration, recordation, public information and records, accounting and processing, and acquisition. While the result would be an over-arching enterprise system, each copyright function must have its own dedicated development cycle to include analysis, requirements, design, build, testing/pilot, deployment and maintenance phases.

It should be noted that these goals would need to be reconciled with the existing systems and future plans of the Library of Congress. But, it seems unlikely if not impossible that the Copyright Office can administer the law effectively into the future without IT investments that reflect the importance and growth of its mandate. In short, the goal should be to evolve the Copyright Office's technology department from a small liaison staff that relies on and is required to advocate to the Library's technology office, to a fully-empowered operation in which technology decisions are measured against the singular goal of furthering the objectives of the copyright law and meeting the needs of the copyright community, including the content and technology sectors that are the Office's customers. Each one of these industries accounts for an economic impact of billions if not trillions of dollars. It seems obvious that a significant percentage of staff in the Copyright Office should be technology experts, including dedicated systems, data, and architecture staff, rather than the current paradigm in which twenty or so staff liaison with central Library offices. Put another way, Library staff do not have the benefit or experience of working in the Copyright Office, and therefore will never have the context or specialized knowledge that is essential to Copyright Office success.

Mobile Capabilities: A digitally-integrated Copyright Office would include interactive and mobile technologies now common in the global marketplace. Mobile capabilities could be leveraged by external users for ease of copyright application submission and provide for the use of Copyright Office services from smart phones and tablets. In addition, a mobile platform could be leveraged by internal users for improved efficiencies in performing daily work functions; work force benefits might further include a secure, robust, state-of-the-art telework program.

Because the existing user interface is inelegant and inflexible, the Office needs to redesign and deploy a new, browser-agnostic web portal. This redesign effort should also include a fully redesigned copyright.gov that offers patrons the ability to perform all transactional activities and provides for improved searching capability. The new website must be fully under the control of the Copyright Office, not only because the Office cannot risk disruption to its daily business and services, but because the website is the primary tool by which the Register administers the copyright law. The website of the Copyright Office must be geared to and meet all of the needs of its customer base. In redesigning the site, the Office should incorporate the many wizards, tutorials, and other self-help features suggested by the copyright community.

A Better Public Record: One of the ongoing and primary objectives of the Copyright Office is to create and maintain a public database of robust, reliable, and authoritative records of copyright ownership. What constitutes a good public record should focus on

both legal sufficiency and client requirements. In addition, the Office should identify and consider potential enhancements to the public record, such as whether it should allow for samples of deposit copies (*e.g.*, thumbnails of images or snippets of sound recordings), or whether claimants should be granted access to update their own contact information for rights and permissions. Finally, the redesigned public record should be seamlessly integrated with the various efforts to improve copyright data, as described in the section below.

Internal efforts might include conducting an alternatives assessment on existing database products capable of supporting the new public record, as well as conducting a cost-benefit analysis comparing migration to a new, large-scale database versus enhancing current capabilities. Further, the Office could release a Request for Information or Statement of Objectives to industry experts and technology research organizations to obtain input on deployment of enterprise class databases.

Sophisticated Data Management: Because data is such a vital asset to the Copyright Office and because it is fundamental to mission success, the Office needs to develop a long-term Data Strategy, a comprehensive Data Management Plan and a detailed Data Governance Plan. The Office will also want to develop a detailed "as-is" document, consult with industry experts, and develop a best-possible data model.

The project team recommends that the Register establish and appropriately staff a data group to be chaired by the Copyright Office Chief Information Officer. This group would have many functions including but not limited to participating in working groups of data standards-setting bodies and developing API's. One task for the group would be dedicated to data exchange to establish a short-term pilot for secure, bulk submission of registration applications or documents for recordation.

There is significant additional data and metadata in registration and recordation records, electronic deposits, and other sources of data that the Copyright Office has yet to coordinate and exploit, such as within legacy systems or metadata inherent in born-digital submissions. The Copyright Office might wish to seek interim solutions to harvest this data and make it available. Meanwhile, it is essential that the Office establish a secure data warehouse, applying data management and archiving principles and best practices. Once established, this data repository would allow the Office to engage in data analytics, including conducting business-side and partner-side data analytics. As a first step, the Office should undertake a study to analyze the available models and frameworks to gain an understanding of the infrastructure requirements and best practices necessary to administer the law.

Security Practices: Finally, the Copyright Office has a statutory responsibility to securely protect the works it needs to examine for registration, meaning that it must meet the requirements of the law and the concerns of copyright owners. This focus is lacking in eCO but is essential to a next generation system, particularly because the Office will be exploring a variety of new registration procedures for digital works – from the data that should be captured in digital photographs, to the methods for examining works that are delivered as streams rather than as copies to works that change rapidly

and frequently throughout the day, such as news sites. The Office moved this work forward in December 2014, when it published a substantially revised Compendium of U.S. Copyright Office Practices, a key foundation for future practices and regulations in the digital space.

A Commitment to Interim Work: While pursuing enterprise and data solutions for the long term, the Copyright Office must continue with the interim modernization efforts that it has put in motion over the past two years. This report addresses the processes for how digital works should be registered and the options for improving and automating document recordation. In order to enhance its existing offering in a meaningful way, the Office will need an application development strategy, which may include in-house application development capability. In other words, the Copyright Office will need to invest in a common resource infrastructure – primarily servers, in which applications can be developed, tested, staged and put into production.

The project team recommends that the Copyright Office engage in an effort to itemize proposed improvements to the Copyright Office architecture by defining "as-is" and "to-be" models, as well as analyzing the benefits of using commercial off-the-shelf (COTS) products versus in-house development. The Office should engage in a similar analysis to define the as-is and to-be of its infrastructure, with a goal of determining if the existing technology infrastructure is sufficient. Additional infrastructure analyses will have to include a cost-benefit study on a Copyright Office data center versus using existing facilities and deploying infrastructure dedicated to the Copyright Office. A dedicated data center appears preferable for many reasons, but would require a more thorough analysis to identify all possible operational and maintenance expenses, for example, floor space, electricity and cooling for servers, additional servers for storage and backup, software licenses, telecommunications and personnel.

Acknowledgments

The Technical Upgrades Project was a challenging but important exercise. On behalf of the project team, I would like to thank Register Pallante and our Copyright Office colleagues for the opportunity to engage on these issues. While the assignment was ultimately to think ahead, this would not have been possible if we had not had the freedom to engage critically and objectively.

For all aspects of this endeavor, I am indebted to the talented members of my project team. They are: Joanna Corwin, Project Manager; Peter DesRoches, Technical Subject Matter Expert; Susan Todd, Copyright Office Registration Subject Matter Expert; Christopher Reed, (former) Senior Advisor to the Register for Special Projects; Annette James, Business Analyst; and the many other wonderful staff members who graciously volunteered their time, passion and knowledge of the Copyright Office to accomplish the goals and objectives of this very critical project.

Finally, I am grateful to the many members of the copyright community who spent valuable time meeting with us and responding to inquiries. You have educated and

inspired us, and we thank you for sharing your expertise for the betterment of the United States copyright system.

Doug Ament
Chief Information Officer

1. Background

In late 2011, the Register of Copyrights and executive management team of the Copyright Office published a document entitled *Priorities and Special Projects of the United States Copyright Office*. That document outlined ten "Special Projects" which would become core focus areas for improvement by the organization over the course of two years.

This public report is the culmination of one of the ten special projects, entitled *Technical Upgrades to Electronic Registration*, which progressed over time to a broader exploration and shorter title: *Technical Upgrades*. As stated in the Executive Summary, the project should be viewed as an intelligence gathering initiative, not only to evaluate the technological environment in which the services of the Copyright Office are both offered and used currently, but to ensure long-term flexibilities that will be necessary as the Office plans for and implements changes to the way it administers the copyright law.

We hope the report will inform the strategic direction of the Copyright Office as it navigates the twenty-first century and moves to a new generation of services.

2. Project Objectives

Ultimately, the Copyright Office needs to align its IT capabilities and processes with the technological and procedural needs of the copyright community. Thus, a key objective was to work collaboratively with willing stakeholders to identify technologies that could help shape a more effective Copyright Office.

The second objective was to research and analyze original research, drawing upon the technology expertise of the project team in addition to information gleaned from stakeholder interactions.

The third objective was to document outcomes, prepare findings and propose recommendations for the Register of Copyrights and her senior management team. These should in turn drive the information technology aspects of a new strategic plan.

3. Project Approach and Methodology

3.1 Intelligence Gathering and Analysis

The Technical Upgrades project team performed two major intelligence-gathering initiatives to solicit input from the copyright and technical communities. A Notice of Inquiry (NOI) published in the Federal Register (attached as Appendix A) elicited twenty-eight responses. In addition, the project team completed a series of interviews with interested parties, and over the course of many months, the project team met with thirteen external organizations and conducted eleven internal interviews.

The NOI, published on March 22, 2013 in volume 78, issue number 56 of the Federal Register, provided the public with sixty days to submit comments. The NOI highlighted six areas of specific interest:

(1) the capabilities of the Copyright Office's public portals;
(2) the nature of the data captured during the registration and recordation processes;
(3) data and metadata standards within the copyright industries;
(4) data storage practices and policies;
(5) search technologies; and
(6) the possibility of integrating third-party data or databases.

The project team compiled and analyzed all of the information collected from the NOI responses in order to develop recommendations responsive to the public's needs.

The Technical Upgrades interview program spanned twenty months, from November 2011 to July 2013. The project team's interviewers generally met with one organization at a time, to better focus on each party's input and to keep the tone of the meetings intimate and conversational. Further, we informed the parties that their comments would be not be attributed to them, to encourage them to provide honest feedback about Copyright Office systems, and to share openly about their own technological or business challenges.

Collaboratively and with confirmation from subject matter expert (SME) volunteers from within the Copyright Office, the project core team performed business analysis on the intelligence and data captured in the site reports for all of the interviews conducted as well as the comments provided in response to the project's NOI. This analysis established the final list of recommendations offered by the interested parties. Finally, the core team performed a technical review of the recommendations in order to develop an implementation strategy including a high-level timeline and estimate of resources required.

3.2 Resource Utilization

To efficiently complete the series of interviews, the project team was divided into sub teams with specific tasks to accomplish. Senior leadership identified organizations from the copyright community for outreach and participation in the project, including specific contact information. A volunteer team of SMEs conducted background research on each interested party and documented their findings in a site report. The project manager established contact with the parties, scheduled meetings and distributed the read-ahead information from the volunteers to each member of the interviewing core team. The interview was conducted and following each meeting, the business analyst compiled the notes and updated the site report to document all of the data, recommendations, process improvements, and technology proposals (where applicable) provided by the interested party.

By leveraging a sub-team approach, this Special Project was able to minimize the number of dedicated resources by recruiting volunteers within the Copyright Office to perform many of the sub-team activities. Because of the dynamic nature of the project, individual volunteers often played multiple roles simultaneously. In addition, the team structure varied and was tailored to the specific needs of each interview.

4. Participating Parties

There were forty-seven participants in this Special Project offering a variety of perspectives. Interviews and written comments were shared by trade and professional associations, rights management groups, legal practitioners, publishers, content creators and internal staff of the Copyright Office.

The table below lists all of the parties from the copyright community that participated in this Special Project and provides some key information regarding each:

Name	Primary Service	Size, Location	Form of Participation
American Association of Independent Music	Trade association for independent music labels	Approx. 270 members, New York, NY	Comments to Notice of Inquiry
American Bar Association, Intellectual Property Law (ABA-IPL)	Bar association of intellectual property attorneys	20,000+ members Chicago, IL	Comments to Notice of Inquiry
American Intellectual Property Law Association (AIPLA)	Bar association of intellectual property attorneys	Approx. 16,000 members, Arlington, VA	Comments to Notice of Inquiry
American Society of Composers, Authors and Publishers (ASCAP)	Performance rights organization	460,000+ members New York, NY	Comments to Notice of Inquiry
American Society of Media Photographers	Trade association of professional photographers	6,000+ members, Philadelphia, PA	Comments to Notice of Inquiry

Name	Primary Service	Size, Location	Form of Participation
Association of American Publishers	Trade association of book and serial publishers	Approx. 300 members, Washington, DC	Comments to Notice of Inquiry
Author Services, Inc.	Representatives of L. Ron Hubbard	N/A, Hollywood, CA	Comments to Notice of Inquiry
Broadcast Music, Inc. (BMI)	Performing rights management organization	600,000+ members, Nashville, TN	Interview, Comments to Notice of Inquiry
Colton, Robert	Retired Section Head of Recordation	Arlington, VA	Comments to Notice of Inquiry
Copyright Clearance Center	Provides licensing solutions to copyright content	Approx. 250 employees, Danvers, MA	Interview
Copyright Society of the USA	Professional association for the copyright community	N/A, New York, NY	Interview
County Analytics Ltd	Independent consultant, Paul Jessop	N/A, Luton, United Kingdom	Comments to Notice of Inquiry
Dominican University Graduate School of Library and Information Science	Students of an ALA accredited Masters of Library & Information Science program	N/A, River Forest, IL	Comments to Notice of Inquiry
Donahue Gallagher Woods, LLP	San Francisco Bay area law firm	34 attorneys, 3 offices in CA	Comments to Notice of Inquiry
Educational Testing Service (ETS)	Standardized testing and assessment service	2,500+ employees, Princeton, NJ	Comments to Notice of Inquiry
FLASHLIGHT2013	Grassroots copyright awareness effort	N/A, Ocean Beach, CA	Comments to Notice of Inquiry
Graphic Artist Guild	Trade association for graphic artists	Approx. 2,000 members, New York, NY	Comments to Notice of Inquiry
Harry Fox Agency	Rights management organization	N/A, New York, NY	Interview
ISNI International Agency	Maintains the International Standard Name Identifier (ISNI)	United Kingdom	Comments to Notice of Inquiry
ISRC Agencies	Administer the International Standard Recording Code (ISRC)	N/A, Washington, DC	Comments to Notice of Inquiry
LegalZoom.com, Inc.	Online legal documentation service	500+ employees, Los Angeles, CA	Comments to Notice of Inquiry
Library of Congress, Information Technology Services (ITS)	Provides technical support to the Library of Congress	Approx. 300 employees, Washington, DC	Interview
Microsoft	Software corporation	N/A, Redmond, WA	Interview
Morris Music Law	Private practice of Jess E. Morris, attorney	N/A Marina Del Rey, CA	Comments to Notice of Inquiry

Name	Primary Service	Size, Location	Form of Participation
Motion Picture Association of America (MPAA)	Trade association for the motion picture and television industry	Approx. 200 employees, Sherman Oaks, CA	Interview
Music Reports, Inc.	Performance rights organization	Approx. 100 employees, Woodland Hills, CA	Comments to Notice of Inquiry
National Music Publishers Association (NMPA)	Trade association for the music publishing industry	2,500+ members, Washington, DC	Comments to Notice of Inquiry
New Hope Publishers	Book publisher	N/A Birmingham, AL	Comments to Notice of Inquiry
NWReflections, LLC	Fine art photographers	Under 10 employees, Washougal, WA	Comments to Notice of Inquiry
Optimos	Development and maintenance of eCO	Approx. 10, Washington, DC	Interview
Pearson Education, Inc.	Book Publisher	40,000+ employees, Upper Saddle River, NJ	Comments to Notice of Inquiry
Perkins Coie	International law firm	900+ attorneys, Seattle, WA	Comments to Notice of Inquiry
Petruzzelli, Nanette	Retired Associate Register for the Registration Program	Hyattsville, MD	Comments to Notice of Inquiry
ProQuest	Aggregator and publisher, particularly of dissertations and periodicals	Approx. 1,800 employees, Ann Arbor, MI	Interview
Recording Industry Association of America (RIAA)	Trade association for the music recording industry	Over 1,600 members, Washington, DC	Interview, Comments to Notice of Inquiry
SAIC	Staff of eCO Help Desk	Approx. 8, Washington, DC	Interview
SoundExchange	Performance rights organization	Approx. 75 employees, Washington, DC	Interview, Comments to Notice of Inquiry
COPYRIGHT OFFICE, Copyright Acquisition Division (CAD)	Acquiring copyrighted works for the Library of Congress	Approx. 20 employees, Washington, DC	Interview
COPYRIGHT OFFICE, Copyright Technology Office (CTO)	Providing technical support to the COPYRIGHT OFFICE	Approx. 20 employees, Washington, DC	Interview
COPYRIGHT OFFICE, Director of IT	Providing technical direction for the COPYRIGHT OFFICE	1, Washington, DC	Interview
COPYRIGHT OFFICE, Information and Records Division (I&R)	Providing copyright information and products to the public	Approx. 70 employees, Washington, DC	Interview
COPYRIGHT OFFICE, Literary Division (LIT)	Registration and recordation of literary works	Approx. 55 employees, Washington, DC	Interview
COPYRIGHT OFFICE, Performing Arts Division (PA)	Registration and recordation of performing arts works	Approx. 55 employees, Washington, DC	Interview

Name	Primary Service	Size, Location	Form of Participation
COPYRIGHT OFFICE, Receipt, Analysis & Control Division (RAC)	Materials processing and accounting at COPYRIGHT OFFICE	Approx. 80 employees, Washington, DC	Interview
COPYRIGHT OFFICE, Register of Copyrights	Director of the COPYRIGHT OFFICE	1, Washington, DC	Interview
COPYRIGHT OFFICE, Visual Arts Division (VA)	Registration and recordation of visual arts works	Approx. 25 employees, Washington, DC	Interview
Writers Guild America West	Labor union for film, TV and radio writers	Approx. 20,000 members	Interview

5. Findings: Additional Service Recommendations

This section outlines the key issues the copyright community identified as users of the current Copyright Office systems. As well, it documents the suggestions and requests from the participating parties for the Copyright Office to take under advisement. In analyzing the feedback collected through both interviews and comments to the NOI, the project team determined four priorities: (1) the overall user experience; (2) the official Copyright Office public record; (3) copyright data management; and (4) the overall technical posture.

5.1 Challenges with Current User Experience

5.1.1 Feedback from External Users

The Copyright Office implemented the eCO system in 2007 with Library of Congress technology services. It was designed to bring registration processes and some related functions online. Since the initial implementation, the agency has made a series of improvements to address some of the shortcomings of the software. Although it was already aware of many of the deficiencies, the project team wanted a clear assessment from the perspective of external users.

eCO Interface

When users log into eCO, they would like to have a personalized dashboard so they can quickly navigate to the information they seek, such as the status of a particular case or an inventory of all cases in correspondence. An improved user interface might offer customizable columns, places for users to add notes, a more complete audit trail and potential integration with docket management software for attorneys. Users would benefit from sorting capabilities or filters, such as the ability to sort individualized pending queues by categories, such as the type of work being registered. To become a truly twenty-first century Copyright Office, the eCO user interface should offer all the capabilities that users have come to expect from an online service, such as those commonly seen in various electronic commerce applications.

While many updates to eCO have been released since its launch, external users continue to report that the system is cumbersome, requiring too many clicks and too much scrolling. The navigation panel added to eCO was a welcome benefit to registration applicants; however, they would like to see an even simpler interface where they do not have to navigate through so many screens. Several customers indicated they would like the Office to offer fill-in forms and/or have the ability to enter all of the required information into a single screen.

For example, the **Graphic Artists Guild (GAG)** suggests, "*a registration form as a single page akin to a sheet of paper which scrolls down. All questions and content would be visible to the user on one webpage.*"[8]

Users have asked that the system take advantage of profile management capabilities with saved values, which the system would prompt for "auto-fill." Currently eCO allows users to save templates, but the project team also heard feedback about ways in which the templates could be improved so they are easier to save and retrieve, as well as navigate.

Further, applicants would like to see improvements to tracking, status and notifications. The external user interface could be enhanced with a tracking feature such as a status bar or other visual display that would clearly indicate the progress of their applications. An additional feature some would like to see is the ability to submit an inquiry or follow-up on a specific case from within the online system. For example, some users requested the ability to signal the Office with a simple click from within the system when they have not received a certificate for a case marked closed. Currently, users receive system notifications when their application is submitted and when their deposit is uploaded. Some users would like these notifications to be consolidated and several would like to see more information provided about the deposit, such as a listing of all the file names and their sizes. Users also requested additional system-generated notifications not currently in place, such as confirmations of documents received or an indication that a certificate is in the mail.

Our interviews of internal users also generated feedback regarding notifications for external users. Copyright Office staff would like for the system to provide more detailed and helpful information to the public regarding the status of their requests, including receipt acknowledgements. Copyright Office employees who support the registration process spend a great deal of time providing status updates to applicants. If the system provided detailed information, for example, where a claim is in the registration process, this would eliminate the need for applicants to call the Office which would in turn reduce inbound call wait times. Likewise, employees who support the mandatory deposit process would like follow-up requests for demands to be automated and system-generated.

[8] Comment 15, Graphic Artist Guild, at 8.

Regarding registration records, some external users would like the ability to view all of the correspondence, notes and application versions pertaining to a case together as a single file history which they can access at any time (such improvements may also streamline internal processes when the Copyright Office provides copies of records to members of the public, *e.g.*, for litigation). When a claim is pending and in correspondence, external users cannot view the application as submitted in order to understand a Registration Specialist's request. After a claim has been closed, external users cannot access the case file to determine how a previous application was filed for a similar situation. Also, registration records are not linked to previous or subsequent registrations, documents filed or corrections/amplifications.

