

Technology: Pied Piper or playground bully? or, creating meaningful measures using emerging technologies: separating the reality from the myths

Christine M. Koontz

Florida State University, Florida Resources and Environmental Analysis Center, USA

Abstract

This paper discusses rapidly changing technology's blessing and curse, for library practitioners, administrators, and researchers. Four projects are reviewed within the real-world context of a decade plus of library research. The projects were primarily affected by the: 1) lack of availability of technology; 2) lack of standardized data collection methods; 3) resistance by the profession to utilize available technology; and 4) lack of professional agreement on policy and standards issues. Finally, the realities and the myths of technology's role in data analysis and data collection will be discussed.

Introduction

Rapidly changing technology remains a blessing and a curse for library practitioners, administrators and researchers. For example, user-related technology such as the Internet, while providing information access hitherto impossible, is also inherently difficult because of lack of technology, for library managers to count and measure. Identifying user market areas using geographic information systems software through user address data can provide precise geographic market areas, yet local library policies regarding any aspect of user privacy limit utilization of the data on a national or standardized basis. And while software packages for libraries burgeon for data collection, data collection standards are disparate, nationally and internationally, impeding large-scale standardized database development. Further, personnel and funding constraints within libraries and research communities may limit research when both data and technology are available. This paper discusses the implications of the blessing and curse of rapidly changing technology for library practitioners, administrators and researchers, within the real-world context of the author's decade plus of library research.

First – Practicing Librarians: Is Rapidly Changing Technology a Blessing or a Curse?

Some years back I stumbled upon an article, written rather tongue in cheek, regarding how, with each new medium of information introduced, librarians as facilitators go into a tailspin. "What," I said, "librarians? Why

libraries are the primary medium to transmit all media formats!" Continuing to read, I realized it was true. The writer (whom I have long forgotten, and therefore, offer my apologies to) noted that when paperbacks first hit the stands, uproar and fear, and discussions by fellow librarians across the country was unending. A book with no hardcover? Who could have thought? Surely this would lead to comic books in the library.

The same round of uproar and discussion continued with the introduction of microfilm and microform. And continued when records gave way to 8-track and cassette tapes, and tapes to CD-ROMS, filmstrip projectors to films then videos. Most recently it seems every medium gave way (or at least stood aside) for the Internet. The World Wide Web of the Internet is organized daily, and whipped into shape by search engines and subject directories, which embody the reference librarian of old (and these devices are probably created by one!)

So how did, and do, these new media affect libraries? Obviously, for the better, since libraries survive and succeed by mirroring the information needs and wants of users' serving it up in the users' most desired media. These new media blend into and out of library collections and services, finally purged through the weeding process. And so the cycle goes.

So is rapidly changing technology a blessing or a curse for librarians? Probably a curse in transition, but a blessing in the end for meeting our users' changing needs.

Second – Library Management: Is Rapidly Changing Technology a Blessing or a Curse?

On the library management side, new media and technologies for better management also create uproar and non-ending discussions. First, pre-cast Library of Congress catalog cards dramatically changed technical services. Second, the largest and most recent change was the demise of the card catalog, which evolved into the online public access catalog (OPAC). The OPAC continues to transform and some can be found in a wireless state on a Palm Pilot screen.

While these two changes were not rapid, the transition was labor-intensive and reliance on experienced staff catalogers shifted to for-hire vendors. Innately, ven-

dors and librarians have different goals. Librarians strive for customer satisfaction through service, and vendors seek profit and market share. This difference in goals leads to problems. For example, the type of customer data available from current vendors is limited due to lack of demand by library managers, and additional expense for vendors to add customized features for a few. Thus, rarely do most current circulation systems offer even the simplest customer profiles such as type of use by demographics, volume of use by class level, or user address data by zip code. Further there is little doubt that, even if libraries of all types demanded this type of customer data, collectively it would be too small a percentage of the vendor's market.

Automated circulation systems and barcoded books vastly changed public services operations. Remember, prior to this, we wrote our name and address on the library card? Everyone in town had the option of knowing what everyone else was reading. But librarians professionally guarded the hand-written cards, never considering selling users' tastes and preferences, unlike today, where vendors sell our users' subject searches and search strategies to private sector marketers. Technology creates glass windows where there were none. Librarians struggle to keep the curtains down.

