Digital Sport History: History and Practice

Jennifer Guiliano

Department of History, IUPUI, Indianapolis, IN, USA

503A Cavanaugh Hall

425 University Blvd

Indianapolis, IN 46202 USA

1 (317) 274-8560

guiliano@iupui.edu

Jennifer Guiliano holds a position as Associate Professor in the Department of History and affiliated faculty in both Native American and Indigenous Studies and American Studies at IUPUI in Indianapolis, Indiana. She received a Bachelors of Arts in English and History from Miami University (2000), a Masters of Arts in History from Miami University (2002), and a Masters of Arts (2004) in American History from the University of Illinois before completing her Ph.D. in History at the University of Illinois (2010).

This is the author's manuscript of the book chapter published in final edited form as:

Digital Sport History: History and Practice

This article traces the history of sport historian’s use of digital technologies for the purposes of scholarly research and production. The article begins by discussing how quantitative and social sport history served as an early form of digital sport history that was enabled by statistical software and the popularization of spreadsheets for researchers’ use. The article then discusses how the proliferation of the internet created the opportunity for new forms of scholarly production while also enabling internet cultures and social media as areas of interest for sport historians. It locates the mass digitization effort of the late 1980s and 1990s as pivotal for sport history as researchers could more easily identify and access digital sport heritage around the world. It then highlights how those digitized objects have served as fodder for digital sport historians to analyze the sporting past. The article closes by looking briefly to the future of digital methods and calls for their embrace in methodology courses, professional associations, and sport history publications.

Keywords: sport history, digital, theory, practice, digital humanities

Introduction

Each day, 2.5 quintillion bytes of data are created. More than 3.7 billion people surf the internet. Sixteen million text messages are sent per minute. Four point seven trillion photos are stored each day. The deluge of both digitized and born-digital materials is simply unceasing. From the digitization of analog physical materials, to the recovery of materials stored on early media formats like floppy disks, to the harvesting of web and social media platforms that document the hundreds of thousands of sports

---


2 Born-digital materials originate in digital form in contrast to digitized materials which are reformatted from analog to digital form. Common born digital materials include digital photographs, electronic records, and content created on the internet and via software platforms.
forums and events, sport historians of the future will certainly have to confront digital artifacts and platforms when they write sport history. In the last twenty-five years sport historians have begun to fully integrate their analog research methods with digital opportunities and methods. Full-scale digital repositories allow you not only access documents and artifacts-on-demand but to annotate, analyze, combine, and remix them into new forms of scholarship. Catalogs and search engines assist in identifying resources. The computer and its associated software can make organizing and producing research more efficient than previously able.

Digital Sport History + Quantification

The starting point for most digital sport historians is the act of turning on one’s computer, connecting to the internet, and opening a browser window. We eagerly jump onto a search engine and begin querying away. However, the roots of digital sport history lie not in the internet but rather in quantitative and social histories of the 1980s and early 1990s. In 1980, Steven A. Riess published Touching Base: Professional Baseball and American Culture in the Progressive Era. A cultural history of players, owners, and spectators in Atlanta, Chicago, and New York between 1900 and 1920, Riess surveyed living players who had played in the major leagues via a mailed questionnaire. He augmented the data returned with information gleaned from newspapers to assemble a dataset of players and their social and professional relationships. Riess then weighted the dataset to address issues of over and under-representativeness aligning his sample to the National Baseball Library’s authoritative listing of baseball players and their demographic backgrounds (particularly their level of educational attainment). His adoption of the Statistical Package for the Social Sciences (SPSS), a computer program that allowed for statistical processing, facilitated the work that that today’s digital historians commonly utilize when assembling and processing

---

their datasets. Unique in his utilization of statistical software, Melvin Adelman and Don Morrow would both lament that sport historians had not embraced quantification en masse given the proliferation of records that lent themselves to quantitative methods. The lone bright spot within sport history were those who explored the economic aspects of sport. Digital tools and methods were slowly creeping into practice through the adoption of the spreadsheet to track research and create datasets. Spreadsheets served as the most common form of quantitative analysis in part because they allowed sport historians to examine data over time, craft comparative analyses, and craft mathematical and graphical representations of information. Frequency distributions, charts, graphs, and line plots of change over time were all enabled by sport history’s rich numerical record that could be culled from newspapers, game reports, business records including financial and salary documents, and player rosters. Here sport history mirrored the larger historical discipline that embraced dataset creation and analysis to discuss social and economic history. It also reflected the influence of sociology of sport scholars who were trained in quantitative methods and who were interested establishing the existence of sporting cultures.