For example, the **American Association of Publishers (AAP)** recommends that the Office "*organize records into complete file histories, i.e., if a member of the general public searches for a particular work (by title, author, ISBN, etc.), all documents associated with that particular work should be part of the file.*"[9]

In general, the external user experience needs to be improved and the system more flexible. Users would like eCO to be compatible with a wide range of web browsers.

For example, the **American Society of Media Photographers (ASMP)** notes that "*[t]his is a significant problem for an industry that relies heavily on non-PC based browsers, such as Apple's Safari . . . . There is a real need to upgrade the Copyright Office's system to incorporate cross-browser compatibility.*"[10]

Likewise, the **American Bar Association, Section of Intellectual Property Law (ABA-IPL)** suggested "*[a] potentially low-cost improvement to the eCO claims interface could be to integrate 'Tips' for practitioners with respect to each of the fields that need to be filled in.*"[11]

Other respondents suggested resources such as a "common mistakes" document or "do's and don'ts" compiled by registration staff. In addition, participating parties suggested the online application system provide instructions and definitions in plain English rather than "legalese" to the greatest extent possible.

Staff suggested numerous improvements to the phone system that would better serve the public. A series of automated questions could be used to route calls to staff with areas of specialty; call wait times or place-in-queue announcements could be useful to callers; and other recorded information could be played to callers on hold such as answers to frequently asked questions, Copyright Office news and other helpful information. Additional upgrades to the system might address call monitoring for improved quality control and recordkeeping.

[9] Comment 6, Association of American Publishers, at 8.
[10] Comment 5, American Society of Media Photographers, at 3.
[11] Comment 2, American Bar Association, at 2.

Registration Claims

Under the Copyright Act, the application, fee and deposit copy[12] are the three necessary components in a copyright registration claim and each process would benefit from technical improvements. Beginning with application submission, users would like the ability to preview the completed registration application, the ability to edit the application from this preview, and the ability to both print and save the preview copy. Currently, users have to drill into fields to see all of the information provided on an application, must navigate to specific sections of the application to make changes, and cannot save or easily print a copy for their records. Multiple users complained that printing a preview of the application requires at least seven pages that cannot be formatted properly and are hard to read because of the font size. Some expressed the desire to see better alignment between the paper and electronic application, such as consistent terminology or a correlation between the two regarding the information requested.

Next, the applicant must pay the filing fee for the registration. The project received several comments that this process could be streamlined. Suggestions included developing an integrated payment option rather than directing applicants to pay.gov, or at least providing additional payment options aside from pay.gov. Some applicants filing on behalf of numerous claimants would like the ability to create individual receipts and the ability to link to various deposit accounts, or have multiple account profiles link to a single deposit account.

The final step in the process of submitting a claim is to provide a deposit copy. Many users indicated that the deposit upload functionality should be improved. While the current system has been upgraded, users still complain about the upload process being awkward, slow and too often resulting in system "time outs." The Office needs to expand the list of acceptable file types and provide the necessary software for staff to view them. In addition, many external users requested the ability to view the files once they have been uploaded; this is not a feature available in the present system.

Comments from **LegalZoom.com, Inc.** included "*the case summary page does not allow users to view the deposit copy once it is uploaded or check it for completeness before it is submitted, and does not provide immediate confirmation of the upload or submission.*"[13]

It is worth noting that despite the virtues of making deposit files available through the Office's online web portal, such features raise important security considerations, including a risk of improper disclosure, either inadvertently or as a result of malicious intent. These kinds of security issues will need to be addressed not only as technical issues but also as legal and regulatory issues, and it is possible that the Office will adopt a variety of rules over time through appropriate public processes.

---

[12] The term *deposit copy* refers to the copy of the work being registered and which is being "deposited" with the U.S. Copyright Office for examination in conjunction with the application.
[13] Comment 18, LegalZoom.com, Inc., at 1-2.

Beyond electronic deposits, the project also received some comments regarding improvements to the shipping slips used to transmit physical copies and requests to improve options for mailing physical copies. Employees who facilitate mandatory deposit would like publishers to have label printing capabilities for demand submissions.

Automation Issues

All of the foregoing addresses feedback on the existing system, which is currently limited to registration functions. As much as external users would like to see improvements to the existing registration process, they would also like to see the Office create new automation features. This includes new registration options as well as other Copyright Office services, in particular document recordation.

Regarding registration, users would like all registration processes to be automated and all methods of registration available online. Currently, many group registration options including the group registration of published photographs and group registration of databases cannot be processed through eCO. Some other non-basic registration options are also unavailable through the electronic registration system at the present time, such as supplemental registration, renewal registration, and GATT registration. Users would like to see the Copyright Office complete its transition to an electronic filing system by automating all remaining registration applications. Although the Register prioritized recordation as a separate special project, the Technical Upgrades project provided a vehicle for stakeholders to underscore their discontent.

Thus, **AIPLA** writes, "*[w]e think that an electronic system for recording assignments is an imperative improvement.*"[14]

And the **Recording Industry Association of America (RIAA)** said, "*the inability of registrants to file transfers of copyright ownership, license agreements and other relevant documents with the Office electronically deters copyright owners from filing relevant documents with the Office.*"[15]

Some requests for an automated recordation system include the ability to record documents on a mass basis and the ability to record a document directly against a given registration. Other users envision a notification system wherein any documents recorded against a registration will prompt an action requiring approval from the original copyright registrant. Finally, internal users requested that the document numbering process be automated.

Many of the technical issues involving recordation implicate questions of law. The Copyright Office pursued the legal issues on a separate track, culminating with a major study entitled: *Transforming Document Recordation*, published in January 2015. The

[14] Comment 3, American Intellectual Property Law Association, at 3.
[15] Comment 27, Recording Industry Association of America, at 2.

study was conducted by the Abraham Kaminstein Scholar in Residence, a new position that attracts independent scholars to work in the Register's Office, and the incarnation of yet another special project, in which the Copyright Office seeks to supplement its own expertise with programs that invite the participation of academic institutions and their faculty members.

Although recording a copyright document (*e.g.*, an assignment, a license, a security interest) offers a copyright owner important benefits, it is not required, with one critical exception. Authors (and as applicable, heirs) who seek to terminate their earlier transfers (*e.g.*, an exclusive license to publish) must provide notice to the transferees and record the notice with the Copyright Office within a statutorily prescribed timeframe. These notices may be filed under Sections 304 or 203 of the Copyright Act, pursuant to the requirements of those provisions.

The public record for termination notices is something authors, heirs, businesses and successor businesses rely on to assess copyright ownership, liability and expiration of copyright term.

As **Broadcast Music, Inc. (BMI)** observes: "*With terminations playing an increasing role in ownership determinations, virtually every member of the content industry will at some time likely require the means to verify the validity of termination notices. . . . It is therefore crucial that the digitization and uploading of termination notices to the Office's website for public review becomes and remains up-to-date*."[16]

There a variety of other Copyright Services that the public would like to see automated, all of which affect the legal rights and/or economic interests of private parties. Among these are filing of notices of intent to use, particularly in bulk; a seamless online system for DMCA filings; and a repository of Copyright Office review board decisions. In general, stakeholders recommend that all Copyright Office services be accessed through a single web-based portal. Through this main entry a user could self-guide to various resources, but the goal would be to have an interface that is easy to use and easy to navigate. A major redesign of copyright.gov could provide these capabilities, assuming it is done with significant customer feedback and an eye to the commercial standards and expectations of the copyright marketplace. To be clear, the Copyright Office has implemented refreshes to its website now and again, since it launched the website nearly twenty years ago. What we are discussing in this report, however, is an overall change to architecture, service and interface, to enhance both the user experience and overall interoperability.

5.1.2 Feedback from Internal Users

Copyright Office employees are a separate but equally important constituency. Most employees have adapted to or been trained into the current system, but they have developed or requested workarounds, and they have been instrumental in pressing for

[16] Comment 8, Broadcast Music, Inc., at 4.

the periodic updates and small improvements undertaken so far. However, these internal users still struggle with many aspects of eCO and are perhaps the first in line when it comes to calling for additional capabilities.

Just as with external users, employees would benefit from individual, customized dashboards at eCO login. Specialists from various divisions indicated that they spend too much time navigating within the system and drilling into records to obtain the information needed. Not enough of the data is available at the top level and they spend a lot of time clicking into applets and expanding fields. A more streamlined interface would create greater efficiency and would ease frustration. Specialists and technicians develop individual work patterns and sometimes work within their own areas of expertise. They would benefit from a customizable interface suited to their daily work routines, rather than a "one size fits all" model.

Employees would also like to see improved methods of tracking, including the ability to identify the location of a specific deposit or application. The current system utilizes barcodes that have partially addressed the need for tracking deposit copies. However, the current workflow does not accommodate bar code tracking throughout all Copyright Office processes. Both time and space considerations limit the success of a system where bar codes must be laser scanned at each location. Recommendations from staff included RFID (radio frequency identification) technology for enhanced, seamless asset tracking.[17]

Employees could perform their jobs more efficiently if the Copyright Office had better system integration with the software used. In particular, the registration staff would like to see a more fluid connection between eCO and email. The email feature in eCO lacks the functionality users have come to expect from an email program. For example, staff does not directly receive notifications about undeliverable email. If this functionality were available in eCO, they could take immediate action to find alternate methods to communicate to a remitter. Nor can a staff member set up an "Out of Office" notification to alert applicants when he/she will be unreachable for a period of time. External users also indicated there were difficulties in corresponding with staff, especially through email.

A member of the **AIPLA** indicated "*[m]y BIGGEST complaint is that it is very difficult to talk to copyright examiners. . . . The email inquiry from the examiner is apparently not connected to their personal email but instead an electronic case file. Depending on how the message was coded in their system, your reply email may or may not reach the examiner.*"[18]

[17] RFID technology transfers data over electromagnetic fields wherein readers identify and track tags attached to objects. RFID is used to track anything from packages to livestock and is currently used by many libraries.
[18] Comment 3, American Intellectual Property Law Association, at 3.

Registration staff stressed they would like the ability to link emails to cases in a many-to-one relationship. In order to preserve the record where multiple cases require correspondence, staff must decide between sending a separate email from each case and sending one email and then manually entering notes into the remaining cases; both options are time consuming. Because examiners often address multiple numbers of cases in an email attached only to one case, specialists in the Records Research & Certification section have difficulty identifying and producing complete records for members who seek such materials for litigation or other purposes.

Similarly, Copyright Office registration staff would like to see external users have the ability to mass update pending applications. A significant update to the current system has been the ability for a specialist to return an electronic application to an applicant for amendments. However, these alterations must be completed individually by drilling into one claim at a time. Under the current configuration, changes as simple as updating an address or correcting the answer to the “work made for hire” question quickly become time and labor intensive.

Finally, staff in the Registration Program requested better utilities for deposit examination. Many would like the Copyright Office to standardize the file formats received including offering conversion tools within eCO to assist applicants in providing an acceptable file type. Currently, eCO does provide a wide list of acceptable file types but the system does not validate these file types. Since there are limited restrictions on what types of files may be uploaded, registration specialists receive many files they cannot view and must correspond to request a replacement. In addition, employees would like external users to have the capability to upload larger files, especially for motion picture claims and full resolution photographs. Certainly, registration staff would like to have the latest versions of primary software for examining capabilities. Too often the versions of software at their workstations lag significantly behind those used by the creative communities.

5.1.3 Additional Ideas and Features

While many of the suggestions to the Technical Upgrades project are in response to the current eCO system, the community also shared several forward-looking solutions. An upgraded copyright system would not just address improvements upon the past but should include additional capabilities not fully utilized at present. In addition, the House Judiciary Committee is currently reviewing Title 17. Potential revisions to the copyright law could result in need for technical capabilities that go beyond the examples that follow.

Of particular importance to practitioners and attorneys is the ability for third parties to certify the application. As noted previously, eCO does not allow a user to preview a completed application before submission, let alone save or share a file copy. Lacking this capability, attorneys are unable to show their clients the prepared application or have their clients sign off on it. Creative businesses are similarly affected, such as in-house professionals at publishing houses or film studios. These filers would like the

parties they represent to be able to review applications before submission. The community strongly suggested that any new copyright system contain this feature.

Similarly, copyright owners would like to see options for receiving electronic certificates. The current system provides for electronic applications and deposit copy upload, but issues paper certificates. Many parties would prefer to receive or have continuous access to an electronic certificate as well. Some have suggested that the Office move to electronic certificates with the option to purchase a "hard copy" paper certificate.

At a greater level of sophistication, the Copyright Office could offer digital certificates that would be accepted in courts by taking advantage of digital signature capabilities including signed hashes and digital fingerprints. Thus, not only would an applicant be able to receive an electronic certificate by email for their records, but official certified copies of the certificate could also be electronic using digital signature technology. If such a capability were offered, an attorney could log into eCO to submit certified digital copies of certificates to courts directly. Of course, digital certificates raise policy and legal questions that would require further review and, potentially, proposals for legislative or regulatory adjustments.

As noted above, applicants indicated they would like to be able to view the files they uploaded as electronic deposits. Expanding on this, both internal and external users suggested that portions of the deposit copy could become part of the available public record at the copyright owner's option. This would require very careful deliberation and consultation with stakeholders, but the point is a good indication that the Office should be exploring new paradigms.

For example, **BMI** noted that a sample or snippet of an MP3 file could be made available. "*With regard to musical works, BMI believes that, if a copyright owner wishes to authorize uploading short digital samples of works to the Office website and incorporating such samples as part of the registration records, samples would be useful in the identification of copyright works . . . .*"[19]

This goal was further supported by the visual arts communities, including the American Society of Media Photographers and the Graphic Artists Guild whose expanded vision included the use of image recognition technologies to search records.[20]
NWReflections, LLC, an independent, small-entity photography studio wrote, "*it would be extremely valuable to have the ability to identify specific digital files within a registration . . . It would be incredibly valuable to be able to query that repository," and "[o]ptimally that search also would display a web rendering of the uploaded file.*"[21]

[19] Comment 8, Broadcast Music, Inc., at 2.
[20] *See* Comment 5, American Society of Media Photographers, at 5; Comment 15, Graphic Artists Guild, at 10.
[21] Comment 23, NWReflections, LLC, at 2.

Respondents noted that this additional feature would not only help them in their own record keeping but could also help connect users of copyrighted content with rights holders. Additional security measures would have to be put in place to protect against infringement, such as watermarking.

Many participants in the project expressed a strong interest in the Copyright Office utilizing mobile technology. Internal and external users alike noted that the registration system should be compatible with mobile devices, especially since so much creative content is now being created on cell phones, tablets, and other devices. Authors and creators should be empowered to register their works at the moment of creation, from their mobile device. In addition to being able to register works, users should be able to search copyright records, research copyright information or check the status of a pending registration from their mobile device. In an age where things as complicated as banking and finances can be transacted from a mobile device, the Copyright Office needs a mobile platform.

Mobile devices may also be used to support internal Copyright Office functions. Currently, barcode readers used to track copyright deposits must display on a monitor attached to a desktop computer. However, mobile devices that would display as well as scan would expedite workflow. Selection officers in their rounds to Registration and CAD (Copyright Acquisition Division) must have access to a desktop computer for research at each location they visit. However, these could be replaced with a single hand-held device, integrated to the various Copyright Office and/or Library systems.

Both the wider copyright community and staff would like the Copyright Office to incorporate alternative means of correspondence in a twenty-first century communications strategy. Currently, the Office uses phones, email, letters and fax. However, users would like to see the Office take advantage of chat and instant messenger technologies.

ASMP notes, "*[m]ost users of today's computers are accustomed to, and expect, almost instantaneous support, whether by telephone, instant chat, online FAQ's or even user forums linked to the main website.*"[22]

Or, as with many online customer support centers, applicants can leave a phone number and expect to receive a call within moments. Using profile management tools, applicants could indicate their preferred method for being contacted by the Copyright Office, including days and times. Internal staff would welcome this information as well for managing their workloads. Internal users also suggested that communication be automated to the extent possible. And of course, any new copyright system must be fully compliant with Section 508 of the Rehabilitation Act as well as the Americans with Disabilities Act.

[22] Comment 5, American Society of Media Photographers, at 3.

Under the Register's Special Project for Public Outreach and Copyright Education, the Copyright Office has begun utilizing webinar and videoconference technologies.

The Technical Upgrades project received a similar recommendation from **Education Testing Service (ETS**), which files high volumes of copyright registrations for secure tests. Currently, such applications must be presented in person. ETS notes that using instant messaging, video calls and online meetings "*would eliminate our travel time and travel expenses through the use of desktop sharing and conferencing.*"[23]

Finally, Copyright Office employees recommended enhanced foreign language support. Comments ranged from improvements to the copyright registration certificate and public record to how Office staff interacts with the public. Currently, eCO does not recognize diacritical marks and thus cannot accurately convey names or titles with accent marks. This should be amended so that the public record accurately reflects the work and author to which it pertains. Additionally, staff suggested that the Office adopt live translation for phone support and offer foreign language support and translation services for the public record.

5.2 Challenges with Existing Public Record

According to the participating parties, the second major focus area for system improvement is the public record. It concerned the project team to hear that the Copyright Office is not the definitive source for all copyright information, and that a few parties indicate they search other databases before visiting copyright.gov. This is a fracture in the Copyright Office's mission. Stakeholders would like to see more robust copyright records updated more timely and with richer data.

The **ABA-IPL** notes, "*[t]here are sometimes significant delays in the appearance of registration records in the Copyright Office's online database following the issuance of corresponding registrations.*"[24]

And **AIPLA** writes, "*[a]s to the searching system, my main quarrel is credibility. I know I have searched for items that should be there, and come up empty. I have little confidence in the present search system.*"[25]

Similarly, the **American Association of Independent Music (A2IM)** writes, "*[w]e believe the Copyright Office database should become a key searchable source for copyright information so that creators' works are easily identifiable and do not become orphan works.*"[26]

[23] Comment 13a, Educational Testing Service, at 1.
[24] Comment 2, American Bar Association, at 2.
[25] Comment 3, American Intellectual Property Law Association, at 6.
[26] Comment 1, American Association of Independent Music, at 1.

Metadata is a major issue. That is, in addition to records being more current, as a result of faster processing times, participants expect additional metadata in a twenty-first century system. Those who create and use visual arts works specifically cited the inclusion of metadata in images that could be used in copyright records. For example, metadata is automatically captured and stored by many digital cameras and the software used in digital photo editing. Some image metadata is not currently required on copyright registration applications, however its inclusion would make for a richer record.

As for metadata that is required, supplying a solution whereby that metadata can migrate from the image to the application would save applicants from having to manually key it in. The Tech Upgrades project team notes that metadata is a major issue where music rights are concerned, and it was therefore an essential discussion point in the Copyright Office's recently policy study, *Copyright and the Music Marketplace*, published in February 2015.

Various respondents had specific comments about what additional data fields to include in copyright records, with a number agreeing that information about change of ownership or recordation of transfer would be valuable, including updated contact information. Much of this data (such as change of address information) is gathered by the Copyright Office and is publicly available through examination of the physical records, but is not included in our online records.

Author Services, Inc. stated, "*there is not enough information to determine the ownership if there is a transfer and the contact information of the parties and/or owner and/or new owner are not available as part of the online record. It must be obtained either by visiting the Copyright Office in person and pulling the record or by researching the owner's name using other online research sources.*"[27]

Beyond making updated contact information available, several participants added they would like the ability to directly update contact information through the system. Other suggestions for additional data points in copyright records include: a note about whether a termination notice had been filed against a work, an indication as to whether the work was registered upon appeal, and whether a work was currently in a legal dispute.

Finally, there was considerable demand to have all of the Copyright Office records available online, including those prior to 1978. The Copyright Office is well aware of this concern and is addressing it through the Register's Special Project for Public Access to Historical Records. Since the funding was first obligated in 2009, the Copyright Office has digitized 35.8 million pre-1978 registration records as well as 661 record books (Copyright Catalogue entry). Currently, the Office is digitizing 2.4 million registration applications. Much more challenging are the means by which the Office will make these aging records searchable, especially since so many of them contain data that was initially captured with handwriting, including pencil.

[27] Comment 7, Author Services, Inc., at 3.

Copyright patrons need the search functionality of copyright records to be greatly enhanced. The project team heard that the existing Voyager system crashes, times out and cannot handle large volume requests. Put differently, there is an upper limit to the number of records returned on a given search. Even if more records exist that match the requested data, they will not be provided. Hence researchers that may need to inventory all of the registrations owned by a particular party cannot currently do so.

Other times, searchers are overwhelmed with the results they receive and have requested that the Office implement sorting and filtering capabilities. In this scenario, a searcher would conduct an initial search and then apply filters to the results either to narrow their results or to compare results based on alternative criteria. Examples given by participants included filtering by format of work, date of first publication, nation of publication, and so forth. The current Voyager system offers certain search limits but they cannot necessarily be combined, cannot be applied to results after the fact, and are generally just not as flexible as the community would like to see.

Music Reports, Inc. commented, "*[w]hen search results are returned by the Office's system, they tend to include numerous irrelevant results. Specifically, the publicly searchable records of the Office outside of the Office itself do not permit a researcher to filter the results by subject matter type. . . . Implementing search filtering by work type would be a useful improvement to the Office's public-facing search capabilities.*"[28]

Additional capabilities suggested include the ability to search by partial names and even more nuanced search strategies such as combinations of names/partial names, date ranges, and key words all at once. The Voyager system does offer some Boolean search options where users may construct advanced searches, but not to the level of granularity many users require.

5.3 Inadequacies of Current Copyright Data

As previously discussed, the establishment, preservation and maintenance of the public record are core missions of the Copyright Office. Both the Office and the copyright community would like to see more robust, relevant and useful records. Improving the public record begins with gathering additional relevant data. The project team recognized early on that data would play a vital role in technical upgrades to the Copyright Office.