Automated circulation systems opened up a whole new array of quickly accessible user data, creating an opportunity for libraries to more easily measure the impact of services upon the communities of people libraries serve. As a result, output measures were born. No longer did libraries simply measure inputs, i.e., holdings, square footage, circulation, reference questions, FTEs, and program and attendance counts. Now these uses could be quickly computed to library use numbers per capita, i.e. circulation and program attendance per capita of population served. Yet true user data per capita are rarely available at the outlet or branch level. This hinders knowledge of who uses each branch, what services and materials are used, volume of type of material/service used, etc. This delimits vital customer data for libraries, data that no successful retailer would or could think of going without.

Vendor-created scanned barcodes in books and other materials greatly facilitate check in and check out procedures, shelf reading and weeding and, most recently, in-library material counts. At present, these packages are expensive (\$8-10,000) and therefore many libraries do not own the software or system. And since there is no single barcode standard, software packages for different applications are bought by individual libraries, resulting in inconsistent barcodes. This lack of standardization and availability severely inhibits development of comparable performance measures.

The Internet just further confounds the growing efficiency and effectiveness of developing standardized performance measures, by obfuscating through its own largess, how to count the who, and for-what, the when

and from-where people are using the world wide web. For library management, rapidly changing technology pushed by profit-motivated vendors often dilutes control over individual library activities that were once the stronghold of manual operations. Technology facilitates operations but at what cost to accurate relevant data?

Finally – Library Researchers: Is Rapidly Changing Technology a Blessing or a Curse?

By and large for library researchers, technology that facilitates research – not research about technology – can never change fast enough. I would like to provide a brief insight and review of research, which originated when there was little technology, and research that was furthered by or could have been furthered by technology. This research is built upon data initially collected manually at the branch or outlet level in the public library field. I would like to review four projects that were affected by the fact that: 1) technology was not available; 2) technology was available but data collection methods were not standardized; 3) professionals were resistant to utilizing the available technology; and 4) technology created a starting point, but progress relied upon agreement on policy and standards issues.

Project Number One: Koontz' Dissertation. (Koontz: 1990)

The purpose of my dissertation (in brief) was to answer the question my marketing mentor, Dr. Persis Rockwood, posed to me. How can public library management site library facilities effectively? After fingertipping through thousands of catalog cards, I found there was little location research on location and siting. I selected 13 geographically dispersed library systems, and ended up with only six systems that fit the criteria (each counted the same library use measures) comprised of approximately 100 branch outlets. The study demonstrated a methodology for constructing location models for diverse urban environments.

The models utilize variables including:

- 1) Relevant demographics (Demographics included: population, sex, race, age, family life cycle (represented by US census categories of households with social security, public assistance, or a female head of household); owner occupied housing, income, education and vehicles per housing unit);
- 2) Spatial (Estimated size of market area and distance between library facilities. A geographic area contains the customers/users of a particular firm/library for specific goods or services.);
- 3) Library use (circulation, reference transactions and program attendance); and

- 4) Library “attractiveness” units (hours of operations, registrations, square footage and holdings).

The models utilize these variables to: 1) estimate geographic dispersion of the population and size of market areas, 2) quantify use levels at an existing facility location, based on the estimated market area and the population characteristics of the estimated market area, 3) forecast use levels at other locations, i.e., for purposes of siting a new facility, or for closure or re-siting, and 4) estimate potential use levels of targeted markets or sub-groups within the population in order to develop new service, and potential new facility locations. (Koontz, 1997) Now in order to do all this, I needed to know exactly where each branch was located so I could map it, and attach the variables to the proximate geographic market.

PROBLEM NUMBER ONE.

The addresses of the branches were not available from a national directory or database. Only the addresses of the stand alone or main facility of the system headquarters was available and the number of branches and bookmobiles within the system. There was little or no interest in branch level data at the state or federal level due to the following reasons: 1) there was not any branch level data available from a central source; 2) public library data was just starting to be collected for the first time ever on a system-wide basis, much less at the branch level; and 3) since there is no legal mandate for a public library (such as schools have) there was little built-in bureaucracy to support data collection on any level. Yet it should be noted that the majority of user activity, data for other industries, is of course, at the branch level. [The US National Center for Education Statistics published, through the US Government Printing Office (GPO), the first E.D. Tabs Report, Public Libraries in 50 States and the District of Columbia: 1989. The first compilation of data from 8,699 public libraries in all 50 states and the District of Columbia based on national standard data elements and definitions. The Public Library Data Statistical Report, Public Library Association, was published in 1988, till present, offering system-level data on a sampling basis. These publications did not include branch level use data at present, still do not, and there is no future plan by either group to make outlet level data available.]