Digital Sport History + the Internet

Computers, the internet, and the content that that proliferated beginning the early 1990s

---

4 His utilization of an appendix to elucidate his methodology separate from analysis is one that continues to dominate scholarly production in digital history, much less digital sport history.
would dramatically shift sport history as it began to intersect with new media and communication technologies. The widespread adoption of personal computing and the expansion of the Internet in the 1990s enabled historians to share and communicate much more rapidly. E-mail, listservs, file transfer protocols that allowed for the sharing of large datasets, and messages boards transformed sport history as it allowed sport historians to pool resources and communicate about potential avenues of research. It also facilitated discussions of the teaching of sport history. The Internet would also serve as abundant new research archive for the study of sport as it saw the digitization of analog materials and the launch of born-digital sport in the new digital world.

Organizational websites, fan sites and forums, sports blogs, fantasy leagues, and


even encyclopedias that incorporate sport history, have all proliferated since the wide adoption of the internet. Fiona McLachlan and Douglas Booth provide an extensive overview of how sport historians have approached the internet as an archival site. Using swimming as a case study, they argue that sport historians needed to recognize the Internet as a purveyor of sources, a context through which we can theorize about sport, and a form of historical representation that is itself subject to continual meaning-making.

Perhaps unsurprisingly, the largest body of digital sport history literature is situated around blogging as both content and scholarly practice. Andrew McGregor, Andrew Linden, and Lindsay Pieper have posited that blogging on the internet offers immediate access to sporting communities and audiences that would otherwise be left by the wayside. McGregor argues that blogs “serves as an increasing important meeting ground for scholars and interested publics” that allows sport historians to communicate to willing audiences. Linder and Pieper analyze who the authors of sport blogs are, what types of content are they creating, and whether the content aligns to scholarly practices in sport history. Using a corpus of six hundred blogs drawn from six sites, they argued that the content of sport history blogs were most frequently exposition on current events (37.5%) and reflections (25.4%). The corpus also revealed problems that mirrored the state of sport history and the academy: most contributors to sport history blogs were male with most content about men’s sports and men’s teams. It is not only

19 A useful summary of how sport historians encounter all three Internet activities is Noah Cohan, “New Media, Old Methods: Archiving and Close Reading the Sports Blog,” *Journal of Sport History* 44, no. 2 (August 8, 2017): 275–86. Cohen uses born-digital internet content in the form of blogs to analyze blogging as a communication medium. He also conducts an auto-ethnographic reflection of his own blog creation and its relationship to academic standards.
22 Linden and Pieper, “Writing Sport Online.” 262.
blogging that the expansion of the internet has enabled within sport history, it is also social networking.

Analysis of social media platforms like Facebook and Twitter have also received attention from sport historians. In 2015, Mike Cronin asked, “how do we, indeed should we, engage with a Twitter and other social media?” Using his own work on the Gaelic Athletic Association and Ireland’s built sporting environment as a case study, Cronin argues that sport historians have to distinguish between three impulses in using social media: 1) social media as a network of potential individuals who can provide research assistance; 2) social media as an advertising platform that allows you to gather research materials from its community; and 3) social media content about sport generally. The latter two are, in Cronin’s articulation, in flux. As Gary Osmond argues in his analysis of Lesbian, Gay, Bisexual, and Transgender athletes on Twitter, the platform’s development and uptake by users parallels the growth of “out” athletes. As a result, it offers an opportunity for contemporary and recent sport historians to engage with social memory research. His conclusion, that Twitter helped perpetuate memories of Australian rugby league player Ian Roberts, though was tempered by methodological concerns about social media research in general: What is the legal and ethical implications of Twitter archives? How does a historian retroactively gain permission to use a “tweet” in publication? Does the inability to systematically examine every tweet within a specific research query limit the conclusions that sport historians can reach? The question of limitations is one that Holly Thorpe explores in her analysis of Facebook memorial pages lamenting the deaths of surfer Andy Irons and freestyle skier Sarah Burke. She argues that virtual memory carries with it many of the same concerns of power, authority, and agency that non-virtual mourning practices do. Ultimately like Osmond, Thorpe concludes that sport historians must grapple with the

