June 2005

People Profile: Paul Streby

Editor

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

Part of the Library and Information Science Commons

Recommended Citation
DOI: https://doi.org/10.7771/2380-176X.4888

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.
The “invisible Web” presents special difficulties. For example, if a library has hyperlinks to ten gray list Websites, exactly what percentage is that of Who-Knows-How-Many? The best solution here is to rely on librarians’ careful searching of invisible Web directories, their professional judgment, and consultation with experts whenever necessary. A more qualitative analysis may be called for here.

Numerical Standards For Print and Other 20th Century Concepts

One problem we ran into with the 4th (1992) edition of the WLN Conspicuous was the overgenerality of its numerical standards. For example, it indicated that for a current collection level rating of 3a (lower division undergraduate level), a division should contain between 8,000 and 12,000 monographs, but that “will vary according to publishing output.”

“Will vary?” Is that a way of dodging the hard questions, like the “results may vary” caveat that appears in ads for the latest fad diet? When I looked at some publishing stats in The Bowker Annual Library and Book Trade Almanac, I began to wonder. In 2002, to use a recent example, there were 458 academic titles published in “chemistry;” and 12,645 published in “literature and language.” That’s a heckuva lot of “will vary!”

It turned out, though, that there was a relatively simple and objective way to translate that into workable numbers. I compiled five years of publishing statistics from the Bowker Annual table “North American Academic Books;” divided the total number of titles in each subject by the total number of titles in all subjects, multiplied by 24 (i.e., the number of divisions in the conspicus) and called the product P. Multiply P by the 4th edition WLN numerical standards, and you have some parameters that take account of the greater or lesser importance of monographs in the various fields of learning.

For example, from 1991 to 1995 (i.e., the most recent stats I had), there were 307,431 academic books published or otherwise made available in the United States, of which 16,082 were sociology titles. Thus, we get the following equation: P = (24 x 16,082) ÷ 307,431 = 1.255. Multiply P by the WLN parameters, and instead of 8,000-12,000 books being adequate for a 3a lower division undergrad sociology collection, you would want between 10,040 and 15,060 monographs in sociology.

My system wasn’t perfect. The subjects used in the Bowker Annual don’t match up perfectly with the 24 divisions in the WLN Conspicuous. For instance, unlike the WLN Conspicuous, the Bowker Annual combines math and computer science into a single category, so I had to estimate publishing output for each subject separately using other data in that venerable annual.

Also, publishing output can vary wildly over time; the number of computer science titles has increased dramatically since 1900. Arguably, five years of publishing data is not enough; perhaps ten years or even more would be better.

But as the conspicus project grew, I became a firm believer in the maxim that the perfect is the enemy of the good, and that my method seemed good.

Good Enough For Government Work?

Maybe not. Some people think numerical formulas create an illusion of precision where none is now possible (and arguably, never was). The 5th (1997) edition of the WLN Conspicuous “departs from any formula approach that uses numbers of monographs and journals to determine collection depth indicators.” Numerical formulas are problematic because of each discipline having its own “publishing culture,” the blurred line between ownership and access of online journals, and the inappropriateness of measuring collections around the world by North American Anglophone publication standards. Now, “the use of comparisons in determining collection depth indicators is encouraged.”

I think these are valid concerns, but that they make a case not for jettisoning numerical formulas, but for adapting them to local conditions and particular disciplines. Comparing collections is indispensable, but without some basic numerical guidelines, you risk making subjective judgments based on cherry-picked data or methods of interpretation.

I noted above how to solve the problems of electronic resources and differences among disciplines in their publishing output. The geographical question is a little trickier. Obviously, North American publishing stats are of less interest to libraries in Singapore or Algiers than they are in Michigan or Manitoba. The important thing here is that libraries obtain an appropriate portion of the monographs published in the market(s) from which they purchase. Foreign publishing statistics could be used for libraries that because of their size or location must purchase foreign materials (e.g., the National Library of Andorra) or do so to support masters’ and doctoral research collections.

Where to get foreign publishing statistics? For periodicals, there is Ulrich’s Periodicals Directory. If you want to know what magazines and journals are coming out of Taiwan, Tonga, or Tuvalu, that’s the place to go.

For monographs, there are: Global Books In Print (English-language titles) and Spanish Books In Print, both from Bowker; Indian BIP, Arabic BIP, and African BIP from various publishers; publication yearbooks for Japan and Korea; and probably some other minor markets covered. There doesn’t seem to be a near-end solution on page 58.