BOOK REVIEW

Finding the Thinkable Thoughts

Index, A History of the: A Bookish Adventure from Medieval Manuscripts to the Digital Age
Dennis Duncan (W.W. Norton 2021), 352 pages

Beth Hirschfelder Wilensky*

He who knows where knowledge may be had is close to having it.[1]

When John B. West developed his legal classification system in the 1880s,[2] he likely didn't anticipate that his system of indexing the law would persist into the twenty-first century. But persist it has, and flourished, most prominently in the form of Westlaw's headnote and key number system. West's index built on several precursors,[3] but it was his version, West's *American Digest*, that established the foundations of the classification system that most modern legal researchers use today, in one form or another.

Others had earlier attempted to create classification systems that enabled lawyers practicing in the common law system to find "the law"—i.e., precedents relevant to a client's case.[4] But it is no accident that West's system, the most thorough and enduring one, arrived when it did. The second half of the nineteenth century witnessed a surge of interest among intellectuals in indexing written works generally, culminating in an 1877

* Clinical Professor, University of Michigan Law School.

1 Dennis Duncan, Index, A History of the: A Bookish Adventure from Medieval Manuscripts to the Digital Age 228–29 (2021).

2 For a detailed description of the early development of the West reporter and indexing systems, see Michael O. Eshelman, *A History of the Digests*, 110 L. Libr. J. 235, 237–49 (2018).

3 *Id.* at 241–45.

4 "Without digests, claimed Frederick C. Hicks, the law librarian at Yale, 'the whole fabric of the common law would long ago have broken down.'" Eshelman, *supra* note 2, at 239 (quoting Frederick C. Hicks, Materials and Methods of Legal Research 251 (1st ed. 1923)).

conference of librarians who gathered with the intent to create a Universal Index of knowledge.[5] West's index was a natural outgrowth of this larger movement.

This history—of the late-nineteenth-century fascination with indexing—and much more is told in charming detail in Dennis Duncan's *Index, A History of the*.[6] Anyone interested in how we classify, find, and use information is likely to be intrigued by the book's recounting of the life of this tool. The index is nearly ubiquitous in non-fiction materials today, yet we tend to take for granted what a remarkable tool it is. And that includes, of course, lawyers and other legal researchers. While the book does not address legal indexes specifically, its discussion of the history, purposes, and substance of the index calls to mind modern legal indexes like Westlaw's headnote and key number system.

The early history of the index

Index begins by walking the reader through early attempts at categorizing information in the late Middle Ages. That period saw the growth of two institutions—universities and religious orders—that each inspired a need for texts that were accessible to their users.[7] As a result, the first materials to be indexed were the Bible and similar religious tracts on the one hand, and the works of Greek philosophers on the other.

The book then turns to a lively history of the traditional index's raw ingredients: alphabetization and page numbers.[8] Most of us give hardly any thought to these two simple inventions,[9] but they are as essential to indexing as the invention of the wheel was to transportation: obvious in hindsight, but revolutionary in their time. Similarly, it is surely no accident that West developed his own index in conjunction with his National Reporter System:[10] an index only functions if the body of work it refers to

5 Duncan, *supra* note 1, at 209–12.

6 *See generally id.*

7 *Id.* at 51.

8 *Id.* at 19–47 (alphabetical order); 85–112 (page numbers). "The page number has become the universal referencing unit, the second basic ingredient—along with alphabetical order—of pretty much any book index in the last 500 years." *Id.* at 98.

9 "With barely a thought we know how to use a table where alpha order is the sole organizing system (as in the old residential phone books), or where it works in tandem with another specialized or context-specific categorization (as in the old Yellow Pages, where entries were grouped first by trade, then alphabetically within these). It's a system with which we are completely familiar, something so deeply ingrained, something we acquire so early, that it might seem self-evident." *Id.* at 25–26. With respect to the first printed page number, the book explains that it "will revolutionize the way that we use books. And in doing so it will become such a commonplace that it will almost disappear from view, hiding in plain sight at the edge of every page." *Id.* at 86.

10 Eshelman, *supra* note 2, at 247.

has standardized embedded location tools—volumes and page numbers in F. Supp. 3d, for example.

