The Record of Business and the Future of Business History: Establishing a Public Interest in Private Business Records

David A. Kirsch

Abstract

Long-term preservation of and access to business records of the sort that business historians have relied upon for decades and that supported the development of applied fields, like business strategy, is increasingly threatened by their growing value. The expansion of entrepreneurial capitalism following the end of the Cold War has produced a new era of business ascendant across much of the globe. At home, we are reminded of President Calvin Coolidge's observation that "the chief business of the American people is business." Not since Coolidge first uttered these words more than eighty years ago have they rung more true. And if the business of America is business, then surely the history of America is the history of American business. Yet, if we have witnessed a new gilded age in American industry, the evidentiary record of these events may disappoint future scholars, policy makers, and the interested public.

The article is organized as follows: first, the problem is described. The section "Record of Business at Risk" explores these factors. The next section, "The Digital Archive of the Birth of the Dot Com Era" reviews some of the efforts that have been supported by the National Digital Information Infrastructure Preservation Program (NDIIPP) of the Library of Congress to develop the outlines of an effective response, including summaries of the major digital preservation projects in which I have participated. The conclusion describes possible future scenarios and the policy contexts that might determine eventual outcomes.

LIBRARY TRENDS, Vol. 57, No. 3, Winter 2009 ("The Library of Congress National Digital Information Infrastructure and Preservation Program," edited by Patricia Cruse and Beth Sandore), pp. 352–370

The Record of Business at Risk

Why are business records more at risk now than in the past? Information technology is implicated at every step, but technological change is not the only cause of the threat. The sources of the problem include entrepreneurship, litigiousness, and shareholder capitalism itself, each of which are indirectly affected by changes in the underlying technological landscape.

Increasing interest in and pursuit of entrepreneurship will lead to decreasing persistence of business records. The rise of entrepreneurship and its growing importance to the future of the American economy is well documented (Schramm, 2006; Baumol, Litan, and Schramm, 2007). As one indication of this trend, Shane (2003) reports that there are more businesses created in the United States than marriages. Even accounting for the supposed "Dot Com Bust" in the early part of the decade, Metrick (2007) notes that outlays by U.S. venture capitalist funds have increased approximately tenfold from 1991 to 2002 when twenty billion dollars was invested in venture capital-backed companies. Entrepreneurial ventures are, by definition, created de novo. At inception, no basic practices are inscribed in the firm, and such organizations are said by Stinchecombe to suffer from a "liability of newness" (1965). Records management is merely one of a variety of practices that are lacking in a new venture, and only if the venture survives and grows to maturity is the firm likely to develop policies in this area. Lacking formal research on this issue, anecdotal evidence suggests that records management is very far down the list of priorities of new entrepreneurs, and with good reason. The search for financing, customers, suppliers, employees, and, above all, profits, must come first. However, entrepreneurship research has shown that most new ventures fail within their first several years of existence (Shane, 2008), likely well before records management practices have been established. Little is known about the fate of the records of failed startups in general, but based upon the general research on business plans described later, many new ventures disappear without a trace. Within the subset of business plans we were able to preserve in the Business Plan Archive (http://www.businessplanarchive.org), there was a significant fraction for which we were unable to find any other record, digital or otherwise, beyond the business plan: there were no stored pages in the Wayback Machine (www.archive.org), no identifiable public records, and no mention of the firm in searchable media. And even firms that introduced important innovations before ultimately failing (and will therefore leave a small historical footprint) may warrant additional preservation efforts. For the moment, the recent increase in entrepreneurship has not been accompanied by a concomitant increase in resources dedicated to the preservation of the records of these firms.

The rise of shareholder capitalism also portends the loss of business records. At first pass, one may wonder how the shareholder revolution and the persistence of business records are connected, but the relationship is straightforward and revolves around the existence of "a public interest in private records." Following the model of a traditional business archive like the Hagley Museum in Delaware or the Baker Library at Harvard, scholars would access business records that had been discarded or deaccessioned by the original records producers (General Motors, DuPont, etc.). To the extent that records from these individual businesses were used to produce generalizable findings about the workings of business, scholars extracted residual public value from private records. The original owners of the records—that is, the shareholders of General Motors and DuPont—did not receive compensation for the full value that their records created. In fact, it is quite likely that the shareholders were not even aware that private records were being put to a public purpose. Often the transfer of traditional records occurred decades after their production when corporate records managers and business archivists would agree to deaccession collections that had, by then, acquired evident historic value. In this traditional model, paper business records were inadvertently or accidentally preserved inter alia within the confines of the producing organization and then handed over to independent archivists when it became clear that a public interest existed and outweighed any residual private value inhering in the documents. The shareholder revolution, heralded as an essential advance for common stockholders, lays bare the contradiction inherent in the traditional model. If a proposed action does not directly benefit shareholders, then it should not be supported by responsible management. Because the private benefits of contributing to an archive are questionable at best, we can expect that in the future fewer organizations will donate business records.

