Climbing the Digital Everest: The Journey to Digitize the Nineteenth Century

Ray Abruzzi
Gale, ray.abruzzi@cengage.com

Follow this and additional works at: https://docs.lib.purdue.edu/atg

Part of the Library and Information Science Commons

Recommended Citation
Abruzzi, Ray (2013) "Climbing the Digital Everest: The Journey to Digitize the Nineteenth Century," Against the Grain: Vol. 25: Iss. 1, Article 42.
DOI: https://doi.org/10.7771/2380-176X.6443

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.
Climbing the Digital Everest: The Journey to Digitize the Nineteenth Century
by Ray Abruzzi (Director, Strategic Planning, Gale, part of Cengage Learning)

Most may not liken digitizing nineteenth-century manuscripts and playbills to facing life-threatening conditions while climbing the largest peak in the world, but the nineteenth century was, without a doubt, one of the most exciting and revolutionary periods in our history. In many ways, it’s the equivalent of the “Digital Everest” to historians, archivists, and curators alike.

When one of the earliest Everest-attempters George Mallory was asked the question, “Why climb Mt. Everest?” he famously answered, “Because it is there.” In our case, we chose to climb the Digital Everest because, well, our customers asked us to.

And like any good mountaineer will tell you, make sure you bring the right tools and equipment and never climb alone. At Gale, we took this advice to heart.

The Idea and The Reality
It was shortly after Eighteenth Century Collections Online (ECCO) launched in 2003 that customers began asking when we would do “the same thing” for the nineteenth century. ECCO had changed the face of digital scholarship and there was no going back. We saw the logic and the opportunity. Eight years later, we are delivering on that promise, and what transpired in that time happened in neither the sequence nor the manner we had planned.

Soon after stating we would tackle the nineteenth century, we realized that the scope and scale of such an endeavor was simply too large an undertaking with the technologies then available. The amount of publishing in the nineteenth century dwarfs that of the eighteenth century (thanks to the steam-driven printing press, increased literacy rates, and many other factors). Existing bibliographies did not begin to cover the scope of material available. How could we approach the mammoth goal of comprehensively digitizing the nineteenth century — the books, the manuscripts, the images, the newspapers, the pamphlets, and more? Well, we didn’t — at least not immediately.

Instead, we put our efforts into several other ventures. These include 19th Century British Library Newspapers, 17th and 18th Century Burney Collection Newspapers, The Making of the Modern World, and a second part of ECCO. These relatively smaller programs met with success over several years, and we continued to develop archives around single content types (generally either books or newspapers) until 2008, when Gale took another step forward with Slavery and Anti-Slavery: A Transnational Archive (SAS).

SAS was launched as an ambitious five-year publishing program that relied on expert advisors to select content and on new technology to integrate multiple content types. No longer was...
While Gale has always sought user feedback in the development of our electronic products, user-driven product design is something much more deliberate and goes way beyond simple surveys and other forms of market research. Extensive, in-person interviews result in the creation of personas—essentially, iconic figures that represent the needs, workflows, and goals of the variety of users that will rely on our products—in this case, World Scholar—to do their work. A constant loop of user-testing and feedback helps our development team create advanced tools and features that fit into the workflows of our customers while remaining intuitive and user-friendly.

What is AGILE development? In 2001 a group of developers established what they called a “Manifesto for Agile Software Development,” which stated:

“We are uncovering better ways of developing software by doing it and helping others do it. Through this work we have come to value: Individuals and interactions over processes and tools; Working software over comprehensive documentation; Customer collaboration over contract negotiation; Responding to change over following a plan.”

Embracing AGILE at Gale meant restructuring the major of our development staff into more or less self-contained teams, all working side-by-side, with a common vision and a shared understanding. Gone are cubicles, gone is the “waterfall” development approach, gone is a series of hand-offs that eventually lead to a product. The result is a faster development cycle that centers on the specific and evolving needs and goals of our user “personae” rather than on meeting a pre-conceived list of requirements or solving a discrete set of problems. The somewhat unnerving part of the AGILE process, at least from the view of a publisher and product development team, is that the product is never “done” — it continues to change, evolve, and incorporate new tools that meet the needs and outcomes users are looking for.

Enter Nineteenth Century Collections Online (NCCO)

By 2010 we had a fresh plan for product development, and we understood the needs of our customers far better than before. We slapped on our climbing gear and felt ready to tackle NCCO once again. But NCCO is not ECCO. The sheer volume of publishing in the nineteenth century, not to mention the globalization of publishing, is enough to make any digital publisher anxious. While the British Library’s bibliographic program for the many disciplines of the nineteenth century, there is no single, comprehensive global bibliography listing all of the key research materials across such a broad range of content types. What to do about this? Based on expert guidance from our customers, we decided instead to take a modular approach, developing a program of archives that approaches the nineteenth century from a variety of topics and themes. This decision led to the development of our global advisory board.

Smart Climbers Never Climb Alone

The Global Advisory Board for NCCO—which is still growing—is made up of scholars and librarians of different backgrounds and specializations who were brought together to define Gale’s NCCO Archives program at the macro-level. The board advises on topics, themes, concepts, and regions to be covered; recommends institutions, associations, and scholars to provide the evaluation of concepts; consults on product features, design, and user experience; identifies subject matter experts (scholars, librarians, specialists) to define editorial criteria and content selection for particular collections; and conceptualizes and commissions accompanying content, such as introductory essays and headnotes. In essence, the Global Advisory Board helps Gale shape the NCCO program in a manner that makes the nineteenth century manageable for Gale and for customers alike.

Of key importance, too, are the libraries and archives that supply content for NCCO. Curators at these institutions know their collections inside and out, and they are uniquely qualified to advise us on the content of the publishing in the nineteenth century.

NCCO’s innovations are not limited to content management and discovery. Another breakthrough involves a new series of tools for textual analysis.

The “Term Clusters” tool helps users discover related content, and our “Graphing Tool” exposes the occurrence of words and concepts over user-defined periods of time. These tools let researchers view search results as data sets, represented in distinct manners that allow for different data points, query structures, and visual displays. NCCO also introduces comprehensive subject indexing to the content — another first in the industry. The subject terms are derived from a taxonomy we developed especially for the nineteenth century that contains tens of thousands of common terms along with countless place names and personal names. In addition, Named Users, Annotations, and Tags enable researchers to create a user account, store documents, create annotations for their own personal use, and create and share tags for themselves and for other users. NCCO content and citation tools are also optimized for Zotero, the popular tool for collecting, annotating, organizing, and sharing research sources and outcomes.

The Summit

With NCCO now in the market, the work continues: To upgrade the end-user experience on a continuing basis, and to source new archives from digital archives imbedded in digital libraries.

I’m not sure if we’ve reached the summit yet, but we are close, and unlike the Everest attempts and conquerors, we’re not just making history, we’re preserving it for generations to come. The excitement is only beginning.

Author’s Note: On Friday, Nov. 9, 2012 at the 32nd Annual Charleston Conference, Ray Abruzzi co-presented “Climbing the Digital Everest: the Journey to Digitize the Nineteenth Century” along with Simon Bell, Head of Strategic Partnership and Licensing, The British Library, and Caroline Kinbell, Head of Licensing, The National Archives, Kew (United Kingdom).

<http://www.agilemanifesto.org/>