Resource Review:

Media History Digital Library and Entertainment Industry Magazine Archive

James Steffen¹

EMORY UNIVERSITY

jsteffe [AT] emory.edu

The field of film and media studies is finally starting to benefit from the recent boom in digitized collections, if not to the same extent as other disciplines such as history and literature. This is partly because the field is neither as established nor as widely dispersed within academia as those disciplines, but also because of the greater complexity of questions around copyright and ownership. Many of the core industry-related artifacts that scholars study (or would like to study) were produced and remain owned by corporations, and fall well within terms of existing copyright law.

Arguably, the two most important digital collections today for studying media industries are the *Media History Digital Library* (MHDL) and the *Entertainment Industry Magazine Archive* (EIMA). Both are primarily historical in focus. The MHDL is an open access project headed by David Pierce (Library of Congress) and Eric Hoyt (University of Wisconsin). The underlying collection is hosted on the Internet Archive, accessible through a main web portal and two companion search interfaces, Lantern and Arclight. The EIMA is a collection licensed and packaged by ProQuest and hosted on their standard platform. The fundamentally different origins and underlying philosophies behind both resources have a major impact on the finished products. While both collections have various advantages and disadvantages, when taken together they open up tremendous possibilities for historical research on media industries, as anyone who has spent hours scrolling through microfilm reels of *Variety* will readily understand.

The MHDL is a continually growing and evolving entity, not just in terms of the underlying collection of digitized materials, but also in terms of its multiple interfaces and the features they offer. The bulk of the MHDL's holdings consists of periodicals out of copyright, with some runs more extensive than others depending on a specific periodical's copyright status and which issues the MHDL has been able to acquire or access for scanning. Thus, for

instance, it contains *Broadcasting* from 1931 to 1963, while coverage of *Variety* went up to only 1949 at the time of review. (They are in the process of adding volumes through the 1950s and eventually up to 1964.) The MHDL also contains a growing number of books and ephemera. One notable example is a Warner Brothers Vitaphone brochure from 1927. On the whole, its richest holdings focus on the Hollywood studio era, but its increasing diversity of content makes it useful for scholars working in other areas, such as radio and television, early cinema, film technology, and global cinemas.

Unlike ProQuest's highly standardized interface used in the EIMA, the MHDL offers more of a distinctive, intentionally curated web presence. Its main portal emphasizes browsing by title, but the holdings are also organized into overlapping thematic collections such as "Hollywood Studio System Collection (1914-1964)," "Fan Magazine Collection (1911-1963)," "Broadcasting and Recorded Sound (1896-1964)," and "Technical Journals Collection (1916-1965)." Each issue or volume links out to the Internet Archive interface, where users can read an issue online, search inside it, or download it in full.

The Lantern interface (see Figure 1) is closer to the basic ProQuest search interface and emphasizes keyword searching, with facets on the left that allow one to filter results by date, title, collection, format, and language.

Figure 1. Media History Digital Library Lantern interface.

Its options for each result are "Download This Page and Item Details," "Read in Context," and "Read in Context with Search Highlighted." The latter function is slower, but it both flags appearances of a term within a given volume (or document) and highlights it on the page. It is important to note that Lantern does not contain article-level indexing, and thus, it also

does not offer the kind of automated citation generation offered by large information providers such as EBSCO and ProQuest. This is not so much a criticism as an acknowledgment of different aims, as well as different financial and staff resources.

The third main interface, Arclight (see Figure 2), is a data visualization tool that allows users to track the prevalence of search terms over time and compare them with other terms.

Figure 2. Media History Digital Library Arclight interface.

While the tool is intriguing and potentially very useful, one has to take into account often-significant differences in the coverage of individual periodicals and other materials, which affect how results are counted and visualized. Over time, this problem should be mitigated as the MHDL continues to add to its corpus of searchable materials.

ProQuest's EIMA database is akin to the company's other historical newspaper collections. A fixed collection, it offers full text access to fifty periodicals in film, broadcasting, music, and to a lesser extent, theater. Coverage focuses on complete runs of periodicals up to the year 2000. (It is a self-described "archive.") For the purposes of media industries researchers, probably the most valuable of these periodicals are *Variety* (1905–2000) and the various incarnations of *Broadcasting/Broadcasting and Cable* (1931–2000). Other noteworthy titles include *Billboard* (1894–2000) and the London-based fan magazine *Picturegoer* (1911–1960).