Finally, American litigiousness also plays an important part in this saga. The costs of saving business records come in two flavors. The first is plain vanilla and includes such predictable expenses as storage, curation, and migration (for digital formats). To borrow from Harry Potter, these costs—while not inconsequential—pale in comparison to the Bertie Botts of legal liability that are mixed into the vanilla. The costs associated with these surprises are unknown and therefore especially frightening to a corporate world desperate to predict and thereby control every aspect of economic activity taking place within the boundaries of the firm. From a statistical viewpoint, reasonable observers agree that the liability risk associated with any given record is vanishingly small. However, if a corporation did something wrong, inadvertently or otherwise, and the act was documented in a surviving business record, there might exist a ticking time bomb of legal liability, a toe jam flavored Bertie Bott Any-Flavor Bean buried in the Record of Business. If such a record exists, we know four things. First, its existence and contents would be of great interest to aggrieved parties, potentially including the government. Second, the cost of scouring the Record of Business to find the specific offending record would be high. Third, were the firm to know of the existence of such a record, that fact would compel the firm to disclose the wrongdoing or compound its crime by concealing or actively destroying a business record known to be sought in discovery. Fourth, historians and other scholars would also have an interest in knowing of the existence and contents of the record, as part of the landscape of facts onto which historical agency is mapped. For almost any conceivable constellation of facts, the potential private costs associated with factors one through three swamp any potential public benefit that might arise from factor four. The benefit accrues to society in the form of hard-to-measure historical, cultural knowledge, while the costs—potential and real—are incurred entirely by the firm and its shareholders. Under these circumstances what results is a set of policies that have come to be known as document retention policies, policies that define different categories of records and how long each document type is expected to be preserved. Even proponents of these policies admit this is a misnomer if ever there was one. The reality is that document retention policies dictate timetables for records destruction, and their sweeping scope and breadth insure that the company cannot be charged with arbitrarily seeking to destroy any specific record or conceal any specific act. The specific treatment of, say, employment records relative to tax records, or administrative records relative to accounting records, will vary from firm to firm and state to state. But their purpose is clear: every record produced within the boundaries of the organization should be saved until a certain date and then destroyed. These policies are the proximate causes of the threat to the Record of Business.

Far from criticizing any firm or individual for implementing document destruction policies of the type described above, I marvel that corporate America has been so slow to adopt them. For decades, scholars benefited from the fact that organizations did not really know what was happening to their records. The Record of Business is the repository of organizational memory, the seat from which all-important organizational knowledge is managed. Though no one in a position of authority wishes ill to history, neither can a decision to spend scarce resources for the benefit of a non-shareholder be squared with fiduciary requirements. Because history is not a recognized shareholder, there is no mechanism by which the interests of history can be factored into firm-level decision making. History, like Tennessee Williams' Blanche DuBois, has always depended upon the kindness of strangers. Unfortunately, such kindness must now take place in broad daylight and in full view of the shareholders whose checkbooks pay the costs. Today, the CEO, general counsel, or other corporate officer who fails to destroy records the moment after they cease to be worth more to the corporation than they cost to maintain (or one moment longer than set forth in the company's document retention schedule) will be doing so in explicit violation of that officer's fiduciary obligation to shareholders and should be replaced with someone who will. Business records have survived by a wink-and-a-nod arrangement. Firms maintained their records during the initial, primary-use phase of the information lifecycle, and then variously retained them under more or less active managerial regimes. Some eventually made their way into an archive; others did not. Serendipity determined the fate of any given item. The effect of the forces described above is that this informal equilibrium is no longer stable. The checks that had previously moderated the risk to business records are being eroded, and the emerging records management regime will ruthlessly reward firms for leaving nothing to future scholars.

Each of these factors—entrepreneurship, shareholder capitalism, and the culture of litigiousness—is clearly connected to the transformed technological landscape. Information technology has both created entrepreneurial opportunities and reshaped the processes of economic coordination by which such opportunities are exploited (Lamoreaux, Raff, and Temin, 2003). From Web casts of quarterly earnings reports and the U.S. Securities and Exchange Commission's EDGAR online database to the rise of the Motley Fool and online stock trading, shareholders have more access to information about how firms deploy resources than ever before, thereby furthering the advance of the shareholder revolution. And technology has increased the efficiency with which all professional services are provided, even the plaintiff's bar.

But technological change contributes to the problem in other ways. The digital revolution has raised the stakes, increasing the value of effective records management (a.k.a. knowledge management) and extending the primary-use phase of the information lifecycle, while simultaneously drawing attention to the risks associated with records that outlive their value to the firm. In the sanitized parlance of contemporary techno-utopia, "friction free" capitalism will no longer leave a trail of dust and crumbs for future generations to interpret. Instead the frictions and inefficiencies that produced these traces will be obviated, the incidental copies tracked along with every other valued economic output and eventually destroyed. Digitization simply lays bare the contradictions that allowed the Record of Business to be preserved in the past. The fact that a single sheet of paper attesting to a potentially embarrassing event committed by a wayward firm has ever survived is testament to the failure of corporate records managers to efficiently manage the archived burden of past generations. With paper—where the default preservation setting is, in the medium term, retain—the Record of Business survived until it was lost, lost to the records managers that supervised it, lost to the office clerks that ignored the "save until" date on the paper file, lost to the CEO who could not be bothered with worrying about fiduciary obligations extending down to the file room. For every record that was "lost" to the corporation, there was a non-zero statistical probability that the record would be "found" for history, appearing at some future date in a curated, archival collection. Without a numerical estimate of that probability, I propose that the effect of digitization is to reduce it by several orders of magnitude; thus, the Record of Business that will be lost to business and found to history will be vanishingly small for the very reason that the technology which made the knowledge encoded in the record valuable enables its corporate creators to track its progress through the administrative digestive tract of the firm. The fact that instances of information loss have posed such visible and costly challenges to corporations and other organizations (see, for instance, Fidelity Investments [Wong, 2006], the U.S. Veterans Administration (Greenemeier, 2006), and AOL (Hafner, 2006)) underscores the value attached to records of business, and items of value will not be misplaced, lost, or otherwise overlooked.