Therefore, I had to call system administrators or branch managers of 13 library systems (450 outlets), and often the individual outlet, just to identify where each library was located (cross streets), so I could attach the variables to them (latitude and longitude). This was extremely time-consuming, but without this precise information there would be no way to continue, hence prohibiting my research.

PROBLEM NUMBER TWO.

Because there is no standard way that geographic branch market areas are determined, the branches are simply subsumed within what is called the legal service area of the system. These legal service areas are mostly arbitrary city and county boundaries, though some use school district boundaries. These legal service areas identify the population that the system is legally ascribed to serve, but in no way reflect who really uses each of the outlets in the system. Various methods of defining branch market areas include: assigning two or three census tracts to each branch; utilizing existing government planning districts; placing a radius of varying widths around each outlet; or, using the experience of the staff to determine proximate geographic area served and an estimate of the population in that area (VanHouse, 1987).

While retailers have long utilized customer zip code and address data to determine geographic customer markets - which provides much greater accuracy - library professionals have been reticent to do so, due to privacy issues. Yet, as early as 1972, it was suggested that libraries utilize user address data to better determine user branch market areas (Coughlin, 1972). I used the radius method of various sizes based upon estimated populations served to standardize the methodology. (These estimates were from the librarians, or staff. I used SELS to estimate the radii that the population encompassed.) I used compass and pencil based upon casting radii upon estimated population figures, and a GIS (geographic information system) type product that would estimate population. GIS is best defined as a computer system that maintains databases much like other computer systems, except the databases include geographic references such as address, latitude and longitude, census tract, voting district or a unique library drawn service area. It is possible to examine, for example, which library is nearest another library, which transit line is close to a new or proposed site, which addresses of users lie outside the designated library service area, and the demographics of the geographic service area.

This was in the mid-eighties, and the primitive geographic information system software (GIS) package I used was available for free from the University of Nebraska. It was free because it was 1989 when I procured the package, and the Census data was from 1980. The SELS package (Siting Evaluation and Location System) would identify census variables around a point, i.e., the geocode of the public library outlet by latitude and longitude. There were no digital maps available (cheaply) at that time. SELS required using paper census maps, identifying the location, and all the statistical calculations had to be performed outside the SELS package.

With paper maps and precise cross streets and locations, I could graphically identify topographical barriers to library use, and estimate the distance between facilities. With the automated census data, I could identify population characteristics proximate to each outlet. I could also utilize important research that others had conducted at the outlet level, including the impact of the size of the library, distance between the library facilities, and demographic characteristics associated with all aspects of the library use, including mode of transportation, groups affected by distance to the library, etc. (Koontz, 1997: 57-58)

PROBLEM NUMBER THREE.

While I had 13 library systems willing to work with me, only 6 could be included due to the other 7 not collecting standard library use categories. For example, all 13 systems did not collect reference transactions. Only those systems that collected circulation reference statistics and program attendance could be included. While library visits, in-library material use, and inter-library loan were of interest, these were not counted in a standardized or methodical fashion by all of the libraries, and therefore could not be included.

In summary, it was lengthy and arduous because the needed technology had yet to arrive. There was a lack of precise address data for branch outlets; lack of sophisticated GIS, which was an expensive and rapidly changing technology; lack of a standardized method of determining branch market areas; and a lack of standardized library use categories that all public libraries collect. Technology - holding hands with standardization of collected library use categories and an awareness of the importance of branch level data - could have greatly facilitated this research.

Researching at the branch level presented a significant finding: analysis of the demographic and library use data - showed where counted, there was higher in-library use, higher reference transactions and greater program attendance in branch markets serving predominately minority populations (this has been coined *majority-minority*.) I could not have detected this if the data was aggregated at the system level. These populations had lower circulation, which often led to an assumption that the libraries were less used because circulation is traditionally recognized as the primary use. Unfortunately, in-library material use and library visits are not required use counts. Due to this, thousands of libraries appear little used, as opposed to the truth: that they're just differently used. This was an important finding that led to the next project.

