October 2016

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

Ray Abruzzi
Gale, ray.abruzzi@cengage.com

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

Part of the Library and Information Science Commons

Book Reviews — Monographic Musings

Column Editor: **Debbie Vaughn** (College of Charleston) <vaughnd@cofc.edu>

Column Editor's Note: *Though Boss's Day is only observed once a year, it is always appropriate to address one's management skills through personal and professional development. Regular reviewer **Wm. Joseph Thomas** offers his opinion on **Library Management Tips that Work** as well as suggestions for other library management titles worth pursuing. Many thanks to **Joseph** for his contribution, and happy reading, everyone! — **DV***

Smallwood, Carol, ed. *Library Management Tips that Work*, ed. Chicago: ALA, 2011. 978-0-8389-1121-1. 208 pages. $55.00.

Reviewed by **Wm. Joseph Thomas** (Head of Collection Development, Joyner Library, East Carolina University) <thomasw@ecu.edu>

Carol Smallwood's latest book aims to provide "concise how-to case studies of successful managers employing innovation." *Library Management Tips that Work* partially meets this goal. Forty-eight bite-sized chapters are divided into five parts: The Manager Role, Running a Library, Information Technology, Staff, and Public Relations. Thirty contributors from a variety of backgrounds provide two chapters each, with an average of three pages per chapter. Some of the pairs of contributions are more closely related than others — for instance, the two chapters on time management, the two on avoiding discrimination, or the two

on planning for emergencies. Others, though, are only loosely related, and some do not seem to be related at all.

The chapters that work best are those that are generally applicable to any library type; these also display several themes present throughout the book: clear communication, documentation, planning, and partnership, are all necessities for library leaders. The Information Technology chapters share timely tips on deploying technology for library purposes; for example, the noteworthy "Why a Wiki? How Wikis Help Get Work Done" lays out helpful examples on how to choose wiki software, how

libraries are already using them, and how to prevent wiki "fade-away." While the sections on The Manager Role and on Staff are decidedly useful in a variety of settings, the section on Public Relations really only addresses public library directors and their boards.

Many of the chapters have helpful call-outs, and some have figures, appendices, and/or references. One handy bit of advice that emerged from a call-out was the suggestion to have "manager handbooks" to ensure consistency in policies and manager actions in order to avoid discriminatory practices. This book, as another reviewer points out, is perhaps best suited for directors of small public libraries. Overall, though, *Library Management Tips that Work* is uneven. Library managers looking for practical strategies to implement in their own libraries can find them on the Web or in other published literature. For a stronger general introduction to management issues and techniques, there are other books like *Be a Great Boss: One Year to Success or Managing Library Employees: A How-to-do-it Manual*. 🖎

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

Gale publishing a collection of (just) newspapers, or (just) monographs, or (just) manuscripts; rather, *SAS* brought together in a single place multiple content types, and *SAS* made them all work together seamlessly. **Gale** tackled a new challenge in 2009, embarking on a plan to bring a long history of scholarly publishing via our major reference imprints, **Macmillan Reference USA** and **Charles Scribner's Sons**, together with aggregated secondary sources (journals, videos, maps, etc.). The tricky part followed when we had to match them alongside primary sources on the same scale as our existing digital archives. The result is **Gale's** *World Scholar* program. The first installment, focusing on Latin America and the Caribbean, was published in April 2011, after two years of groundbreaking work. *World Scholar* is a new kind of library/classroom resource that integrates more than a million pages of digitized primary sources — dating from the 15th century into the early 20th — with secondary sources and topic portals devoted to major areas of interest in Latin American studies. Following shortly will be *Gale World Scholar: The Middle East*.

World Scholar also represents **Gale's** first major foray in two complementary, synchronous methodologies of product creation: user-driven product design and "AGILE" development.

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 personae — 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,"[1] 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 majority 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 there are numerous bibliographies covering 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 — putting together 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 head notes. 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 on editorial matters. **Gale** has collaborated with over 300 institutions over the last years to bring together over 150 million pages of primary source content in dozens of different projects. Following consultation with our Advisory Board, customers, and partner institutions, the first four Archives for 2012 were identified as follows:

- *British Politics and Society*
- *Asia and the West: Diplomacy and Cultural Exchange*
- *British Theatre, Music, and Literature: High and Popular Culture*
- *European Literature, 1790-1800: The Corvey Collection*

Gale has taken a "mosaic" approach to the **NCCO** program, realizing that the sheer scale of the undertaking means that we will never publish everything created in the nineteenth century. Rather, we aim to be "provisionally comprehensive." By this we mean that the program intends to focus on the major issues, events, and topics of the long nineteenth century, with four major archives to be published every year, each devoted to a particular theme or field of research. For 2013, we're already focusing on archives related to Science, Women, Photography, and Colonialism in Africa.

With the above matters more or less "settled," **Gale** then decided to make the user experience far richer than ever before. First, we decided to capture images at 400 DPI: an ideal balance between image quality and speed of retrieval. Capture is usually performed onsite at the source institution. We have the ability to capture any size document ranging from small pamphlets to large drawings and even enormous maps. In cases where the source material is brittle or could be damaged by handling, we work directly with the source institution to ensure preservation of the original artifact. In practice, this often means that **Gale** funds extensive conservation efforts. For the source institutions, this means the creation of a digital surrogate, which increases access to the documents and saves the originals from wear and tear.

A case in point: On a visit to a partner library, I was lucky enough to be in the archives, assessing materials and opening boxes that had not been opened for perhaps a century. I came across a collection of plays that had been damaged by smoke in the Theatre Royal, Drury Lane fire of 1809. The conservator, ever vigilant, spotted not just smoke damage but mold, too — mold being the bane of archives. Following that visit, **Gale** had the documents cleaned of mold and had the smoke-damaged pages inlaid with "Japanese" paper — and now they are available to researchers worldwide through **NCCO**.

Much of **NCCO** is manuscript content. This kind of material is by its nature unique, and no technology exists (yet) to "read" manuscript documents electronically so they can be made searchable. This means they tend to get "lost" in an online product when they are competing for attention with printed works. Our customers value manuscripts very highly, so **Gale** decided to hand-key selected text (places, names, and dates) from handwritten documents and to field the resulting data, enabling manuscripts to be searched in ways never before possible. In the process we've already added two million searchable terms to handwritten materials that would otherwise be searchable only through metadata and browsing.

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 enables 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 for integration into the program.

I'm not sure if we've reached the summit yet, but we sure are close, and unlike the Everest attempters 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 Kimbell**, Head of Licensing, **The National Archives, Kew** (United Kingdom).*

Endnotes
1. *http://agilemanifesto.org/*