Thus, we can now state the problem with some clarity: In the past, business records survived inadvertently. The digital revolution, consistent with changes in the structure of the economy and the clearest dictates of shareholder capitalism, will destroy the Record of Business not by accident, but on purpose by making it manageable and valuable for a discrete, but limited period of time beyond which its value to the corporation falls below the cost to maintain it. On that day the Record of Business will be erased.

The Digital Archive of the Birth of the Dot Com Era

This section describes the efforts that we have undertaken, in partnership with the Library of Congress, the Alfred P. Sloan Foundation, and other NDIIPP participants, to address the issues identified above. Believing that the recent past offered both a fertile moment in the history of business and a bonanza of potentially valuable "at risk" digital business records, we focused our efforts on the creation of a new repository, the Digital Archive of the Birth of the Dot Com Era (hereafter shortened to the Dot Com Archive or DCA). Established in June 2002 in the depths of the post-bubble technology recession, the Digital Archive of the Birth of the Dot Com Era set out to identify, collect, and preserve a representative collection of born-digital business records and related digital ephemera from companies that sought to exploit the commercialization of the Internet during the 1990s. The DCA includes several distinct yet overlapping collections intended to provide future scholars a broad, representative picture of technology entrepreneurship in the Dot Com Era. It also includes several company-specific collections that will allow scholars to immerse themselves in the messy details of individual technology ventures. In addition to describing the collections, we include summaries of initial research findings, accounts of the processes by which the collections came into the DCA, and some of the challenges attendant thereto.

Business Plan Archive

The Business Plan Archive (BPA) was the first effort in this area and contains business planning documents and information from more than three thousand Dot Com Era-technology companies. The Business Plan Archive accepts contributions from individuals and organizations, both through the website and by traditional means. In general the collection is organized by firm. Research establishing the overall size of the target population suggests that the BPA will likely constitute the largest such collection from this period (Kirsch, 2007; Kirsch & Goldfarb, 2008). Capturing the language, look, and feel of these early Internet ventures, the BPA offers a valuable research context for scholars interested in exploring this period.

Contributions to the Business Plan Archive have come from two sources: individual entrepreneurs and private equity investors. Individual entrepreneurs have tended to contribute planning documents relating to the firm or firms they started, while investors have contributed larger batches of documents, subcollections ranging in size from tens to hundreds and even thousands of records at a time. Because our objective was to look for the "blueprints" of Dot Com Era companies, business planning documents are defined broadly as any material that helps understand the assumptions, strategies, and tactics that underlie a startup. These materials might include: brief executive summaries or PowerPoint presentations that entrepreneurs typically distributed widely to potential funders or partners; full, detailed business plans; spreadsheets or other financial documents that were distributed with the business plan; and related memos, e-mails, screenshots, or published articles that shed light on startup strategies.

To address possible bias in the sample arising from the open collection methodology, substantial effort has been made to establish the representativeness of the holdings. Focusing on specific subcollections submitted by single large investors, we demonstrated that the firms in our sample did not differ significantly from the overall population of firms seeking funding during and immediately following the Dot Com Era (Goldfarb, Kirsch, & Pfarrer, 2005). Exploiting the initial value of the records for social science rather than historical inquiry, research papers contributing to entrepreneurship, management, and finance have been published or are in press. Future users of the BPA will be able to refer to these early works as quasi-finding aids for the larger collection, and the expanded information about the BPA firms collected for these papers (i.e., outcome, entry and exit date, nature and extent of outside investors) enhances the ability of subsequent users to make additional contributions. Reporting on a 2007 paper in the *Journal of Financial Economics* (Goldfarb, Kirsch and Miller, 2007), the *Wall Street Journal* referred to the Digital Archive as "Smithsonian Institution of the Internet Bubble," an accurate, if perhaps optimistic analogy (Gomes, 2006).

Turning briefly to the contents of the Business Plan Archive, several observations are in order. First, the diversity of types of materials contributed was surprising. Based on prior research on venture capital funding of technology startups, our initial assumptions were that business plans are relatively stable artifacts with highly conserved structural elements and that every entrepreneur prepares and submits a similarly structured request when seeking funding. Analysis of the funding requests revealed much more variety: less than 40 percent of the companies submitted what we would call a formal business plan, and the plans themselves demonstrated considerably variability. The average plan contained less than two-thirds of the elements prescribed in typical entrepreneurship textbooks, and almost half did not explicitly indicate the amount of funding they were seeking, something we assume would be relatively important to a potential investor. Second, after exhaustive study, we were able to find little connection between the objective content of a business plan and the outcome of the proposed venture. In other words, it does not appear that good business plans lead to more successful ventures in any direct way that we could measure. Instead, we found that social ties more accurately predicted successful acquisition of venture funding, suggesting that at the margin, entrepreneurs should focus on building their social networks with potential investors rather than perfecting their written business plan. These findings—reported in a forthcoming paper in *Strategic Management Journal* (Kirsch, Golfarb, and Gera, forthcoming)—have far-reaching pedagogical implications for the way entrepreneurship is taught. Continuing research aims to extend these findings to look more broadly at the types of signals entrepreneurs send to potential investors (i.e., beyond simply submitting a business plan or reaching out through social networks).