PROJECT TWO: EXPANDED NATIONAL RESEARCH

After graduating and landing this current job in a GIS (geographic information systems) applications center, I settled in to use this technology to solve library problems. In 1996, FSU applied for and received a grant to

further the above hypothesis that use was different in majority minority library markets and that it was not reflected completely in circulation figures (Koontz). The purpose was to provide a snapshot of non-circulation use in the library, including material uses, observed uses, and librarian assisted uses.

This project not only used GIS and U.S. Census Data to identify the 3500 majority minority markets, but incorporated handheld computers to facilitate counting in-library use by scanning barcodes of previously defined categories of use, including in-library materials use, observed use, and librarian assistance. This standardized one-barcode data was uploaded and mailed to our center for analysis.

This reduced the need for extra manpower in the field to analyze data, collected one week per quarter at these small libraries, and thus gained staff support. Finally, this project elicited a clear picture of all uses, and underlined the critical need to collect in-library use data in low-circulation libraries. Enhanced technologies, specifically GIS to process outlet level data based upon a growing body of research, gave this project national impact. Yet diffusion of this methodology is hindered by lack of a mandate at any level of government or professional organization to collect these types of use that occur in the library. This project led to projects three and four.

PROJECT THREE: LIBRARY APPLICATION OF GIS

This research was designed to solve one of two remaining problems illuminated in projects 1 and 2. First, at present, there is still no standardized approach to determining branch market areas other than those manual and experiential methods discussed earlier. Yet there is technology available, GIS, to use to geocode user address data based upon library circulation or registration records. We received a small research grant from ALA to illustrate the utility of using GIS to conduct this type of research (Jue and Koontz, 1999). User data from three library systems was geocoded for best determining actual market areas. This methodology far surpasses any other method in accurately depicting geographic market areas.

While the three library systems in this project were able to provide this type of data, there is no assurance that all library systems can use this approach due to lack of software programming to extract user address data from the library's database and policies or legalities against extracting this data to a vendor or researcher to put together such a map. This resistance from the profession emanates from old user privacy issues, laced with misinterpreted portions of the American Library Association's Library Bill of Rights, i.e., "a person's right to use a library should not be denied or bridged because of origin or age." There is a feeling that the dot on the map portrays address and all aspects of user characteristics. I challenge this on

the basis on sanity. But this research project did not solve the basic underlying dilemma, which led to Project Four.

PROJECT FOUR: COMPREHENSIVE GEOMAPPING OF PUBLIC LIBRARY LOCATIONS

Our center was recently awarded a grant to map and geocode all estimated 16,000 outlets (Jue and Koontz, 2000-2001). The outlet addresses are collected by state under the direction of the Federal State Cooperative System, and many (approximately one third) have addresses that could be accurately geocoded (a dot placed accurately on a map as to the most accurate location).

This project replicates Project One, my dissertation. Yet while this project utilizes and benefits from new technology and increased manpower, it was still constrained by lack of standardized data collection, lack of accepted methodology for geographic market determination, and lack of outlet data from a central source.

Four students are calling each of 5 to 6,000 outlets for which better location information is desired. While I began pushing for a nationally geocoded public library base map seven years ago, to be honest, it would not have been as cheap as it is today: GIS was more expensive, geocoding was more expensive, etc. The enormous breakthroughs remain to be seen when this map is finally available in early 2002. This map will provide a first-ever, base map for researchers from all fields to utilize to identify the impact of public libraries in many areas of American life.

We hope to see a progression of data layers that will be added, including, initially: relevant demographics associated with library use; local library use figures in population ranges for comparison for same-size libraries; and accurate legal service area delineators. In the future such layers could be added as telecommunication networks, bandwidth, etc. With a first-time-ever accurate base map of public library outlets, the doors open to research that can build upon accuracy, not guesswork. It is just the beginning. Rapidly changing technologies greatly facilitated this body of research that started out manually, but to further this research, important policy decisions must be made by government and library and information leadership, and funding must be prioritized to further develop this landmark database.

Reality and Myths: Part One

The reality is in these projects. Technology:

1. Enhanced the quantity and quality of data that could be effectively amassed, i.e., branch level data including the branch address, library use and demographic data and all other spatial aspects such as distance between the facilities, schools, etc.;
2. Facilitated the collection of difficult to count performance measures;
3. Reduced the amount of time that would be needed to process and analyze data. (No joke, I estimate I would have added four years to my life if I had been using full-blown GIS in my dissertation);
4. Reduced the need for more manpower to conduct research;
5. Made the data digital – and therefore accessible and available in aggregated and disaggregated forms for other researchers.

