EDUCAUSE
CENTER FOR ANALYSIS AND RESEARCH

BI Reporting, Data Warehouse Systems, and Beyond

CDS Spotlight Report

Research Bulletin
April 23, 2014

Leah Lang, Manager, Core Data Service, EDUCAUSE

Judith A. Pirani, Consultant, EDUCAUSE, and President, Sheep Pond Associates

Overview

This ECAR research bulletin series highlights findings from the EDUCAUSE Core Data Service, focusing on a small but meaningful slice of data collected in CDS. These selected highlights are intended to provide context and meaning for CDS benchmarks that may be of especially broad interest, be especially timely, or draw connections between research from ECAR and CDS. The series is featured along with other CDS publications on the CDS website at http://www.educause.edu/coredata and is now available to eligible ECAR subscribers as part of their subscription.

This Spotlight focuses on data from the 2013 Core Data Service to better understand how higher education institutions approach business intelligence (BI) reporting and data warehouse systems (see the sidebar for definitions). Information provided for this Spotlight was derived from Module 8 of CDS, which contains several questions regarding information systems and applications. Responses from 525 institutions were analyzed. Only U.S. institutions were analyzed for this report. In addition, we interviewed four subject-matter experts to gain insights about the current and future state of BI reporting and data warehouse systems in higher education.

BI Reporting and Data Warehouse Systems: Now and Beyond

Data collection, management, and analysis continue to grow more important in higher education in response to issues such as regulatory and accreditation requirements, student success measurement, and operational efficiency expectations. In addition, parent and student use of personal online banking and shopping applications fuel similar capability and data-transparency expectations in areas such as tuition payments and academic performance evaluation.

These pressures promote an ongoing cultural shift in higher education, one that embodies greater measurement, assessment, and accountability. With this change comes an increasing recognition of the tactical need and strategic value of using institutional data. BI reporting and data warehouse systems are two mainstays in an expanding portfolio of tools and capabilities that enable senior leaders and others to unlock the inherent value of institutional data to address compliance and regulatory needs, to enhance users' experience, and to monitor progress on short-term goals and long-term strategies. Perhaps this is why these systems are the fifth (BI reporting) and sixth (data warehouse) most rapidly changing core system areas (see figure 1).[1]