Terms of access to the collection have varied. Initially, the site was open to the public, with donors able to control access rights themselves. It was hoped that the collection would develop on its own as more and more former Dot Com-Era entrepreneurs and investors learned of the opportunity to participate in the archival initiative. This Web 2.0 vision—first labeled "Open Source History" to reflect an inclusive, expansive notion akin to public history—has not proven successful. Of more than 110,000 registered users of the Business Plan Archive website, no more than 100 individuals have contributed to the collections, and the largest contributors have requested anonymity, thereby further confounding our initial open source vision. More than 4,600 users have registered from ".edu" addresses, indicating adoption within the target user community, and many faculty at schools where access to real business planning documents is limited have used the site to introduce students to this topic. It should be noted, however, that while the site was publicly available (from June 2002—December 2007), the vast majority of users consulted it in search of a model business plan. This use, though not entirely unexpected, was not part of the original vision for the site. In late 2007, we were contacted by a company whose business plan had been inadvertently distributed beyond the BPA to other public sites for this purpose. When the company threatened legal action, university counsel recommended that we limit further public access, and since then, only approved researchers have been allowed to use the archive records. Approved researchers are required to submit a research proposal that is reviewed by archive staff and sign an appropriate non-disclosure agreement. No legitimate requests have been turned down though less than fifty such requests have been received. The real risks attending unintended disclosures are limited (see discussion below); nevertheless, program staff have generally concurred with counsel to err on the side of caution.

Brobeck Closed Archive

The opportunities and legal entanglements associated with the Business Plan Archive pale by comparison to the next important collection in the digital archive: the Brobeck Closed Archive (hereafter, Brobeck Archive or Closed Archive).

Brobeck, Phleger, & Harrison was a San Francisco law firm founded in 1926. By 2001, it had become regarded as one of the top two Silicon Valley firms representing newly emerging dot com ventures, as well as established corporations such as Cisco, Sun Microsystems, and Nokia. Brobeck expanded heavily during this period, serving thousands of technology clients and incurring upward of $100 million in debt that (among other causes) would result in its dissolution in February 2003 and subsequent filing for bankruptcy protection (September 2007). Over the course of its seventy-seven years, Brobeck amassed a fortune of historical client records, both in paper and, later, in digital format. Upon Brobeck's dissolution, the firm's Liquidation Committee asked a third-party vendor (Gallivan, Gallivan, & O'Melia, a leading electronic discovery consulting firm) to make a digital backup of the firm's entire network, which former Brobeck partners and clients then used to retrieve their records. The backup contains a diverse range of client data, as well as information related strictly to the administration of the partnership—minutes of partner meetings, operating and financial information about the firm, billing and accounting records, and other digital ephemera.

Brobeck represented several thousand technology companies during the 1990s. Given the focus of the larger digital archive on the activities of firms founded in this period to commercialize the Internet, we believed that the Brobeck corpus would include the records of many hundreds of target firms and set out to explore the possibility of preserving the Brobeck records as part of the Digital Archive of the Birth of the Dot Com Era.

When first contacted about the Brobeck Archive, many casual observers dismissed the effort as a lost cause. How, they wondered, could we save materials that retained legal privilege and confidentiality without abrogating those rights? Some went further: Even if we could save some of the Brobeck records, what good would they do? To whom would the records be of any legitimate interest? Indeed, our initial approach was agnostic on these questions. Far from being convinced that preserving the records of the failed law firm was necessarily a good idea, we embarked on the first phase of the project with an experimental orientation: What should happen to entrained records when a failing law firm represented hundreds if not thousands of failed clients? Who is the party responsible for managing records when the responsible party, the law firm itself, fails?

While the Business Plan Archive was initially conceived and implemented at the University of Maryland, the Brobeck Archive required we engage the larger community of digital preservation, including the Center for History and New Media at George Mason University (http://chnm.gmu.edu), the National Digital Information Infrastructure and Preservation Program of the Library of Congress (http://www.digitalpreservation.gov), the Estate of the failed law firm, Gallivan, Gallivan, & O'Melia (http://www.digitalwarroom.com), and a blue-ribbon advisory council including experts in privilege, confidentiality, legal ethics, bankruptcy, venture law, and the management of data enclaves.

Within a year of launching the exploratory phase of the Brobeck effort, project staff concluded that the records could be saved, but access to and disclosure of the contents would need to be carefully controlled. Because the Supreme Court has held that privilege and confidentiality are perpetual, the permanent repository must preserve the underlying rights inhering in the legal records. In effect the existence of a permanent repository will maintain these rights by acting as if the firm had not failed. Most fundamentally, we believe that clients never had any reasonable expectation that their records would be destroyed—only that they would be shielded from public disclosure. In fact, most law firms (including Brobeck) have policies requiring client records to be kept indefinitely, even (and often especially) if clients request they be destroyed. Brobeck's dissolution interrupted this indefinite preservation, and our designated archive will respect confidentiality by preserving client records *in the same manner* as did Brobeck. Appointing an indefinite third-party repository for confidential materials is accepted practice in the medical community, where bankrupt private practices have named local hospitals as repositories for patient records. The result of this philosophy is a closed archive that preserves information but has no interest in providing access. Rather, the proposed closed archive will have arrangements with traditional archival institutions, to which it will forward documents no longer in need of confidentiality protections. Examples of such situations are (a) after receiving the consent of a former client; (b) after making a specific determination that a document was never confidential; (c) aggregated statistical data; or (d) after changes in the applicable confidentiality laws. Authenticity, provenance, and other basic archival issues are thus overlaid on an underlying set of client rights (confidentiality and privilege, not to mention basic property rights like copyright), third-party rights (privacy of contract), and obligations to observe the rules of ethics that govern legal practice. Archivists are already familiar with various manifestations of these issues; however, the Brobeck Archive, like the Business Plan Archive poses several novel challenges that continue to occupy project staff.

