Business Data: Issues and Challenges from the Canadian Perspective

by Eun-ha Hong and Linda Lowry*

Introduction

This paper explores the issues and challenges that we have faced as Canadian academic business librarians when working with business data. As this is an exploratory study, we hope only to start a discussion among data librarians about some key challenges facing the academic community related to supporting the teaching and research use of business data. Our paper begins with a brief discussion of general data trends, followed by a detailed exploration of business data trends and trends in Canadian business education. We discuss challenges and issues related to working with business data from both the collections and reference service perspectives, including the pros and cons of providing business data services and support within the library environment. We conclude by suggesting some measures that both academic business librarians and data librarians can take to address some of these challenges.

General data trends

Halliwell observed that over the last twenty-five years, Canada's quantitative researchers have benefited from a substantial increase in data supply, but there has been an even larger increase in the demand for data[1] While he is referring primarily to the supply of and demand for government-produced data sets (from Statistics Canada or other government agencies), his observations could also apply to the demand for data for academic research, particularly in the social sciences. Davis and Vickery recently identified several current trends that indicate the growing importance of data sets (over journal articles) as a unit of information currency, including the transformation of data sets into valuable economic commodities, an increase in legislative initiatives to protect data sets, the sheer growth and manipulability of data sets, evolving publication expectations placed on authors, and public/private partnerships developing around data sets[2]

Business data trends

As Canadian academic business librarians, we have also observed an increase in the demand for business data. Before we explore business data trends in more detail, it would be prudent to define what we mean by business data. Business researchers either gather their own data (via surveys or other means) or rely on secondary sources of data. These secondary data sources may be divided into two types: social science data and commercial data. An example of a social science data set frequently used by business researchers is Statistics Canada's CANSIM database. Researchers seeking international data also rely on data from international governmental and nongovernmental organizations such as the International Monetary Fund or the Organization for Economic Cooperation and Development. Most of these data producers are not-for-profit organizations who offer their data sets to universities at little or no cost. On the other hand, commercial data sources are used to obtain accounting and financial data such as company financials and stock market data. An overview of the core numeric business databases used in Canada can be found in *table 1*. These commercial data producers include data aggregators such as Thomson Financial's Datastream or Standard & Poor's Compustat, the University of Chicago's Center for Research in Stock Prices (CRSP) and financial industry sources such as the Toronto Stock Exchange. Many of these commercial data producers are profit-seeking organizations, so while they may offer an academic discount, their primary target market is not the university sector and their data is often very costly to acquire.

Database Name	Company Financials	Equities	Fixed Income	Derivatives / Futures	Macro-eco-nomics
CFMRC (TSX)		Canada			
CRSP US Stocks		US			
Compustat North America	US Canada				
Datastream Advance	Global	Global	Global	Global	Global

NOTE: Company financials includes financial statements (balance sheets, income statements) & financial ratios. Equities includes stock prices (open, close, return, volume, beta).Fixed Income includes bonds and treasury bills. Derivatives / Futures also includes commodities, options & warrants. Macroeconomics includes items such as interest rates, exchange, GDP

Table 1 Core numeric business databases, data types, and geographic coverage

Trends in Canadian business education

In this section, we examine the current situation in Canadian postsecondary business education and identify some key trends witnessed at our own institutions. An increasing number of research-oriented graduate programs in business are being offered at Canadian universities which are driving the growth in demand for business data. These programs include thesis-based master's degrees in business economics, business administration, and management, and future plans for PhD programs. Increasingly, Canadian business schools are seeking accreditation from the Association to Advance Collegiate Schools of Business International (AACSB), partly as a means of staying competitive in the national and international marketplace for both students and faculty, and in order to increase the emphasis on research[3.] There is an increasing focus on the research productivity of business school faculty. This is partly in response to the results of a study conducted by Erkut in 2002 to measure the output and input of Canadian business school research which found, among others, that the paper output of Canadian schools is relatively low and declining, there are significant differences among Canadian business schools in research output and impact, and a few 'stars' produce most of the impact[4]. Perhaps in response, some Canadian business schools have implemented research performance management incentive reward systems to improve the quantity and/or quality of publications[5.] Canadian business schools are also engaged in bidding wars for newly minted accounting and finance PhDs who do empirical research, as there is intense competition from the private sector in these fields[6]. In addition, faculty and students are increasingly demanding remote access to business data, as well as 'one-stop-shopping' research portals such as Wharton Research Data Services, where scholars can access data sets from multiple vendors within a uniform interface[7..]