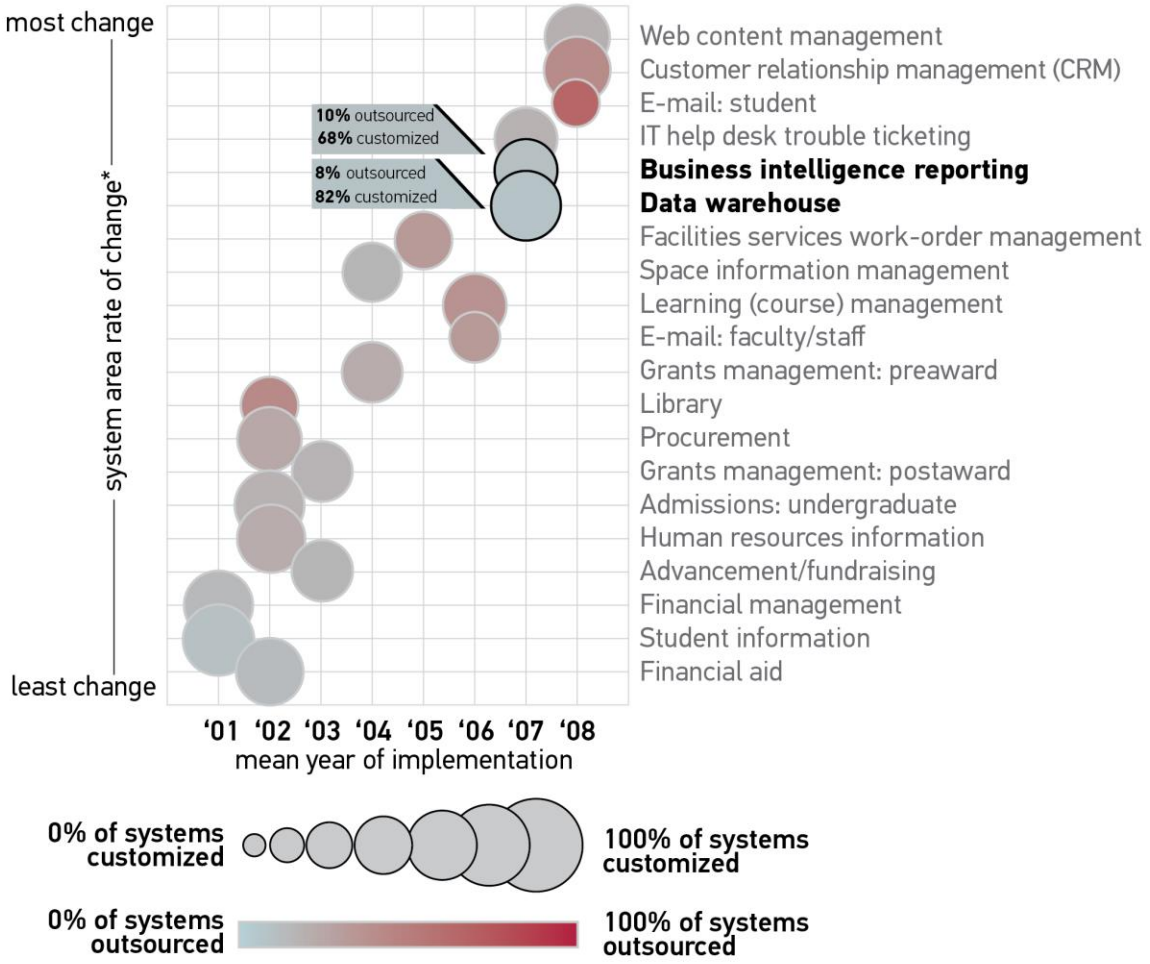

Figure 1. Characteristics of Core Information Systems

Current Snapshot

Institutions with BI reporting and/or data warehouse systems are poised to innovate in their use of institutional data. According to CDS, in FY 2012/13, BI reporting and data warehouse systems were relatively common at U.S. institutions (see figure 2). Approximately 80% of U.S. institutions had some type of BI reporting system, and 71% of U.S. institutions had some type of data warehouse system; both system types are more pervasive at larger schools than smaller schools. The presence of these systems has increased in the past three years, and it appears this growth will continue in the next three years. CDS data from FY 2010/11 show that only 62% of institutions had a data warehouse/BI system. In the next three years, an additional 8% of institutions intend to implement a BI reporting system, and an additional 12% intend to implement a data warehouse system.

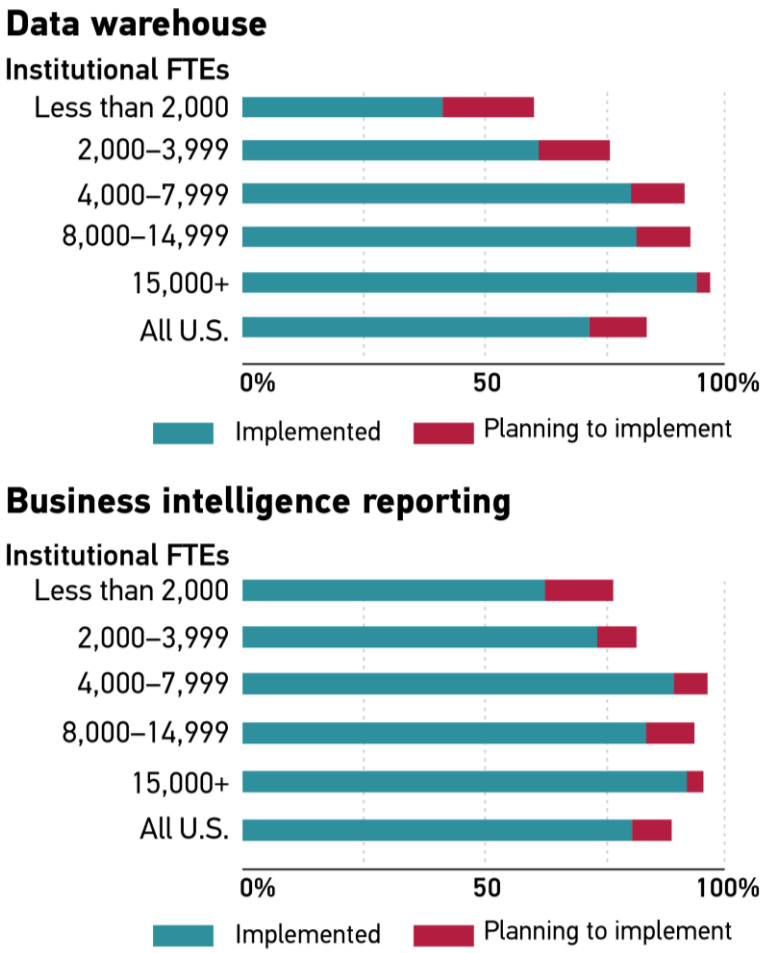

Figure 2. Percentage of U.S. Institutions with Systems Implemented or Planned Implementations

As institutions implement new systems, the debate over centralized versus decentralized management will continue. Most institutions with existing systems rely on centrally managed systems to enable better data management in a number of ways: a central extract, transform, and load (ETL) layer for information access; a common toolset for reporting; and "de-siloed" information for common data sets. About 62% of U.S. institutions have BI reporting systems provided by central IT; 59% of U.S. institutions have data warehouse systems provided by central IT. The theme of central control is echoed with an overwhelming preference for in-house management of both BI reporting (90% of U.S. institutions with a BI reporting system) and data warehouse systems (92% of U.S. institutions with a data warehouse system) over vendor or cloud-based solutions. In fact, these systems are the least likely to be outsourced out of our list of core systems (see figure 1; BI reporting outsourcing ranking = 19, data warehouse outsourcing ranking = 20).

