COMMUNIA

Comments of Communia Association to the Library of Congress RFI for Third Party Digitization Initiatives

The mission of the International Communia Association[1] is to foster, strengthen, and enrich the public domain. To fulfill its mission, the Communia Association and its members raise awareness in, educate about, advocate for, and offer expertise on and research about the public domain. The Association was formed upon the completion of the EU-funded Communia Thematic Network[2], which operated from September 2007 until February 2011. That Thematic Network gathered over 50 members from academia and civil society researching and promoting the digital public domain in Europe and worldwide.

We applaud the Library of Congress (Library) for taking the initiative to solicit expertise and capacity from outside vendors to assist in fulfilling the priority to increase public access to the Library collection. The Library should push for broad, unencumbered public access to the materials as soon as possible. From the sample project list, we assume that the majority of the materials to be digitized will be public domain. We offer a few suggestions for strengthening the RFP and related policies below.

The Library notes, "The Contractor shall not claim copyright in the digitized copies of the original Library materials" ("RFP," *Data Rights*, p. 22). This is aligned with the views of Communia, which has communicated via the Public Domain Manifesto[3] and policy recommendations[4] that digital reproductions of works that are in the public domain must also belong to the public domain. It would be beneficial for these digitized copies to be marked as being in the public domain using a tool such as the Creative Commons Public Domain Mark[5]. The Public Domain Mark enables works free of known copyright restrictions to be labeled in a way that allows them to be easily discovered on the Internet.

Concerning digitization formats, the RFP mentions that the contractor will be responsible for "providing the Library with a copy of the digitized materials in the format designated for the individual project materials, as agreed upon between the Library and the Contractor" ("RFP," *C.1.1.1.B Scope*, p. 3). The Library should commit to requesting that digitized materials be transferred in formats meeting open standards definitions so that the materials can be made available to the public (either immediately or after the embargo) in such a way that does not require specialized software or hardware to view and manipulate the digitized materials. The Library should give preference during the

---

[1] http://www.communia-association.org
[2] http://www.communia-project.eu
[3] http://publicdomainmanifesto.org/manifesto
[4] http://www.communia-association.org/recommendations-2
[5] http://creativecommons.org/publicdomain/mark/1.0

assessment process to bidders that pledge to make the digitized materials available to the Library in open standard formats.

The RFP states, "In return for access to the Library's materials, the contractor may market and resell, for a limited period, access to the digitized collection to cover the costs of digitization" ("RFP," *B.1 Cost*, p. 2). The Library should consider bids offering a wide variety of access models. For example, some digitizers might be in the position to offer immediate, free ad-supported access (instead of selling access on demand).

The Library should clarify the statement that contractors must "make materials widely available" ("RFP," *C.1.1.1.E*, p. 3) to mean that bidders should be unambiguously required to provide public access on the internet (free or paid). Making available only at limited locations like the Library of Congress, public libraries, or through university subscriptions should not be considered to be widely available. Bidders should also be required to include the terms of this access in their proposal and the Library can then judge submissions based on the terms that a bidder offers. Broader access will be a factor in favor of the bidder. This is aligned with one of the goals of this digitization in the first place, which is to increase public access to the collection and related materials. And in general, the Library should prioritize bids that provide free public access sooner than those that have longer embargo periods.

The sample project list shows that there are several interesting and important collections that the Library is aiming to digitize, and it would be great for the public to have broad access to these materials. At the same time, these projects appear relatively limited in size and scope, and most likely only represent a tiny percentage of the potential collection of public domain materials held by the Library that could be digitized. In future RFPs, the Library should consider how to leverage the expertise and capacity of digitizers to scan not only these small, interesting, and impactful collections, but also the vast (yet less visible) trove of public domain materials that comprise the bulk of the Library's collection. One technique employed by the UK National Archives doing similar digitization work was to require that contractors digitize "less attractive" materials bundled alongside those with higher commercial value to ensure that these other materials get digitized.

The Library requires the vendor to provide a set of core metadata. The Library should be authorized to release this metadata into the public domain using the CC0 public domain dedication[6], thus aligning with an open metadata model used by several large libraries around the world, including The British Library, Harvard Library, and soon to be used by Europeana[7]. Open metadata can promote the development of innovative business models and interesting applications, as well as improve the accuracy and verifiability of digital collections[8]. Europeana and the Digital Public Library of America

---

[6] http://creativecommons.org/publicdomain/zero/1.0
[7] http://creativecommons.org/weblog/entry/29133
[8] http://pro.europeana.eu/documents/858566/2cbf1f78-e036-4088-af25-94684ff90dc5

(DPLA) have agreed to collaborate on content and data sharing. The Library should consider providing open metadata—and if appropriate, access to digitized copies of public domain materials—to new initiatives such as the DPLA.

The clause requiring bidders to demonstrate at least three projects in the previous five years ("RFP," *M.2 Factor 4*, p. 43) might make it difficult for competent newcomers to be awarded contracts. For example, an up-and-coming company might be able to better leverage a new technology or innovative scanning process, but might not be able to show the requisite three-project experience in the last five years. Depending on the digitization needs for an individual project, the Library should apply some discretion so as to properly match expertise and capacity with the respective project. In addition, other potential coalitions or community initiatives could be hampered from leveraging a wealth of interested and motivated individuals to assist in this work. The Library should consider involving volunteers and other community organizations willing to assist in the digitization and quality control work. For example, Wikimedia France partnered with the Bibliothèque nationale de France to process high definition files of public domain texts[9]. In any case, the Library should prioritize partners that can provide the broadest public access with minimal embargo period.

One final concern is what happens once the embargo period comes to an end and the Library then has the ability to provide free public access on its website. The Library should develop a strategic access plan and secure the necessary funding so that the materials can be properly archived and made publicly available without delay once the period of exclusivity has come to an end.

We appreciate being able to provide feedback on the proposed digitization RFP and look forward to the progress on this exciting project to increase public access to the Library collection. We'd be happy to answer any follow-up questions you may have.

Communia Association
Contact: Timothy Vollmer (tvol@creativecommons.org)

---

[9] http://outreach.wikimedia.org/wiki/GLAM/Case_studies/National_Library_of_France