

The story of Hein's dramatic transformation in response to the coming of the electronic age

By Joe Gerken

In December 1995, Kevin Marmion, then vice president of William S. Hein & Co., drafted a three-page memo titled "Electronic Law Review Project." The memo succinctly spelled out the potential benefits of an online law review library—to subscribers, to the law reviews, and to Hein. It also sketched, in broad terms, how the product would work. This description was quite prescient with regard to the necessary components of such a system, as well as pricing and royalty calculations that might be employed. A creature of its time, the Marmion memo envisioned monthly delivery of content via CD Roms while speculating, "If on-line services were used [delivery] could be done daily."

Four and a half years later, in May 2000, Hein went online. The ensuing

13-plus years have witnessed a truly remarkable growth in both the contents and customer base of HeinOnline. This is the story of the invention of this groundbreaking product, from first conception to implementation to its



Kevin Marmion

present place as a leader in online legal literature. This story is based on interviews with Marmion, now president of Hein, Hein Senior Vice President Daniel Rosati, and Rich Marisa, a computer expert at Cornell University.

The story can actually be told in two ways. On the one hand, it is the story of the dramatic transformation of a traditional print legal publishing company in response to the coming of the electronic age. However, HeinOnline can also be seen as an outgrowth of the strategies and values that have always driven Hein

publishing—in other words, as the logical continuation of the Hein tradition. The interaction of these two narratives is what makes this such a fascinating story.

Why HeinOnline?

HeinOnline came about because of a unique concatenation of demand and potential. Because of their close relationship with academic law librarians, people at Hein were acutely aware that legal publishing, and hence research, was quickly moving from a reliance solely on print sources to include electronic sources, as well. In particular, the emerging dominance of Lexis and Westlaw demonstrated the utility of online access.

Many of the law librarians that Hein staff spoke with had expressed concern over burgeoning shelf space devoted to law reviews. There was a time when most

law schools published only one law review. Over time, more specialized law reviews focusing on issues such as civil rights or environmental law were published, as were interdisciplinary journals. With this dramatic expansion in titles, librarians were confronted with serious space issues.

The early 1990s saw a series of technological developments that, taken together, made a product like HeinOnline at least theoretically feasible. Four innovations in particular were significant: (1) scanning software to convert print text to an electronic image, (2) increasing computer storage capacity, (3) programs that allowed for collecting, sorting, and retrieving discrete documents, and (4) the internet. The potential was there to create an online law journal library, though a good deal of innovation would likely be required.

Hein was uniquely placed to build a digital law journal library. Upon purchasing Fred Rothman Co. in 1998, Hein had become the largest American distributor of print law journals, with more than 700 titles. There were also intangibles. Since its inception in 1961, Hein had focused on law libraries, particularly academic law libraries, as its primary customer base. That longstanding tradition has been premised on librarians' trust in the quality of Hein's products. Hein also had a longstanding practice of listening to its customers. Indeed, as discussed momentarily, many of Hein's most significant innovations had been initially suggested by librarians. Having its "ear to the ground" proved to be critical to Hein's decision to develop HeinOnline. The company had also built solid relationships with law reviews. Although editorship of these journals changed with each school year, the editors were quite aware of Hein's tradition of treating law reviews equitably and of being responsive to their needs.

Developing the Technology

By the time of Marmion's memo in late 1995, it was evident that an online law journal database was feasible. One key technological component was scanning technology. In the early 1990s, Xerox developed Docutech, technology that created electronic facsimiles of print pages. Hein first obtained Docutech in 1994 to enable it to produce short runs of print journals and books. Previously, it was necessary to run off hundreds of copies of an item to make it economically feasible. With Docutech the pages were stored electronically so that a print run of even one copy was feasible. This capability—of converting print to electronic pages—was, of course, central to the notion of an online library.

From the beginning, Hein was

committed to offering an image-based product rather than one that relied on inputted electronic characters. To Marmion, this was a "no brainer." While the cost of an image-based system was estimated at about 25 cents per page, the cost of an inputted system was estimated at \$2 to \$3 per page. Those costs would have to be passed on to customers, moving the subscription price of HeinOnline beyond the reach of many libraries. Also, Hein had access to the print contents of more than 700 law reviews, making scanning of the images of pages much more economical. Indeed, many journals had already been scanned for printing purposes using Docutech.

