STANDARDS FOR INDEXING: *Evolution Since 1989*

BY JUDI GIBBS, MLS

The nature of indexing makes related standards somewhat stable, and yet they do evolve—and have evolved over the 27 years since publication of *ISQ* began. The evolution reflects changes from a number of directions, including:

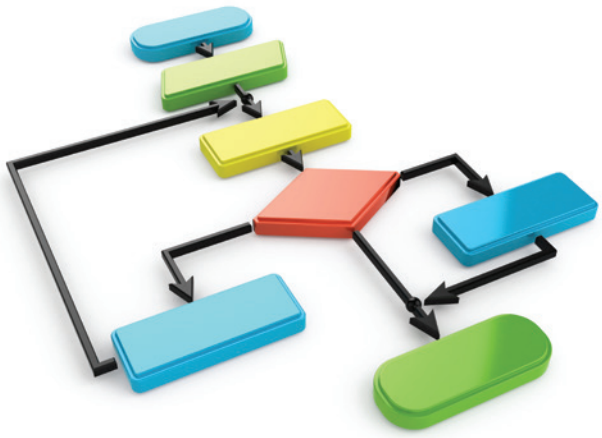
- » global financial pressures that impact the entire publishing industry
- » changes in the literature that sets the standards
- » indexing software improvements that make error scanning easier but also raise the bar
- » websites and e-books

This article discusses these changes. It also touches on international standards. Finally, it notes the potential for e-book indexes to increase functionality beyond that traditionally provided by back-of-the-book indexes.

Before getting into those details, it is important to explain what distinguishes indexing from online searching or a concordance. Online searching finds every occurrence of a term or phrase in the search bar. A concordance is an alphabetical listing of words in the book—without context or cross-references. Indexing, on the other hand, requires reading, analyzing, and creating a site map to the content. A good index picks up significant mentions; provides main entries, subentries, and cross-references; and reflects relevant context, dates, and chronologies. (Top Hat Word and Index, n.d.)

A solid index makes a book usable. It also helps sales. Many readers check the index before deciding whether to buy a book. Amazon recognizes the role of indexes in sales by including them in its "Look Inside" feature.

Standards are enormously important in keeping the quality of indexes high despite financial and other pressures, but those standards are flexible when appropriate. Indexers often say, "It depends." Specifications for entries and subentries are flexible. For example, a single subentry (known as an orphan



A good index picks up significant mentions; provides main entries, subentries and cross-references; and reflects relevant context, dates, and chronologies.

subentry) is usually considered poor indexing, but some style specifications require that author entries have the title as a subentry—even if it results in an orphan subentry. By contrast, a cross-reference to a nonexistent term is never appropriate, nor is a circular cross-reference.

Literature on Indexing

In the 27 years since publication of *ISQ* began, a number of standards have been in place around the world. The sidebars provide details. Indexers, however, need guidelines with greater breadth and depth than these standards offer. Since 1989, the array of literature filling that void has evolved and expanded enormously.

The Chicago Manual of Style has long been the bible of the publishing world. First published in 1906, it is now in its 16th edition. The first edition had no chapter on indexing. The 16th edition has a 50-page chapter that establishes the basic requirements for indexing; publishers change specifications to reflect their own preferences. The 15th edition changed the way that locators are truncated. It also changed the sorting of "St." so that it sorts as "st" rather than as "saint." Many publishers continue to prefer the 14th edition's sorting specifications and its more concise locator truncations that save space in an index where every line wrap counts (University of Chicago Press).

The American Society for Indexing (ASI) has been a leader in publishing books that expand beyond the constraints of *The Chicago Manual of Style*. In 2015, ASI issued *Best Practices for Indexing*, which is available online at no charge (American Society for Indexing, 2015). Every indexer should be familiar with the standards in *Best Practices*. Editors and publishers will also find it a useful resource.