The book's discussion of different methods of indexing bears particular resonance for modern legal researchers—and especially for those of us who teach legal research to law students. The two primary methods are organizing around words (i.e., a "concordance") and organizing around subjects (i.e., a cascading index like the West key number system)—what the book describes as "matching letters versus identifying concepts."[11] That's the precise distinction between Boolean searching and digest searching. And of course, the latter—"identifying concepts"—has historically required editorial judgment from an actual person, unlike (for the most part) "matching letters."[12] Duncan comes out strongly in favor of the importance of human intervention and subject indexes. (This law professor feels similarly, as my students can attest.) Here is Duncan:

> The limitations of *unimaginative* indexing, of the simple string search, become starkly apparent if one tries to locate the parable of the prodigal son, that famous tale of mercy and forgiveness, using a Bible concordance. The parable does not contain the words *forgiveness* or *mercy*, or, for that matter, *prodigal*.[13]

The same is true of legal concepts, which are often not captured by specific and unique words. A classic example is of the multiple ways that a court opinion might refer to someone under the age of majority: "juvenile," "child," "minor," "infant," etc. A researcher who can't anticipate all of the possibilities is likely to miss key authorities. An even thornier problem for concordance searches is that some concepts are difficult to describe in a way that won't produce an unwieldy number of irrelevant "hits." For example, imagine trying to create a useful Boolean search for this question: "Can a jury verdict on one charge be voided if it is inconsistent with another charge?"

The subject index: complexity and complications

The counter to these problems is the subject index, an index that is created by a human user who classifies the material into sections, subsections, sub-sub-sections, and so on. The obvious benefit of such an

11 Duncan, *supra* note 1, at 258.

12 Before the computer age, matching letters to identify words often required editorial judgment about *which* words were worthy of indexing. The advent of computer searching has enabled the ability to index *every* word.

13 Duncan, *supra* note 1, at 260.

index is to render findable such information that a word search might otherwise bury. But—as skilled users of legal indexes and their close cousin, the Table of Contents, know—a good subject index has power beyond that. A subject index tells a detailed story about the relationship between complex ideas. The wise researcher uses it not just to locate a specific concept, but to understand that concept in the context of related concepts. Thus, this quote from the index of an early sixteenth-century historical work: "Read, dear reader, the following table, / And soon under its guidance you will hold the entire work in your mind."[14] Similarly, in the legal realm, even the *name* of West's paper-bound index, "West's Analysis of American Law,"[15] conveys that it is much more than a finding tool for cases; the organization system itself *constitutes analysis*. At 2,116 pages in its current edition,[16] it probably is not one that a legal researcher could read and thus "hold the entire work in [their] mind." But skilled lawyers know that sitting down to browse through the index to a section will open up areas of inquiry and suggest connections between concepts that they would not otherwise have discovered or thought of on their own.

At their best, then, subject indexes suggest to readers new ways of finding and thinking about the source material. But there are drawbacks. Duncan describes the key concern in his description of an early attempt at a universal index created by the thirteenth-century poet Robert Grosseteste:

> His grand *Tabula* is . . . what we now call a subject index, an index of ideas, and as such it is alive to the play of synonyms, able to identify a concept even where the text does not explicitly name it. It is also, then, a *subjective* index, the work of a particular reader, thinking and parsing their reading a certain way. Concepts are slippery things. We make a choice when we say that a text is *about* something; that, say the, story of Noah's Ark is about forgiveness, or anger, or rain.[17]

Thus, if a subject index rests on choices, those choices can be biased, designed with only some users in mind, inattentive to the ways the source material might be understood in future years, or inadequate in any number of other ways. Here is Duncan again: "Indexes are the work of individuals, they are linguistic and therefore human exercises, steeped

14 *Id.* at 118.

15 West's Analysis of American Law (2023).

16 *Id.*

17 Duncan, *supra* note 1, at 51–52.

in the same paradox, redundancy and subjectivity as all language uses."[18] Legal indexes are hardly immune to these same concerns.