What is the current context for postsecondary business education across Ontario's twenty universities? Most offer an undergraduate degree in business as well as an MBA program. However, fewer than half of these have doctoral degree programs in business administration, which has implications for business faculty and their access to research support and resources. A 1996 study by Tompkins, Hermanson and Hermanson on expectations and resources associated with new finance faculty positions revealed important differences in expectations and resources between three types of institutions: AACSB-accredited doctoral schools, AACSB-accredited non-doctoral schools, and non-AACSB schools[8.] They found substantial differences in data resource availability, whereby new faculty at doctoral institutions could expect plentiful research resources and excellent database access, compared to new faculty at non-doctoral institutions who could expect 'reasonable' research resources, including database access, and new faculty in non-accredited, non-doctoral business schools who could expect moderate to poor

research resources and poor database access[9.]

As noted earlier, data demand in the academic disciplines of finance and accounting can often only be met by commercial data sources with high costs that are often difficult to obtain given the funding situations of the researcher's parent institution. A recent study explored the effects of database choice on international accounting research[10]. In accounting and finance research it is the previous literature, specifically benchmark papers, that set the standard for database choice. In the United States, data comes primarily from Compustat and CRSP, but the choice is less clear-cut in international accounting research due to factors such as a lack of tradition and a wider choice of databases[11] Due to budgetary constraints, universities will only invest in databases where there is an assurance of usage and therefore the type of research conducted by PhD students and faculty is driven by database availability.

Business data issues from the collections perspective

Librarians involved in developing data collections encounter five main business models: the institutional membership model (e.g. ICPSR); the serials continuation model (i.e. by subscription); the one-time payment model; the ad-hoc arrangement; and the lack of a business model (where no process exists to deal with the needs of academic libraries)[12] While business data producers typically follow the serials continuation model, many business school faculty mistakenly believe that the one-time payment model is in place, thus resulting in mismatch between funding sources (one-time research grants) and funding needs (ongoing subscriptions). It has been our experience that some faculty who were conditioned to use particular databases for research as doctoral students are now demanding that their institutions acquire these, even if their own institution does not offer a doctoral program in their discipline. In other cases, new hires are being promised new business databases by business school deans who may not realize the financial implications of such promises.

Many libraries lack a formal collection development policy for numeric data which often limits the level of support for numeric data files[13] Ideally, business data collection development decisions are made proactively and collaboratively such that faculty and librarians work together to investigate and select data sets for purchase in order to meet current and/or future teaching and research needs. However, our experience tells that this is not always the case, as many data sets are purchased unilaterally by business faculty without any library involvement. These unilateral decisions may be partly due to a lack of understanding among business faculty of library collection development practices, or may arise in cases where the library is unwilling or unable to commit funds to purchase a particular database.

So, how prevalent are numeric business databases in

Ontario academic libraries? We compared the holdings of four core numeric business databases (CFMRC, Compustat, CRSP, Datastream) across 17 Ontario universities by examining library web sites (see table 2). The most frequently held database was CFMRC, followed by Compustat. Only five universities held all four of these products. Interestingly, we found universities with doctoral-level business programs that do not hold any of these four products. While it is hard to know why they do not, it could be related to the nature of their programs (perhaps lacking a focus in accounting or finance), or these data products may be available in the business school, but not listed on the library's web site. The latter reason, although rare, further complicates the support that librarians are able to provide for business data as the librarians they may not be able to access these data sets from the library.

CFMRC: 12 universities subscribe	5 universities have 4 of these
Compustat: 9 universities subscribe	3 universities have 3 of these
CRSP: 8 universities subscribe	4 universities have 3 of these
Datastream: 8 universities subscribe	5 universities have none of these
NOTE: Based on a review of the databases holdings of 17 Ontario universities with business or management programs as listed on library web sites.	
Table 2 Numeric Business Database Holdings in Ontario Universities	

Business data issues from the reference service perspective

In this section, we continue our exploration of business data by examining issues from the reference service perspective. A recent study by Bennett and Nicholson that investigated the interaction between business library services and research data services in academic institutions described numerous models for the location and provision of both types of services.[14] A snapshot of business librarianship at Ontario universities revealed that 13 of 20 universities employed one or more business librarians. Of these, five universities (all with doctoral programs in business) had a separate business library. The lack of a standard service model for academic business librarianship leads to inconsistent levels of service across the academic business community. For example, there is no consistency in the hiring practices of librarians, particularly business librarians. Unlike faculty, where the student-faculty ratio informs the need for additional faculty positions, increases in student-librarian ratios do not often result in additional librarian positions. This means that in some universities there is no dedicated business librarian, even when the size of the business student population warrants the need. This is a concern, because unlike the humanities or social science disciplines where subject librarians can easily apply their expertise to other disciplines when necessary, business

subjects require expert knowledge - even more so when librarians are dealing with business data.

