Open Library Society, Inc.
EIN: (redacted)
Attachment to form i1023
Thomas Krichel
30 March 2009

Part I, line 10

We currently have not and do not expect to have gross annual receipts of more than $25000. Our operating costs are very minimal. We work purely on our own goodwill efforts.

Part IV

The activities of the society are difficult to explain in simple terms. Fundamentally, we are collecting digital document and data relating to these document. At them moment, we are mainly working on the second aspect. We propose to collect data via the Internet, in a collaborative fashion, and to provide the results freely to the public. By providing the results freely, we hope that we can encourage contributions in kind to the collections.

The data that we are currently collecting mainly relates to the description of scholarly documents, and of events, persons and institutions involved with academic documents.

Let me begin by describing a simple example of a service that I already run independently from the Open Library Society. I hope that this helps to convey some idea regarding the kinds of services that the Open library society hopes to provide once it is fully developed.

When a new academic paper is written, it needs to reach an audience of people interested in the subject. One way to do this is to allow individual people find it on the web, for example by using Google queries. Another, more interesting way is to have an editor looking at a large number of recent papers, decide which papers are relevant to a particular subject area, and which papers are not part of the subject, and to send out news regarding these recent papers to an email list or an RSS feed. The difference between the work of the individual Google user, who works in isolation to find the papers that are in their area of interest, and the editors who filter papers is essentially that the editors' work is reusable by other people. Their data collations can be exported to digital libraries running on the web. With the description of an academic paper, the library can point to the fact that this document was placed into a particular data set by an editor who collected data in a certain area of research.

So the society's work is in describing contents. At this time, the work of the society concentrates on trying to improve the description of academic authors and of academic institutions.

The first project of the society is to create a list of all academic and research institutions in the world, ariw. It resides at http://ariw.org. It provides a second simple example of what the society's work is about. The site has a list of academic institution names. They may be in

different languages. There may be abbreviated names. It contains a link to the homepage of the institution and an identifier. The society has a collaborator who maintains a part of that data. The data from the collaborator is merged with the society's data periodically.

The society will be looking for volunteers to maintain that data. Even if the society were to receive funds, they will not be spent on the actual maintenance of data but rather on creating software that would make the data easier to maintain or would integrate data from other contributors to the society's data holding. The way the http://ariw.org web site is is built is commensurate with the aims of the society. Rather than end-users coming to the site, the society emphases that the site be easy to copy so as increase the use of the data by others.

The grand present and future project is AuthorClaim at http://authorclaim.org. This is a free author registration system for authors of all disciplines. Author registration was pioneered by myself in 1998. At that time I opened an author registration system for the RePEc digital library that I had been working on since late 1992. This is called the RePEc Author Service. The principle is quite simple. Authors contact the RePEc author service. They give some information about themselves, such as name and affiliation data. Then they search in the publications data of AuthorClaim the papers that they have written. The result is a profile of their contributions to the literature. In RePEc, which is a digital library for economics, over time, around 80% of top 1000 authors, as established in an independent list, have registered over time with the RePEc Author Service. AuthorClaim,org, a project of the society, is trying to achieve a the same success across all disciplines. As authors register with the system, the create a machine-readable profile with a list of their works. An example for a profile, as on the web, is my profile at http://authorclaim.org/profile/pkr1. The true machine interface profile is at ftp://authorclaim.org/k/r/pkr1.amf.xml. The aim of AuthorClaim is to have all academic authors go through the procedure to claim their papers.

Now you may ask well what is the point of all this? Well one of the major issues in research evaluation is to find out who has written what. Author names can help, but they only go so far, and nowadays a shorter distance than they used to, because there are an increasing number of Asian authors with very similar names. The lack of a precise association with between names and authors makes it difficult to build additional services, such as author rankings. Furthermore, the lack of associations with institution names makes it difficult to rank institutions. This is important because we believe that a clearer and easier ranking will push authors and institutions to make their research results more widely and freely available. That will happen when better usage methods are available.

There is a huge amount of work to be done to set up such a system. The society has been formed to support it over my lifetime and beyond. AuthorClaim is freely available to authors, publishers and libraries. It is free for authors. They will register and say what papers they have written. It is free to publishers. They furnish metadata about their publications, whether they be open access or toll-gated. It will be free for libraries. They will use AuthorClaim data to build author-centered content aggregates. The word "will" here is important. Currently they don't do it. We have a chicken and egg problem of few authors being registered, therefore no

libraries using the data, and no authors signing up as a result. This is one of main motivations of creating of the society, to build an organization that sustains this effort, to avoid the impression that this is a fly-by-night operation.

The system is already running at http://authorclaim.org. It has over 17,000,000 document records. There are 11 registered authors. This is underwhelming, but we have to start somewhere. Over the coming years, we will work to improve the interface, so that it becomes quite easy for people to register even if they have very common names and may find that thousands of documents initially match. This will be the beginning of a revolution in scientific documentation.

Part VI, line 1a

We provide digital services over the web, that will allow access to the documents we curate. At the moment we provide two services, ARIW at http://ariw.org, a set of documents about academic and research institutions. They can be used by individuals who want to gain access to a list of institutions by country, for example. The second service is AuthorClaim, a service that allows authors to claim that they have written documents described in a database.

Part VI, line 1b
We provide digital services over the web, that will allow access to the documents we curate. At the moment we provide two services, ARIW at http://ariw.org, a set of documents about academic and research institutions. This can be used by organizations wishing to merge our data with other data they have regarding institutions. The second service is AuthorClaim, a service that allows authors to claim that they have written documents described in a database. This can be used by organizations wishing to provide listings of authors and their documents.

Part VIII, line 4a

We do not currently have any fundraising programs as there are neither staff employed nor volunteers to do this. However, we hope to carry out fund raising in the future using email, personal solicitations, foundation or government grant solicitations; or by accepting web site donations.

Part VIII, line 10

We are currently working mainly on aggregates of factual data for which, under current legislation, we do not think we can claim to be the copyright holder. We may in the future, build collections of digital documents and not only organize the collections, but also hold copyright over some or all items in the collection. However, our mission prohibits the sale of access rights to the digital documents, because our aim is to make such documents available to the public at no charge, not only for use but also for reuse. The purpose of us holding the copyright is to prevent others from claiming copyright and use that claim for financial gain on documents that we want the public to have free access to.