Indexing in the computer age

Index travels from early hand-written concordance and subject indexes across the span of centuries, ending in the computer age. In yet another theme that parallels the development of legal research tools, the book discusses the ways in which computing has changed indexing, and the ways indexing is still very much the same. When it comes to creating a simple concordance, computers have the upper hand; they can quickly scan a large volume of text and identify instances of any words the end user choses. And yet, the end user must still know what words to choose; the computer can't replace the researcher's own thinking.[19]

When it comes to creating a subject index, human involvement is essential: "It is a job of deep reading, of working to understand a text in order to make the most judicious selection of its key elements."[20] Reading that line brought to my mind the difference between Lexis's headnote system and Westlaw's headnote system: Lexis's system primarily pulls language directly from court opinions, relying heavily on technology to automate the sorting of snippets into topics.[21] Westlaw, on the other hand, uses attorneys to summarize key points of law from court opinions, which enables it to sort cases into its index "even where those cases may use atypical language."[22] That is one reason I have generally found Westlaw's headnote system to be more effective.[23]

Even so, when humans get involved in creating an index, we can muck things up. Susan Artandi, an early developer of a computer-assisted indexing project explained that "the terms must be known to be indexed."[24] That, in turn, creates new problems: "Terms which appear for the first time in primary sources are missed . . . because they are not yet included in the dictionary."[25] This observation anticipates the way

18 *Id.* at 181.

19 At least not yet.

20 *Id.* at 250.

21 *See generally* UF Law, *Headnotes in Lexis Advance*, Lawton Chiles Legal Info. Ctr., https://guides.law.ufl.edu/legal-research/lexisheadnotes (last visited May 13, 2024).

22 *See* Maggie Keefe, *Free vs. Westlaw: Why you need the West Key Number System*, Thomson Reuters, https://legal.thomsonreuters.com/en/insights/articles/using-the-west-key-numbers-system (last visited May 13, 2024).

23 *Index* itself contains an example of the benefits of human indexing over computer indexing: As the author explains, the book has *two* indexes, one produced by a human and (a partial) one produced by a computer. Duncan, *supra* note 1, at 254. The human-created index is plainly better.

24 *Id.* at 246.

25 *Id.*

that indexing can ossify ideas, including legal ideas. Legal indexes exert a hegemony that impedes the development, recognition, and acceptance of new ideas or ways of thinking about the law. As one pair of scholars has observed, legal indexes "function like DNA; they enable the current system to replicate itself endlessly, easily, and painlessly."[26] As a result, "Their categories mirror precedent and existing law; they both facilitate traditional legal thought and constrain novel approaches to the law."[27]

And of course, because computer indexes still rely on human intellectual effort, they have not eliminated human biases and related problems described above. In fact, as *Index* points out, modern concerns about biases built into (for example) "the black of Google's algorithm"[28] mirror the concerns of "the eighteenth-century pamphleteer who discovered [an indexer] serving up anti-Tory propaganda in the back pages of [a volume of] Tory history."[29] At least back then, a careful reader would recognize that a human was to blame. The computer age has made it all too easy to assume that a technology-produced index is bias-free.

The index endures

And the age of technology has changed things not just for the indexer, but also for its audience, i.e., the end user. This brings us to a final recurring theme of *Index*: How much can a good index do? In particular, should we worry about what aspects of thinking, reading, and research it replaces? And again, this isn't a concern that magically sprang up in the computer era. *Index* describes how, almost as soon as indexes became commonplace, some writers started complaining that indexes would make for lazy readers, who would read just individual pieces and not the whole thing: "The printed index was only just coming into its own, and already alarums were being sounded that indexes were *taking the place of* books, that people didn't read *properly* any more. . . . "[30] That concern echoes the admonishment of law professors to students everywhere: the headnote is an extremely useful tool, but not a replacement for reading the underlying opinion.

But *Index* also shares a more optimistic take, one that might give us reason not to despair. The book quotes a sixteenth-century indexer who

26 Richard Delgado & Jean Stefancic, *Why Do We Tell the Same Stories?: Law Reform, Critical Librarianship, and the Triple Helix Dilemma*, 42 Stan. L. Rev. 207, 208 (1989).

27 *Id.* at 208.

28 Duncan, *supra* note 1, at 232–33.

29 *Id.* at 233.

30 *Id.* at 118.

insists that, while some careless users might rely on the index in place of reading the complete text, "the quality of those books is in no way being impaired, because the excellence and practicality of things will by no means be diminished or blamed because they have been misused by ignorant or dishonest men."[31] Similarly, when an "ignorant and dishonest" attorney relies on just the headnotes, it is the attorney, and not the headnote system, that is to blame for the poor legal analysis that is likely to result. The rest of us—the knowledgeable and honest—will continue to benefit from what a remarkable tool the index is, one that is indispensable in making the law findable.

31 *Id.* at 112 (quoting Hans H. Wellisch, *How to Make an Index – 16th Century Style: Conrad Gessner on Indexes and Catalogs*, Int'l Classification 8 (1981).