Article-level indexing is easily EIMA's most valuable feature. Furthermore, the individual pages of multipage articles are combined into a single entry. The interface offers typical ProQuest features such as a bibliographic citation generator; email, print, and save functions; and the ability to visualize search results over a specified time span.

To compare the functionality and user experience of both databases, I chose a random issue of *Variety*—December 29, 1943—since both EIMA and MHDL contain it in their collections. In the MHDL, this particular issue was scanned from microfilm and is therefore not quite as clear as EIMA's paper-based scan (see Figure 3). Nonetheless, the MHDL scan was still readable and certainly easier to work with than microfilm. (With the notable exception of *Variety*, the majority of scans in the MHDL are in fact paper-based and of very high quality like those in the EIMA.) When you download the MHDL issue, you get the entire month of December as a single file. The PDF is fully searchable thanks to embedded optical character recognition (OCR), and you can also download the underlying dirty OCR as a text file.

Figure 3. Sample page scans from *Media History Digital Library* (left) and *Entertainment Industry Magazine Archive* (right).

In the EIMA, individual issues are accessible through the “Publications” menu. When you select “Browse This Issue,” it takes you to an article list—a handy feature if you know what you are looking for. To look at the actual scan of the article, you have to click on “Full Text.” If you want to browse a full issue, you have to click “Browse This Issue” again above the page image. The EIMA thus reflects one of the main limitations of the standard ProQuest interface: it heavily emphasizes retrieving individual results over browsing and exploration. As a result, students are unlikely to find or use the issue browsing functionality. This contributes to an atomized research experience that to a certain extent discourages, or at least deemphasizes, open exploration, serendipitous discovery, and seeing individual documents within a broader context. In general, the interface is more locked down than the MHDL, as one might expect, since it is a commercial database that licenses third-party content. For instance, search terms are highlighted on the page of a document, but downloaded PDFs are not searchable. One also cannot access the underlying OCR.

For a comparison of keyword searching, I looked in *Variety* for instances of “Palladium,” a common name for theaters, and limited results to the year 1943. EIMA yielded 134 results (that is, individual articles) and listed the titles of each article. MHDL’s Lantern interface required using the facets on the left to limit by publication and year, as mentioned earlier. It yielded 105 results. Instead of article titles, it offered brief OCR excerpts from the text—also very useful, though in a different way from the article titles. Without closer investigation, it is difficult to know what to make of the discrepancy in search results, though it is worth noting.

Finally, it is important to mention that both resources have two significant gaps for media industries researchers: *The Hollywood Reporter* and *Daily Variety*. Of these, the MHDL has the 1933–1934 volumes of *The Hollywood Reporter* and coverage will be expanded at a future date. The most extensive source for *The Hollywood Reporter* available to libraries remains microfilm, but even in that case, only the years 1934 to 1941 were ever offered. Currently, historical coverage of *Daily Variety* is only available electronically through the *Variety Ultimate* package, but its cost and access model make it beyond the reach of many researchers and academic libraries. Given the underlying complexities for making both titles available, this gap is unlikely to be resolved anytime soon.

To conclude this comparison, the MHDL is free and thus widely used by researchers who may or may not be affiliated with an academic institution—including filmmakers and technicians, as I learned at a recent conference for film librarians. Open-access resources like the MHDL fulfill a critical role in addressing economic and institutional imbalances of access to information. Academic institutions and individual researchers should consider ways to support projects like this both in terms of logistics and finance, in the interest of global citizenship and as part of their broader educational mission. Practically speaking, such projects will have to exist for some time alongside subscriptions to commercial products such as ProQuest's EIMA. The concentration of resources in a handful of large content providers who (understandably) emphasize the interoperability of their products results in tradeoffs between standardization and stability on the one hand, and flexibility and innovation on the other. With regard to flexibility, innovation, and overall user access, the MHDL has a basic structural advantage. Even so, the EIMA has the advantage of more complete and recent coverage of two major trade publications, *Variety* and *Broadcasting & Cable*, to say nothing of its high-quality, article-level indexing—something that requires significant staff resources.

Finally, both products, as valuable as they are, face a larger, long-term problem. In my experience as a librarian and teacher, both undergraduate and graduate students in film and media studies are increasingly less interested in researching historical topics. They tend to gravitate toward current or very recent topics. This raises the question of how much use such digitized collections will receive by future generations of scholars. For now, we can do our part by vigorously maintaining historical inquiry within a larger media studies curriculum, but that seems only a partial solution.

---

¹James Steffen is the film and media studies librarian at Emory University. He is the author of *The Cinema of Sergei Parajanov* (2013). His research specialties include Soviet cinema and film censorship.