24 Cronin, 102.
contextualization of virtual and social media as a form of representation.27

Digital Sport History + Cultural Heritage Digitization

In the closing paragraphs of the foreword to the 2010 Companion to Sports History, Peter N. Stearns writes, ‘one of the defining features of good sports history, along with its relationship with kindred disciplines and its deep interest in linking to other historical facets, involves its commitment to high-level analysis.’28 For Stearns, the kindred disciplines were sociology and kinesiology that embraced the cultural turn in history that invigorated questions of race, gender, class, ethnicity, and lived experience. This conclusion was supported by S.W. Pope and John Nauright who articulated the future of sports history as global in nature. This was accurate; yet it also elided an even greater turn in sport history that would surpass quantitative history: digitization and the representation of digital surrogates. The entre point for most sport historians to digital sport history is through the consumption of digital resources in the form of digital archives and digital libraries. Digital surrogates enable scholars to view Edward III’s Proclamation banning football in 1363, coverage of the first ladies’ Wimbledon Tennis Championship in 1884, and diaries from the 1936 Berlin Olympics.29 Institutional digitization efforts allow researchers to browse the Avery Brundage Papers Collection held by the University of Illinois, the physical culture collection at the Stark Center at the University of Texas Austin, and the Sport Collection at the National Museum of Australia.30 In the late 1990s and early 2000s cultural heritage institutions, often backed by government initiatives, undertook a

27 Thorpe’s work is an important extension of Mike Huggin’s call to embrace the visual turn in sports history within the context of digitized photographic and film archives to expanding the potential source base for historical analysis. See Mike Huggins, “The Visual in Sport History: Approaches, Methodologies and Sources,” The International Journal of the History of Sport 32, no. 15 (October 13, 2015): 1813–30, https://doi.org/10.1080/09523367.2015.1108969.
massive digitization effort to share local cultural heritage with potential audiences across the world.\textsuperscript{31} Europeana, a 2008 initiative of the European Commission, provided access to 4.5 million digital objects from over 1,000 institutions.\textsuperscript{32} They joined a landscape of digitization initiatives already populated by the Internet Archive, the Google Print/Google Books project, as well as national digitization initiatives from the governments of India, China, Norway, Australia, and the US (to name just a few).

Books, manuscript collections, newspapers, and physical objects have all been fodder in the growing landscape of digital cultural heritage. Digitization of The Sporting News, Sports Illustrated, Sporting Life, Baseball Magazine, the Spaulding Baseball Guides, and hundreds of local, state, regional, and international newspapers nudged sport historians to embrace digital consumption as one of the first acts of historical research. Digital historian Roy Rosenzweig noted these opportunities in his 2004 Sport History on the Web: Towards a Critical Assessment article. Sport historians will be astonished at the “gems they will find in a single afternoon of searching” including baseball cards from the Library of Congress, historical newspapers from the ProQuest digital resource, and secondary scholarship available through digital archives and associations.\textsuperscript{33} Martin Johnes and Bob Nicholson illustrate how digitization of newspapers both enables a rapid research process and introduces a number of practical problems that sport historians must consider.\textsuperscript{34} Optical character recognition technologies, search algorithms, and information disambiguation in addition to digital remediation can complicate using digital newspaper databases. Digitized materials available for download as datasets are quite valuable for sport historians.