Of course, another essential component was the internet. While the internet had been around for decades, two significant developments had made an online library feasible. For one, speed of transmission had been increasing. A number of key developments increased bandwidth exponentially, notably, the use of fiber optics. Ultimately, the development of wireless fidelity, or wi-fi, would take internet transmission to a new level.

The other significant development was the creation of the World Wide Web (WWW), which enabled greater connectivity of websites. In 1995, WWW was in its infancy, with comparatively few contributing websites. During the five years of HeinOnline's development, WWW grew to be a dominant mode of communication. It was also during this period that increasingly sophisticated web browsers came on the market.

Cornell

Although an online law journal library appeared to be technologically feasible, it was also evident that a good deal of innovation would be required to make such a library functional. In particular, Hein needed a system to organize the articles and make them readily retrievable. A user-friendly interface was important. Hein also was insistent that the system not require idiosyncratic equipment—libraries should be able to access and use HeinOnline with their existing computers. Quality of the image was also important: Hein wanted to replicate as perfectly as possible the look of the print page, both in the onscreen image and on printouts. Last but not least, speed of communication between the patron and the database was essential.

In early 1996, Hein officials took a "tour" of numerous universities to see and discuss digital libraries that they had created. The tour took them to Yale, SUNY Binghamton, Michigan, and Cornell, as well as the Library of

Cornell. At that time much of the innovation was occurring at universities. From the beginning, Marmion and other Hein officials were impressed with what they saw at Cornell. Together with the University of Michigan, Cornell was bringing up the Making of America collection. The computer experts at Cornell seemed knowledgeable and excited about discovering new ways to benefit the research community with advanced technology. Clair Germaine, law library director at Cornell, had affirmatively contacted Hein when she first heard about the online journal library idea. Sarah Thomas, university librarian, was also enthused about the idea, as was Steve Worona, assistant vice president for information technology.

Both sides stood to benefit from a cooperative agreement. Hein would, of course, obtain from Cornell the technological ability to run HeinOnline.

Cornell would get funding to develop that technology, which would have applications that would benefit scholarship and the university. Worona asked Rich Marisa, an IT specialist at Cornell, to take the lead on the project. Marisa had followed the digital library work going on in Cornell's Computer Science



Rich Marisa

Department and the Cornell Library and was enthused about the Hein project.

Reflecting back on the developmental years, Marisa stressed that cooperation between staff at Cornell and Hein was exemplary. The parties shared a "nexus of interest"—they very much wanted to see this project succeed and envisioned it as a significant contribution to scholarly research.

Following an initial grant for exploratory research on a user interface, which included some testing by law students at Cornell, the relationship was sustained by agreements, the terms of which tracked the changes in technology used to host the collection.

As the HeinOnline project was beginning, researchers in the Cornell Computer Science Department had developed an open-source architecture called Dienst for structuring and serving electronic documents. Cornell IT produced an implementation of Dienst targeted to law journals for this project. In applying this program to HeinOnline, Marisa worked with the Hein team to determine the metadata elements that would be needed to categorize and retrieve items and to write software to efficiently capture that metadata as articles were added. Dienst was used to define sets of pages as unique articles, group those articles into journal issues and volumes, and expeditiously identify the articles that patrons needed. It also facilitated the addition of new libraries

with a variety of content. While the software in HeinOnline's system developed over time, the Dienst document model remains recognizable in the current implementation of the system.

Image quality was a critical factor in system design. Hein was insistent that the page as displayed on the computer monitor and as printed out should be "identical" to the page in the print journal. A limiting factor was the computers and printers used by most law libraries at the time. Ideally, 300 dots per inch (for text) or 600 dots per inch (for black and white images) would be needed to recreate a sufficient print image. However, in the late 1990s, most library monitors had much lower resolution—in the range of 72 dots per inch. And many customers' printers at that time offered much lower resolution or had limited memory, making high-resolution printing difficult. Marisa's approach was to create lower-resolution images "on the fly" for different devices and manipulate grey levels to maximize image quality; he characterized this approach as "a little artistic." Over time, as the quality of monitors and printers available to libraries improved, this issue was much simplified.