5.3.1 Data Exchange

Of all the suggestions participating parties shared with the project team, perhaps the most frequent recommendation was that the Copyright Office pursue business-to-business capabilities to enable data exchange and batch processing, including the use

[28] Comment 20, Music Reports, Inc., at 3.

of application program interfaces (APIs).[29] Both internal users and external users of the current copyright system, as well as interested parties from the copyright and technical communities, saw a great opportunity for the Office to enhance its records and its services with business-to-business technology. On the simplest level, many registration and recordation customers would like to see batch processing capabilities so that they may directly export data to populate registration applications or provide the titles for recordation. Currently, data is manually keyed in on a title-by-title basis, both by external users and internal users. Batch processing capabilities would not only create a more effective, efficient system but could lead to greater volumes of registration and recordation.

Batch processing or business-to-business transactions would not simply benefit applicants who submit high volumes of claims. Historically, this has been the vision of how batch processing would be utilized, but in fact this capability would service the entire community. Certainly high-volume applicants have much to gain from being able to directly export data to the Copyright Office. However, even individual creators and small businesses would benefit. The Office has already heard that users would like a simpler interface and that the community would like to see more data in copyright records. To accomplish this without asking users to spend even more time keying in data, the Office should develop an API. In some cases, such as digital photography, that data is already stored within the copyrightable work itself and simply needs to be harvested.

In order to register a work for copyright protection through eCO, creators go to the Copyright Office website and enter the required information into successive screens. However, this is an outdated model while modern approaches use an online transaction model. As rapidly evolving technology continues to permeate our lives, transactions must become more fluid. The Copyright Office needs to be adaptable and our systems must be able to receive data in the various methods people wish to provide it. Web services or web APIs offer this flexibility. This approach may also improve the quality of the data; since it requires fewer keystrokes, there is less potential for error.

An API would also allow for direct interfaces with other services. Copyright subsists from the moment of fixation. Ideally, creators could register their works at that moment and from that creative space. In other words, the Copyright Office could publicly offer an API with which software developers could create a rich tie-in whereby authors can submit registration applications from their word processor, their photo editing application, their sound recording program or their graphic design suite. Tie-ins could also be created with parties throughout the copyright community. Content creators do not always register their work with the Copyright Office but may be involved in other registries at performance rights organizations, writers' guilds and licensing services. An API would allow these copyright owners to register their works through seamless interoperability and third-party services. Again, in the twentieth century model, affiliates of the Copyright Office may have directed their patrons to copyright.gov through a link

[29] *See* discussion of APIs *infra*, section 6: Enabling Technologies.

on their websites. But in the twenty-first century, creators can enter their information in one location and that data would be shared through an API approach.

These activities would not only grow the Copyright Office database and promote registration and recordation, but also put it in a position to offer data verification. Some of the participants in this project do not file registrations or record documents, but do search copyright records to look for matches within their own information. For example, performance rights organizations that manage registries for the purpose of collecting and distributing royalty payments need copyright data. Searching and matching exercises are currently performed manually, but with data exchange capabilities, organizations could automatically compare their information to the official public record maintained by the Copyright Office.

In fact, use of APIs may not only generate additional copyright registrations and recordations to enhance the value of the public record, but could grow the entire copyright community and promote the national copyright system itself.

SoundExchange shared this vision with the project: *"Rights management organizations and other data aggregators could utilize APIs in order to learn about newly copyrighted works, either by generating a request to the Copyright Office that initiates a response containing the information, or by subscribing to a data feed maintained by the Copyright Office that notifies subscribers when new works are successfully ingested into its systems. Finally, third party developers and other services could utilize these APIs in support of new and innovative applications for collecting and disseminating information regarding copyrighted content across the supply chain."* [30]

5.3.2 Data Model and Data Standards

Several of the interested parties interviewed suggested the Copyright Office, working with the copyright community, begin with creation of a standard data model. This model would determine the structure of the data collected, stored and transmitted by the Copyright Office and specify the relationships and constraints between data points. Before implementing any system improvements, the Office should consider which data elements to include and under what schema. Former Associate Register for the Registration Program, Nanette Petruzzelli, writes, "*[a]lthough the Office is a department of the Library of Congress, the Office now creates (unlike the Library) records of works which give copyright information as opposed to bibliographic (library) information. It is this copyright information, in the form of consistent, accessible and legally meaningful data that should make the Office a one-stop shopping site for the copyright/protectibility status of so many works.*"[31]

[30] Comment 28, SoundExchange, Inc., at 6.
[31] Comment 26, Nanette Petruzzelli, at 3-4.

As stated at the outset of section 5.3, the wealth of information and activity regarding data standards could warrant its own study. However, a high-level discussion of data standards follows. Along with improving the external user interface and enhancing the public record, the most commonly cited suggestions were to develop an API and adopt data standards.

Parties participating in this special project emphatically recommended that the Copyright Office utilize global data standards and take a leadership role in supporting the use of standards. Whether the standard identifies a work, identifies a party or organization, or establishes a format for messaging, the digital world runs on standards.

For example, the **ISRC Agencies** write, "*[g]iven the increasing importance of both digital distribution and electronic recordkeeping with respect to all manner of copyrighted works, we believe the Office would be remiss if it failed to position itself now to collect information that will be of increasing importance in the digital age.*"[32]

Interested parties would like, at minimum, for the registration applications to contain optional fields for standard identifiers that would be validated during the examination process by registration staff, such as ISBN or ISSN. Some parties urge the Copyright Office to consider requiring these fields and recommend validation be automated. Certainly, they advised that any system improvements should be constructed anticipating changes in industry standards. Ideally, the system would be flexible enough to allow for identifiers to be added post-registration to enhance existing records. An advanced system would be able to validate the requested data standards. Improperly formatted codes would be flagged and a notification could be sent to the user. Many codes include a country identifier in their structure or contain a check character. Validations for these aspects can be built into the system and compared against the application data. Some envisioned a truly sophisticated system that could check for duplicate identifiers, offer suggested resolution for problem identifiers, or even refer an applicant without a code to the service provider. In the case of the International Standard Recording Code (ISRC), the Copyright Office could become an ISRC manager and provide codes to applicants who do not have one.

Data standards would enrich copyright records and searching copyright records would become significantly more successful. An added benefit would be that they create an opportunity to exchange data with registries around the world.

County Analytics Ltd writes, "*[b]ecause standard identifiers are in general global in scope, the Copyright Office data would become interoperable with worldwide databases.*"[33]

The following list is not exhaustive but merely highlights some key data standards recommended to the project team by participating parties:

[32] Comment 17, ISRC Agencies, at 4.
[33] Comment 10, County Analytics Ltd, at 9.

International Standard Book Number (ISBN), International Standard Serial Number (ISSN): Copyright registration applications currently include an optional field for applicants to provide ISBN and ISSN information. However, parties are not able to search copyright records by ISBN or ISSN. In addition, these fields are neither required nor verified during registration.

International Standard Recording Code (ISRC): Used to identify sound recordings and music videos, the ISRC is widely used by digital music services, performance rights organizations and others.

SoundExchange advises, "*[i]f the Copyright Office collects ISRCs at the point of registration, then the public can use ISRCs as a defined connection point between third party databases and the Copyright Office's records. The public will be able to search the Copyright Office's databases more easily . . . . This, in turn, strengthens the public's trust in and reliance on the Copyright Office as a repository of valuable information*."[34]

International Standard Musical Work Code (ISWC): Used to identify musical compositions, the American Society of Authors, Composers and Publishers (ASCAP) recommends that the Copyright Office include ISWCs in registered works.[35]

International Standard Name Identifier (ISNI):
Uniquely identifies individuals and/or organizations to prevent disambiguation. Separate ISNIs are provided for an individual, his pseudonym and his corporation.

The **ISNI International Agency** notes, "*the use of an international standard identifier such as ISNI will, if included in the public database held by the Copyright Office, allow its interoperability with databases whether crowd-sourced or operated by collective management organizations or commercial entities. This will allow innovative services to be created, generating better awareness and consumption of copyrighted works, thereby promoting 'the progress of science and the useful arts.'*"[36]

Online Information Exchange (ONIX): A standard format for capturing bibliographic data related to books.

AAP notes, "*[i]n conjunction with publishing more digital content, many publishers already submit some ONIX metadata on copyright ownership to third party entities in the publishing supply chain. Hence, publishers are, in general, optimistic about the possibilities of bulk data transfers using ONIX . . . .*"[37]

Common Works Registration (CWR):

[34] Comment 28, SoundExchange, Inc., at 3.
[35] Comment 4, American Society of Composers, Authors and Publishers, at 7.
[36] Comment 16, ISNI International Agency, at 2.
[37] Comment 6, Association of American Publishers, at 8.

A standard format for registration information designed to facilitate information exchange by music publishers.

ASCAP explains, "*[t]o ensure . . . that the entire world musical works repertoires are aligned, works registration follow CISAC*[38]*-agreed registration standards, referred to as 'Common Works Registration.'*"[39]

Digital Data Exchange (DDEX): A standard format for XML messages to exchange metadata.

SoundExchange recommends that "*the DDEX working group could define a profile for use with the Copyright Office's databases*."[40]

Incorporating internationally accepted standard identifiers would strengthen copyright records and generate dynamic opportunities for their use. But additional metadata may also be considered. Parties recommended the Copyright Office create its own identifiers as needed, such as naming conventions for the various types of registration applications or unique customer identifiers. Within eCO, each service request has a unique number but that number does not tell the user – internal or external – about the case.

Simple conventions could distinguish applications by the type of work being registered at various levels of granularity, or it might indicate whether the application is a basic registration, a group registration or an appeal. Also, eCO customers are identified by their login or, if they have a deposit account, their account number. However, the various employees of a given publishing company each have individual logins, so there is no easy way to query cases throughout the organization. Unique identifiers for parties could be used in a number of ways to search and sort cases. Finally, it was recommended that an upgraded system would allow users, particularly external users, to provide optional identifiers of their own such as docket numbers for attorneys, release dates of sound recordings, universal product codes for commercial goods, etc.

5.3.3 Data Repository

A data warehouse with robust analytics and business intelligence capabilities would be a valuable investment for the Copyright Office. The reporting capabilities of the current system are very limited. Hence the staff is not sufficiently enabled to work at their greatest productively and managers at all levels are not empowered to make vital business decisions, potentially missing opportunities. Building a data repository would begin with centralizing the multiple data sources and sets into a single location. For

[38] Confédération Internationale des Sociétés d'Auteurs et Compositeurs (CISAC) (an international creators' rights group).
[39] Comment 4, American Society of Composers, Authors and Publishers, at 5.
[40] Comment 28, SoundExchange, Inc., at 7.

example, some parties would like to see connections between registration and recordation data. Another component to the data repository is to provide necessary long-term storage solutions for both data and electronic deposits. Such measures would not only be responsible, proactive stewardship of valuable copyright information, but would improve the Office's data exchange capabilities.

The Copyright Office data warehouse would effectively and efficiently centralize all of the Office's existing systems containing copyright information, but may also include new databases for crowd-sourcing and third party data. While participant parties respect the official copyright records and wish for them to remain clearly identifiable, some suggested that the Office also provide additional information sourced from third parties that appeared alongside, but did not replace, the official legal record. Increasingly, crowd-sourcing is proving to be a successful means for growing vast bodies of data. However, the benefit of this method is that the information is provided by the public and does not compete for resources. Such data could be valuable to the community but must be clearly demarcated and distinguished from the official and authoritative records created by the Copyright Office. Alternatively, the Office might wish to devote resources towards validating data received from third party sources in order to incentivize participation. If an organization were to share its data with the Office and receive our endorsement, it would buttress that organization's efforts. Parties presented various strategies to the Tech Upgrades project team that should be further vetted and explored for collaborative opportunities.

5.4 Outdated Architecture & Infrastructure

The project team met with interested parties throughout the copyright community not only to solicit information about their experience with the Copyright Office's business, but also to establish a better understanding of their businesses. Parties shared technological solutions that have proven successful for them, as well as information about the challenges they faced.

Many of the organizations participating endorsed cloud implementation strategies as a recommended alternative to maintaining and updating their infrastructure. Benefits include cost savings – especially for reducing expenditures on software licenses – greater efficiency and improved system elasticity. Several parties indicated that they have been using cloud solutions for years after weighing its benefits with risk.

The U.S. Patent and Trademark Office shared that they are partnering with Amazon Web Services for a proof of concept for a patent system project. Additional methods of reducing operations and maintenance budgets parties shared with the project include outsourcing and remote support. Of course, as discussed previously, the registration system must maintain adequate security for electronic works, and transitioning to systems where much of the hosting and processing is effectively outsourced will require careful consideration of the security risks associated with offsite hosting, server management, and the like.

As a best practice, the Copyright Office should first develop an independent IT enterprise strategy and plan reflective of its needs to support the creative industries. Another best practice found throughout the copyright and technological communities is to provide a high level of system redundancy that leads to a higher degree of reliability for our line of business. There is a vital need for redundancy to provide full system backup, thereby reducing risk to loss of data. In addition, these improvements would enhance the Office's availability by minimizing system down time, whether due to planned outages for maintenance or unplanned failures or crashes. Greater availability would also better accommodate our users, particularly international users, and could potentially enhance Copyright Office revenues.

The interested parties interviewed suggested that as the Copyright Office proceeds with implementing upgrades consistent with the enterprise strategy, it would be wise to develop iterative releases to accomplish a phased plan of improvements. In addition, they recommend the plan be communicated to all users, in particular external users. Copyright Office customers may need time to adapt and strategize how to alter their own business processes as legacy systems are replaced. Awareness of the Office's implementation plan could promote buy-in, provide reassurance and perhaps gain patience from the public.

Of significant interest is to deploy highly secure systems and mechanisms to meet known threats and risks to the copyright community. The Copyright Office is well aware of the ongoing need to protect our systems for a number of reasons including safeguarding records that may contain personally identifiable information, or preventing unauthorized distribution of copyrighted works. However, even our own internal staff expressed interest in greater measures to protect against potential piracy of valuable electronic deposits. External users, as well, voiced this concern.

For example, **ASMP** shared, "*[i]n recent months, there has been a lot of media coverage of vulnerability to cyberattacks, and security of data has become a major concern. Like almost every web-based application, eCO appears to need a significant upgrade in every security aspect in light of these recently apparent threats.*"[41]

Issues surrounding personally identifiable information in registration records are not new to the copyright system, but take on new importance in the digital age. In considering the appropriate scope of public information, the Office must weigh the value of incorporating certain information in the record to facilitate the identification of copyright owners, against the potential risks that such information could be used for nefarious purposes.

Finally, the next generation Copyright Office will also require improvements to its remote workforce capabilities these should be incorporated into planning in conjunction with client-focused upgrades.

[41] Comment 5, American Society of Media Photographers, at 3.

6. Enabling Technologies

6.1 Technologies

This section includes a description of technologies that may be considered to improve support of the U.S. Copyright Office and its interested parties within the copyright community. The technologies considered were either identified by the Special Project team or were discussed with interested parties during the interview phase of the project. It is expected that derivative technologies from those included in this section will also be considered as more detailed requirements are obtained in the deployment of a modernized Copyright Office.

6.1.1 Application Programming Interface

Technology Description

At the broadest interpretation, an Application Programming Interface (API) is a mechanism that specifies how software components should interact with each other. Traditionally, APIs were developed and distributed by organizations to be included in external applications that were developed by other companies. For example, Adobe Corporation developed an API that when included in Microsoft Word maintains the “look and feel” of Word while providing the added capability of creating PDFs.

Organizations that create APIs allow for market expansion and also benefit end-users by providing an integrated solution. There has been considerable expansion of the API concept as the Internet and mobile computing technologies continue to proliferate. APIs can be deployed in various forms. Although the traditional form of an API is still being developed and distributed for use in other applications (internal and external to organizations), web based APIs have seen significant growth.

In this context, web-based APIs are installed on servers, which do not require distribution and end-user installation. Commonly, web based APIs are installed to provide business-to-business functionality and mobile device transactions. Web-based APIs are also referred to as “web services.” For example, Pay.gov offers a web service to government agencies to process customers’ payments. This service allows the agencies to develop their own user interface to accept payments that are processed, transparent to the user, by exchanging data securely between two systems.

Technology Benefit

The Gartner Group discussed the importance of APIs in the banking industry at their Gartner Symposium/ITxpo 2012. Although not all industries operate like banks, the copyright community has similar needs to the banking industry when it comes to sharing data and functionality securely. Gartner said that, “apps, not applications, enable a new

style of engagement with customers—one that is focused on providing needs-based and context-aware services."[42] The Gartner Analysts believe that APIs will allow banks to deliver services based on customer need relevant to the customer's context. For example, if the customer is used to working with a specific financial software package, APIs will provide a layer of abstraction so customers will not be forced to utilize a different software package when interacting with an external bank. In addition, providing an API interface to the bank will allow third-party software developers to market solutions based on customer need.

Another significant benefit of APIs is that they provide other opportunities to expand market space and provide enhanced services to their customers and partners. This includes:

- Partner connectivity (B2B)
- Mobility programs
- External developer ecosystems
- Cloud integration
- Internal application data exchange

It is clear that server-side APIs are critical in supporting business-to-business (B2B) operations. There are well-established standards that support secure web-based communications between disparate organizations. This capability has enabled B2B to flourish. However, security, scalability, performance, complexity, regulatory compliance and integration are still challenges to organizations. Therefore, deployment of APIs require IT governance and reliance on service-oriented architecture (See section 6.1.4, Service Oriented Architecture, below) to obtain the greatest value.

APIs have also enabled organizations to interact with mobile devices that run applications or "apps" designed specifically for the format of the mobile device (phone, tablet, etc.). Both client-side and server-side apps have been created to provide data exchange between mobile devices and organizations. As more and more devices and systems exchange data via APIs, an ecosystem has emerged that allows users to work with data through multiple platforms. For example, a transaction such as form submission could be created via a mobile device, updated via a browser on a workstation, and be viewed from a software packaged installed on a workstation. Apple has been successful using this approach with their devices. Their iPhone, Apple TV, iPad, and workstation platforms all interact with each other. Users that start with one of their devices begin to see benefits as they add additional devices that are part of the Apple ecosystem.

With the emergence of cloud computing, APIs are being created by organizations to provide cloud service providers with the ability to resell software and services to larger audiences. For example, a sales application from one company can be integrated with

[42] Press Release, Gartner Group, Gartner Says Banks Should be Banking on APIs and Apps, Not Applications (Oct. 30, 2012), http://www.gartner.com/newsroom/id/2217415.

a marketing application from another company and hosted as a service on the Internet. A detailed discussion of cloud computing is presented in 6.1.2.

Lastly, APIs are also being created for internal use as well. The current approach most organizations are taking is to develop applications as services to provide a higher level of integration and re-use of custom software. For example, an application could be created to share copyright data in real-time. The application could also have further functionality that provides reporting, querying, and other business capabilities. These services can be delivered in the form of APIs and part of the organization's Service Oriented Architecture (see section 6.1.4, below).

6.1.2 Cloud Computing

Technology Description

Cloud computing is emerging as a viable alternative to the traditional approach of standing up and hosting applications within an organization.

The National Institute of Standards and Technology (NIST) defines cloud computing as, *"a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g. networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction."*[43]

Traditionally, organizations have purchased and deployed infrastructure (networks, hardware, software, etc.) and dedicated the IT resources to specific applications. For example, an accounting application would require a specific number of servers, storage, network resources, etc. Internal or contracted staff would be assigned to specific applications or hardware and software to maintain the datacenter.

Cloud computing differs from a traditional datacenter approach in that specific hardware and software is not purchased for a specific application and/or customer. Instead, using a shared platform model, infrastructure is deployed to support multiple applications and IT resources are provisioned out as a service. Consumers do not need to know nor be directly concerned with the underlying infrastructure that supports their IT systems. Rather, consumers are mostly concerned about service delivery. There are several characteristics of cloud computing that differentiate the capability from more traditional IT implementation. This includes the following:

- On-Demand – Ability to provision IT services immediately without needing to necessarily purchase hardware and software. Deployment of IT resources typically occurs almost immediately or within hours.

[43]Peter Mell, Timothy Grance, *The NIST Definition of Cloud Computing*, NIST Special Publication 800-145 at 2, Sept. 2011, http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf.

- Internet/Intranet Access – Applications or services can typically operate over the Internet or Intranet. This allows multiple types of devices such as mobile laptops, phones, tablets, etc. to access resources.
- Resource pooling – Hardware, software, network, storage, and other IT resources can be pooled so that multiple user groups can share the same underlying infrastructure while maintaining their own security boundaries.
- Elasticity – Applications and computing resources can be expanded and contracted based on need. This provides for rapid scalability based on demand.
- Measured Service – IT services can be measured and charged based on utilization. This characteristic has also provided the ability to use computing as a utility. Much like electricity or water is charged, computing can be charged based on utilization. Resource utilization can be monitored and reported to provide transparency to the consumer of the services. However, this presents a different cost model than many organizations are accustomed to.

NIST has defined several service models for cloud computing. A service model describes the capability that the cloud service provides. The three service models include:

1. Software as a Service (SaaS) –This capability provides software to users that is typically accessed through a web-browser. This includes applications such as email, database applications (*e.g.*, contact management), customer relations management (CRM), etc. Customers do not manage the software, operating systems, network, servers, storage, etc. under a SaaS service model.
2. Platform as a Service (PaaS) – This capability allows consumers to deploy their own custom applications in a cloud environment. Consumers are responsible for managing their own applications but do not need to manage the underlying operating systems, network, servers, storage, etc.
3. Infrastructure as a Service (IaaS) – This capability provides consumers with the most flexibility but requires consumers to be more involved with the management of their environment than any of the other service models. Consumers are not required to manage hardware but they are required to install and manage operating systems and applications.