Murray G. Phillips, Gary Osmond, and Stephen Townsend conducted leveraged text analysis methods to explore and analyze newspaper coverage of Muhammad Ali, women’s surfing in Australia, and homophobic language in coverage of Australian sport. They conclude that the digital method of distant reading is dependent upon the quality of the original analog material, the quality of the digitized surrogate, and the accuracy of the algorithm used in optical character recognition. These are common problems of textual analysis that Amanda Regan faced in her work on *Mind and Body*, a monthly publication on physical education that ran from 1894 to 1936. Regan downloaded every issue and used optical character recognition to turn the printed issues into plain-text files. The plain-text files of *Mind and Body* then operated as her own personal research archive, allowing her to experiment with text analysis methods to analyze the corpus. Borrowing linguistic analysis methods from computer science, she ultimately used MALLET, a statistic analysis tool that identifies topics and trends within corpora, to uncover sixty themes that dominated *Mind and Body*. These results then allowed her to understand shifts and continuity without having to read and characterize every single article within the publication.

Digitization enabled the identification of sport history sources in far-flung locales through digital catalogs, finding aids, and digital repositories. Digitization also allowed researchers to create their own personal digital archives that could be manipulated and shared electronically to support research communities. Rwany Sibaja, for example, uses the Omeka content management system to craft a personal research


archive on Argentinean soccer. Jennifer L. Shaeffer in her work August 1973 and March 1974 political protests in Buenos Ares by creating her own digital repository of primary source materials that she could augment through timeline and map tools. Liz Timbs crafted her archive on the 2010 World Cup by gathering “openly accessible texts, images, sounds, and videos that capture fans’ perspectives and experiences at World Cup stadiums and fan parks.” Ari de Wilde utilized digitized photographs of a six-day long cycling race at Madison Square Garden to create a panoramic digital representation of the Garden from the perspective of someone who had attended the race. Culling through materials has become a challenge for sport historians as the flood of analog to digital and born-digital materials continue.

**Digital Sport History + the Future**

Mass digitization for all its potential has reproduced many of the biases and problems of analog archival materials. Similarly, as Dain TePoel has pointed out, digital sport history is complicit in the consumptive practice of digital media and internet technologies that have steep ecological consequences. Even as sport historians have embraced the consumption of digital cultural heritage resources and the utility of the computer and the internet for the purposes of research, the discipline has been much slower to recognize its environmental impacts. Similarly, the embrace of

---


analytical, productive, and experimental capabilities of digital sport history that result in anything other than a written peer-reviewed article or monograph has been slow to emerge. Most digital sport history arrives not in its digital form but as a scholarly article that justifies the digital product. In part, this is a result of the glacially slow pace of historical methods courses to embrace digital history as a relevant methodology. But slowness also illuminates core questions that are largely unresolved for digital sport historians: Can a digital sport history project exist without a peer-reviewed companion article? What knowledge does a non-digital sport historian need to adequately engage with and review a digital history project? What explanation must a digital history project do to validate as a quality piece of sport history scholarship?

These questions aren’t trivial. There are dozens of digital methods that might be utilized by sport historians. Digital sport history scholarship to this point has concentrated in social media analysis, digital surrogate and repository creation, distant reading coupled with text analysis, and geo-spatial, geolocative, and image-based visualization. Possibilities are limitless, though, for digital sport history as digital methods become more commonplace in history programs. For us to embrace that future though, we must be willing to meet digital sports history where it originates: in the digital realm. We cannot ask scholars to continue producing analog companion works to support their digital sport history projects. We must instead transition our professional structures including our associations and journals to incorporate digital projects as de facto forms of scholarship. Digital project demonstrations at annual meetings, born-digital publications enabled by editors of press series and flagship journals, and the inclusion of peer-review of digital projects without hesitancy would go a long way to moving digital sports history from the periphery to the mainstream of our scholarly practice. And, as TePoel cautions us, we must do so with full awareness of the consequences of digital scholarship for the world around us.


Bibliography


Klugman, Matthew. “‘Get Excited, People!’: Online Fansites and the Circulation of the Past in the Preseason Hopes of Sports Followers.” In *Sport History in the Digital*


https://doi.org/10.1080/09523367.2015.1086343.


https://doi.org/10.1080/09523367.2015.1090976.


http://imbiza.matrix.msu.edu/.