Indexing standards require careful attention to names, which creates one of the more difficult issues in indexing. Prefixes are handled differently in different languages. Names and titles of royalty and nobility are always challenging. In 2012, ASI published *Indexing Names* (Bridge, 2012), a title that belongs on the bookshelf of every serious indexer and many other publishing professionals.

Many fields have additional specialized challenges. ASI has published Index It Right, volumes 1 – 3 (Zafran, 2014; Perlman, 2010), which covers a number of fields, and *Indexing Specialties*, which includes volumes for medicine, law, and scholarly books (Wyman, 1999; Kendrick, 2001; Towery, 2005).

Indexing Software

Advances in indexing software since 1989 have reduced the challenge of complying with standards. For example, NISO Technical Report 03: TR03-1999 provides instructions for sorting letters, numbers, and symbols, and author Hans Wellisch underscores the need for these sorting standards. Indexing software includes settings for letter-by-letter or word-by-word sorting, and it makes it easy to force-sort symbols and numbers. The software also makes it possible to truncate locators according to specifications in either the 14th or 15th edition of *The Chicago Manual of Style*, and it makes it easy to set preferences for cross-references. Thus, contemporary indexing software greatly eases the challenge of ensuring adherence with the standards.

The three major indexing software packages—SKY, Macrex, and Cindex—all make it possible to set specifications for format, cross-references, locator truncation, sorting, and other features. Error scanning picks up orphaned subentries, circular or nonexistent cross-references, overlapping locator ranges, and other issues. It makes spotting and correcting these routine errors much easier and saves an enormous amount of time.

New versions of the software continue to provide increasingly robust functionality and error scanning—which further increases their role in ensuring standards adherence.

Current Indexing Standards

Several domestic standards on indexing are in place around the world. In Australia and New Zealand, for example, indexers use AS/NZS999; inChina, GB/T 22466-2008; in France: NF Z47-102; and in Germany, DIN 31630.

In the United States, in the early days of *ISQ*, indexers were using a standard that was already a few years old, ANSI Z39.4-1984, Basic Criteria for Indexes. When the standard came up for revision, its working group aimed to expanded it to include:

- electronic indexes (the 1984 standard only addressed print indexes)
- displayed and non-displayed indexes
- nonhuman- as well as human-created indexes.

However, some of those voting on the ballot found the proposed new standard too broad, among other problems. Instead of making the changes requested, which would have left the 1984 document almost intact, the broader document was instead published as NISO Technical Report 2: NISO-TR02-1997, Guidance for Indexes and Related Information Retrieval Devices (Anderson, 1997). Based on the six years of work that went into the standard, the document was aimed at expert indexers, with the authors noting that those who were less familiar with the work could use the items listed in the bibliography for further assistance.

As mentioned in the article opposite, in 1999, NISO published another technical report, NISO Technical Report 03: TR03-1999, Guidelines for Alphabetical Arrangement of Letters and Sorting of Numerals and Other Symbols (Wellisch, 1999), which comprised a companion to NISO-TR02-1997. In addition to covering indexes, it provided instructions for bibliographies and dictionaries. The report covered the sorting of Arabic and Roman numbers and, unlike previous works on alphabetization, treated symbols other than numbers as significant. In the report, author Hans W. Wellisch notes that there are no international standards for alphabetization, and every language needs its own standard.

Work on the International Organization for Standardization's ISO 25964 – The international standard for thesauri and interoperability with other vocabularies—began in 2008 with the aim of expanding ISO 2788 and ISO 5964 to include networked information retrieval. The project leaned on related documents and working groups: British Standard BS 8723 (BSI, 2002), ANSI/NISO Z39.19 (ANSI/NSO), and W3C standard SKOS (World Wide Web Consortium, 2009).

ISO 25964, which is currently undergoing systematic review, has two parts.

Part 1, Thesauri for information retrieval, was released in 2011, and Part 2, Interoperability with other vocabularies, in 2013. Part 1 of the standard covers all aspects of developing a monolingual or multilingual thesaurus. Part 2 was innovative and encouraged "high-quality information retrieval across networked resources that have been indexed with different vocabularies. It explained how to set up mappings between the concepts in such vocabularies and other forms of complementary use."