These challenges are reflected in the court order that officially sanctioned the creation of the Brobeck Closed Archive. On August 9, 2006, in recognition of the extraordinary efforts of the project team and the historic nature of the Brobeck records, Judge Dennis Montali of the U.S. Bankruptcy Court, Northern District of California, San Francisco Division, authorized the creation of a Closed Archive allowing a large subset of these records to be preserved at the direction of the Library of Congress.[1] The court approved a high-level set of principles—the "Closed Archive Methodology"—that establishes the guidelines under which the Brobeck Closed Archive will operate. Certain categories of records, which we have no expectation to ever have reason or right to access, will be separated from the collection and destroyed (see fig. 1). For finer grained distinctions among different types of client and law firm records, general bankruptcy practice was followed. We were required to privately and publicly notice all parties of the expected creation of the Closed Archive. However, rather than requiring former clients and other rights holders to "opt in" to an open archive, the methodology instead allows rights holders to "opt out" (at any time) but assumes that many clients will not be able to respond and therefore impounds all remaining materials in the Closed Archive and places strict limits on access to and disclosure of specific, attributable materials. This arrangement was necessitated by our desire to maintain the integrity of the overall corpus and the fact that so many of the rights holders had ceased to exist. Indeed, the project mailed out approximately 11,000 copies of the court notice; more than 3,400 (30 percent) were returned as "undeliverable," implying that this subset of former clients may *not* have received the intended notice. For the sake of comparison, had Judge Montali limited the collection to those clients who opted in, we would today have open access to the records of less than fifteen clients.

We also recognized that even if the bulk of the Brobeck Archive would need to remain off-limits to historians, potentially into perpetuity, the scale and breadth of the collection could support social science research to answer a host of interesting questions without requiring that specific, confidential information be disclosed. The closed archive methodology approved by Judge Montali expressly envisions such research under an access model similar to that employed by the U.S. Bureau of the Census' Research Data Centers. As with the Business Plan Archive, project

	Traditional Archive	Closed Archive	Destroyed
	Currently accessible under prevailing archival and legal rules	Currently confidential; potentially accessible in the future under prevailing archival and legal rules	Never accessible under prevailing archival and legal rules
Documents	Court-approved partnership documents Consenting clients' documents	Non-consenting clients' documents Non-court-approved partnership documents	Objecting clients' documents All personal clients' documents Sensitive employee documents
Databases	Database rows or e-mails relating to consenting clients or court-approved partnership	All databases	None
E-Mail		All e-mail	None
Network Share Drives	None	All network share drives	None

Figure 1. Closed Archive Classifications

staff is working to establish the representativeness of the Brobeck Archive and account for any systematic distortions arising from the pattern of opt outs received (see fig. 2). Current plans call for the creation of an experimental access site in 2008–9 to enable social science research. Under this model, access will be restricted to archivists or scholars who have signed strict non-disclosure agreements and whose proposals will have been vetted by our advisory council. Initially, access will take place in an on-site, non-networked, institutional setting, and only for enumerated purposes. Archivists administering the Closed Archive will log search queries and document retrievals to ensure that users are complying with the narrow boundaries of their approved access and will only allow aggregated or redacted data to leave the secure area. We believe that this solution is workable and balances the need to safeguard confidentiality and privilege while still permitting approved scholarly access.

The data in the Brobeck Archive fall into a variety of technical types, which we have condensed into four categories. Figure 3 presents a graphical overview of these categories:

1. *Managed Documents.* These are ordinary computer files (Word or Excel documents, etc.), which were centrally managed by Brobeck, such as attorney work product.
2. *Relational Databases.* Databases were used by Brobeck to track files and financial information such as billing records or client contact information.

Matter Type	# Matters	Total Billed Hours (>10)	Opted Out (>10)	%
General Corporate Rep	3582	926,694	68,442	7.39%
General litigation	2349	808,824	88,242	10.91%
Mergers & Acquisitions	**1774**	**682,618**	**54,495**	**7.98%**
Product liability litigation	1941	615,562	35,575	5.78%
Securities litigation	735	480,845	66,319	13.79%
Technology / IP Litigation	742	407,533	51,145	12.55%
Labor	2022	386,908	80,437	20.79%
Public offering / Company side	330	357,178	34,648	9.70%
Corporate IP	3313	348,112	60,531	17.39%
Venture Finance / Company Side	1240	261,830	9,885	3.78%
Public offering / Underwriter Side	620	227,589	57,012	25.05%

Interpretation: For Matter Type "Mergers & Acquisitions," there are 1774 total matters with more than 10 billed hours. Of these, clients representing 7.98% of the total billed hours have asked to "opt out" of Closed Archive.

