

## NRHS Bulletin Digitization Project

The *NRHS Bulletin* – in publication for over 80 years – has published a substantial volume of historical material. However, much of that information is locked in paper copies of the magazine that are generally not accessible to members and historians. Digitization and distribution technologies have vastly improved over the past 10 years and NRHS management believes that now is the right to embark on an effort to digitize the *NRHS Bulletin* – formerly – carrying the title *National Railway Bulletin* - and make it available to members and the public.

### Process

There are four steps to the digitization project. They are outlined below:

Step 1: Collect a copy of each National Railway Bulletin and NRHS Bulletin that we have published since the 1930s. This was done many years ago by the National Office Staff in Philadelphia.

Step 2: Scan all of the magazines. With somewhere between 20,000 and 30,000 pages of material, this is not a small job and is probably best done by an external vendor. External vendors have specialized equipment and experience to ensuring that the scans are done properly and everything is formatted uniformly without damaging the originals. Jeff Smith – NRHS Bulletin Editor - has all the source files for the post-1998 issues so we would not have to scan anything published after 1998, we can generate PDFs off the electronic files.

Step 3: The scanning process generates an image of the original. To make the PDFs searchable so that information can be found, a process called optical character recognition must be applied to each PDF. This is software that the PDF is run through that extracts data and indexes it so that it can be searched. This process would be done by Bulletin Editor, Jeff Smith.

Step 4: The PDFs must be distributed in some manner. There are a few options here depending on what strategy the NRHS desire to pursue.

### Distribution

The distribution piece is where the most variability in cost exists. There are a few options for distributing the magazines.

Option 1: Put the PDFs on few DVDs and sell them. The challenge with that is that PDFs are easily copied and shared making it very difficult for the NRHS to protect its intellectual property. It also doesn't provide the ability for the user to easily search the collection. They would have to search each individual PDF making its practical use as a research tool questionable. NRHS management does not recommend this approach.

Option 2: Produce a piece of software that would provide the search interface and prevent the actual PDFs from being shared, copied or uploaded to the Internet. The drawback to this approach is forward compatibility with new operating systems. Over time, the archive will

become difficult to use. It also doesn't allow additional magazines to be added to the collection nor does it permit the material to be used on any device without a DVD Player. Today that includes many laptops and desktops as well as all tables and phones. NRHS management does not recommend this approach.

Option 3: There are a number of online distributors to select from Zinio, Apple, Google, Amazon and others. For discussion purposes, I'll use Zinio as an example. Model Railroader used Zinio to release their archive just a few months ago. NRHS would upload the content to Zinio and it would be viewable only through Zinio's viewer which works on nearly any platform (Mac & PC browsers, iPads, Android, Microsoft Surface, etc.). The platform allows new magazines to be added to the archive. NRHS could sell individual copies of magazines or monthly/annual access to the entire archive. Regardless of the distribution vendor we choose, this is the approach that NRHS management recommends since it protects our intellectual property, is easy to add to new magazines to the system and is constantly being updated to support new browsers and devices.

### **Costs**

There are two main categories of cost. The first is for the scanning. Scanning the magazines will cost 50 cents a page. Given that we have somewhere between 20,000 and 30,000 pages to scan we are looking at a cost between \$10,000 and \$15,000. This would be a one-time cost. There would never be a need to re-scan the magazines. 50 cents a page is an industry standard rate. I contacted Kalmbach to see how they accomplished the Train Archival DVD that was released a number of years ago and they paid 50 cents a page to have their scans done.

The optical character recognition step of the work will be completed with software that the NRHS has already purchase for the production of the Bulletin (Adobe Acrobat Pro). There would not be any costs associated with the optical character recognition portion of the project.

There would be distribution costs, however, these have not explored in detail with any of the vendor. Generally, there will be setup costs to get the files into their system and formatted properly to display correctly. Then there will be a fee per subscription sold.

### **Requested Board Action**

Today, NRHS management is asking the Board to authorize the expenditure of up to \$15,000 for the scanning of the NRHS Bulletin to be done by an external vendor (Digiscribe - <http://www.digiscribe.info/>) overseen by Bulletin Editor, Jeff Smith. Once the scanning is done, Bulletin Editor Jeff Smith will present recommended distribution options and costs to the Board.