As the library grew, navigation became more challenging. Marisa created a program that would create live tables of contents for journal issues and volumes with links to each article. This significantly simplified browsing for the patron. The availability of optical character recognition (OCR) software enabled HeinOnline to implement a number of enhancements. For example, when a patron accessed an article, he or she could place a cursor over a citation recognized in the image text and click to link to articles elsewhere in the HeinOnline collection. Once the Hein team created a database of such links, it was possible to create an "articles that cite this page" feature. This enhancement was of value to researchers as a way of identifying related articles on a topic. It was also welcome among authors wanting to see how influential their writings had been. Later, when the Congressional Record library was added to HeinOnline, character recognition was employed to create a table for converting pagination from the daily edition to the permanent edition and vice versa.

Searching became more important as the collection grew, as well. Early full-text indexes, first for volumes and then collections, were custom built at Cornell, and, as technology improved, the search strategy was replaced several times. The Hein team now generates a full-text index for the entire 100-million-page collection for each month's content update using open-source tools.

Cornell hosted the HeinOnline data at its server farm. Initially, the process for loading the data scanned at Hein was a bit cumbersome. Once a week, Federal Express would deliver a box of CD Roms with the articles in compressed format. Roughly four volumes of a law review would fit on each CD. Marisa would take each CD, decompress the contents, and load it on the server. As scanning ramped up and portable disks became available, the CDs were replaced with 250 GB-discs with roughly 80 times the capacity. Eventually, increased bandwidth and lower cost made it optimal to transfer the bulk content via the internet.

Economies

As time went on, computer-related costs for operating HeinOnline fell significantly. Data storage costs dropped as computers with significantly higher storage capacity were developed. With increased bandwidth capacity, the cost of data transmission also fell. In addition, scanning technology improved, as did OCR technology, allowing for much higher rates of converting print pages to electronic format. Thus, the per-page cost of scanning was significantly reduced. In part because of these reduced costs, Hein was able to dramatically expand the contents of HeinOnline, adding numerous new libraries without significantly increasing its subscription price.

Hein (Re)Defines the Content of its Law Journal Library

From the beginning, the idea was to start small and learn by doing. Thus, Hein planned an initial offering of select law reviews. In his 1995 memo, Marmion suggested that HeinOnline begin with "second and third" tier law reviews. He reasoned that "they are the most expensive [for libraries] per issue/page cost" since they do not have the readership of the top journals.

However, during the planning stage, participants contended that Hein should start with the "best" journals. This was based in large part on the expressed preferences of law librarians who had discussed the online journal concept. This latter approach carried the day. Thus, the first offering consisted of the "top" 25 law reviews. It took some time to decide which journals were the "best," drawing from a number of sources for this assessment, including an article by Olaui Maru, "Measuring the Impact of Legal Periodicals," in the *American Bar Foundation Research Journal*. Circulation figures were also consulted. Once the 25 titles were selected, it took almost five years to get them to all agree to participate in HeinOnline. That five-year period happened to coincide with the

time it took to work out all of the technical issues.

Almost immediately after HeinOnline went up, an effort was made to expand the list of contributing journals. The library grew at a rapid rate. By February 2006, more than 800 journal titles were offered. By 2007, every "flagship" law review published by an accredited U.S. law school was in the collection.

At its inception, HeinOnline was envisioned as a source for older volumes of law reviews. Indeed, initially, it was colloquially called the "old law journal project." The plan was to go from volume one up to the 1980s, i.e., the years when Lexis and Westlaw began coverage of most journals. This decision was based on the assumption that law libraries likely did not want to pay for duplicate coverage of the same content. Also, law review editors were much more willing to agree to having their older content on HeinOnline; however, they were concerned that giving Hein access to more recent volumes would diminish their worth to Lexis and Westlaw.

