

Big Data's Role in Information-Centric Organizations

INFORMATION PROFESSIONALS CAN EXTRACT VALUE FROM BIG DATA TO SUPPORT EFFORTS WITHIN THEIR ORGANIZATIONS AS A WHOLE AND WITHIN THEIR LIBRARIES AND INFORMATION CENTERS.

BY ELAINE TEAGUE, MSLS, AND JULIE LEGEROS, MLS

Big data. Everyone is talking about it; in fact, it seems to be the buzzword du jour. But, what is big data—or, more to the point, what's the big deal about big data? What does big data mean for information professionals, and what impact will it have on us?

Gartner, a leading information technology research and consulting firm, defines *big data* as “high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making” (Gartner 2014). Thomas H. Davenport, author of the recently published book *Big Data @ Work*, puts it more simply: big data, he says, is “perhaps the most sweeping change in what we do to get value from data since the 1980s” (Davenport 2014).

Just how big is big data? Enormous and getting more so, according to a report published by IDC, a provider of market intelligence to IT and telecommunications firms. “The digital universe in the United States—the digital bits created, replicated, and consumed each year in the country—is expected to grow from 898 exabytes to 6.6 zettabytes between 2012 and 2020, or more than 25 percent a year, which means it will double about every three years” (Gantz and Reinsel 2013). IDC also forecasts that the market for big data technology and services will grow at a 27 percent compound annual rate, to \$32.4 billion through 2017 (IDC 2013).

While there is no disputing that the quantity of data is immense (be it text, voice, images, or video), the real revolution is that we can now do something with this data using analytics and there-

by extract some meaningful value from it. A recent Bain & Company report revealed that companies that use analytics insights perform better financially than their counterparts that do not use analytics, are three times more likely to execute decisions as intended, and are five times more likely to make decisions faster (Wegener and Sinha 2013). When you get down to brass tacks, data without context is just data.

So, how are organizations using big data analytics? Retailers are analyzing customer activity to guide product placement decisions; colleges and universities are using predictive analytics to determine the success rate of students; and governments are developing models to predict criminal activity.

Our employer, SAS, lives and breathes big data and analytics, both for our customers and our employees. The



ELAINE TEAGUE is the manager of the SAS Library & Information Services. She has worked in corporate libraries for more than 30 years and has been a member of SLA all that time. She can be reached at Elaine.Teague@sas.com. **JULIE LEGEROS** is a senior information and research librarian at the SAS Library & Information Services. She has worked in special libraries for more than 30 years and has been a member of SLA for 23 years. She can be reached at Julie.Legeros@sas.com.



authors, as information professionals, play an integral role in providing context and meaning to big data and thereby increasing its impact.

Aligning Our Services

What role can you play to help support and use big data within your organization? You might be surprised at how much you are already doing. For example, are you providing access to online databases that allow users to view articles from magazines, journals, or proceedings? Are you answering research questions about big data? Are you providing a collection of print books or e-books that users can borrow?

It can be easy to get overwhelmed by big data. At SAS, our library staff consists of four information professionals, and we (like many of you) have had to consider the possibility of staffing reductions. We decided to focus on the services we provide that are aligned with our organization's big data and analytics goals. This meant we had to re-examine and redefine our library's mission statement, which allowed us to shift our services to meet corporate goals and objectives. As a result, we were able to introduce new services that can help the library, and SAS, remain relevant.

For example, we listen to all executive-level internal Webcasts, prepare a written summary of initiatives that were presented, and implement services that align with these initiatives. We also partner with our Competitive Sales Intelligence group to stay up to date with the activities of our major competitors, and we conduct research on lesser-known big data competitors.

We perform a monthly analysis of items that are borrowed from our collection and the searches that are conducted in our online catalog and make the necessary purchases to supplement our collection. For example, a significant interest in Hadoop (open source software for big data) led us to review our print and e-book collections and beef up our selection. In addition, we created an online Snapshot (subject guide) on Hadoop that provides our

users with links to analysts' reports and external thought leadership.

We offer an online avenue for employees to recommend books, resources, or services that will help others do their jobs more effectively or gain needed skills. To market and advertise our services and resources, we publish blog posts regularly on the SAS intranet and provide a mechanism for employees to receive posts through a news feed.

We publish daily posts on the SAS internal Facebook page (The Hub) that provide links to external thought leadership on topics that support SAS' industry-specific solutions. More than 10,000 employees have joined The Hub, so this platform provides high visibility for the library. (In fact, the library is one of the 10 most active Hub groups.)

We offer a weekly "Library 101" orientation that showcases the ways the library can connect employees with the information they need. This orientation is available for new employees as well as those who need a refresher.

We participate in a monthly new-hire sales training program that allows us to exhibit the services and resources we offer. We also provide regular orientations and training for departmental and divisional employees worldwide.

We subscribe to Lynda.com so employees can take courses on topics such as Hadoop, big data, or analytics using an "anytime / anywhere" model. (Self-paced online training is in high demand at SAS and is critical to our success in developing, selling, implementing, and supporting our software.)

We continue to expand our collection of e-books so employees can tap into the most current and authoritative content on big data and analytics by reading books online or downloading them to the device of their choosing. (According to PricewaterhouseCoopers, e-books will account for 22 percent of all global book revenues by 2017.)