REFERENCES

Anderson, James D. 1997. NISO-TR02, Guidelines for indexes and related information retrieval devices.
http://www.niso.org/publications/tr/tr03.pdf

ANSI/NISO. ANSI/NISO Z39.19-2005, Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies.
http://www.niso.org/standards/z39-19-2005/

BSI . BS 8723-4:2007, Structured vocabularies for information retrieval. Guide. Interoperability between vocabularies.

Wellisch, Hans 1999. NISO-TR03, Guidelines for alphabetical arrangement of letters and sorting of numerals and other symbols.
http://www.niso.org/publications/tr/tr02.pdf

World Wide Web Consortium. SKOS Simple Knowledge Organization System Reference. W3C Recommendation, 18 August 2009.
Latest version available at: http://www.w3.org/TR/skos-reference

Websites and E-Books

The years since 1989 have seen the rise of two previously nonexistent forms of publishing: websites and e-books. Some e-books are much like a website, with a number of related files that users can navigate (Ream, 2012a). Large websites and nonfiction e-books need an index with the same features as back-of-the book indexes. One difference is that cross-references are usually hyperlinks that readers click to reach the target.

Many publishers convert a PDF to e-book format, with links from page numbers in the index to page breaks in the e-book. ASI has set a higher standard of precision by pinpointing the links to specific paragraphs. ASI books already in print are being converted to e-books, and new ASI publications will likely be available in both hardcopy and e-book format. All ASI e-book indexing will reflect this level of precision (Ream, 2012b; Ream, 2014).

The process of pinpointing the links to specific paragraphs is referred to as embedded indexing. It is more time-consuming and costly, but it is also relatively easy to update for later editions. It pays off for software manuals and other publications that are updated every few months, as well as for publications that are segmented differently for different uses. The embedded indexing makes it possible to generate an index for just those sections. For example, a textbook with different chapters used for different classes can have separate indexes for each of those classes. Multiple uses for a single body of material is referred to as sole sourcing.

Many publishers convert a PDF to e-book format, with links from pages numbers in the index to page breaks in the e-book. ASI has set a higher standard of precision by pinpointing the links to specific paragraphs.

The globalization of the publishing industry has created pressure to trim expenses whenever possible. Many presses now require authors to assume the cost of indexing.

The Future of Indexing

The globalization of the publishing industry has created pressure to trim expenses whenever possible. Many presses now require authors to assume the cost of indexing. Some authors do their own indexing, often without much related background or experience. Other authors go shopping for an indexer, sometimes with suggestions from their publisher. Presses that still pay for indexing may be tempted to seek low-cost services wherever they can find them, although there are many fine indexers who can create excellent indexes for a reasonable rate. Financial pressure may also result in very limited space for indexes. Some excellent indexes have to be cut extensively to fit the available space.

With globalization, international standards are essential—particularly for e-publications—and the International Digital Publishers Forum (IDPF) is the body responsible for EPUB standards. The ASI Digital Trends Task Force (DTTF) was an integral part of the IDPF and under the leadership of ASI members Dave Ream and Jan Wright created the international standards for digital indexing that were adopted by the IDPF in EPUB 3.0.

Indexing standards were also one focus of the 2016 TC 46 meeting in New Zealand, which recognized that related standards ISO 999:1999, Information and documentation—Guidelines for the content, organization and presentation of indexes and ISO 5963, Documentation—Methods for examining documents, determining their subjects, and selecting indexing terms are out of date. At the Wellington meeting, an ad hoc group, appointed for analysis purposes, proposed that these standards should be merged through a revision of ISO 999. Once the revision of ISO 999 has been published, it is proposed that ISO 5963 will be withdrawn. The new version of ISO 999 will also need to be updated to reflect use with the Internet, electronic documents, online collections, and big data (Shuter, 2016).