The above realities are myths – if the maxims below are not true.

Truth #1: Technology is useless without a purpose. That is why it lives and dies. In this day and age more is embodied in less; smaller and smaller packages contain more and more data. Mainframes are a thing of the past for all but the largest entities. The 8-track had its day; now music is binary. By the way, or BTW, happy birthday to the personal computer (PC); it is 20 years old this week. Can you remember life before PCs and email? The PC and email are useful. To procure technology because it is new, and learn demonstration packages with no purpose in mind is a worthless activity. And that research can now be replicated nationwide on the Internet and further gain in value, but only if the original research is of quality. Otherwise, technology is useless.

Truth #2: Technology is useless if it makes something that is ugly and has no content pretty. Having sat through too many PowerPoint presentations with no content... I do not want to rag on and on about this. But refer to the hand-hewn radii in Figure 1. While ugly, these hand-drawn circles communicated that to conduct research we must know exactly where a library branch outlet is located in order determine which customers use it. This led to actually procuring user address data and plotting it with GIS, changing the circle into an actual custom-made geographic market area.

And finally, for the first time ever an accurate national database of geocoded public library branch outlets is being developed to further and facilitate policy and research. Not bad for an ugly hand-drawn circle full of top notch data.

Truth #3: Technology is a tool for people. People are not a tool for technology or to increase vendor profits. We must shape, develop and demand vendor products that meet 100% of our needs. Our customer database needs to be as slick as American Express. We need to be able to data mine our customers so we can better meet and predict customer needs. We must stop pussy-footing about false privacy issues that do not exist. A dot on a map, signifying address, does not tell you what they read, how, and with whom. It is a spatial guidepost of a market.

Truth #4: Technology is misused if it hinders calculating all the results. While automated circulation systems greatly enhanced knowledge of library use and user characteristics - e.g., number of juveniles checking out science fiction, number of adults utilizing genealogy services, and circulation per capita of various services and materials - it took the emphasis off important hand-counts such as reference transactions, program attendance and in-library use. Without full counts of all use, we are shortchanging ourselves. My analogy - what if McDonald's only counted all the hamburgers sold through the drive-through and none eaten in the restaurant?

As we rush from inputs to outputs, outputs to outcomes, web hits to web use, we must be vigilant that technology is our tool, not our master. We must be technology's bully, and technology's Pied Piper. Technology should never serve either of those roles. We must join together as a profession, unify, and create standardization of important library use counts. Some counts are more important for some markets than others, and not everything must be counted every day; some use can be sampled. With the help of technology we must keep our data disaggregated and develop reliable accurate standardized databases and base maps to build our research upon. We must shove aside false ideas regarding the parameters of privacy of users and demand that vendors supply us with software that will provide top quality customer data. Then, and only then, will we be the true Pied Pipers of our villages, bringing all those good folks to the library to find their (customized!) pot of gold.

References

- Coughlin, Robert, et al. (1972) *Urban Analysis for Branch Library System Planning*. Westport, CT: Greenwood Press. p. 103
- Jue, Dean K. and Koontz, Christine M. (2000-2001) *GeoMapping of Public Library Outlets Project: Phase One*. Funded by the National Commission on Library and Information Services and the National Center for Education Statistics. Florida State University, GeoLib Program, 2001-2002
- Jue, Dean, K. and Koontz, Christine M. and (1999) *Library Service Area Determinations through User Address Data and Use of Geographic Information Systems*. Loleta Fayan Award, American Library Association
- Jue, Dean K. and Koontz, Christine M. *Market-Based Adult Lifelong Learning Performance Measures for Public Libraries Serving Lower Income and Majority-Minority Markets*. Field Initiated Studies Educational Research Grants, Office of Educational Research Grants, Office of Educational Research and Improvement, U.S. Department of Education
- Koontz, Christine M. (1997) *Library Facility Siting and Location Handbook*. Westport, CT: Greenwood Press
- Koontz, Christine M. (1990) *Market-Based Modeling for Public Library Facility Location and Use-Forecasting*. Unpublished dissertation. Tallahassee, FL: Florida State University
- Van House, Nancy, et al. (1987). *Output Measures for Public Libraries*. Prepared for the Public Library Development Project, 2nd ed. Chicago, IL: American Library Association. pp. 35-36