It soon became clear that law librarians did not view HeinOnline as merely duplicating Lexis and Westlaw's versions of the journals. Professors and other serious researchers would request "print" copies of articles since they were more readable and had supporting footnotes on the same page as the text. Also, virtually every law review required that its cite checkers go to the print version of articles. Initially, this meant that cite checkers had to literally go to the print volumes and photocopy pages. However, editors soon realized that a printout of an article from HeinOnline was, for all practical purposes, a "print" version. The bottom line is that within a few years of its inception, librarians were clamoring for HeinOnline to include journal content up to the most recent issues.

Initially, many law review editors were reluctant to cede HeinOnline their more recent content. However, some factors worked in Hein's favor. Hein had a long history of working with law journal editors who had come to trust and respect the company. Also, HeinOnline quickly earned a reputation as a reliable and much-utilized source for the journals. Editors also learned that putting their content online had little if any impact on print subscription numbers or on Lexis and Westlaw royalties.

Journals also came to appreciate and rely on the royalties that Hein paid. Hein has had a consistent formula for calculating those royalties, which is based on either hits or pages, whichever yields the higher number. In addition, journals were guaranteed a minimum royalty

payment even if no one printed or downloaded articles. Common sense argued that readers' use of any journal, and hence royalties, would increase appreciably if more recent volumes were posted.

This negotiation had to be accomplished one journal at a time. As some of the more prestigious journals agreed to give Hein recent content, other journals decided to follow suit. Also, each generation of law review editors was a bit more used to relying on online sources for legal information. Sometimes, a compromise was reached, with journals holding back the most recent year or more. Hein was flexible with these arrangements on the theory that as editors saw that adding recent content benefitted them, they would shorten the lead time for posting material on HeinOnline.

Marmion observed that there has been a significant shift in law journal priorities in the years since HeinOnline was introduced. Where previously there was a focus on "selling" the journal to customers (i.e., on subscription numbers), more recently, the emphasis has been on "selling" the journal to potential contributors. One way of doing so is to show authors that their work will be read and cited more if they contribute to a particular journal. No doubt, Fred Shapiro's family of articles on "most cited" articles, authors, and journals contributed to this shift of emphasis. Also, many schools have come to do analysis of citation statistics as an element in hiring and tenure decisions. Cumulatively, this shift of emphasis has benefitted HeinOnline since, clearly, access to articles is facilitated and readership is thus increased by posting articles on it.

Addition of New Libraries

Once it became apparent that HeinOnline was going to have a viable subscription base, it began considering other libraries to add. In September 2002, the first offering of the Federal Register went up, covering 1970 to 1974. Soon thereafter, the full run was available. The Supreme Court Library, including United States Reports, was also added during this period. Initially, Marmion was skeptical that this set would be used much given the availability of Supreme Court decisions on Lexis and Westlaw. However, Laura Gasaway, then director of the University of North Carolina Law Library, pointed out that the text of those decisions did not always correspond to the full text in the print U.S. Reports. As it turns out, this library has been quite well received.

In 2003, the Treaties and Agreements Library was added, including the invaluable resources developed by Igor

Kavass. These online sources significantly streamlined treaty research. Other libraries followed, including the Legal Classics Library and the Historical Code of Federal Regulations, both in 2005, and the Statutes at Large and U.S. Presidential Libraries in 2006. In June 2007, HeinOnline began adding the full contents of the Congressional Record and its predecessors, Annals of Congress and the Congressional Globe. Shannon Hein, vice president for sales, was the first to suggest this truly monumental enterprise. The sheer volume of this potential library was initially daunting, but Hein decided to go through with it. The library has also met with a very positive reception. Librarians have particularly appreciated the feature that enables them to convert citations from the Daily Edition to the Permanent Bound Edition and vice versa. Phil Berwick, director at Washington University Law Library in St. Louis, initially suggested this feature, which turned out to be feasible using OCR software.

One indication of the growth of HeinOnline libraries is the number of pages available on the database: 13 million in September 2004, 33 million in December 2007, 70 million in July 2011, and 100 million in November 2013. As mentioned above, Hein has been able to add this content without an appreciable increase in subscription price in part because the unit cost of creating and maintaining its online libraries has significantly decreased.