The wave of the future includes EPUB indexes. A 2011 program by the ASI Heartland Chapter captured the potential of e-books. In her summary of the meeting, Carol Reed noted that e-books have the potential for greater functionality than

traditional back-of-the-book indexes. A prototype that the DTTF developed allows readers to:

- browse the linked index as a chapter in the e-book
- highlight a section of text and see the index entries corresponding to that section—a semantic "map" of topics
- type a term in a search box from any page. Search results will show both the index entries and full-text search results. Both types of results will display snippets of the text, so readers can decide if a link is relevant before following it. A scrollable section of surrounding index entries lets readers browse for better terms and see cross references.

Reed notes that:

> *Some of the best practices for indexing e-books will certainly differ from best practices for indexing print books. For instance, while five locators might be acceptable for an entry in a print index, it's far too many when you're following links. E-book indexes will work better with greater specificity, so single-locator entries and more subheadings are appropriate. Small screen sizes will also dictate formatting, heading length, and the number of subheadings to some extent* (Reed, 2011).

The next 27 years are likely to bring many additional changes to augment those we have seen in the 27 years since the publication of *ISQ* began. It will be interesting to see the tools that indexers working three decades from now will use, and the challenges they will face.

JUDI GIBBS
Judi Gibbs, MLS, is a freelance indexer. She has also been a librarian and a medical and technical writer. Her work as a writer often included usability testing of documentation. Her indexes reflect her usability experience and her experience as a reference librarian in university settings and special libraries. She is currently secretary for the national board of the American Society for Indexing and the ASI representative to NISO. Her pro bono work includes volunteering as a librarian for the Danny Schneller Foundation, a non-profit dedicated to providing medical updating of healthcare professionals caring for developmentally disabled clients.

ACKNOWLEDGEMENTS
I acknowledge the invaluable contributions of Charlee Trantino, ASI President, 2014 – 2015, and a past winner of the Wilson Award for Excellence in Indexing. Charlee's knowledgeable and thoughtful suggestions added enormously to this history of the evolution of indexing standards over the past three decades.

REFERENCES

American Society for Indexing. 2015. *Best Practices for Indexing.* http://www.asindexing.org/wp-content/uploads/2015/11/ASI-BestPracticesIndexing2015.pdf

Bridge, Noeline (ed). 2012. *Indexing Names.* American Society for Indexing.

Kendrick, Peter and Zafran, Enid (ed). 2001. *Indexing Specialties: Law.* American Society for Indexing.

Perlman, Janet and Zafran, Enid (ed). 2010. *Index It Right! Vol. 2.* American Society for Indexing.

Ream, David K. 2012. "Anatomy of an EPUB e-book," *Key Words* 20. pp. 85, 106.

Ream, David K. 2012. "Converting a Legacy Print Book to an EPUB with Pinpoint Index Linking," *Key Words* 20. pp. 83 – 84.

Ream, David K. and Pilar Wyman. 2014. "ASI Publications Revamp: eBooks and an Index Mashup," *Key Words* 22. pp. 156 – 157.

Reed, Carol. "Dave Ream on E-book Indexes." American Society for Indexing, Heartland Chapter, 2011. http://www.heartlandindexers.org/dave-ream-e-book-indexes.html

Shuter, Janet. Personal communication, June 1, 2016.

Top Hat Word and Index. (n.d.) The Penny Farthing Commuter. https://tophatwordandindex.wordpress.com/

Towery, Margie and Zafran, Enid (ed). 2005. *Indexing Specialties: Scholarly Books.* American Society for Indexing.

University of Chicago Press. *Chicago Manual of Style.* 1st, 14th, 15th, and 16th editions.

Wyman, Pilar (ed). 1999. *Indexing Specialties: Medicine.* American Society for Indexing.

Zafran, Enid (ed). *Index It Right!* American Society for Indexing. Volume 1, 2005; Volume 3, 2014.