Figure 2. Initial Impact of "Opt Out" on Overall Brobeck Corpus, by Matter Type

3. *Microsoft Outlook Data.* These objects are Personal Storage Table (PST) files containing employees' e-mail and calendars.
4. *Network Share Drives.* These are personal storage areas that allowed employees to store files that were not centrally managed by Brobeck.

The Brobeck records offer rare glimpses into the world of high-tech entrepreneurship that produced the Internet revolution. The collections also offer the opportunity to preserve sensitive, "at risk," born digital archival materials while preserving the privacy and confidentiality interests of the parties involved. Recently, multiple organizations have made unwelcome news for their failure to adequately protect confidential personal information and/or respect limits of personal privacy. The Brobeck corpus offers an opportunity to develop, test, and evaluate potential technical and institutional solutions necessary for the creation of a functioning closed archive. Though initial efforts have yielded promising results that do not risk disclosure of confidential information (Goldfarb et al., 2008), in general, the approved methodology must still be translated into operating guidelines and procedures that will govern the operation of the closed archive in years to come.

Dot Com Archive and Related Firm-level Collections

The third component of the Digital Archive of the Birth of the Dot Com Era includes additional holdings that capture the experiences of individual Dot Com-Era workers and the detailed records of individual firms

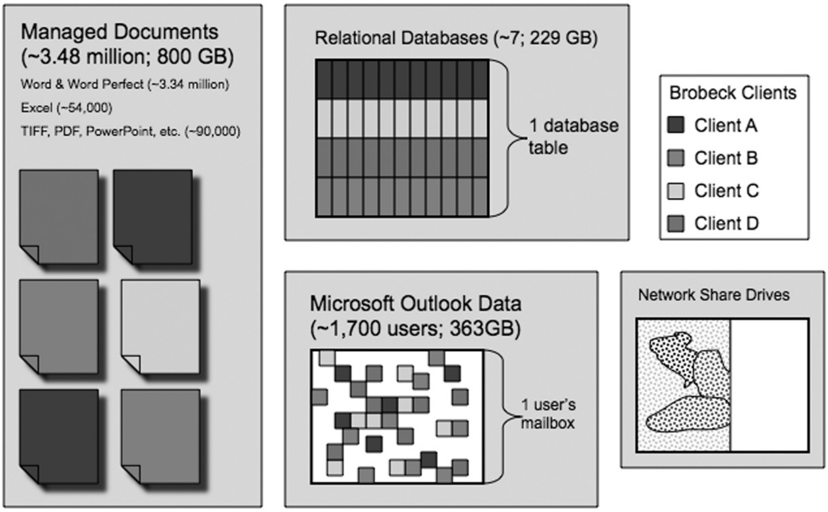

Figure 3. Categories of Records in the Brobeck Corpus Prior to Opt Out

whose founders (or their successors) donated collections to the archive. Firms whose records have been collected in this manner include Scient Corp., Broadband Sports, and OneSoft. These collections more closely approximate traditional business archives in as much as the preserved materials reflect the position(s) of the donor(s) and other collection-specific biases. Significantly, all of the major parts of this collection come from failed firms, a seeming fluke that may portend greater changes in the archival landscape that are discussed below.

Taken together the Digital Archive of the Birth of the Dot Com Era includes more than 6.4 million characterized digital objects, ranging from multi-gigabyte databases to individual e-mails and memos. Because the Brobeck collection is the largest single subcollection and because managing the rights associated with these materials is especially complicated, our initial efforts have focused on figuring out how to preserve these confidential legal materials. The Digital Archive permits systematic exploration of a host of new questions, requiring detailed, historically nonpublic data. These data provide a window into the Internet boom, an extraordinary period of technology entrepreneurship and industry creation. Understanding the dynamics of such episodes will help future generations learn to, at the very least, make *new* mistakes. Finally, the data enable scholarly conversations across multiple disciplines, some of which will converge to produce significant interdisciplinary knowledge. This confluence of factors is rare in archival and library science and underscores the scope of both the challenge and the opportunity before us.

Conclusion: Codifying the Public Interest in Private Records

The concluding discussion reconsiders the relationship between the initial challenge and our response to date: Are we doing enough? Can we save the right types of business records in the emerging preservation regime? What more can be done? By whom? What are the possible resulting scenarios?

In general the efforts described previously are mere baby steps. Even if we succeed in preserving every document and digital object on which we have set our sights, the Digital Archive of the Birth of the Dot Com Era will preserve only one tiny piece of the larger world of business history. Our efforts demonstrate only that *some* business records—from the Dot Com Era, from startup ventures, from failed companies, from clients represented by a huge bankrupt law firm—can be saved under certain, limited circumstances. But to be frank, since our project began, if the preservation outlook has improved in some quarters—geospatial data, Web harvesting, social science data sets, and many other domains that have been successfully transformed by the considered efforts of our digital preservation colleagues—the situation for business records is getting worse not better. Changes in the Federal Rules of Civil Procedure, growing awareness of the immense costs of e-discovery, wider adoption of records retention schedules, and tighter control of the flow of information within firms are all pointing to the end of the Record of Business. As we charged forward, our destination receded, leaving us even further from the goal of being able to faithfully preserve a representative sample of born digital business records from the age of business ascendant than when we started. Yes, the Digital Archive of the Birth of the Dot Com Era is an existence proof: records can be saved. These collections are new types of scholarly objects that may eventually be as important and valuable as a fundamental scientific breakthrough or a new statistical technique. But it should not be this hard. It should not take an historic boom and bust to produce this preservation opportunity. Events such as we have recently seen are few and far between in the history of entrepreneurial capitalism and will not likely be repeated for decades. So the good news is that we have preserved an important set of materials from a singular moment in business history. The bad news is that we have not succeeded in changing the tilt of the preservation landscape. If we have won a battle, we seem less and less capable of winning the war.