NIST has also defined deployment models of cloud services. A deployment model describes where the physical infrastructure is deployed and who manages it. NIST includes four deployment models:

1. Private Cloud – IT infrastructure is deployed for a single organization but may be used by multiple business units. For example, a single government agency may deploy a private cloud within its datacenter that would only support business units within the agency. IT infrastructure is purchased, owned, and managed by the agency. Each business unit would share the common IT resource and be billed based on what it consumes.
2. Community Cloud – IT infrastructure is deployed for exclusive use by a specific community of consumers with the same mission, security requirements, policy,

etc. For example, Amazon Web Services (AWS) provide a cloud service that can only be used by government and organizations responsible for managing government systems. In this case AWS is a third party responsible for the community cloud. However, a community cloud may also be deployed and managed by multiple organizations that share use of the community cloud resources. The infrastructure may be deployed on or off-premises.

3. Public Cloud – IT infrastructure is deployed for use by the general public. The infrastructure is owned and managed by a commercial business or government organization. The infrastructure is located at the public cloud service provider's location but may be used by organizations outside of the cloud provider's location
4. Hybrid Cloud – IT infrastructure is a mix of two or more cloud deployment models (*i.e.*, Private, Community, and/or Public). For example, an organization may deploy a private cloud to support their internal line-of-business applications, and use a public cloud service provider to host applications that support the general public.

Technology Benefit

Cloud services are growing at a rapid rate because of the benefits that they bring. Gartner predicts strong growth in public cloud services, with a growth of 18% in 2013 alone.[44] Gartner predicts that the market will grow an additional 38% by 2015. However, this does not necessarily apply to all the deployment models (*i.e.*, private, community, and hybrid cloud). Business and security challenges have not kept up with the technology growth. This includes establishment of common criteria for service level agreements (SLA), and accreditation of secure environments for use by government agencies. Nonetheless, there are compelling reasons for both private and public sectors to consider deploying systems using cloud technologies.

Most organizations in public and private sectors are already in the process of deploying private cloud solutions since they present the least amount of risk while still attaining gain. The main reason for this is that private clouds are completely under control of the organization where security risks can be contained and where SLAs can be more easily adapted to reflect cloud operations.

A Gartner poll by Forbes showed that 75% of respondents intend to deploy a private cloud strategy by 2014.[45] Most organizations are seeing benefits of private cloud solutions when deploying IaaS and PaaS service models. Because the first step to

[44] Gary Flood, *Gartner Tells Outsourcers: Embrace Cloud Or Die,* INFORMATIONWEEK (July 31, 2013, 2:12 PM), http://www.informationweek.com/cloud-computing/infrastructure/gartner-tells-outsourcers-embrace-cloud/240159246.

[45] Louis Columbus, *Hype Cycle for Cloud Computing Shows Enterprises Finding Value in Big Data, Virtualization*, FORBES (Aug. 4, 2012, 12:19 PM) http://www.forbes.com/sites/louiscolumbus/2012/08/04/hype-cycle-for-cloud-computing-shows-enterprises-finding-value-in-big-data-virtualization/.

developing a private cloud requires deployment of virtualization technologies, companies quickly realize benefits as they move to a private cloud solution. Virtualization refers to the creation of a virtual (rather than physical) component of IT infrastructure. For example, in the past if a system required five servers, five separate physical computers would need to be deployed to support the application. Virtualization allows the capability to install all five servers, virtually, on one or a few physical machine(s). This is possible because of the advent of the faster processing power of today's computers. In this example, the five servers operate autonomously on a single piece of hardware. There are several benefits of virtualization that include energy efficiency, lower cost of ownership (less hardware, maintenance, etc.), faster deployment of infrastructure, increased up time inherent to virtualization technologies, and improved disaster recovery.

Deployment of public, community and hybrid cloud deployments are moving at a slower pace. However, SaaS is rapidly gaining adoption by commercial and government organizations. Gartner forecasts more than 50% of enterprises will deploy applications based on the SaaS service model by 2015. The reason for this is that there are numerous benefits to cloud services. In 2011, the U.S. White House released its Federal Cloud Computing Strategy[46] that listed several benefits of cloud computing. The report included the following benefits:

- Efficiency
    - Improved asset utilization (server utilization > 60-70%)
    - Aggregated demand and accelerated system consolidation
    - Improved productivity in application development, application management, network, and end-user
- Agility
    - Purchase "as-a-service" on a metered basis (*i.e.*, only pay for what is used)
    - Near-instantaneous increases and reductions in capacity based on demand
- Innovation
    - Shift focus from asset ownership to service management
    - Leverage private sector innovation
    - Faster deployment of new technologies

[46] Vivek Kundra, U.S. Chief Information Officer, *Federal Cloud Computing Strategy,* (2011) http://www.whitehouse.gov/sites/default/files/omb/assets/egov_docs/federal-cloud-computing-strategy.pdf.

The Federal Government has seen success deploying cloud solutions, following the SaaS service model. This includes deployment of Microsoft's cloud service that supports 120,000 staff. The service includes Exchange, SharePoint, Office Communications, and Live Meeting. The Library of Congress has also launched a successful Cloud project in support of their National Digital Information Infrastructure and Preservation Program (NDIIPP). The project allowed for perpetual access to certain digital content.

In considering future infrastructure and architectural options, the Copyright Office must contemplate migrating to cloud solutions in some form or fashion so as to realize the many benefits of cloud technology. In fact, in a Library of Congress Survey report, the Office of the Inspector General concluded that the Library needs to develop an enterprise wide plan to evaluate use of cloud computing and achievable cost savings.[47] Figure 1 shows the maturity level of the various service models. Based on the analysis in the report, it is clear that cloud technology offers some cost benefit to the Library of Congress, including the Copyright Office.

Figure 1 shows the maturity level of the various service models. Based on the analysis in the report, it is clear that cloud technology offers some cost benefit and the project team would agree that this is a good planning focus for endeavors specific to the Copyright Office.

[47] LIBRARY OF CONG. OFFICE OF THE INSPECTOR GEN., LIBRARY-WIDE INFORMATION TECHNOLOGY SERVICES: THE LIBRARY HAS AN OPPORTUNITY TO BE MORE PROACTIVE IN ADOPTING MORE EFFICIENT COMPUTING TECHNOLOGIES, Survey Report No. 2013-IT-101 (2013) http://lcweb2.loc.gov/master/libn/about/office-of-the-inspector-general/documents/rpt2013marCloudComputingServerEfficieny.pdf.

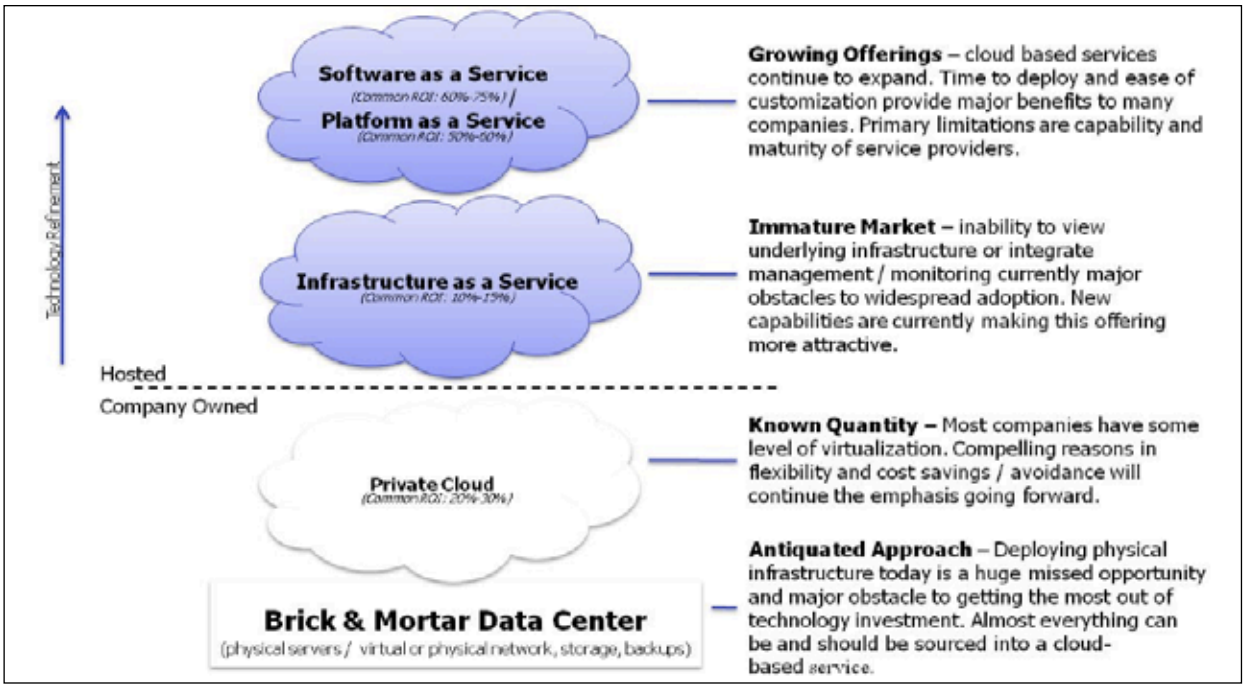

Figure 1 - Cloud Service Maturity (Source: Library of Congress, Office of the Inspector General)

6.1.3 Integrated Solutions / Business-Driven Ecosystems

Organizations have recognized that one-size-fits-all solutions are often times not possible for all users. For example, internal users focusing on financial information may not be interested in information regarding case management. In other cases, operations management may be interested in both financial and case management. External users may require a completely different system interaction experience altogether. Because of this, organizations are developing solutions that provide internal and external system developers with the flexibility to provide targeted and integrated solutions.

By leveraging Service Oriented Architecture technology (see section 6.1.4 discussed below) and web services (see section 6.1.1), systems are now being delivered that will provide a multitude of interfaces for end-users. This has enabled third parties to develop solutions targeted to specific user groups and thus allowing business-driven eco-systems to emerge. This section provides additional details on integrated solutions and how business-driven eco-systems can benefit both internal and external users.

Technology Description

Over time, organizations have developed a significant number of systems and solutions that include multiple vendors, complex configurations, and interfaces all designed to meet a user's needs. This approach has caused exponentially rising maintenance costs and has forced organizations to consider other approaches in order to better manage integrated solutions. With the advent of Service Oriented Architecture, web-services,

and expanding cloud solutions, organizations are beginning to look at the ecosystem paradigm model.

An IT ecosystem relies on more than one organization, vendor, and solution provider to meet the demands of users. For example, instead of developing or purchasing a point solution[48], organizations are investing in ecosystems. Ecosystems include solutions from multiple providers and may operate in multiple hosting environments (*i.e.*, in-house, cloud, external organization). This includes solutions that are designed to integrate with one another providing a rich experience to end users. This capability continues to grow as cloud solution providers begin to offer solutions as part of a SaaS service offering.
For example, a case management system may be developed or purchased by an organization that provides core functionality, relying on other solutions for value-added services. External organizations or third party vendors may develop functionality such as advanced querying, rendering (*e.g.*¸ printing), authoring tools, etc. This approach allows the market place to develop functionality based on market and user demand.

Technology Benefit

Integrated Solutions/ecosystems provide users with the widest selection of technologies. Although users may require core functionality from a single organization, users will have a richer user experience when third party solution providers offer additional functionality. By building interfaces into applications, organizations responsible for providing core functionally will benefit from greater expanded functionality in areas that are not within an organization's core competence. For example, organizations that are strong at forms processing could develop or select a case management system while relying on third parties to develop additional functionality (*i.e.*, authoring tools).

By developing an ecosystem, multiple vendors are able to provide competing products to appeal to various user needs. Cloud computing is a good example of rapid ecosystem deployment that enjoys reduced project risk, costs, and rapid deployment. By integrating various cloud applications and service offerings, users will begin benefiting from an integrated solution. As additional cloud-based applications and services are added, they are more easily integrated into a complete solution.

6.1.4 Service Oriented Architecture

A Service Oriented Architecture (SOA) is a key "building block" to the delivery of other technologies discussed within this document. SOA is a software architecture that is based on applications that provide functionality as services. A service is a self-contained logical representation of a repeatable function. For example, an application

[48] A software solution that provides a specific type of function such as finance, membership management, etc. The solution is typically a best-of-breed solution.

that authenticates users can be used in multiple applications that require users to be authenticated. Services can be combined to provide a full suite of functionality to an organization and its users.

Technology Description

A SOA is a collection of services that communicate with each other. This can either be simple data transfer between two services or it may involve several services working in tandem to perform an activity. SOA requires services that are well-defined, self-contained, and have the ability to operate independently without requiring the context or state of other services.

Web Services or Web API is a prime example of connection technology used in SOAs. Ecosystems also rely heavily on SOAs by integrating multiple services to fulfill a specific business function. For example, either several departments within an organization or third parties may develop or procure software written in different languages (following SOA principals) but still have the ability to interface with one another. Organizations can pick and choose which software modules to integrate without requiring software re-write. An SOA always includes the ability to re-use software modules.

Technology Benefit

As mentioned above, an SOA is an essential building block that is required in the delivery of other technologies recommended within this document. This includes application program interface (API), cloud computing, integrated solution/business driven ecosystems, and mobile computing. Deploying these technologies without leveraging an SOA would be extremely difficult if not impossible as well as inefficient and costly. For this reason alone, an SOA provides immediate benefit to organizations.

As illustrated in other sections, SOA allows simultaneous use and easy mutual data exchange between software modules written by multiple organizations and/or vendors and in multiple programming languages. SOA also allows software reuse resulting in lower development and maintenance costs and higher quality. In addition, organizations have realized a time savings in deployment of systems.
SOA benefits include:

1. Agile Software Development – Organizations can respond more quickly to new business requirements and leverage existing services.
2. Cost Reductions – Organizations reduce software development costs by taking advantage of software reuse, which greatly reduces software development costs.
3. Increased capability with lower return on investment (ROI) – Rapid deployment of new business functionality is possible by leveraging software that is already available, thus reducing the ROI.

6.1.5 Mobile Computing

Mobile computing continues to grow in the market place. Gartner Group has predicted the traditional PC market to decline 7.6% in 2013 as consumers transition to tablets and ultra-mobile devices.[49] Mobile computing includes hardware devices such as smart-phones, tablets, mobile computers, and other portable computing devices. This section will focus on how mobile computing is emerging as a first tier mechanism to interface with business and government.

Technology Description

A mobile computing device (MCD) is any device that utilizes components designed for portable use. An MCD comes in several forms that include phones, tablets, and ultra-portable screen/keyboard devices that look similar to laptops. An MCD must be capable of operating, executing and providing services and applications like a typical PC but may be limited in screen size and/or input capability (screen based keyboard). MCDs have similar hardware and software components as used in PCs, such as processors, random memory and storage, Wi-Fi, Bluetooth, etc. However, they differ from PCs in that they are specifically built for lower power consumption.

Lower-cost devices in combination with accessibility (smaller, portable) have spurred the growth of system access via mobile computing. Mobile computing encompasses more than the end-user devices (*i.e.*, phone, tablet, etc.) itself.
Mobile computing also includes infrastructure and software systems optimized for mobile devices. While there will still always be users who interface with systems using a traditional PC, organizations will need to provide the capability of accessing their systems using a mobile device.

Organizations that deploy mobile systems for internal use must also deploy infrastructure that helps manage devices and provide sufficient security. This includes deployment of Mobile Device Management (MDM) software. MDM software provides organizations with the capability to secure software and systems accessible through the mobile device. In addition, MDM software provides monitoring tools for mobile devices to ensure mobile access policy is being followed. Finally, MDM provides the ability to distribute applications, data, and configuration settings over-the-air or by connectivity to the internal network.

Mobile devices are being used by organizations for access to internal systems and by end-users to access business systems. For example, tablets are being used for email access, presentations, or access to inventory systems. End-users have a plethora of applications for everyday tasks (email, mobile texting, access to weather, news, etc.). In addition, organizations are beginning to develop interfaces into their business

[49] Press Release, Gartner Group, *Gartner Says Worldwide PC, Tablet and Mobile Phone Combined Shipments to Reach 2.4 Billion Units in 2013* (Apr. 4, 2013) http://www.gartner.com/newsroom/id/2408515.

systems that are designed specifically for mobile devices. For example, airline reservation systems have developed the necessary software on servers to provide airline booking information in a format that can easily be read by users of the mobile device.

Technology Benefit

The Gartner Group predicts that the decline in PCs will continue beyond 2013. This will require organizations to adjust their delivery model to accommodate mobile computing. As shown in the table below, The Gartner Group predicts that mobile phones will see the most growth over the next four years.[49]

Worldwide Devices Shipments by Segment (Thousands of Units)

Device Type	**2012**	**2013**	**2014**	**2017**
PC (Desk-Based and Notebook)	341,263	315,229	302,315	271,612
Ultramobile	9,822	23,592	38,687	96,350
Tablet	116,113	197,202	265,731	467,951
Mobile Phone	1,746,176	1,875,774	1,949,722	2,128,871
Total	**2,213,373**	**2,411,796**	**2,556,455**	**2,964,783**

Organizations recognize the need to deploy mobile capabilities to meet customer demand. This has resulted in developing software and web interfaces optimized for mobile device utilization. Some of the benefits realized by organizations include:

1. Improved Business Productivity – Provide immediate access to systems anywhere and anytime.
2. Reduced Operation Costs – Increase visibility into business systems faster (no need to be at the office/desk) to correct issues or logistics problems.
3. Improved Customer Relationships – Provide additional opportunities to connect to systems to obtain information anytime, anywhere.
4. Add functionality – Organizations can develop applications that take advantage of unique mobile computing hardware (*i.e.*, cameras, recording capability, etc.).

6.1.6 Big Data

"Big Data" is a term used in information technology to refer to extremely large data sets that are stored in large database systems and rely on high-performance systems to capture, store, and analyze data to provide useful Business Intelligence (BI).[50]

Technology Description

Organizations have always relied on data collection and analytics to provide business intelligence to help executives make business decisions. As data capacity and processing speed has increased and the cost of both has decreased, larger data sets are being analyzed.

Organizations have the ability to collect much more business data today than in years past. Big Data requires large processing power which is becoming readily available to organizations. Big data also requires vast amounts of disk storage involving hundreds if not thousands of computers processing data in parallel. With the advent of cloud computing, leveraging external resources for data storage and processing is becoming a viable and cost effective method of big data analysis.

Gartner defines Big Data as, "*high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making*."[51]

The main reason is Big Data requires more than just computing power but specialized software that can analyze large amounts of data. The software required includes specialized database software that can handle larger amounts of data than traditional relational database management systems (RDBMS).

A major proponent of Big Data deployment is Netflix. Netflix deployed a Big Data solution to determine the quality and reliability of video streams to their customers.[52] In addition, they utilize Big Data to help programmers determine what their customers are most likely to watch. As the company grew, Netflix began to run out of capacity in their internal datacenters so they decided to move to a cloud provider (Amazon Cloud). Netflix decided on a highly scalable, high-performance, data processing software known as Hadoop. A combination of cloud computing resources and Big Data processing software allows NetFlix to grow or scale back as their needs change.

---

[50] Business intelligence is a set of theories, methodologies, processes, architectures, and technologies that transform raw data into meaningful and useful information for business purposes. WIKIPEDIA, "*Business Intelligence*," http://en.wikipedia.org/wiki/Business_Intelligence.
[51] Gartner Group, *IT Glossary*, "*Big Data,"* http://www.gartner.com/it-glossary/big-data/.
[52] Joel Schectman, *Netflix Uses Big Data to Improve Streaming Video,* THE WALL STREET JOURNAL (October 8, 2:30 PM) http://blogs.wsj.com/cio/2012/10/26/netflix-uses-big-data-to-improve-streaming-video/.

Technology Benefit

According to Gartner, big data maturation will occur in the next five years.[53] However, organizations both in the private and public sectors are benefiting from early deployment. At a very minimum, organizations are developing their deployment strategy of Big Data technologies.

In order to take advantage of Big Data technology and realize the most gain, organizations will need to deploy a Data Analytics solution (see below). A combination of both will provide organizations with the business intelligence to not only gain efficiencies but also meet their customer needs, thus increasing growth.
Big Data technology provides several benefits:

1. Expand Market Space – Help organizations determine customer needs and develop new markets.
2. Increase Efficiency – Help organizations analyze operations, products, and services to determine where efficiencies can be made.
3. Tailor Services/products – Analysis of Big Data allows management to refine services/products that better meet their customers' needs.
4. Improve decision-making and reduce risk – Analysis of Big Data allows the ability to do predictive analysis on past and real-time data. This information helps reduce risk in the decision-making process.

6.1.7 Data Analytics

Data analytics (DA) is the process of examining data in order to draw conclusions about that information under review. DA is used by organizations to make better business decisions. Oftentimes DA is used in combination with Big Data (see above) as a necessary aid to process a large amount of data.

Technology Description

Data Analytics (DA) is the process of reviewing, normalizing (removing duplicate, invalid, or corrupt data) and modeling data to provide conclusions and useful information in the decision-making process. Data analytics is used by many organizations to make better business decisions. Data analytics focuses on inference or the process of deriving a conclusion based on all or a subset of captured data. DA is used in the sciences and information technology. For IT, DA is used to examine operations and processes. The analysis is used to determine whether the systems in place are operating efficiently and accomplishing an organization's objectives and goals. DA also helps organizations predict and improve services to their customers.