Testing, Roll-Out, and Growth of the Subscriber Base

Hein relied heavily on feedback from potential users during the years when HeinOnline was under development. Students taking Advanced Legal Research at Cornell Law School had an opportunity to try out HeinOnline during its developmental period. Senior Vice President Dan Rosati and Executive Vice President Dick Spinelli went on a "road tour" of numerous academic law libraries, using a laptop computer to demonstrate the product. Hein representatives also met with librarians and other potential customers during the annual meetings of AALL, the American Association of Law Schools, and New England Law Libraries. These contacts yielded numerous constructive suggestions regarding both the content of the online library and the functionality of the webpages. They also generated considerable buzz about the product. Thus, by the time HeinOnline went online, many librarians were at least

intrigued by the product and, in many instances, eager to try it.

HeinOnline went online in early 2000. On May 8, 2000, New York University Law Library became HeinOnline's first subscriber. Library Director Kathy Price told Spinelli and Rosati that she was eager to subscribe: "I want to be known as the first HeinOnline subscriber." Let the record so reflect. The venerable Jenkins Law Library in Philadelphia was the first nonacademic law library to subscribe, and Arnold and Porter was the first firm library.

A turning point came at the July 2000 AALL Annual Meeting. Hein set up a computer and projector at its booth in the Exhibit Hall. Rosati and Spinelli met with a steady stream of librarians, doing dozens of group demonstrations each day. By the end of the conference, they were exhausted but exhilarated. Hein subsequently won the 2001 AALL Best New Product Award. By March 2001, HeinOnline had 100 subscribers. This figure grew to 150 by June. By December 2001, 80 percent of ABA-accredited law schools had subscriptions. The subscriber base continued to grow steadily—by September 2005, HeinOnline had subscribers in more than 3,200 locations in 75 countries.

The Hein Way

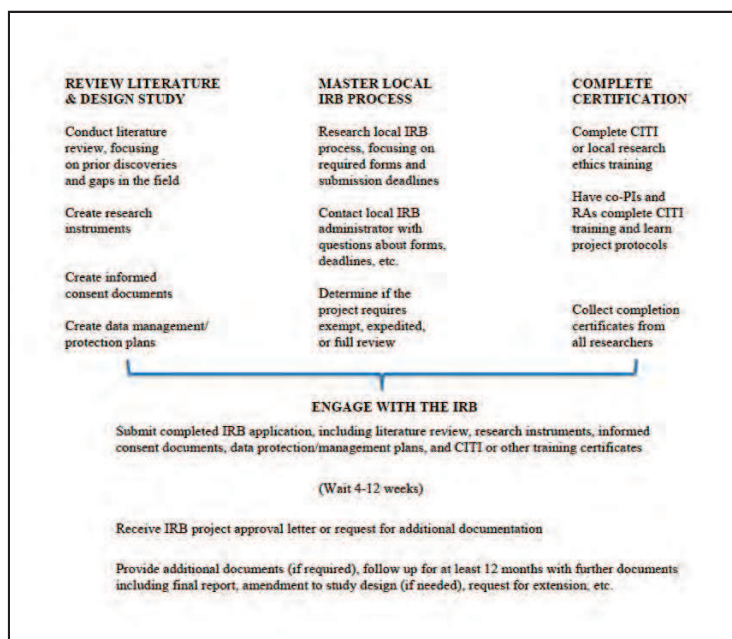
HeinOnline revolutionized law review distribution. The product also marked a dramatic reinvention of the Hein company. From a traditional print publisher, Hein, in a matter of years, became a publisher whose predominant format was electronic. In a sense, the mission of the company had been redefined. One the other hand, one might also say that the mission of Hein remained exactly the same—putting law journal literature in the hands of researchers.

The key components of HeinOnline's success derive from principles that guided the company since its inception. Prominent among these is "listen to your customers." In Hein's case, that has typically meant "listen to librarians." It was librarians who first convinced Marmion and Hein staff that the time had come for journals to go online. This was by no means the first significant idea that librarians had suggested to Hein. The company was the first legal publisher to print government documents on acid-free paper, an initiative that was sparked by librarian input. Hein's movement into microform publishing was also due in large part to librarian requests for material in that format.