Big Data and Your Operations

In addition to putting context and meaning around big data for your organization as a whole, you can examine and analyze the big data in your library or

information center. You might be surprised to realize that you are already collecting and using big data in your operations. Think about the Gartner definition of big data at the beginning of this article: "assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making." Then think about the massive amounts of data that libraries and information centers gather on a daily basis, such as circulation, acquisitions, cataloging, and other types of data from an integrated library system, usage data from online databases, Website visits and downloaded documents, and research requests and reference questions.

In our library, we collect and analyze the following types of data, paying particular attention to the value they can provide:

- We use SAS software (SAS® Enterprise Guide®) to generate textual and visual data to help with collection development. We analyze which titles, subjects, and keywords are searched most often and which materials are frequently borrowed. These analyses, in turn, help drive our purchasing patterns.
- We also use SAS software to assist our marketing efforts by identifying employees who are frequent library users, then targeting our outreach efforts to those users. By manipulating this data, we can drive usage of our resources and services.
- We use Web analytics software to monitor Website traffic and generate textual and visual reports. These insights have allowed us to continue offering shared access to industry and analyst reports, thus saving SAS money on employee purchases.
- Using an in-house electronic resource management (ERM) system, we manage more than 80 online database contracts and 9,000-plus employee accounts. Through a Web-based reporting application, library staff and company executives can view cost and

usage data in real time. The data that is generated is used to solicit employee testimonials and justify renewals.

- Library users can submit research requests using a Web client interface. Through a back-end report wizard, library staff can generate reports of the completed research requests by user, division or department, geographic location, and other data points. With SAS® Enterprise Guide®, we can generate visual and textual data that enable us to gather user-specific anecdotes on the value that the research provided to SAS.

By analyzing our data stores, we were able to demonstrate that the research projects we completed, the print and e-books that we loaned out, and the Snapshots that were used all contributed to the rollout of SAS In-Memory Statistics for Hadoop, a new SAS solution. As a result, we were able to trans-

form big data into useful insights.

In our more than 60 years of combined corporate librarianship, we have seen a lot of changes. We are thankful that typing and filing catalog cards is a thing of the past. We have online databases instead of directories in multi-volume sets. Our worldwide employees can access library resources any time of day. As Thomas H. Davenport reminds us in *Big Data @ Work*, big data provides us with an opportunity to embrace this sweeping change in what we do to get value from data. So, let's get it! **SLA**

REFERENCES

- Davenport, Thomas H. 2014. *Big Data @ Work: Dispelling the Myths, Uncovering the Opportunities*. Boston: Harvard Business Press.
- Gantz, John, and David Reinsel. 2013. *The Digital Universe in 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East-United States*. Framingham, Mass.: International Data Corporation.

Gartner Inc. 2012. "Gartner Says Big Data Creates Big Jobs: 4.4 Million IT Jobs Globally to Support Big Data By 2015." Press release. 22 October.

_____. 2014. "IT Glossary: Big Data." Web page. Accessed 25 March 2014 at <https://www.gartner.com/it-glossary/big-data/>.

International Data Corporation. 2013. "New IDC Worldwide Big Data Technology and Services Forecast Shows Market Expected to Grow to \$32.4 Billion in 2017." Press release. 18 December.

PricewaterhouseCoopers. 2013. Global Entertainment and Media Outlook 2013-2017: Consumer and Educational Book Publishing. Web page. Accessed 25 March 2014 at www.pwc.com/gx/en/global-entertainment-media-outlook/segment-insights/consumer-and-educational-book-publishing.jhtml.

Wegener, Rasmus, and Velu Sinha. 2013. *The Value of Big Data: How Analytics Differentiates Winners*. Boston, Mass.: Bain & Company.

EDUCATING DATA WORKERS FOR THE 21ST CENTURY

What training and skills are needed to prepare the workforce of the 21st century to work with big data? According to Gartner (2012), the demand for data and analytics workers will reach 4.4 million jobs globally by 2015, but only one-third of those jobs will be filled. The emerging role of data scientist is meant to fill that skills gap.

Our employer, SAS, is a leader in business analytics software and services and the largest independent vendor in the business intelligence market. SAS is looking for individuals who have excellent programming skills and the ability to deal with very large volumes of data. At SAS, a job posting for a data scientist requires a master's degree in statistics, mathematics, computer science, engineering, or the physical sciences. Also required are two years of related experience, such as analyzing data and building analytical models.

Customers of SAS have stated that they need more people who are capable of leveraging world-class business intelligence systems. To address this need, SAS partnered with North Carolina State University to create a master's degree program in analytics, which is housed in the university's Institute for Advanced Analytics. It was the first graduate degree program in analytics offered in the United States.

Since the program's inception in 2007, SAS has provided strong support in the development of the curriculum.

The program has been so successful that each graduate receives an average of 10 job offers. More than 50 higher education institutions worldwide are now offering graduate degree programs in analytics.

"As leading organizations continue to recognize the value of advanced analytics tools and skills in increasing their competitiveness, we will see increasing need for this talent," says Radhika Kulkarni, vice president of advanced analytics at SAS. "Without the right tools, people and process to gain insight from the data, it is easy to starve for information while simultaneously drowning in data."

The rising importance of analytics can be seen at the secondary school level. A high school in Charlotte, North Carolina, the Phillip O. Berry Academy of Technology, is helping high school students prepare for a world of big data. By offering an SAS programming course, the school is preparing students for college and careers using analytics.

The commitment of SAS to education and the importance of analytics was recently strengthened with the announcement of SAS Analytics U. This program provides free SAS software for academic research and educational purposes within higher education. With this program, students will be able to experience the depth and range of SAS analytics in their research as well as increase the pool of next-generation data users.