[53] *Big data won't be mature for at least five years, Gartner predicts,* INFORMATION AGE (Aug. 19, 2013) *http://www.information-age.com/it-management/strategy-and-innovation/123457283/big-data-won---t-be-mature-for-at-least-five-years--gartner-predicts*.

For example, DA has been used to help management determine help desk efficiency by analyzing data such as call volume, call duration, number of support technicians and end-user feedback.

Organizations that provide online services use DA to examine web site traffic or system use patterns to determine which services are more or less useful. For commercial organizations, this information can also be used to determine which product or service would most likely be the most profitable. Most DA systems include information dashboards supported by real-time data sources.

Technology Benefit

DA is an essential technology for managing large amounts of data (see Big Data above). DA provides organizations with information to help make them more competitive and target products and services that provide the most return on investment. Organizations use DA as one method for predictive analysis based on data captured by their systems. Government and commercial organizations continue to deploy DA as a tool to increase profitability and efficiency. As previously mentioned, in order to make sense of Big Data, DA is almost an essential tool. In fact, The Gartner Group believes that DA will grow because of organizations deploying Big Data solutions.

Gartner reported that organizations are seeing real and measurable return on investment from deployment of Big Data.[54] According to The Gartner Group, 42% of respondents to a survey indicated that they had either invested big data technology or were planning to do so. Because of this, The Gartner Group predicts that DA will grow along with this.

6.2 Deployment Challenges

Organizations deploying new technologies face challenges. Making a shift to new technology is typically accompanied by its own set of unique hurdles. Understanding the 'what' and 'how' of the work a particular organization is currently performing today ("as is" processes) is crucial to anticipating and managing most major challenges associated with deploying new technology.

According to Joel Barger, Project Director of Power Consulting, Meridium, "Many times the adoption of new asset management technology and processes fail because employees do not realize the need for change or understand the benefits associated with the change. This will require investments of time and resources in asset

[54] Douglas Laney, Frank Buytendijk, *Invest in Information and Analytics to Benefit From Big Data*, GARTNER (Mar. 8, 2013) http://www.gartner.com/id=2363115.

performance education to provide a broad-based understanding of the flow of asset performance information, as well as more specific technology training." [55]

In the case of early adopters of new technology, an organization is likely to be exposed to the problems, risks, and annoyances common to early-stage product testing and deployment. One risk in particular is that early versions of new technology may be prone to bugs and/or malfunction, and might cost the organization more at the time of release. On a positive note, the newer technology will in all likelihood become more efficient and less expensive within a short period of time following new release.

6.3 Adoption of New Technology

Not all the enabling technologies shown within section 6 are considered "new" technology. For example, Service Oriented Architecture has been successfully deployed for more than ten years now. In contrast, Big Data is relatively new and just beginning to grow. Most organizations should take into consideration their approach in the adoption of new technology when developing their deployment roadmap.
Some organizations consider the extension of the diffusion process to determine their adoption of new technology approach. The diffusion process was originally published in 1957 by Joe M. Bohlen, George M. Beal and Everett M. Rogers at Iowa State University and was related to application to agriculture and home economics. However, the same sociological principles apply to IT technology.

Figure 2 illustrates the adoption lifecycle model according to the diffusion process. There will be less adopters of technology (2.5% and 13.5% in the chart) early on and at the trailing edge of the cycle (16% in the chart). Most adoption is made in the middle of the life cycle, either early majority or late majority.

[55] Joel Barger, Managing Change: How to Effectively Deploy New Technologies and Processes in Your Company, Meridium APM Advisor (February 2012) http://www.apmadvisor.com/archivearticle.asp?is=60&ord=1.

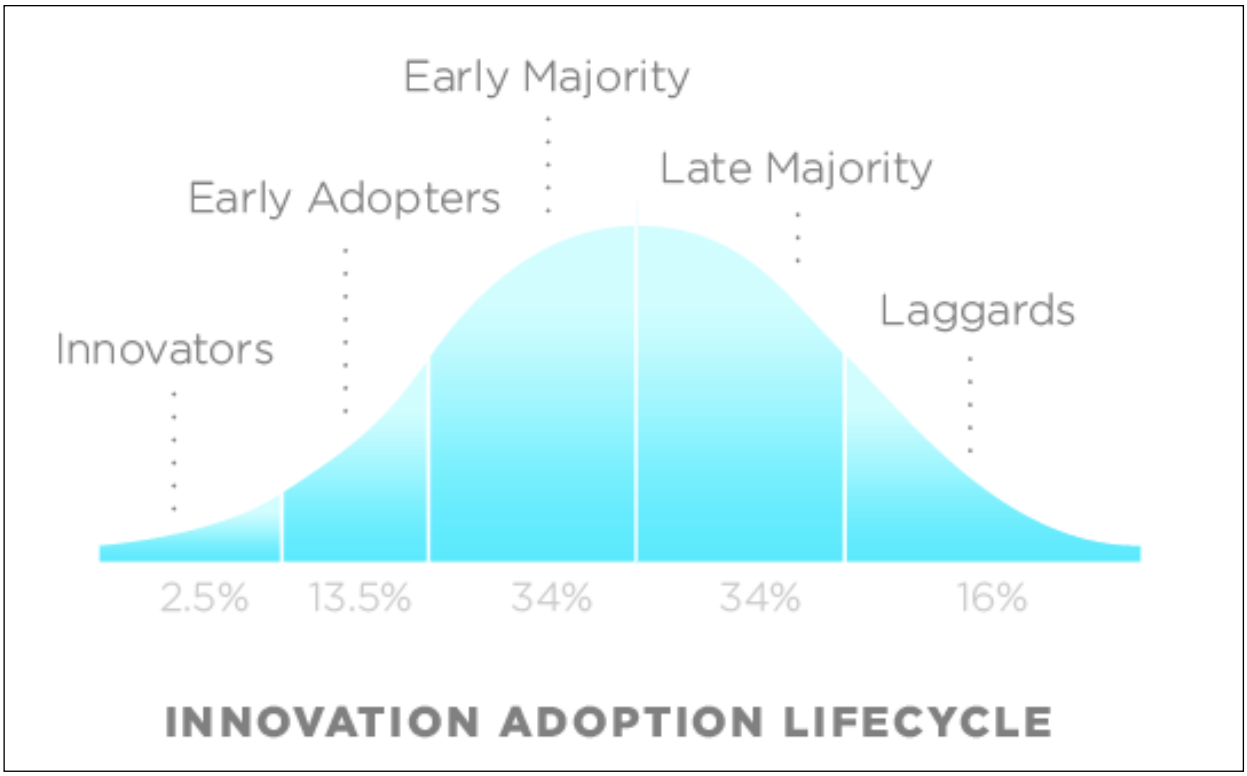

Figure 2 - Adoption Lifecycle (Source: Wikipedia)

In technology, innovators and early adopters take the most risk when deploying new technologies, whereas “laggards” take the least amount of risk but often do not take advantage of the most gain in the technology. Characteristics of the various adoption approaches include:

- Innovators – Large scientific and research and development organizations that have large investment money and are more risk-oriented
- Early Adopters – Large finance investment and highly competitive commercial organizations with less investment money
- Early majority – Large and medium-sized commercial organizations that require efficiency but can withstand some risk. Early majority organizations have a moderate amount of investment capital
- Late majority – Large organizations and government agencies that require efficiency but cannot withstand risk to stability and have minimal capital investment for technology. Oftentimes, they wait until technology costs have come down due to market competitiveness
- Laggards – Very conservative and/or small organizations that cannot tolerate any risk and have little capital investment money for technology

Making a determination on what model best fits the organization, evaluating the technology maturity, and determining the organizational IT priorities will help develop an organization’s IT deployment strategy.

6.4 Cost/Benefit Considerations

Like any change being considered in an organization, IT technology implementation requires that cost/benefits be taken into account. At its most basic level, an analysis will determine whether the cost associated with deploying a technology provides sufficient benefit, whether tangible (*e.g.*, cost savings over time) or intangible (*e.g.*, enhanced service to customers).

Each enabling technology presented in section 6.1 is associated with varying amounts of benefits and costs depending on several factors:

- Costs
  - In-house/contract labor
  - Cost of software, hardware, etc.
  - Cost of deployment (time to deploy)
  - Lost revenue/time due to deploying new technology (risk associated with deploying new technology)
- Benefits
  - Reduced operating costs
  - Increased efficiency
  - Added services to customers
  - Increased availability

Although this is not a complete list, organizations should always perform a thorough cost/benefit analysis when deciding if and when a specific technology should be deployed.

7. Modernizing the Copyright Office

The Office's technology infrastructure impacts all of the Office's key services and is the single greatest factor in its ability to administer copyright registration, recordation services, and statutory licenses effectively. This report thus provides a number of recommendations that, if adopted, could significantly improve the Office's operations and interactions with the public. Specifically, this section will cover recommendations established by the project team for improving the technology of the Copyright Office. This section will also outline proposed technical enhancements which will provide the underlying architecture and infrastructure to achieve some of the proposed recommendations. These upgrades should position the Office in evolving and improving the technology capabilities across the enterprise of the organization. The recommendations are presented in four main areas of focus, responsive to the project's Findings as presented in section 5.

7.1 Enhance User Experience

As noted in *Priorities and Special Projects*, the Copyright Office set out to discover areas for improvement to the existing registration system user experience so that we could identify "*what kind of interface [would be] optimal*."[56] Improving the user interface and user experience, both internally and externally, was a major issue during the data gathering phase of the special project.

Since eCO registration was adopted, the user community, both internal and external, has had significant experience with the existing user interface. Internal and external users commented that the user interface was "clunky," hard to navigate, too browser specific, and overall not well-designed. They would like to see a simpler application process such as a fill-in form or one that guides a user through a series of defined steps for completing an online registration (*e.g.*, the TurboTax interface model). Users recommended numerous improvements to the eCO interface such as individual customization capabilities (*e.g.*, changing screens to use only their applicable activities), improved profile management capabilities (*e.g.*, saving case submission histories, abilities to login and retain user specific settings and pre-populated data) along with limited abilities to upload and poor integration with pay.gov or Deposit Accounts.

The feedback of interface inadequacies is not new to the Copyright Office – the project team is aware that the registration system user interface and overall user experience is not optimal. Prior to engaging in this special project, the Office received numerous recommendations for improvements through our eCO help desk, visitors to the office, through user surveys and through staff. Even though user feedback channels were much less formal prior to the special project, the Copyright Office tracked and maintained a running list of recommendations for future improvement considerations. However, aside from a handful of minor tweaks to the user interface, the current architecture remains largely unaltered, which hinders the Office's ability to make substantial user interface improvements in a cost effective manner. As a result, the Copyright Office has focused on maintenance – keeping the core system up and functioning properly.

The registration architecture is based on a customized Siebel Customer Relationship Management (CRM) platform. This results in a product requiring both licensing fees and software maintenance expenditures.

Enhance User Experience Recommendations

To summarize, these are the steps that the Project Team recommends, on a schedule and budget that would need to be determined:

[56] Maria A. Pallante, *Priorities and Special Projects of the United States Copyright Office, October 2011-October 2013*, 13 (2011) ("*Priorities and Special Projects")* http://www.copyright.gov/docs/priorities.pdf.

Involve Copyright Community

- Assemble internal and external user groups, to meet regularly with representatives from the Office with the objective of communicating users' needs and desires from Copyright Office systems, and for the Office to effectively communicate resource conditions and constraints
- Meet with remitters across industries and at various levels of expertise to gain an understanding of their processes prior to submission so that Office processes and systems are adjusted to partner expectations

Improve User Interface

- Enlist the services of experienced web designers and completely redesign and deploy a new copyright registration system user interface
- Gather the full breadth of requirements and determine "must haves" versus "nice to haves." At a minimum, consider incorporating the following features:
    - Consider the possibility of using fill-in forms for application completion
    - Take advantage of profile management capabilities
        - Individual user customizations of UI will be retained and tied to user profiles
        - Users will not be required to repeatedly input static information
        - Organizations can more easily manage organizational parent/child relationships
        - Develop authorization for users to remit on the organization's behalf
    - Offer views of the entire completed application prior to submission and final registration certificate
    - Offer the ability to transfer applications between parties for certification prior to final submission to the Office
    - Improve "Service Request" naming convention to provide pertinent information (*e.g.*, type of claim, receipt date, remitting organization, etc.)
    - Improve file upload capability to allow significantly increased file sizes
    - Provide tracking of individual copyright cases so the users can easily determine status of submissions throughout the entire process
    - Improve integration with pay.gov and/or deposit accounts
    - Enforce section 508 compliance for user interface redesign
- Consider what, if any, functionality from existing Copyright Office systems can be easily incorporated into any redesign and deployment efforts
- Deploy a solution that is non-browser specific
- Redesign and re-architect copyright.gov to offer patrons the ability to perform all transactional activities

Empower System Users

- Identify existing Copyright Office services which could be offered as "self-help" thereby freeing up staff resources within the office to focus on other activities while simultaneously creating an "on-demand" capability for external system users

- Employ terminology that is easily understandable to both the copyright community and individual users
- Deploy a point-in-time tracking capability throughout the entire work process

New Features

- Consider the policies and feasibility of using electronic copyright registration certificates to replace existing paper certificates
- Consider the policies and feasibility of offering self-certification capabilities for certified copies through use of digital certificates and file validations
- Improve historic case management capabilities for organizations so that they are able to identify past submissions
- Develop rudimentary analytics (*e.g.*, how many registration claims are in a pending queue, how many documents were recorded in a given month) for organizations based on previous submissions
- Engage in an analysis effort to implement artificial intelligence for the examination of registration submissions beginning with claims in an individual photograph

7.2 Enhance Public Record

Along with Registration and Recordation functions, providing an accurate, complete, and up-to-date public record is one of the primary services the Office offers to the public. The Register of Copyrights noted in *Priorities and Special Projects* that the Office seeks to "*improve the nature, accuracy and searchability of USCO public databases.*"

In our meetings with the copyright community, stakeholders confirmed that the existing public record is incomplete and that the search capability is inadequate for locating copyright ownership information. They told the project team that the existing system often crashes, times out, and cannot handle large volume requests. Most importantly, they conveyed that data is missing from our existing public record and migration of new data to the public record can be too slow. They further indicated that much of this information is available from other internet sources, so they tend to use those sources before the Copyright Office website.

The public record databases should show accurate, complete and up-to-date information such as chain of ownership and ownership contact information. The Office should consider whether it will allow the public to update some data directly, such as change of address data. For certain works, the record associated with a copyright could contain snippets of deposits or thumbnails (possibly watermarked), perhaps under an opt-in system. Additionally, some of the community would like more information about a given copyright registration such as whether a termination notice was filed, if registration was made upon appeal, or if the work is in litigation. They would like to see the complete ownership record, and chain of title, in one place. Finally, they would like more data, specifically metadata, for items such as images which could be relatively

easily extracted from the uploaded deposits. Of course the nature and scope of the public record, though implemented through information technologies, is fundamentally a question of law and policy. What information to make available to the public, and on what terms, is largely driven by the Office's obligations under Title 17, which is currently the subject of a comprehensive review by Congress and about which the Office is continually consulting stakeholders.

In addition to the inadequacies of data within the public record, the copyright user community indicated that the search capabilities within the public record were not optimal. Users "search" expectations are more sophisticated that in years past; they expect search abilities which are common among most large data sets such as sorting and/or filtering or faceted searching - a technique for filtering the data search in order to obtain more refined results, (*e.g.*, format of work, date of first publication, nation of publication, subject matter, etc.). Users of the public record would also like the ability to save searches so that they do not have to be recreated each time they use the records and they would like the system to return all of the results without a maximum limit (which is currently not available).

While certain limited improvements to the copyright public record may be possible within the current environment, significant enhancements/improvements are dependent on detailed and agreed-upon system requirements. As is the case when developing any IT solution, initial discussions will help to define the business, functional and technical expectations of a new copyright public record. Once firm requirements are established and accepted, determinations can be made as to whether the existing Voyager database system is capable of supporting the requirements of the proposed copyright public record.

7.2.1 Public Record Database

To achieve the overall objective of improving the Copyright Public Record, the Office should first consider whether Voyager is the best database system to suit the needs of the Copyright Public Record. The current copyright public record database is powered by a Voyager database which is licensed and primarily operated (with limited consideration to copyright system needs) by the Library Services division of the Library of Congress. In 2007 the Copyright Office migrated the electronic copyright public record from a Library of Congress mainframe system to the large-scale Voyager database controlled by Library Services. The arrangement with Library Services offered one individual Voyager database instance along with some limited amount of customization to address Copyright Office requirements. Based on what the Office knows now, the Voyager product is designed with bibliographic cataloging capabilities in mind (*i.e.*, item loan, library cataloging, overdue item tracking, and interlibrary loan of collection material) and is not necessarily adaptable to Copyright Office needs.

7.2.2 Data Sources

The Copyright Office must consider what data sources currently populate the public record and should populate it in the future. The copyright marketplace creates and

manages a number of rich data models for a variety of content, including music, films, and books. The project team believes the Office should consider these and other verified data sources and incorporate them, as appropriate, into the public record.

Recordation

Significant improvements to the public record, while certainly achievable, are dependent on other major Copyright Office initiatives such as the pending effort to automate the activities of the Recordation Section. While eCO data is migrated to the Public Record on a daily basis, data from the legacy CORDOCS system (and legacy CORCATS system) migrate only once a week and are largely reliant on manual processes. As it stands today, work flows and work processes for recording documents are primarily performed in a manual fashion. Once the workflows and processes have been reengineered and automated, data migration must happen seamlessly and can occur on a much higher frequency. Until then, data migration for recording, indexing, and accessing copyright documents will remain less than effective.

Public Access to Historical Records Project

Another major initiative of the Copyright Office is the Special Project for Public Access to Historical Records. This project is related to, yet separate from, the ongoing effort to scan and move to long-term storage approximately 60 million Copyright Office paper records ranging in date from 1870 to 1977. The scanning effort was partially funded through a multi-year appropriation, and digitization of the historic Copyright Card Catalog – approximately 35.8 million cards – was 99% completed in fiscal 2014, with only fragile cards remaining. Most of the pre-1978 paper records have no back-up copies and comprise an irreplaceable record of late 19th as well as 20th century copyright ownership.

Although the scanning project achieves the important goal of creating high-quality digital preservation copies of the records, researchers still have to visit the Copyright Office in Washington, D.C. to search the records. USCO is currently experimenting with Optical Character Recognition (OCR) technologies to determine the best option to integrate these records into the Voyager system. Ideally, data from the cards and other scanned paper records, once extracted and indexed, could be integrated and linked with records in the Copyright Voyager database to provide for a single, comprehensive search interface covering the entire time period from 1870 to the present. “Smart” searching that would automatically find and display multiple records reflecting chain of title of a single work (*e.g.*, original registration, renewal, one or more assignments, etc.) would be highly desirable. Testing OCR of 1971-1977 cards is underway.

7.2.3 Credibility of Data

As stewards of the official public record of copyright ownership information, the accuracy and validity of data contained in the records is of paramount importance to the Copyright Office. The Office must therefore undertake significant efforts to not only validate the data which feeds the public record but also the processes which create the

data. It would be wise to routinely validate the data feeding the official public record and also to ensure that we refrain from permitting sources other than those internal to the Copyright Office (*e.g.*, collective management organizations or other third-party registries) to contribute.

Enhance Public Record Recommendations

To summarize, these are the steps that the Project Team recommends, on a schedule and budget that would need to be determined:

Involve Copyright Community

- Establish a project team of both internal Copyright Office users as well as external experts or "power" users, which are tasked with overseeing the effort of creating a new official public record
    - The team should develop the detailed requirements of a new official copyright public record with improved search capabilities
    - Establish data mappings and data flow to all relevant data sources which will contribute to a new Copyright Office public record and develop data integration plans
- Resolve policy issues surrounding requirements for populating, updating, and maintaining a new Copyright Office public record
    - Resolve policy issues such as whether or not the Office can provide sample images of the deposited copy associated with a copyright record
    - Consider ways of improving "official" public record update capabilities (users desire an easy ability to update information in their own records such as updated contact information)
- Ensure public record redesign is seamlessly integrated with the "Data" considerations in section 7.3 of this document

Internal Considerations

- Conduct an alternatives assessment on database products capable of supporting the enhanced requirements of the copyright public record
- Conduct a cost/benefit analysis for migrating to, and long term support of, a dedicated, large scale database which houses the copyright public record versus enhancement capabilities of the current incarnation of the public record
- Release RFI or Statement of Objectives to industry experts and technology research organizations (*i.e.*, Gartner, Forrester, etc.) to obtain input on deployment of enterprise class databases.
    - Consider the possibility of investing in a new robust, large-scale database solution that is designed to accommodate the requirements of a new official copyright public record
    - Consider partnering with known search utilities in providing search engine capabilities for the public record database

7.3 Improve Data and Information

As an office of public record, the Copyright Office realizes the criticality of maintaining useful, reliable information. The project team understands that quality records are dependent on quality data and quality data is necessary for providing quality information. We further understand that the Copyright Office needs to drastically improve what data and information it wants to capture. With this in mind, the project team recommends the Office focus on capturing the "right data" so that we provide the most valuable information to the copyright community. As the Register noted in *Priorities and Special Projects*, we need to understand "*[w]hat kind of information should be captured . . . what kind of metadata . . . the office should pursue*."[57]

Quality data is the underlying element which makes information useful and valuable, and is therefore the key contributor to offering an effective and searchable copyright public record. As mentioned previously, many users, both internally and externally, feel the existing data (and records) of the Copyright Office are largely inadequate and inaccurate for fulfilling intended purposes – keeping track of who owns rights to particular pieces of copyrighted works.