Can we change the landscape? I suggest the answer is yes. At rock bottom, the issue is the lack of recognized public interest in private records. The only real solution, therefore, is to establish such an interest. Congress can and should create it. According to the same logic by which we enjoy the benefits of the National Register of Historic Places, there should be a mechanism by which a firm can ask a public body to acknowledge the broader value of its private records, place such records in public hands and thence be inoculated against further private costs—discovery, litigation, etc.—arising from the fact that the records were *not* destroyed.

In the end, firms are composed of people, and in general, few people lack the ability to see value in history. The risks that firms possessing potentially interesting historical records are trying to insure against have the following attributes: they are unknowable *ex-ante*; there is a low probability that they will occur; and they have likely low but ultimately unknown costs. Here is where we see the confluence of the factors described at the outset.

The entrepreneurial economy naturally tends to disperse archival records. Failed companies leave few traces and complicated, sometimes conflicting, lines of archival provenance. For instance, the individual author of a business plan may own the copyright to the plan because he or she wrote it before the company it proposed to establish was created. Months or years later, when the company it described ceases to exist, the assets of the firm may be assigned for sale, liquidated in bankruptcy or even abandoned, but who owns the copyright to the business plan? Who owns the intellectual property described in the plan that was subsequently assigned to the company, developed to a further degree and then failed? The case of a failed firm in the Brobeck Archive is even more complicated. If a dutiful researcher wants to obtain consent to examine the records of a failed venture, whose permission does he need: The entrepreneur who lost the company when his or her investors lost confidence and removed the founder from active management? The investors who liquidated the company, selling bits and pieces for salvage? The individual employees whose private e-mails are embarrassingly but incidentally entrained in the preserved records? Or the lawyer whose firm failed, effectively abandoning the records on a street corner? None of these agents has acquitted themselves with honor. The organization itself—the properly responsible entity—no longer exists, but according to the Supreme Court, its privilege and confidentiality are perpetual. The records of the entrepreneurial economy are thus orphan records of a sort. Managing the risks associated with other types of orphan works has shown signs of progress in recent years. Thus, maybe there is hope for business records too.

In some circles, critics have conjured the specter of a "Digital Dark Age" to describe the weak regime for the preservation of cultural materials in the transition from the pre-digital to the digital age (Kuny, 1998; Rosenzweig, 2003). As valuable, born-digital cultural resources slip through and around the pre-digital preservation network, new institutions dedicated to cultural preservation are adapting to meet the challenge. Other contributors to this volume represent a cross-section of these efforts; scholars and the public alike can take great satisfaction in the vision and breadth of these initiatives. Under the good auspices of the National Digital Information Infrastructure Preservation Program of the Library of Congress, the National Science Foundation, and a host of other forward-looking organizations, archivists, and librarians are waging an under-resourced, rear guard action to keep the technological horizon in sight, if not quite at hand. In the few years since these efforts commenced, multiple technological shocks—YouTube, Google Maps, MySpace, Facebook, Twitter, to name just a few—have forced participants to adapt to this changing landscape, and the expanding network of partners catalyzed by the Library of Congress (and others) has done just that. In general, I am much less worried today about the advent of a digital dark age than I was several years ago when these initiatives were conceived and launched.

However, in the area of digital business records, this article has advanced a more troubling possibility. Despite the best efforts of the community of archivists and librarians to preserve representative and valuable subsets of digital culture, the Record of Business is at risk of loss for the very reason that other categories of born-digital records may yet be preserved: because the records have value. As businesses increasingly discover the value of their own records, practicing what researchers and consultants have come to refer to as *knowledge management*, organizations are dedicating more and more resources to records maintenance and management. In the short term, where paper records might have been haphazardly managed in the past, expanded digital records are now being actively curated and exploited by progressive managers. ERP, CRM, and other data-intensive management systems have increased the efficiency of supply chains (SAP), sales and customer management (Salesforce.com), IT resource planning (VMWare), and other critical corporate functions. In each instance, previously diffuse data of unknown economic value are gathered, ordered, and packaged to support core business activities. The increased value of data justifies increased investment in data management, extending the primary use phase of the data lifecycle.

In a brief number of years, the transformation will be complete. Business records will have gone from loosely managed collections that are weakly and haphazardly persistent in paper to highly structured and actively managed but ephemeral databases of ones and zeroes. This process (threat) has been identified in other (i.e., nonbusiness) settings, where it is associated with incidental loss due to the technical and institutional challenges of digital preservation. The digital challenge is real and should not be underestimated. Understanding the multiple dimensions along which the digital transformation strains existing institutional mechanisms will require the goodwill and coordinated efforts of multiple, overlapping sets of participants and stakeholders. In the domain of business, however, the situation is different. Here, the challenge is not that records will succumb to incidental loss. Rather, the risk is that there is a class of records that will be lost to human history in its entirety. The Record of Business will be lost not by accident, but on purpose and in full view of relevant stakeholders, shareholders, and other fiduciaries. Yeah, shareholders should cheer as these records are actively destroyed.