Variations in the presence or absence of business and data librarians have influenced each university's ability to provide support for business data. Our own informal investigation of service models to support the use of business data for teaching and research in Ontario found that in some academic institutions it is the business librarian who provides research support, trains new users and troubleshoots problems, while in other academic institutions it is the data librarian, if there is one. Oftentimes, there is cooperation between business and data librarians where they exist within the same institutions. However, in some institutions, business database support is not provided by the library; instead, someone with business database expertise, for example a faculty member or researcher, trains new users and acts as a champion for business data. Just as university libraries vary greatly in terms of size, staffing levels and budgets, the data needs of their faculty and student clientele also vary. Business faculty and doctoral students have similar needs which can be quite complex in contrast with the fairly straightforward needs of MBA or undergraduate business students.

Pros and cons of business data services

One might ask, should libraries be providing these types of services? Or is it the responsibility of the business school? For example, in the library literature, we have found two completely contrasting views on the Compustat database: (1) "[Compustat's] format is more appropriate as an instructional aid and/or laboratory application than as a library reference product"[15], (2) "Compustat is an indisputably valuable research and education tool".[16] The arguments against libraries providing business data services stem from the basic fact that numeric databases are not a medium that libraries are used to taking care of (compared to bibliographic databases). Most librarians (business or otherwise) do not have the background or capability to understand and support numeric business data users. In addition, these databases are costly to subscribe to, difficult to use and time consuming to support, as users can have complex questions that take a long time to resolve. Even if a library offers a data service, the data librarian is likely to concentrate on social science data and may not be familiar with business data. Finally, it may not be possible to offer good service in particular library environments, as it can be especially challenging for librarians such as ourselves who are 'solo' business librarians in a general or central academic library, and who have few, if any, colleagues to provide backup reference support in our absence.

That being said, there are good reasons for advocating for libraries to provide these services. According to Rebecca Smith:

"libraries no longer have a monopoly on the provision

and distribution of information. As we face competi-tion for information stewardship, it is the quality of service we provide that will determine our fate... Users can already go around the library for some of their information needs... if users continue to perceive academic librarians as knowledgeable and service ori-entated the library will remain central to a university's mission."[17]

The business librarian, and the library in general, risks being marginalized by the business school and business data users. Business schools that bypass the library are usually the ones that have paid for the data themselves; usually when there is cost-sharing between the library and the business school, the library becomes more involved in providing services. Cynthia Lenox identified the following success factors for supporting Compustat in academic libraries: the degree of involvement of the teaching faculty and incorporation into the curriculum; continuing product improvement and ease of use; and the amount of computer support and training offered.[18]

Academic libraries wishing to provide reference support for business data users face two hurdles: low levels of data and statistical literacy amongst reference librarians coupled with a lack of business subject expertise. Data librarians know that the acquisition of statistical knowledge is important in order to support general data users, and have advocated for statistical literacy among librarians.[19] Data and statistical literacy, and financial literacy in particular, have also been identified as key skills to be acquired by business information professionals, not only within the business information community but from the business school.[20] While some information professionals expressed doubts about the legitimacy of their role in promoting and using numeric databases, Allen Foster advocated not only data retrieval skills, but a broad understanding of the nature of the data retrieved and how to manage that data.[21] With these skills, business information professionals can broaden their role to include establishing policies, training users, providing reference services and liaising with vendors.[22]

Conclusion and some suggestions
At the very least, academic business librarians are urged to become statistically literate, particularly if their library is not equipped to support numeric business databases through traditional data library services. Data librarians are urged to become financially literate, so that they can be capable of working with financial data. That way, if and when the library decides to offer such a service, they will be ready, as numeric business data comes with a steep learning curve. Canadian academic business librarians and data librarians must both recognize that the demand for business data services is only going to grow as Canadian business schools realize that "becoming a distributor of knowledge generated elsewhere is not an option for Canadian business schools; we must strive for research

leadership in business just as our fellow researchers do in science and medicine".[23] Such changes can be successfully met if business and data librarians embrace this change because, according to Marydee Ojala, "transformational librarians look at the changes in the information world and take advantage of those changes to enhance their roles".[24]

In addition, we advocate for a strong partnership between academic business librarians and data librarians to support business data use. Bobray Bordelon, speaking on the topic of collaboration between social science librarians and data librarians, suggested that since no one librarian can know enough to meet all the users' needs, subject and data librarians should form partnerships.[25] We believe these new academic business librarian-data librarian partnerships could take place within the same institution, or across organizational boundaries. Just as the Data Liberation Initiative supports the use of Statistics Canada data across Canadian universities, a similar 'community of practice' could be created to support the use of core numeric business databases. Perhaps IASSIST might be the natural home or sponsor for such a community of practice?[26]

Finally, we advocate for better relationships between commercial business data vendors and academic users, including better academic price discounts, better products, better documentation, provision of usage data, and more training. This has already been advocated in the IASSIST community with respect to international economic data.[27] A critical mass of academic business data users will be needed to convince commercial data vendors to change their business models and their products to better suit the academic market.