Figure 3 is a conceptual diagram illustrating how data may be obtained from various sources to provide a consolidated copyright data set which can be used as a reliable central repository by the copyright community. A Copyright Office Public Repository would accept feeds from multiple sources and be subject to change of control and governance to ensure accuracy. In addition, since it is a duplication of data, loss of data would be eliminated should the dataset be compromised.

[57] *Priorities and Special Projects*, at 13.

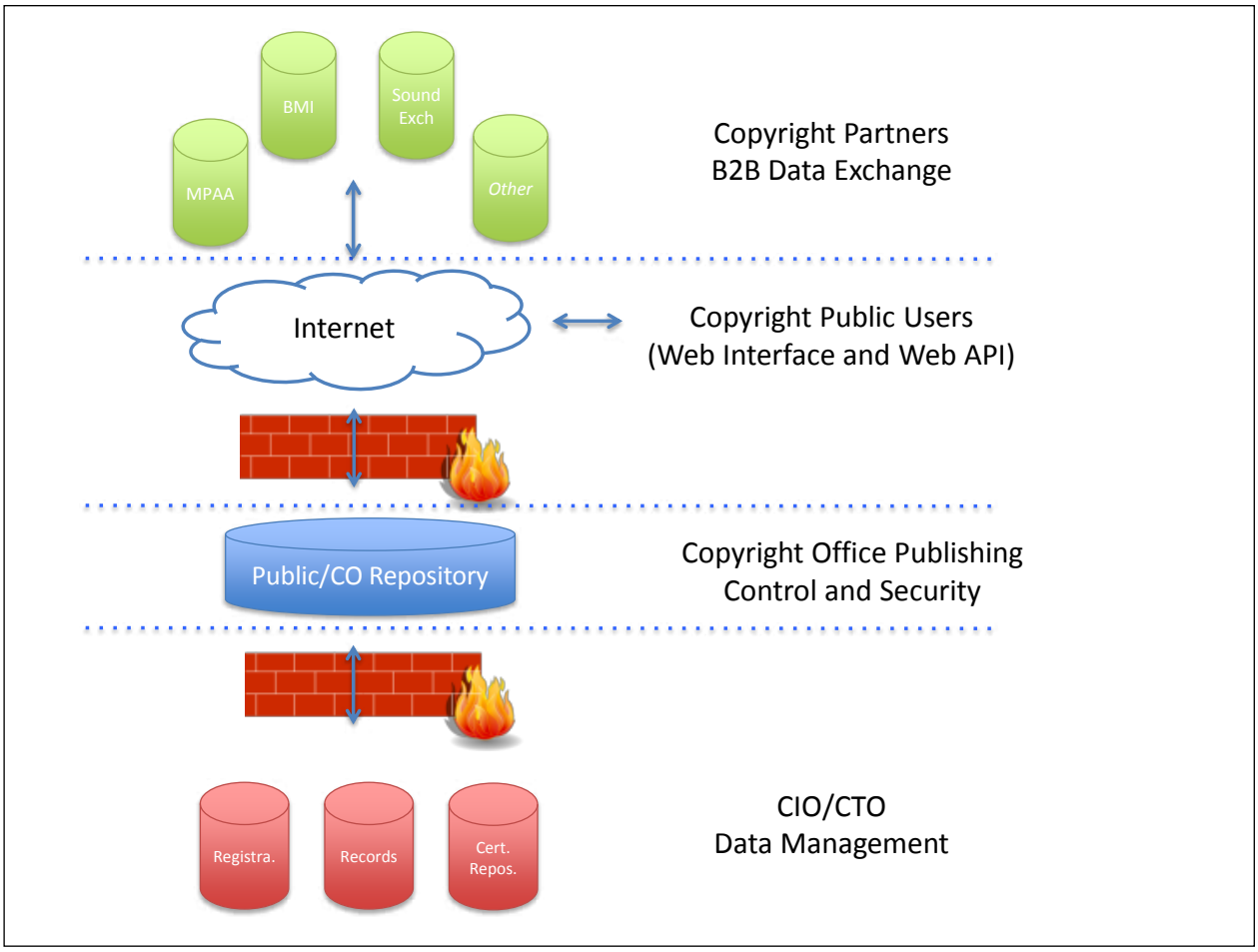

Figure 3 – COPYRIGHT OFFICE Data Model

A targeted effort which focuses on all aspects of data, specifically the establishment of effective data standards, should be a primary goal of the next generation Copyright Office and should be considered a key to the success in fulfilling the needs of the copyright community. The Office must commit to effectively managing its current data as well as determining what data it needs moving forward. Consideration must be given to expectations and requirements for the future. The Copyright Office must identify both the user community requirements and business requirements so that we enable a more informed copyright community.

There are many questions to be answered but foremost among them are: what is the universe of potential data elements, what data is available and from where, what data standards should the Copyright Office adhere or establish, whether the data exists currently and who, if an entity other than the Office, maintains the data. The Office needs to establish who we receive data from and who we send data to; the Office must develop a plan for integrating the data as well as how the data will be preserved or archived. Finally, the Copyright Office needs to better understand how it will manage the data, as well as how it plans to migrate data to future technologies which today may not exist.

7.3.1 Data Strategy, Data Management & Data Governance

Overall, it is critical that the Copyright Office invest appropriate resources in the creation and management of a thoroughly vetted Data Strategy, Data Management Plan and Data Governance Plan. The objective of a detailed Data Strategy simply outlines where the Office is now, from a data perspective, and targets where we want to be in the future. An effective Data Management Plan clarifies such information as what data currently exists, where the data is maintained and archived, who owns the data, and how the data is validated and used.

In an organization such as the Copyright Office, data must be treated as one of the most, if not the most, valuable asset of the organization. Quality data must be maintained using documented and agreed-upon data governance standards which clarify what rule sets are used in managing the data (*i.e.*, what happens in the event of data conflict, which data supersedes similar data, who is authorized to update data, who has access to what data, etc.).

It is common for large organizations to find themselves managing huge volumes of data which can vary significantly in size, content, and established levels of support. Often the data is stored in disparate systems with little to no thought being given to managing the data effectively. Some neglect to leverage the information which may be gleaned from the amalgamation of data in the form of analytics. Many organizations which house these huge amounts of data do so without an effective Data Strategy and Data Management Plan which should minimally address such information as:

- What data exists
- Who owns the data (data authority)
- What data should be captured
- How the data is maintained
- Who has access to specific data
- What the rules are for validating, managing, and archiving the data
- Specifications of data standards employed
- How data is exchanged and with whom (as well as whether they are a trusted party)
- What measures are taken to ensure data validity
- How to ensures all policy issues and records management mandates are adhered to
- How to structure the data
- What are the relationships of the data sources
- Specification of data access control (who is authorized to update the data)

Data is an organizational resource with intrinsic value internally and to Copyright Office customers. Proper resources must be obligated for effective data management and preservation. Thoughtful consideration must also be given to such elements as governance and security.

Currently, the Copyright Office maintains numerous data stores but lacks all of the necessary policies for managing the data. The Office's data is maintained in many different forms and in many different locations and very rarely does one set of data correlate to another. While there is much data that is missing or which goes un-captured, the Office has great amounts of data which we simply need to do a better job of managing and rendering accessible. Not only will the establishment of proper data management practices benefit the organization significantly now and in future generations but as an office of public record, managing the data fulfills the core mission of the Copyright Office.

At a minimum, with effective data management in the form of a thoughtful data strategy, a well-defined data management plan, and a detailed data governance plan, the Copyright Office must:

- Provide accurate, robust, up-to-date and timely records to our copyright community
- Potentially answer many orphan works questions

In the short term, the Office can do a better job to:

- Leverage the information it already possesses
- Establish effective policies and practices for managing the data moving forward

In the future the Copyright Office can focus energies on partnering with copyright community organizations, many whom have expressed an interest in sharing data with us. The true benefit of sharing data with others is to enrich the data available so that the copyright community as a whole benefits.

Two elements of the Data Strategy and Data Management Plans are worth highlighting specifically due to the importance that these two areas hold within the plans: Data Warehouse/Data Repository and Data Integration. These two elements will be briefly covered below.

Data Warehouse/Data Repository

A data warehouse with associated data repositories is a key component to be considered when developing the Copyright Office's data strategy. As part of the data strategy, an equally important component is how the data should be managed and governed.

A data warehouse is more logical (data grouping, access controls, etc.) in nature whereas a data repository is more physical in nature (*i.e.*, disk, servers, etc.). For example, a data repository is much like a physical filing cabinet that has drawers, filing folders, etc. A data warehouse may contain multiple filing cabinets that are logically grouped together (*i.e.*, each cabinet contains data that relates to other cabinets) under one roof (see Figure 4).

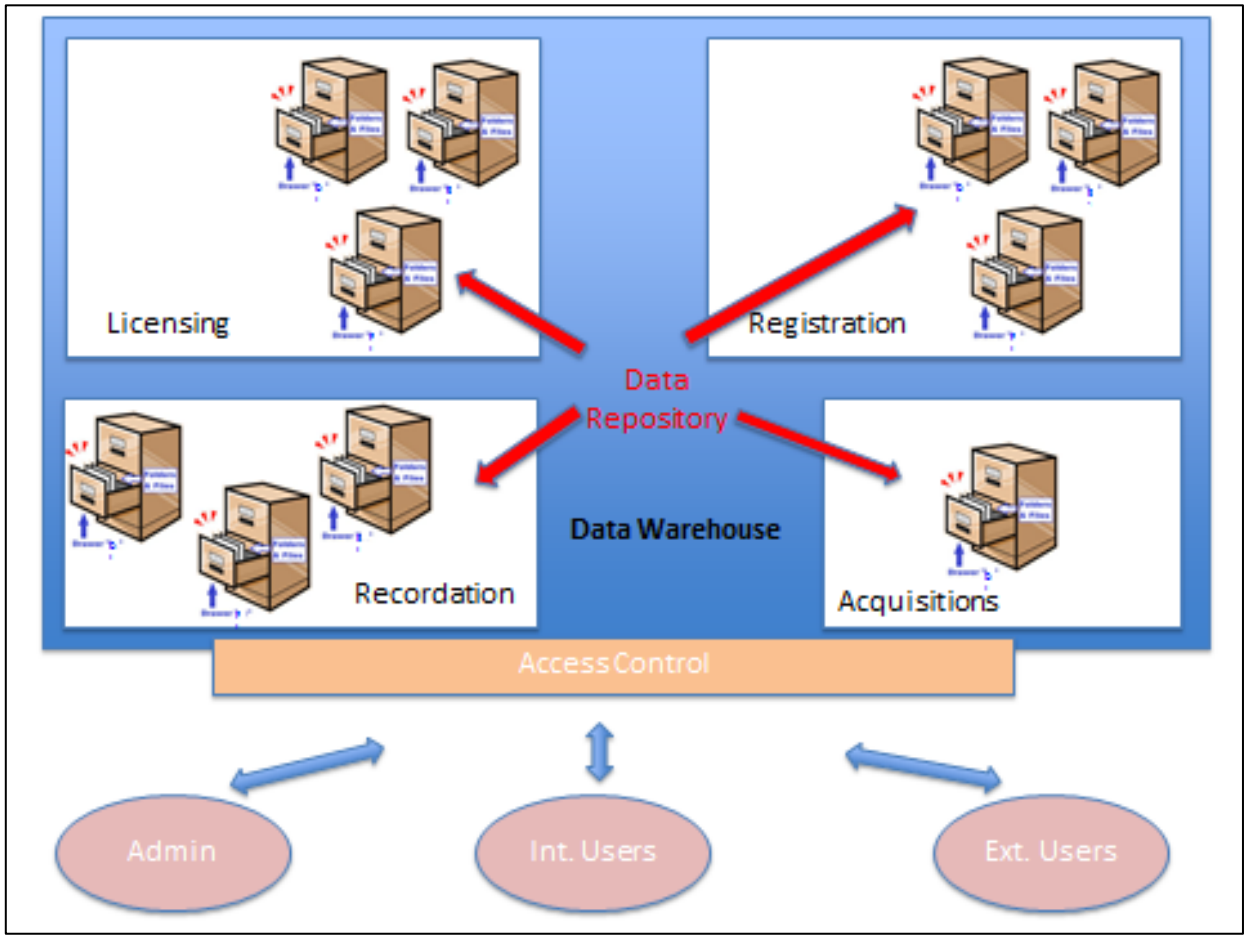

Figure 4 – Data Repository/Data Warehouse

Using this analogy, the Copyright Office needs to develop and manage the warehouse that contains the filing cabinets of all the individual filing cabinets from its systems. The warehouse needs include the necessary management plan, operations, and procedures to store and retrieve information from multiple filing cabinets in an organized fashion. The approach in providing this is covered in section 7.3.2.

Data Integration

A second critical element for consideration when developing data strategy, management and governance plans is that of integrating data repositories together in an effort to possess more accurate and complete data. Data sources can come from both internal and external sources (*e.g.*, performance rights organizations). Before data is integrated and aggregated, significant consideration must be given to where the data originates and the integrity level or quality of the data being aggregated (referenced previously in section 7.2.3 Data Sources and section 7.2.4 Credibility of Data of this document). The net results of merging known accurate data with potentially inaccurate and/or incomplete data will result in an amalgamation of inaccurate data.

The Office understands that the value of integrating data is significant but the Office, as a trusted authority of record, has a responsibility to use the utmost care and caution when integrating data. The Copyright Office will need to develop integration strategies that provide client access while insuring data integrity.

7.3.2 Data Analytics

Business Intelligence uses Big Data (described in section 6.1.6) in an analytical fashion. Data analytics (DA) is the process of examining data in order to draw conclusions about that information under review. DA helps organizations make better business decisions. In most cases, DA is used in combination with Big Data as a necessary aid to process a large amount of data. As we improve our data posture, the Copyright Office will invariably be collecting higher-quality data. When the organization commits to collecting the appropriate data, organization leaders will be able to make actionable and informed decisions, in real-time, for such things as application submissions by type and productivity levels for a given period, in a particular division or sub-division.

Data analytics could inform point-in-time information regarding workflow trends and potential backlogs. For example, in an instant, management would be able to determine how many visual arts registration applications are in a pending queue or how many performing arts registrations were completed in the previous month. The organization could track in-bound receipts as well as expenditures for any given period of time. Copyright community information could be tracked to determine such things as deposit account balances so that the community could receive more timely notification of reduced balances. Appropriate data analytics allow timely and easy access to meaningful business data. With effective analytics (and supporting infrastructure), the types of information obtained are almost limitless.

7.3.3 Data Exchange

Before discussing data exchange, four conceptual elements pertaining to data exchange must be considered: Data Standards, Data Schemas, Application Programming Interfaces (API's)/Web Services. A general concept of these data exchange concepts are covered below.

Data Standards

Data standards enable organizations who want to exchange data with the Copyright Office to know exactly how to communicate with us. Data standards spell out how data should be organized in order to be exchanged between organizations, often likened to a common vocabulary or predetermined semantics. Data standard formatting agreements can be handled individually between the host and partner organizations or the host organization, in this case the Copyright Office, can choose to publish a data standard that simply allows trusted partner organizations to map their data to that of the Office. In this instance, if the Office chooses to publish its own set of data standards, it can create one or the Office can elect to adhere to an existing set of standards which were mapped out and agreed upon by one of several standards setting bodies.

Such standards have normally been vetted and are widely in use by numerous organizations both nationally and internationally. If the Office chooses to follow OPM A-119 guidance and adhere to existing industry data standards, careful consideration must be given to which data standard(s) it aligns with. There are numerous data standards available from numerous data standards governing bodies. It is important that the Copyright Office take into account the standards used by the various industries that interact most frequently with the Office (*e.g.*, publishing, music, art). Most organizations which operate in a particular industry adhere to industry standard data models (ISDM), so the Office will need to be able to transfer data with relative ease amongst like partner organizations. During our interviews, the project team learned some organizations employ small teams of technologists whose job it is to solely work on accommodating various data standards so that they can easily exchange with various partners.

There are numerous and obvious benefits to establishing or adhering to a set of data standards including allowing organizations the ability to data transfer bulk registrations or documents for recordation, making it easier and faster to update copyright record information (from trusted partners only) thereby improving the overall quality of the data. As an added benefit, data which transfers from system-to-system is usually cleaner since less manual keying is necessary. Establishing data exchange through established data standards can also provide much more robust data even though, as noted previously, the Office does not want to capture data just for the sake of collecting data. Rather, the Office needs to ensure it captures the right data.

Data Schemas

Data schemas help depict the data an organization (or specific system) maintains. The data schema identifies the data elements and depicts the relationships to other data elements. Data schemas are usually published graphically and help to define and depict all the levels of data elements (to include metadata) of a system.

Application Programming Interfaces (APIs)/ Web Services

An Application Programming Interface (API) provides a mechanism for disparate applications to communicate and deliver expanded functionality. An API is a tool that lets systems share existing functionality and allows disparate systems to share data with one another. APIs can benefit the Copyright Office and remitters by enhancing registration and recordation options for remitters but also allowing people to interact more effectively.

In many cases APIs are developed by software companies and distributed so that they can be used in third party applications. (As noted in section 6.1.1 Application Programming Interface.) There has been considerable expansion of the API concept as the Internet and mobile computing technologies have proliferated. Web based APIs are installed on a hosting organization's servers and do not require distribution and end-user installation. Commonly, web based APIs provide business-to-business data

exchange as well as mobile device communications. Hosting organizations publish the method to communicate with their servers based on well-established standards. This removes the burden of organizations having to develop client-side APIs in formats compatible with multiple software languages.

A great example of a successful deployment of a web based API is the service that was deployed by the IRS to allow for electronic submission of tax returns. For example, users who prepare their personal tax returns using a third party application (*e.g.*, TurboTax among others) have the ability to upload their completed tax forms to the IRS web servers. The APIs installed on IRS servers also allow for processing of payments from tax payers. The third party application companies also provide valuable feedback to the IRS on system changes and operational modifications. The Copyright Office would benefit from having similar relations with its data users.

Data Exchange

As noted by the Register in *Priorities and Special Projects*, the Copyright Office would like to consider the "*feasibility of connecting the Office's database of registration and recordation records to private sector data to facilitate licensing and use of copyrighted materials.*"[58]

With regard to data exchange, the primary capabilities under consideration for the Office as a result of an improved data posture would include, but certainly not be limited to, the following:

> ***Improving the accuracy and robustness of copyright records***
> There is far more robust data within the copyright community that the Copyright Office could leverage and benefit greatly from. Likewise, we have significant, critical data of which certain industry partners would welcome having easy access. Currently, external organizations resort to crawling the Copyright Office database through automated means but this is a time consuming and somewhat arduous task. Partner organizations definitely prefer the ability to extract data from the Office through automated data exchange methods. Obviously, exchanging data with certain organizations - both transmission and receipt - will make for richer data for everyone which, in turn, benefits the entire copyright community.
>
> ***Enabling bi-directional data exchange of copyright information (transmission/receipt) with trusted parties***
> The Copyright Office should be cautious when receiving data from external entities and needs to ensure proper security and data validation prior to any data integration effort. The Office may need to consider the practicality of "certifying"

[58] *Priorities and Special Projects*, at 13.

data sharing partners to ensure they are transmitting quality data that has been validated before any data integration is considered. At a minimum, the Office absolutely must consider the rule set which governs who can exchange, when, how, and what before integrating into the Copyright Office public record.

Enabling update to an unofficial copyright public record

The Copyright Office should consider the possibility of establishing an unofficial database which houses the vast amounts of copyright data from sources throughout the copyright industry, thereby offering more immediate, yet unofficial, results to the copyright community until such time as the information can be validated and possibly incorporated into the official Copyright Office public record database. Through various improvements mentioned previously in the *Enhanced Public Record* section of this document, the Office will be able to vastly improve the time with which it takes to populate the official public record. But even when effective automation supports recently optimized work processes, there still may be minimal, unavoidable delays in updating the official public record. This same data will likely already exist in partner databases and will only require a data transfer to bring our unofficial public record to a more current state. Again, this connected unofficial public record supports the Register's goal of "connecting databases."

Enabling remitting organizations the ability to bulk transfer copyright registration applications and documents for recordation

The Copyright Office understands that the copyright community would like to improve the data exchange capabilities, including but not limited to bulk registration and recordation submission. The project team understands the benefits of such an ability and we agree that such an effort needs to take place sooner rather than later, but we also understand that such a critical initiative should not be undertaken without proper planning. Not only will modifications be necessary for our existing eCO system and storage, but the Office will require well-conceived data management, data governance, and data standards. The Office needs to fully understand the internal requirements as well as those of our business partners, and we need to understand the range of solutions available. Nonetheless, the project team thinks improvements to the data exchange capability of the Office should be a near term priority.

In the copyright marketplace, there are several remitting organizations that retain teams of people who sit at terminals all day filling out registration applications one-at-a-time. With effective data exchange capabilities in place these remitter organizations may require less staff to work on individual file submission. Additionally, the Office will be better equipped to receive applications either through dedicated transmission channels or improved management of resources by possibly allowing bulk registration (or recordation) during off-hours, or improved automated system monitoring to regulate the inbound transmissions. Not only will such a capability allow for less manual labor, improved quality of data, and an earlier effective date of registration (or recordation), the process, once established, is easily repeatable and can be set up for automated transfer.