Alternatively, the sanitized subset of records that are actively preserved will resemble a representative sampling of the past in the same way that letters to shareholders from CEOs published in annual reports of public corporations satisfactorily describe the historical activities of that corporation during a given year or period, which is to say, not at all. The record of business in this view will be the same puerile, pre-chewed and strained, massaged, PR-processed view of events that currently pass for sense-making in the executive suite.

Business has never been more central to American society than it is today. Paradoxically, unless a new social contract is forged between business and society, upon which a preservation regime can be established to save important business records, we will have no pathways by which to understand the triumph of business.

Notes

1. Speaking from the bench, Judge Montali began, "First of all, let me complement the Trustee and you, Ms. Borrey, and all the people you've been working for, for a very, very comprehensive and, you know, well thought out and well conceived approach to this problem, a unique problem." For full transcript of court session, see http://www.brobeckclosedarchive.org.

References

Baumol, W., Litan R., & Schramm, C. (2007). *Good capitalism, bad capitalism, and the economics of growth and prosperity.* New Haven, CT: Yale University Press.

Goldfarb, B., Hoberg, G., Kirsch, D., & Triantis, A. (2008, October 9). Does angel participation matter? An analysis of early stage venture financing. Robert H. Smith School of Business, Working Paper. Retrieved January 15, 2009, from http://ssrn.com/abstract=1024186

Goldfarb, B., Kirsch D., & Miller D. (2007). Was there too little entry during the dot com era? *Journal of Financial Economics, 86*(1): 100–144.

Goldfarb, B., Kirsch D., & Pfarrer, M. (2005). Searching for ghosts: Business survival, unmeasured entrepreneurial activity and private equity investment in the dot-com era. Robert H. Smith School Research Paper No. RHS 06-027. Retrieved January 15, 2009, from http://ssrn.com/abstract=825687

Gomes, L. (2006, November 7). The dot-com bubble is reconsidered—And maybe relived. *Wall Street Journal*, p. B1.

Greenemeier, Larry. (2006, May 29). VA had many security warnings before its 26.5 million-person breach," *Information Week.* Retrieved January 15, 2009, from http://www.informationweek.com/shared/printableArticle.jhtml?articleID=188500807

Hafner, K. (2006, August 23). Researchers yearn to use AOL logs, but they hesitate. *New York Times.* Retrieved January 15, 2009, from http://www.nytimes.com/2006/08/23/technology/23search.html

Kirsch, D. (2007, November 2). Business Plan Archive: Data Overview. 2007 Kauffman Symposium on Entrepreneurship and Innovation Data. Retrieved January 15, 2009, from http://ssrn.com/abstract=1022096

Kirsch, D., & Goldfarb, B. (2008). Small ideas, big ideas, bad ideas, good ideas: "Get big fast" and dot com venture creation. In W. Aspray and P. Ceruzzi (Eds.), *The Internet and American business* (pp. 259–276). Cambridge: MIT Press.

Kirsch, D., Goldfarb, B., & Gera, A. (forthcoming). Form or substance: The role of business plans in venture capital decision making. *Strategic Management Journal.*

Kuny, T. (1998). The digital dark ages? Challenges in the preservation of electronic information. *International Preservation News, 17*, 8–13.
Lamoreaux, N., Raff, D., & Temin, P. (2003). Beyond markets and hierarchies: Toward a new synthesis of American business history. *American Historical Review, 108*(2), 404–433.
Metrick, A. (2007). *Venture capital and the finance of innovation.* Hoboken, NJ: John Wiley & Sons.
Schramm, C. (2006). *The entrepreneurial imperative: How America's economic miracle will reshape the world (and change your life).* New York: Collins.
Shane S. (2003). *A general theory of entrepreneurship: The individual-opportunity nexus.* London: Edward Elgar Publishing.
Shane, S. (2008). *The illusion of entrepreneurship.* New Haven, CT: Yale University Press.
Stinchcombe A. (1965). Social structure and organizations. In J. March (Ed.), *Handbook of organizations.* (pp.)Chicago, IL: Rand McNally.
Rosenzweig, R. (2003). Scarcity or abundance? Preserving the past in a digital era. *American Historical Review, 108*(3): 735–762.
Wong, N. (2006, March 23). HP worker data stolen: Fidelity alerts 196,000 retirees of laptop theft, monitors their accounts. *San Jose Mercury News*, p 1A .

David A. Kirsch is associate professor of strategy and entrepreneurship at the Robert H. Smith School of Business at the University of Maryland. His research focuses on the intersection of problems of innovation and entrepreneurship, technological and business failure, and industry emergence and evolution. In partnership with the Library of Congress, Kirsch has established the Digital Archive of the Birth of the Dot Com Era, an archive that includes collections of business records documenting the early commercialization of the internet. Kirsch's research has appeared in leading management journals and has been the subject of feature stories in national publications including *The Washington Post*, *The New York Times*, and *The Chronicle of Higher Education.* Kirsch received a Ph.D. in history of technology from Stanford University, an M.A. in economics of innovation from MERIT at the University of Limburg, and an A.B. in history and science magna cum laude from Harvard College.