* Contact: Eun-ha Hong is the Business and Economics Librarian at Wilfred Laurier University (ehong@wlu.ca); Linda Lowry is the Business and Economics Librarian at Brock University (llowry@brocku.ca). This paper is based on a talk presented at the IASSIST Conference in Montreal in May 2007.

Footnotes
1 Cliff Halliwell, "Desperately Seeking Data," Horizons 8, no.1 (2005): 31-37.

2 Hilary M. Davis and John N. Vickery, "Datasets, a shift in the currency of scholarly communication: implications for library collections and acquisitions," Serials Review 33, no.1 (2007): 26-32.

3 Margaret McKee, Albert J. Mills and Terrance Weatherbee, "Institutional field of dreams: exploring the AACSB and the new legitimacy of Canadian business schools," Canadian Journal of Administrative Sciences 22, no.4 (2005): 288-300.

4 Erhan Erkut, "Measuring Canadian business school

research output and impact," Canadian Journal of Administrative Sciences 19, no.2 (2002): 97-123.

5 Linda M. Manning and Jacques Barrette, "Research performance management in Academe," Canadian Journal of Administrative Sciences 22, no.4 (2005): 273-287.

6 Gordon Pitts, "Business academics write their own ticket," The Globe and Mail, March 28, 2007: E1.

7 Davis and Vickery, "Datasets, a shift,":30.

8 James G. Tompkins, Heather M. Hermanson and Dana R. Hermanson, "Expectations and resources associated with new finance faculty positions," Financial Practice and Education, Spring/Summer 1996: 54-64.

9 Ibid.,:63.

10 Juan Manual Garcia Lara, Beatriz Garcia Osma, and Belen Gill de Albernoz Noguer, "Effects of Database Choice on International Accounting Research," ABACUS, 42, no.3/4 (2006): 426-454.

11 Ibid., 427.

12 Davis and Vickery, "Datasets, a shift,": 29.

13 William H. Walters, "Building and maintaining a numeric data collection," Journal of Documentation, 55, no. 3 (1999): 271-287.

14 Terrence B. Bennett and Shawn W. Nicholson, "Interactions between the academic business library and research data services," portal: Libraries and the Academy, 4, no.1 (2004): 105-122.

15 Alexia Strout-Dapaz and Dennis Odom, "Intercepting departmental fumbles and running with the ball" (paper presented at the ACRL 9th National Conference, Detroit, MI, April 8-11, 1999), http://www.ala.org/ala/acrl/ acrlevents/strout99.pdf (accessed April 21, 2007).

16 Cynthia Lenox, "Success factors for supporting Compustat in academic business libraries," Business & Finance Bulletin, 109 (1998): 41-49.

17 Rebecca Smith, "Product management: a new skill for reference librarians?" Reference & User Services Quarterly, 39, no.3 (1999): 121-127.

18 Lenox, "Success factors,": 46-57.

19 Ann Gray, "Data and statistical literacy for librarians," IASSIST Quarterly, Summer/Fall 2004: 24-29.

20 Rita Marcella, "View from a business school: an

interview with Professor Rita Marcella," Business Information Review, 24, no.1 (2007): 30-35.

21 Allan Foster, "Online numeric databases: four years later," Business Information Review, 5, no. 3 (1989): 3-12.

22 Ibid.,7.

23 Erkut, "Measuring Canadian Business School research output and impact", 119.

24 Marydee Ojala, "Journeys and transformations," Online, September/October 2006: 5.

25 Bobray Bordelon, "Experience-education-interest: a collaborative approach to data reference and interpretation", presentation at the Digital Library Federation workshop on social science data archives, January 1999, http://www. diglib.org/collections/ssda/ssdaresults.htm (accessed April 29, 2007).

26 For overview of communities of practice, see: Etienne Wenger, "Communities of practice: a brief introduction," http://www.ewenger.com/theory/ (accessed February 5, 2008).

27 Bobray Bordelon, "Cross-national & intergovernmental data: paying for one-stop shopping," IASSIST Quarterly, Fall 2005: 5-7.