Additionally, partner organizations would be free to develop their own interface for populating the data so that they aren't dependent on the existing eCO user interface.

Enabling data exchange with intellectual property law enforcement agencies

Late in calendar year 2009, Congress directed Customs and Border Protection (CBP) to initiate a project referred to as the Intellectual Property Rights (IPR) Recordation Linking Project so as to report on the feasibility of connecting CBP databases with those of the nation's two intellectual property organizations, the U.S. Copyright Office and the U.S. Patent and Trademark Office. The capability of an aggregated database would allow CBP agents to check shipments as they enter the country from known pirating regions of the world.

In the end, it was determined that CBP, the Copyright Office, and the Patent and Trademark Office all lacked sufficient resources to implement the Recordation Linking Project during FY 2009. The final report also stated that until the two intellectual property organizations are given the mandate, funding, and additional resources to complete this project, CBP would be unable to implement the congressional mandate. Such a request, with proper office-wide data posture, should be achievable.

Offering external entities the option o creating their own interface for the submission of copyright applications or documents for recordation

It is no secret that the user interface for copyright registration submission is less than optimal. While the Office hasn't given up on making major improvements to the interface, having an API or a published set of data standards will allow frequent remitters the option of creating their own interface for submitting registration applications and documents for recordation to the Office. Offering this option will hopefully encourage more copyright registration and recordation, and with effective and well thought through data standards, will hopefully improve the amount and quality of data received by the office. Also, having a published set of data standards and/or an API will permit integration with popular applications for content creation, making it more enticing for a user to submit registrations and recordations.

Offering compatibility with technology devices such a mobile technologies (i.e., smart phones, tablets, etc.)

As most are aware, mobile technologies sales are on the rise and PC sales have been on a steady decline. In many cases, the capabilities of mobile technologies are equal to, or in some cases better than, those offered by their predecessor equivalent technologies. Documents can be created on tablets and high-resolution photographs, videos and audio can be captured with advanced smart phones. Not only should the Office consider integrating metadata which is

automatically generated by such devices, the Office should offer APIs that allow for the creation of mobile applications for the submission of copyright registration and possibly copyright record search and update.

Data Section Note: Within the purview of this Special Project, the team facilitated meetings with various data subject matter experts within the Copyright Office, the Library of Congress, and external agencies to gain a better understanding of the types of data standards already being used internally and by current partner organizations. Even though a great deal of additional research should be conducted surrounding the notion of copyright data and data exchange, the team felt that an additional level of research and understanding was necessary to establish effective data recommendations for moving the organization forward.

Improve Data and Information Recommendations

To summarize, these are the steps that the Project Team recommends, on a schedule and budget that would need to be determined:

<u>Commit Copyright Office Resources</u>

- Establish and appropriately staff a data group within the technology office of the Copyright Office to:
    - Participate in working groups of data standards setting bodies
    - Establish, or adopt, and publish data standard(s) for the office which cover all of the possible data exchange entities within the copyright community (and establish the process for maintaining such standards)
    - Establish a data exchange working sub-group whose mission is to manage data exchange partner relationships and technical logistics
    - Establish APIs for use in the copyright industry or by third party software developers eager in developing IT solutions for improved integration with Copyright Office systems (and establish the process for maintaining such APIs)
    - Develop business models to partner with other software providers to include Copyright Office-related products in their software suites.
- Develop a detailed long-term Data Strategy
- Develop a detailed Data Management Plan
- Develop a detailed Data Governance Plan
- Consult on strategy with data management experts
- Invest in the development of a detailed data “as is” document
- Invest in the development of an all-inclusive data model which minimally depicts all of the data available, data relations, etc. to include, but not be limited to, COINS, CORDS, CIS, eCO, CORCATS, CORDOCS, DOCLOG, and BCS
- Document the universe of data that is within the copyright community as a whole and determine, of the universe, which is of value to the Office
- Leverage existing data stores and repositories to make additional data available until a complete, robust data model can be completed

Establish a Data Repository

- Establish a secure data warehouse for all Copyright Office data and apply industry best practice data management and archive principles
    - Inventory the existing data stores and repositories and define future repositories
    - Research and satisfy all applicable federal data and records management directives
    - Identify and implement appropriate data security mechanisms and infrastructure
- Establish the necessary infrastructure to support the future state of the copyright data environment
- Pilot core business-side data analytics capabilities
- Pilot core partner-side data analytics capabilities

Involve Copyright & Technology Communities

- Establish a short-term pilot for secure, bi-directional data exchange of copyright information with trusted parties with the long-term goal of establishing a process by which data is transferred to/from select external entities routinely
- Establish a short-term pilot for secure, bulk submission with select parties with a long-term goal of secure, open and easy bulk submission from ANY external entity
- Engage in discussions with 3rd party application development vendors (including mobile technology developers) who have contacted us (in the past) and who have an interest in establishing copyright application submission software for submission of registration or recordation applications or modifications
- Establish a secure data exchange test with Customs and Border Protection (CBP) with the goal of routine and regular data transfer among any IP law enforcement agencies
- Engage in an analysis effort aimed at identifying requirements and ultimately implementing a robust video upload capability for receipt of digital video from video production partners that allay security or pirating concerns of partner organizations
- Develop an SFTP (Secure File Transfer) capability for large files and easy tie-in to the registration system

7.4 Architecture

In order to enable the improvements mentioned previously, and to manage the IT resources of the Copyright Office in a manner which supports achieving its core mission and statutory responsibilities, the Office will need significant improvements. In the view of the project team, these improvements should be premised upon a new, copyright-specific architecture and infrastructure. This infrastructure would then support such initiatives as a certified digital deposit repository, an enterprise application solution, and a mobile technology strategy. Deploying modern systems should take a holistic approach. In addition, the Copyright Office will need considerable staffing resources. It will need experts to support a systems development strategy and to handle

implementation of new initiatives. This staff will be fully integrated into the business of the Copyright Office as well as the businesses that depend upon the Office.

7.4.1 Enterprise Architecture (EA)

The Copyright Office does not currently have its own well-defined architecture. Rather, it follows the Library's enterprise architecture (EA). In order for the Copyright Office to deliver services to its stakeholders and improve efficiencies, a holistic IT modernization is required. This begins with developing an EA specific to the Copyright Office. EA minimally includes a well-defined practice for the following areas:

- IT Governance
- Systems Planning and Management
- Systems Analysis
- Systems Design
- Systems Development (including methodology, tools, hardware/software standards, etc.)
- Systems Implementation
- Systems Operations and Management

A holistic and complete Copyright Office EA is required in order to guide the Office, serve its twenty-first century customers, meet its business needs, provide IT strategic direction, provide gains in efficiency, and ultimately reduce costs in operation. The EA should organize the structure and views associated with the Office. This includes defining the business architecture that drives the information architecture. The information architecture will prescribe the information systems architecture identifying the data architecture which ultimately will be supported by the hardware, software, and communications systems (see Figure 5).

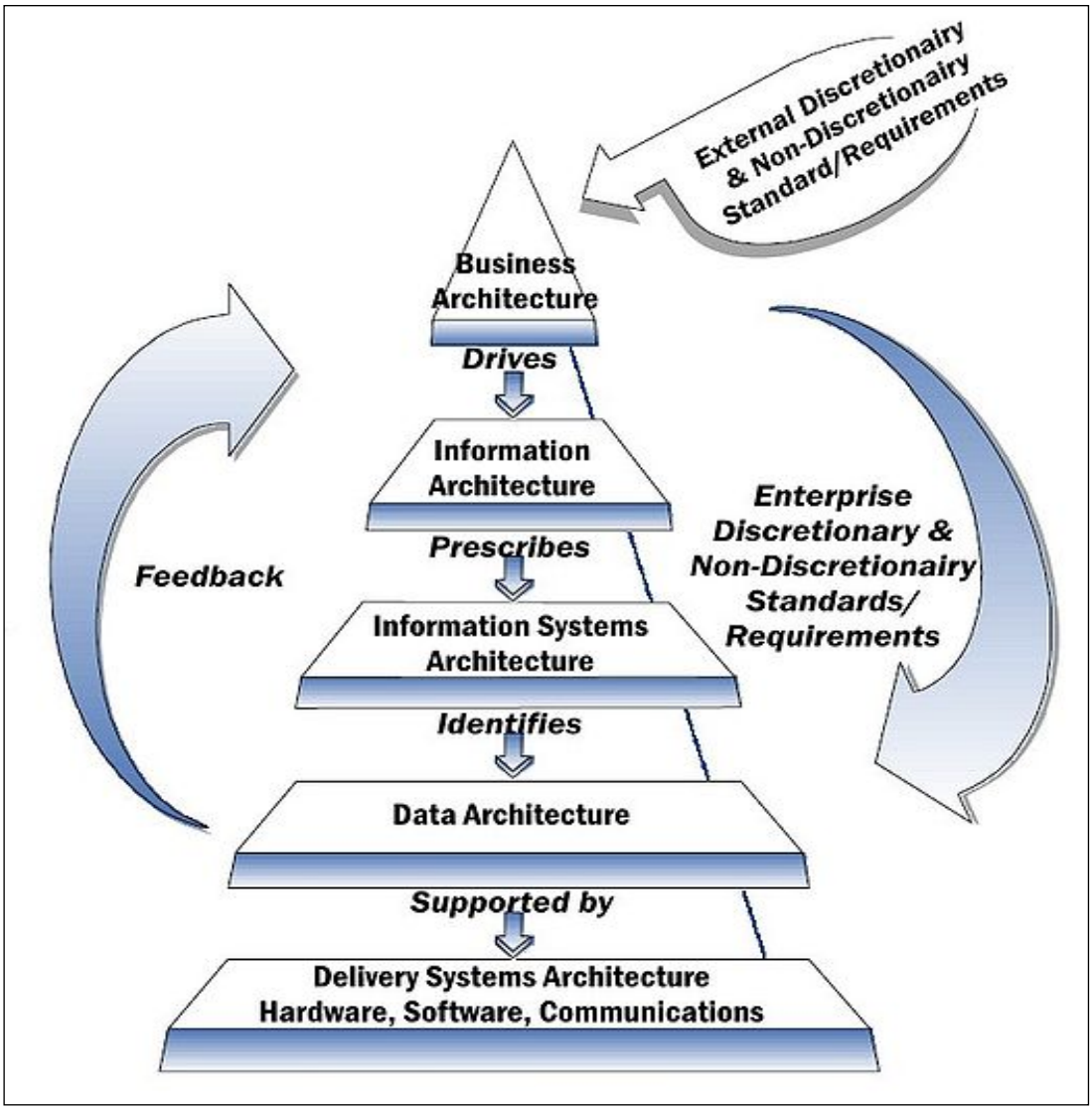

Figure 5 - Enterprise Architecture (NIST 2012)

7.4.2 Technology Infrastructure: Requirements and Alternatives

Currently, the Copyright Office IT systems reside on a combination of Copyright Office and Library of Congress owned infrastructure components; however, the Copyright Office technology staff does not have ownership or control of the overall technology infrastructure that powers the Copyright Office's systems and services. The primary web-accessible system for registration submission and backend case management is the eCO system. Virtually all elements of the eCO system were purchased by the Copyright Office, including hardware and licenses. Several operational legacy systems (*e.g.*, COINS, CORDS, CIS) were originally built by, and are currently supported by, the Library's Information Technology Services (ITS) directorate. During the upgrade of eCO from Siebel version 7.7 to Siebel 8.2 (in 2010), the Office purchased several elements of new hardware to enable such things as application servers, gateway servers, proxy

servers and database logical partitioning. Although the Copyright Office purchased this equipment using its funding (fees and designated appropriations), the Library permits Copyright Office technology staff limited access to the equipment; it allows application level access only. In other words, because the servers reside on the Library's network and because they are located within the Library's data center, it does not permit Copyright Office staff, including the Copyright Office CIO, to control the underlying hardware or operating system controls.

Additionally, Copyright Office staff are not able to establish or enforce IT security policies for these systems (or the data they contain), cannot control maintenance routines such as system backups or data migration, and have limited control of our systems during system outages. The Copyright Office technology staff relies on the Library's resources for administering and troubleshooting Copyright Office system issues. Library IT personnel are not experts in the needs of the copyright system. Moreover, resources are routinely taxed with supporting other departments of the Library (Office of the Librarian, Library Services, the Law Library, the Office of Strategic Initiatives, and the Office of Support Operations). Unlike the Copyright Office, these departments do not administer a federal law, are much less public-facing, do not collect fees for services, and in most instances do not have a separate congressional appropriation.

Because the Copyright Office serves and intersects with a major and global marketplace of creative content, it must operate more like a business, essentially twenty-four hours per day, seven days per week. However, current dependencies with the Library do not currently allow for a "hot" alternate computing site that, in a best case scenario, would allow the Office to balance IT processing needs across primary and secondary sites to achieve high redundancy and high availability.

The Copyright Office is dependent upon the Library's IT shop for basic things as well, including scheduled software releases. If a problem occurs outside of normal operating hours, the Copyright Office must locate a Library technology representative before the situation can be analyzed or solved. Many issues which arise in supporting the eCO system could be resolved in a more timely fashion and be more easily administered if the CIO and technology staff of the Copyright Office maintained primary administrative access to Copyright Office system and network resources. And as noted previously, there are policy and legal reasons as to why Copyright Office data should not be co-mingled with the Library's data.

In order to meet the availability needs of the copyright community, which largely fund the Copyright Office, it is crucial that the Office's website and systems remain open and available as much as possible. This has not been possible under the current IT support arrangement. To solve this, the project team is in favor of, and believes the copyright community expects, a level of direct control over IT decisions, if not the autonomy necessary to serve the copyright marketplace. There are several levels of IT autonomy. At the very least, the Copyright Office could attain a logical IT autonomy. This would entail ownership and control of its own servers and network, even if they remain part of the larger ITS infrastructure.

That said, the project team believes the best solution would entail some level of physical autonomy, as well as logical autonomy. For example, the Copyright Office's IT systems (those it pays for and operates) could be housed within the Library's IT environment, however logically or physically isolated from the Library's equipment, for example in a locked cage. However, due to power and cooling limitations within the current LOC data center, this is not a feasible option. At the highest level of autonomy, the Copyright Office would have its own independent location in which to build and maintain the Copyright Office IT infrastructure. This would likely require some combination of an entirely separate physical space or office building while leveraging certain aspects of a cloud model.

As the Register has indicated, we should "*look to the technology sector*" in improving our technology solutions for the future.[59] To better support the IT needs of the Office, system virtualization and cloud technologies must be strongly considered. The technology industry has leveraged system virtualization for more than a decade and many organizations are now also embracing cloud computing. Virtualization is an underpinning concept that must be mastered (or well-conceived) in order to take full advantage of cloud computing technologies.

For clarification, virtualization leverages network and server resources (*e.g.*, memory, processing power, etc.) by logically segmenting (using specialized utilities) one physical device into many logical devices. For example, one physical server may be logically separated into multiple servers so that the processing power and memory of any one server can be fully utilized. Once virtualized environments are established, spinning-up or ramping-down of needed environments can be performed in very short time period. Obviously cost savings and speed of support are significantly improved in a highly effective virtualized environment.

Cloud computing is one of the more recent advancements in the technology industry even though it has been in place for several years, and is discussed above in section 6.1.2. Certain organizations have adopted cloud computing faster than others. With the enactment of the presidential directive for government agencies to begin moving systems to the cloud (cloud first), a greater number of agencies are successfully leveraging cloud technologies. More cloud providers are available, cloud utilities are more robust, government contracts offices are more adept at issuing cloud service contracts, and agencies are finally beginning to reap the benefits of cloud technology. Cloud models can be implemented as a private model (meaning isolated and walled off to others) and/or a public model (meaning within a cloud provider facility or facilities and likely co-mingled with other cloud customers or subscribers). While not technically a cloud variation, discussion should be had around the concept of co-location and managed hosting solutions and whether or not the solution should be managed by Copyright Office staff or managed by data center staff to the specifications of a service level agreement (SLA). Co-location and managed hosting should be considered as

[59] *Priorities and Special Projects* at 13.

alternatives when conducting an alternatives assessment and cost-benefit analysis for infrastructure support.

Whether virtualized and/or in the cloud, industry best practices include that the infrastructure should be comprised of at least four segregated environments: a development environment, a test environment, a pre-production environment and a production environment. These environments include both the necessary hardware and software needed by the developers. In addition, the Copyright Office must provide skilled, experienced developers.

To achieve the desired operational objectives of high availability, high redundancy, and highly reliable and secure IT solutions, the Office should consider virtualizing its environments while leveraging a cloud solution or managed hosted solutions for data processing and data storage. Along with considering the options for backend infrastructure, the Office must evaluate the options for the entire realm of technology services to include desktop support services, security services, network and telecommunications support services, application support services, telework services, and help desk services.

7.4.3 Secure Repository for Works of Authorship

Approximately 85% of copyright applications routinely arrive through the eCO system, and many of these, around 50%, include digital uploads of the works being presented for examination. It should be understood that, for a variety of reasons, many applicants mail their works to the Office using a specialized barcode linked to their applications. Some of these registered works, approximately 10%, are selected by the Library of Congress for its acquisitions each year. The Register transfers these works to the Library according to provisions in the Copyright Act and applicable regulations. Once transferred, they become part of the Library's collection and are in the Library's custody. The rest of the works registered remain in the Copyright Office's care and responsibility, where among other things they may be needed by courts or parties to copyright litigation.

In the current architecture, digital works which accompany applications, once approved, are ultimately passed through the eCO system and on to a proprietary storage system that is outside the control of the Copyright Office. This system is both owned and operated by the Library of Congress. The Library's IT staff provide very limited, almost negligible, control to Copyright Office's technology staff. The system file shares are co-mingled with other Library of Congress systems, and are only as secure as the inherent storage system configuration settings.

With this current configuration, the Copyright Office cannot provide 100% assurance that a deposit submitted by a copyright owner has not been modified. Nor can it ensure that the deposit has been archived in a manner that is consistent with the requirements of the copyright law or applicable regulations. In fact, in a worst-case scenario, under the current configuration, the Office would have no knowledge as to whether deposits are being accessed, deleted, removed, or appropriated.

The project team has reached the unequivocal conclusion that the Copyright Office should have full control over all works of authorship that are submitted to the Register for purposes of registration, other perhaps than with respect to those works selected by the Library for its collection. With respect to the works selected by the Library, although it is not the purview of this report, the team did hear concerns regarding whether or to what degree the Office can track these works, or make certified reproductions, in case they are needed for litigation. The current system is one that seems to have conflated the Library's needs and the Copyright Office's needs as though they are one and the same actor, despite clear areas of demarcation provided for in the law. To say the least, when copyright owners entrust the Office with their intellectual property, they should feel assured that the data and deposits submitted are maintained with the highest level of integrity and security possible. They are, after all, registering to maximize the legal protection of their works.

There are several policy and pricing considerations for storing of electronic deposits. As the Register points out in *Priorities and Special Projects*, the Copyright Office needs to know "what…the repository standards [are] for acquiring and migrating electronic copyright deposits," as well as how long the Office should "retain deposits and under what practices and cost structures." Consideration must be given to whether the Copyright Office should simply provide the remitter with the exact deposit they submitted or should we (or can we) have the ability to present them with a copy of their work in an updated format? To what level does the Office ensure, or guarantee, the validity of the deposit, and at what cost? Should it be providing file reader capabilities so that all file formats submitted are able to be opened in the future? Should we enable checksum, hashing, or digital signature capabilities upon deposit transfer to the Office? It is the understanding of the Technical Upgrades project team that most of these issues will require a formal rulemaking process. And we would take them into account when designing and implementing a fully functional, Copyright Office-specific, repository system.

7.4.4 Remote Workforce

The Copyright Office participates in the Library's telework program, but because the Office's work is tied closely to customer issues, the technology infrastructure is of paramount concern. Copyright Office staff who currently telework report much increased productivity as well as reduced amounts of stress. There are many additional benefits to implementing a robust telework capability including improved morale and reduced levels of staff attrition. It can be an important recruiting tool, and an important management tool when dealing with inclement weather or natural disasters. Additionally, in the event of an unforeseen emergency such as a COOP (Continuity of Operations Program) implementation or disaster situation, Copyright Office staff disbursed throughout the Washington Metropolitan area (or beyond), may still have the ability to remotely perform their day-to-day functions without any loss of productivity.

Copyright Office staff who telework today experience frequent frustrations related to technological disruption, some of which are caused by VPN connections into the

Library. The Office needs an easy to use but highly secure and highly redundant solution that is platform agnostic and offers access by any number of remote technology systems such as laptops, staff-owned personal computers, tablets, or other mobile devices. Any remote access capability should be a high performance, seamless solution that allows users who login remotely the same experience as logging in from the Office. The organization will additionally need the ability to remotely manage Copyright Office-owned systems and have the ability to push updates to systems without requiring staff to haul machines back into the Office to update applications as they do today. The Office will need a significantly robust and secure telecommunications network to allow multiple, simultaneous connections along with the ability to tie into internet service provider diagnostics so as to isolate issues encountered during troubleshooting.

7.4.5 Mobile Technologies

It is well publicized and documented that personal computer sales are down while mobile technologies, including smart phones, laptops, and tablets, have realized a steady upward sales trend. Keeping up with the latest technology is not the prime reason for embracing mobile technologies, but understanding how our user community may be submitting copyright applications is. The project team believes that the Copyright Office should embrace a mobile device management (MDM) strategy in the future from both the perspective of user community and the internal work processing community. In a voluntary registration and recordation environment, Copyright Office services must be available from as many operational platforms as possible including smart phones, tablets, or even glasses.