From the time he formed the company in 1961, William Hein, Sr.,



Dan Rosati



institutional review boards exist to safeguard the dignity of human beings; (2) the IRB process is complex but manageable; (3) the process requires attention to detail and an investment of time; (4) the work of IRBs often follows a predictable pattern, and local IRB administrators can clarify institutional refinements and requirements; and (5) humane research fulfills the hope of the lawyers, judges, and policymakers who have invested in this issue for more than half a century.

For those who want to dig deeper into the history of research ethics in the U.S. and the creation of IRBs, I recommend Harriet A. Washington's *Medical Apartheid: The Dark History of Medical Experimentation on Black Americans from Colonial Times to the Present* and Susan Reverby's *Examining Tuskegee: The Infamous Syphilis Study and Its Legacy*. ■

Sarah E. Ryan

(sarah.ryan@yale.edu), Empirical Research Librarian and Lecturer in Legal Research, Yale Law School's Lillian Goldman Law Library, New Haven, Connecticut. The author would like to thank Scott Matheson for his helpful comments on a previous draft of this article.



hein online continued from page 20

acted as its chief sales agent. He regularly went on the road to meet academic law librarians face to face. Typically, he did more listening than talking, asking what issues the librarians were having, both in relation to Hein's products and regarding legal literature generally. When William Hein, Jr., took over the company, he continued this practice. Once Hein hired staff devoted specifically to sales, including Shannon Hein, Spinelli, and Steve Roses, their "marching orders" always included listening carefully to librarians and conveying their thoughts back to the company. Thus, it was not a fluke that Hein was quick to pick up on the likely demand for online law journals.

Two related components in Hein's philosophy have been "be open to innovation" and "do not over-extend commitments." The development of HeinOnline was both visionary and pragmatic. It was visionary to appreciate that the moment had come when it was technologically possible to build an online law review library. However, the company was quite pragmatic in each stage of creating HeinOnline. Hein has historically employed a conservative fiscal policy, avoiding unnecessary debt. Hein's print and microfilm publishing operations were still quite profitable, and it had sufficient assets so that it could devote a portion of the profits to develop HeinOnline rather than having to borrow heavily. This enabled Hein to be patient in ensuring the components of

HeinOnline functioned well, since creditors were not looking over its shoulder.

Hein's pragmatism was also reflected in its product development strategy. From the beginning, the plan was to start small, see what worked and what did not, and adjust accordingly—thus the switch from starting with the lower-echelon journals to the top 25, as well as the decision to add volumes up to the present. The choice of additional libraries was also made carefully upon consultation with the customer base.

The decision to use image-based software was also a reflection of Hein's conservative planning. Initially, many librarians asked Marmion, "Why can't we search HeinOnline the way we search Lexis and Westlaw?" He would patiently explain that doing so would have made the product unaffordable. He liked to use a car analogy: "Why are you driving a Chevy and not a BMW?" He would also explain that the real comparison was not with Lexis and Westlaw but with print journals. HeinOnline clearly offered key advantages when that comparison was made. Ultimately, librarians came to see it that way.

One more principle was critical: "Build and maintain trust among your constituencies." In Hein's case, there are two key constituencies—librarians, particularly academic law librarians, who are the core customer base, and law review editors, who provide the "product." For decades, during its

print-only years, Hein had fostered these relationships, responding promptly to complaints and concerns and taking quality control to an almost obsessive level. Librarians' trust in Hein, particularly in the quality of its products, was a key factor in the swift rise in HeinOnline's subscription base. Simply put, librarians expected HeinOnline to work well because it was a Hein product. Likewise, law review editors' trust of Hein greatly facilitated HeinOnline negotiations, particularly with regard to the touchy issue of posting more recent volumes.

The Reality of HeinOnline

There is a third way to tell this story. It is the tale of a small, family-operated company doing battle with the giants of the legal publishing world and carving out its unique niche. It would be easy to romanticize this story as a variation of *The Little Engine That Could*. The reality, though, is that HeinOnline is a consequence of years of hard labor, persistence, and creatively adjusting to setbacks and challenges—undramatic, old-fashioned, workaday values. ■



Joe Gerken

(gerken@buffalo.edu), Reference Librarian, Charles B. Sears Law Library, the University at Buffalo State University of New York