The entire Copyright Office application suite should be developed with leveraging responsive web design and/or adaptive web design in mind so that the user can experience platform agnostic Copyright Office services. Establishing and publishing well defined APIs, data standards, and web services will certainly aid the Office's positioning in the mobile marketplace. This will not only allow external entities to communicate effectively with the Office, but it will allow developers to establish specialized front-end applications which communicate with the Office. In the end, availability of Copyright Office services is significantly improved, data becomes more robust, and inbound submissions are made as easier.

As an example, the Office should be responsive to the growing number of people using their mobile devices to take and share photographs, videos, and songs. With a few clicks, these content creators should be enabled to register their works with the Copyright Office from their mobile devices.

7.4.6 Systems Development Strategy

The Copyright Office needs to consider benefits of moving toward service oriented architecture (SOA) for its application portfolio. This approach provides the Copyright Office with much flexibility when deciding to build or buy applications. The Copyright Office currently has a very limited software development capability due to resources and

focus on COTS products. This has resulted in limited maturity in the areas of software development methodology, process, and standards. However, the Copyright Office has a higher level of maturity in the areas of operations and maintenance in processes. The reliance on COTS has limited the Copyright Office from having flexibility in delivering solutions based on custom-built software.

A significant amount of customization is required to COTS applications in order to meet the Copyright Office's and its stakeholder's needs. The level of customization removes the key advantage of deploying COTS products which includes a reduced level of effort to build and reduced overall total-cost-of-ownership (TCO). In addition, reliance on COTS software and absence of an SOA strategy have caused the following issues:

- Services and application functionality tied to release cycles of vendors
- Customization still requires software development but due to a lack of a formal development capability, applications do not routinely or uniformly follow standards or ability for reuse
- Must maintain licensing contracts and fees, and current version to retain support
- Periodically need to bring in vendor specialists
- Number of support personnel on open market are fewer because they require vendor specific knowledge and are also more expensive
- The need to support and license Siebel and WebMethods products

The Copyright Office should establish a software development capability, so the Office can deliver systems that better meet the Office's and stakeholder's needs in a more cost effective and timely manner. Having an in-house software development capability does not mean that COTS products cannot be leveraged when it makes sense. As Apple and Amazon have done, the Copyright Office should take a holistic approach to development, aimed at creating an integrated business ecosystem (see section 6.1.3) based upon SOA principles. Therefore, as long as a particular selected COTS product follows SOA principals, the components could be used as part of all software development efforts.

7.4.7 Enterprise Solution

Among the most critical conclusions of the Technical Upgrades team is to completely overhaul the eCO system, leveraging SOA. This includes providing additional capabilities other than registration, which is currently the primary function. This enterprise solution should be established as an integrated business solution.

In the view of the Technical Upgrades project team, the best approach comprises a two-phased effort. Phase one would address immediate improvements to the current eCO system, to specifically address the challenges of the current registration system. Several functional and cosmetic enhancements would be implemented to the existing eCO user interface, as well as some adjustments to the supporting infrastructure. These enhancements would improve the system, particularly for the public, and cause minimal or manageable disruption to the internal office workflows and business

processes. These short-term improvements should provide some immediate relief to registration applicants while the Office develops and implements its enterprise solution. Concurrently, the Copyright Office should commence a copyright enterprise-wide solution that focuses on automating the recordation functions, public information and records functions, accounting and processing functions and acquisition and demand functions, followed by a complete redesign and redevelopment of the registration function. The document recordation function would likely be the first initiative, since it is a vulnerable, paper-based process, heavily dependent upon legacy computer systems.

Architecture Recommendations

To summarize, these are the steps that the Project Team recommends, on a schedule and budget that would need to be determined:

Develop and evaluate an Enterprise Architecture

- Engage in an analysis effort to itemize proposed improvements to the Copyright Office architecture by first defining an “as-is” and “to-be” model for the architecture.

Assess Infrastructure

- Engage in an analysis effort to identify proposed specific improvements to the Copyright Office infrastructure by first defining an “as-is” and “to-be” model for the infrastructure.

Evaluate Storage for Works of Authorship

- Engage in an analysis effort aimed at certifiable standards for maintaining and protecting works of authorship submitted to the Copyright Office for registration. The Office would also need to define, develop, and implement an automated system which records transfer and acceptance of physical and digital deposits to warehouse and storage locations.

Develop an Enterprise Solution

- Develop a new enterprise copyright system that focuses on legal and client requirements.

Resources

- To properly develop, implement, and support the recommendations of the Technical Upgrades project, the Copyright Office should address staffing and technology investments.

New Features

- Engage in an analysis effort aimed at implementing a technologically sound, secure, robust, state-of-the-art, upgradeable, and easily and seamlessly managed telework capability.

Closing

In closing, the project team believes that the Tech Upgrades Special Project has been a valuable endeavor. In order to manage the IT resources of the Copyright Office in a manner that supports its core mission and statutory responsibilities, the Office will need significant improvements. The project team has highlighted important concepts and technologies that, if implemented, would optimize key services for customers, including copyright registration, the recordation of copyright documents, and the searchability of public records, and facilitate the exchange of legal and business data with the global marketplace. We hope the recommendations set forth here will inform the strategic direction of the Copyright Office.

APPENDIX A FEDERAL REGISTER NOTICES

committee shall be deemed to be employees of the United States Government for any purpose by virtue of their participation on the committee. Members of the committee will not be compensated for their services or reimbursed for travel expenses.

Authority: The authority for this notice is granted by the FACA (5 U.S.C. App. 2) and the Secretary of Labor's Order No. 18–2006 (71 FR 77560 (12/26/2006)).

FOR FURTHER INFORMATION CONTACT: Paula Church Albertson, Division Chief, Monitoring and Enforcement of Free Trade Agreements, Bureau of International Labor Affairs, U.S. Department of Labor, telephone (202) 693–4789.

Signed at Washington, DC, this *15th* day of *March* 2013.

Carol Pier,

Acting Deputy Undersecretary, Bureau of International Labor Affairs.

[FR Doc. 2013–06630 Filed 3–21–13; 8:45 am]

BILLING CODE 4510–28–P

LIBRARY OF CONGRESS

U.S. Copyright Office

[Docket No. 2013–2]

Technological Upgrades to Registration and Recordation Functions

AGENCY: U.S. Copyright Office, Library of Congress.

ACTION: Notice of Inquiry.

SUMMARY: The United States Copyright Office (hereinafter Copyright Office or Office) is in the process of identifying and evaluating potential improvements and technical enhancements to the information technology platforms that support its registration and recordation functions, including its online registration system. These efforts are part of the Office's ongoing special projects, commenced October 25, 2011 (available at the Office's Web site at *www.copyright.gov/docs/priorities.pdf*). The information garnered through this process has and will continue to inform the development of the Copyright Office's long-term strategic plan, scheduled to commence in October 2013.

At this time, the Office seeks comments regarding existing capabilities and future possibilities. Broadly, the Office seeks comments on (1) how stakeholders use the current online offerings of the Copyright Office, especially with respect to registration and recorded documents, and how the current offerings meet, fail to meet, or exceed user expectations; and (2) how stakeholders would like to interact with the Copyright Office electronically in the future, or, put differently, what online services, or aspects of existing online services stakeholders would like to see. The Office appreciates the comments and suggestions of those who use the national registration and recordation systems to protect their intellectual property, as well as those who regularly use Copyright Office resources to identify copyright owners, investigate the copyright status of works and the public domain, and perform other research, including statistical analysis on aggregated data sets.

DATES: Comments on the Notice of Inquiry and Requests for Comments are due on or before May 21, 2013.

Submission: All comments shall be submitted electronically. A comment page containing a comment form is posted on the Copyright Office Web site at *http://www.copyright.gov/docs/technical_upgrades.* The Web site interface requires submitters to complete a form specifying name and organization, as applicable, and to upload comments as an attachment via a browse button. To meet accessibility standards, all comments must be uploaded in a single file in either the Portable Document File (PDF) format that contains searchable, *accessible* text (not an image); Microsoft Word; WordPerfect; Rich Text Format (RTF); or ASCII text file format (not a scanned document). The maximum file size is 6 megabytes (MB). The name of the submitter and organization should appear on both the form and the face of the comments. All comments will be posted publicly on the Copyright Office web site exactly as they are received, along with names and organizations. If electronic submission of comments is not feasible, please contact the Copyright Office at 202–707–8350 for special instructions.

FOR FURTHER INFORMATION CONTACT: Douglas Ament, Director of Information Technology, Copyright, by email at *uscotechupgrades@loc.gov*; Christopher S. Reed, Senior Advisor for Policy & Special Projects, Office of the Register of Copyrights, by email at *creed@loc.gov*; or call the U.S. Copyright Office by phone at 202–707–8350.

SUPPLEMENTARY INFORMATION:

I. Background

In 2000, the Copyright Office initiated a comprehensive business process reengineering initiative intended to update the Office's technology platform and improve operational efficiency. With the assistance of outside consultants and business analysts, the Office identified opportunities for efficiency enhancements and process improvements. The most significant recommendation was to convert the existing paper-based copyright registration system to an electronic system, which included the development of related new business processes and the automated production of public copyright records.

Funding available for the reengineering effort was limited and decisions made by the Copyright Office and the greater Library were necessarily constrained. Ultimately the Office implemented a commercial off-the-shelf software package. The Office piloted the internal business process functions of the software with a subset of constituents in February 2005, followed by full implementation of the Copyright Office's electronic processing system in August 2007. The public-facing electronic registration system—the system that enables copyright registration applicants to submit materials online—was launched in July 2008. The Office implemented a significant upgrade to its software and hardware platforms in August 2010, and implemented periodic upgrades and enhancements to accommodate the needs of the system's stakeholders—namely those that submit materials for registration, those that search the Copyright Office database for copyright ownership information, and the Copyright Office's staff that process and examine copyright claims.

Today, more than eighty percent of the Office's applications for copyright registration come through the electronic system, demonstrating the copyright community's widespread interest in electronic registration functions. Although the current system, and the periodic upgrades and enhancements, have allowed the Office to maintain a functional electronic platform for many types of works, there is room for substantial improvement. Notably, the Office's recordation services were included in the initial reengineering plan, but were later dropped for budgetary reasons. Recordation processes are, thus, still paper-based and are a top concern of the Copyright Office. Thus, the Office's current systems represent the ''first generation'' of the Office's electronic processing capabilities.

II. Discussion

In recent months, project leaders from the Copyright Office have engaged copyright owners, users of copyright records, technical experts, public interest organizations, and lawyers,

including through professional associations and small businesses to participate in a series of focused discussions on issues relating to the Office's platforms for registration, document recordation, and public access to copyright ownership information. Through these discussions, as well as through its own expert analysis, the Office has identified a number of areas in which the current electronic system could be improved. For example, numerous interested parties have observed that the current user interface for electronic registration is a challenge to navigate. Users have told the Office that it would be helpful to be able to customize the user interface and workflow in order to streamline the registration process to accommodate their own internal workflows. Moreover, users would like to exercise some degree of control over the nature and scope of information they view in a personalized registration system dashboard. The Copyright Office is aware of similar requests from its own staff, many of whom desire customizable workflows to enhance productivity and process efficiency, which would result in improved turnaround times for remitters.

At a global level, the Office is aware that as mobile technology becomes ubiquitous, an increasing number of stakeholders desire to use mobile devices to interact with the Office. To that end, the Office is evaluating the potential to deploy a mobile optimized web interface, ''apps'' that support popular mobile platforms, and the development of an application program interface (API) that can be utilized within third party applications.

The Office has also heard that many of its users would benefit from improved tracking capabilities. Remitters have indicated that the existing electronic registration process is cumbersome and are oftentimes uncertain of their progress within the application process; to improve that aspect of the system, they have suggested that the Office implement a visual representation of the registration workflow and the user's status within it (*e.g.,* a status bar).

Beyond improvements to the registration functions, the Office is aware of opportunities for improvement to its public record search capabilities. Stakeholders have indicated that the Office's search function should be more robust, allowing for more search criteria, refining the display of the search results, adding filters, and generally making the search functionality more user-friendly. Representatives from interested parties also suggested the Copyright Office make it easier to provide updates to the public record to ensure the data maintained is accurate and up to date (*e.g.,* address changes). The Office is thus investigating methods of secure and effective data sharing between interested parties and the Copyright Office in order to determine if such functionality can be implemented in a manner that ensures integrity of the Office's records.

The Office is also aware of the need for long-term, scalable data storage and archiving capability to accommodate the growing volume of digital works that the Office receives. The Office has received recommendations to centralize the various information clusters internally within the Copyright Office to a central data repository and establish a central data warehouse. Implementing such a warehouse presents a series of challenges that the Office seeks to learn more about, including determining scalable infrastructure solutions to accommodate vast amounts of data, analyzing data standards needed to establish a central data model, and evaluating potential data archival strategies.

One recommendation that the Office frequently hears, and one that underlies many of the areas of improvement noted above, is the need for bulk data transfer between the Office and interested outside parties. Such transfer mechanisms would allow more widespread distribution of the Office's records, as well as permit remitters to submit large quantities of electronic material and associated application data to the Office. Such ''system-to-system'' or ''business-to-business'' capabilities are a central area of inquiry for the Office. Interested parties have suggested that the Office expose data portals enabled to facilitate data exchange over standards-based protocols such as ebMS, SOAP, and AS4.

In support of potential bulk data transfer capabilities, the Office is investigating specific data exchange standards, including those that already exist as well as the potential for developing a new standard based upon the needs of the Office's constituents. Interested parties have told the Office that it should continue to take an active role and adopt existing standards that support data exchange between the Office and its stakeholders. This includes defining or adopting metadata standards that support particular industries (*e.g.,* IPTC for photography; ISRC for sound recordings; ONIX for books). Further, standards such as CISAC's Common Works Registration (CWR) and DDEX digital supply chain standards should be considered to help develop the Office's ability to provide better business-to-business data transfers. Interested parties have suggested that the Copyright Office publish a recognized list of data standards so that users are able to establish systems that support more efficient interactions with the Copyright Office.

III. Subjects of Inquiry

The Copyright Office is currently evaluating what the ''next generation'' of its electronic services should look like. Through a comprehensive evaluation of its current technical processing capabilities, and extensive interaction with stakeholders, the Office hopes to develop a complete picture of how the Office currently supports the needs of the copyright community, and where its systems and services could be improved. The Office hopes to achieve a greater understanding of current technical challenges facing the copyright community as well as gain a comprehensive understanding of how the community hopes to conduct business with the Copyright Office in the future. This evaluation process, which is tied to special projects detailed in *Priorities and Special Projects of the U.S. Copyright Office* released by the Register of Copyrights in October 2011, is intended to inform the development of the Office's next five-year strategic plan that will commence in October 2013 and guide, among other things, the technological evolution of the Copyright Office. That plan will, in turn, inform the Library of Congress's overarching strategic plan.

Because the Office's evaluation of its technology platform is intended to be a wide-ranging review of existing capabilities and future possibilities, the Office seeks comments that present conceptual frameworks with concrete examples of future potential applications or services. Broadly, the Office seeks comments on (1) how stakeholders use the current online offerings of the Copyright Office, especially with respect to registration and recorded documents, and how the current offerings meet, fail to meet, or exceed user expectations; and (2) how stakeholders would like to interact with the Copyright Office electronically in the future, or, put differently, what online services, or aspects of existing online services stakeholders would like to see.

Although the Office welcomes comments on the wide range of topics germane to this inquiry, it is particularly interested in comments that address: (1) The nature and capabilities of the Copyright Office's public portals (*e.g.,*

for electronic registration services), including interface-based portals as well as business-to-business portals, or access to Copyright Office services or data through application program interfaces; (2) the nature and scope of information captured during the course of the registration and recordation processes, including that which could be captured through user input, or through metadata harvesting; (3) metadata standards in particular industries that the Copyright Office might adopt or incorporate into its systems (*e.g.,* IPTC for photography; ISRC for sound recordings; ONIX for books); (4) data storage and security standards for electronic copyright deposits, including the development of policies and best practices for data retention and migration; (5) new ways of searching and accessing registration and recordation data and/or registration deposit metadata (*e.g.,* image or music search technology); and (6) the integration of third-party databases of copyright ownership and licensing information (such as those maintained by collective management organizations) and related technologies with data maintained by the Copyright Office.

Dated: March 18, 2013.

Maria A. Pallante,

Register of Copyrights, U.S. Copyright Office.

[FR Doc. 2013–06633 Filed 3–21–13; 8:45 am]

BILLING CODE 1410–30–P

NUCLEAR REGULATORY COMMISSION

[NRC–2013–0020]

Biweekly Notice; Applications and Amendments to Facility Operating Licenses and Combined Licenses Involving No Significant Hazards Considerations; Correction

AGENCY: Nuclear Regulatory Commission.

ACTION: Notice; correction.

SUMMARY: The U.S. Nuclear Regulatory Commission (NRC) is correcting a notice that was published in the **Federal Register** (FR) on February 5, 2013 (78 FR 8195), regarding the applications and amendments to facility operating licenses and combined licenses involving no significant hazards considerations. This action is necessary to correct an erroneous date.

FOR FURTHER INFORMATION CONTACT: Cindy Bladey, Chief, Rules, Announcements, and Directives Branch, Office of Administration, U.S. Nuclear Regulatory Commission, Washington, DC 20555–0001; telephone: 301–415–3667; email: *Cindy.Bladey@nrc.gov.*

Correction

In the FR of February 5, 2013, in FR Doc. 2013–02352, on page 8202, first column, correct the fourth full paragraph to read:

Date of initial notice in ***Federal Register:*** September 4, 2012 (77 FR 53927).

Dated at Rockville, Maryland, this 18th day of March, 2013.

For the Nuclear Regulatory Commission.

Cindy Bladey,

Chief, Rules, Announcements, and Directives Branch, Division of Administrative Services, Office of Administration.

[FR Doc. 2013–06545 Filed 3–21–13; 8:45 am]

BILLING CODE 7590–01–P

SECURITIES AND EXCHANGE COMMISSION

[Investment Company Act Release No. 30427; File No. 812–14114]

Ivy Funds Variable Insurance Portfolios, et al.; Notice of Application

March 15, 2013 .

AGENCY: Securities and Exchange Commission (''Commission'').

ACTION: Notice of an application under section 6(c) of the Investment Company Act of 1940 (''Act'') for an exemption from rule 12d1–2(a) under the Act.

SUMMARY OF APPLICATION: Applicants request an order to permit open-end management investment companies relying on rule 12d1–2 under the Act to invest in certain financial instruments.

APPLICANTS: Ivy Funds Variable Insurance Portfolios (the ''Trust''), Waddell & Reed Investment Management Company (''WRIMCO''), and Waddell & Reed, Inc. (''W&R'').

FILING DATES: The application was filed on January 18, 2013.

HEARING OR NOTIFICATION OF HEARING: An order granting the application will be issued unless the Commission orders a hearing. Interested persons may request a hearing by writing to the Commission's Secretary and serving applicants with a copy of the request, personally or by mail. Hearing requests should be received by the Commission by 5:30 p.m. on April 9, 2013, and should be accompanied by proof of service on applicants, in the form of an affidavit or, for lawyers, a certificate of service. Hearing requests should state the nature of the writer's interest, the reason for the request, and the issues contested. Persons who wish to be notified of a hearing may request notification by writing to the Commission's Secretary.

ADDRESSES: Elizabeth M. Murphy, Secretary, Securities and Exchange Commission, 100 F Street NE., Washington, DC 20549–1090; Applicants, 6300 Lamar Avenue, Overland Park, Kansas 66202–4200.

FOR FURTHER INFORMATION CONTACT: Bruce R. MacNeil, Senior Counsel, at (202) 551–6817, or Daniele Marchesani, Branch Chief, at (202) 551–6821 (Division of Investment Management, Office of Investment Company Regulation).

SUPPLEMENTARY INFORMATION: The following is a summary of the application. The complete application may be obtained via the Commission's Web site by searching for the file number, or an applicant using the Company name box, at *http://www.sec.gov/search/search.htm* or by calling (202) 551–8090.

Applicants' Representations

1. The Trust is a Delaware statutory trust registered under the Act as an open-end management investment company. WRIMCO, a Kansas corporation, is an investment adviser registered under the Investment Advisers Act of 1940, as amended (the ''Advisers Act'') and serves as investment adviser to the Trust. W&R is organized as a Delaware corporation, and is a registered broker–dealer under the Securities Exchange Act of 1934, as amended (''1934 Act''); W&R is the principal underwriter of the Trust.

2. Applicants request the exemption to the extent necessary to permit any existing or future series of the Trust and any other registered open-end management investment company or series thereof that (i) is advised by WRIMCO or any person controlling, controlled by or under common control with WRIMCO (any such adviser or WRIMCO, an ''Adviser''),[1] (ii) is in the same group of investment companies, as defined in section 12(d)(1)(G) of the Act, as the Trust and invests in other registered open-end management investment companies in that same group (''Underlying Funds'') in reliance on section 12(d)(1)(G) of the Act; and (iii) is also eligible to invest in securities (as defined in section 2(a)(36) of the Act) in reliance on rule 12d1–2 under the Act (each a ''Fund of Funds''), to also invest, to the extent consistent with its investment objectives, policies, strategies and limitations, in financial

[1] Any other Adviser also will be registered under the Advisers Act.