CDS Spotlight Definitions

ad hoc query: A nonstandard inquiry to obtain information as the need arises. This contrasts with a query that is predefined and routinely processed.

analytics: The use of data and statistical analysis and explanatory and predictive models to gain insight into and act on complex issues.

big data: Massive amounts of data, collected over time, that are difficult to analyze and handle using common database management tools. Big data can include business transactions, e-mail messages, photos, surveillance videos, and activity logs.

business intelligence reporting: A set of administrative functions and associated software systems that support planning and decision making by categorizing, aggregating, analyzing, and reporting on data resulting from transaction-processing systems.

data warehouse: A central repository of data often created by integrating other data sources and used for reporting and analysis.

extract, transform, and load (ETL): The functions performed when pulling data out of one database and placing it into another of a different type. ETL is used to migrate data, often from relational databases into decision-support systems.

Hadoop: An open-source project from the Apache Software Foundation that provides a software framework for distributing applications on clusters of servers; designed to handle huge amounts of data.

predictive analysis: Methods used to extract information from data and use it to predict trends and behavior patterns.

Sources: CDS Glossary, EDUCAUSE, PCMag.com Encyclopedia, and Wikipedia.

Institutions interested in building analytical capabilities have more to consider than simply implementing systems and deciding where they should be managed. A robust analytic environment also consists of integrating these systems in a way that produces meaningful and actionable analyses. Although many institutions have BI reporting and data warehouse systems, the 2014 ECAR report on higher education's top-ten strategic technologies suggests that few institutions are configured in a way that supports analytic capacity (20% of institutions have BI reporting dashboards in place for analytics; 35% of institutions have a data warehouse in place for analytics).[2] As institutions prioritize analytics, however, efforts are being made to integrate these systems and prepare them for this purpose (59% of institutions are investing in planning and implementation of BI reporting dashboards for analytics; 43% are investing in planning and implementation of data warehouse systems for analytics). These institutions may plan to implement new systems and/or leverage existing systems. A variety of analytic environments could emerge, given the number of current combinations of solution types in place. The most common combination of systems among institutions with either a BI reporting or data warehouse system was a data warehouse solution from an infrastructure provider with a single BI reporting solution (21%; see Table 1).

Table 1. Variety of BI Reporting and Data Warehouse Systems Reported

BI Reporting Solution Type	Data Warehouse Solution Type						Row Total*
	Homegrown	BI/Infrastructure Provider	Admin. Systems Provider	None	Same Solution as BI Reporting Solution	Other	
BI Provider (single)	14%	21%	12%	11%	5%	0%	**63%**
Homegrown	10%	1%	1%	3%	0%	0%	**15%**
BI Providers (multiple)	1%	4%	0%	0%	1%	0%	**7%**
Admin. Systems Provider	1%	1%	3%	1%	0%	0%	**6%**
None	3%	1%	1%	0%	0%	0%	**6%**
Admin. Provider with Third-Party BI Tools	0%	0%	1%	1%	0%	0%	**2%**
Other	0%	0%	0%	0%	0%	0%	**1%**
Column Total*	**29%**	**28%**	**19%**	**17%**	**6%**	**1%**	**100%**

** Certain totals differ from apparent sums due to rounding.*

Even with the range of possible of solutions, customization continues to be an avenue for fine-tuning system configurations to match business processes. Approximately 30% of U.S. institutions have substantially customized BI reporting systems (ranked eleventh in figure 1). Data warehouse systems tend to require more customization than BI reporting systems (46% of U.S. institutions have substantially customized data warehouse systems; ranked third in figure 1). Substantial customization is more common when the data warehouse system is provided outside central IT (66% of distributed systems versus 42% of centrally provided systems). Additionally, older systems tend to be more frequently customized. On average, for each additional year of age of a data warehouse system, the likelihood of substantial customization increases 7.5%. There is not a significant association between age of BI reporting system and customization.

As systems age, institutions begin to consider the trade-offs between replacement and further customization. The BI reporting and data warehouse system areas are relatively young, however, and few institutions plan to replace either their BI reporting system (13% of U.S. institutions) or their data warehouse system (14% of U.S. institutions) in the next three years. Those who are planning to replace systems cite upgraded functionality as the primary reason (84% of institutions replacing a BI reporting system; 71% of institutions replacing a data warehouse system).

Finally, unlike the homogenous learning management system market space, in which 93% of the market uses one of the top-five vendors, the BI reporting and data warehouse landscape is relatively diverse. Only 52% of the market is using the top-five BI reporting vendors, and only 50% of the market is using the top-five data warehouse vendors. Figures 3 and 4 depict the market share for the most common BI reporting and data warehouse solutions at CDS participating institutions.

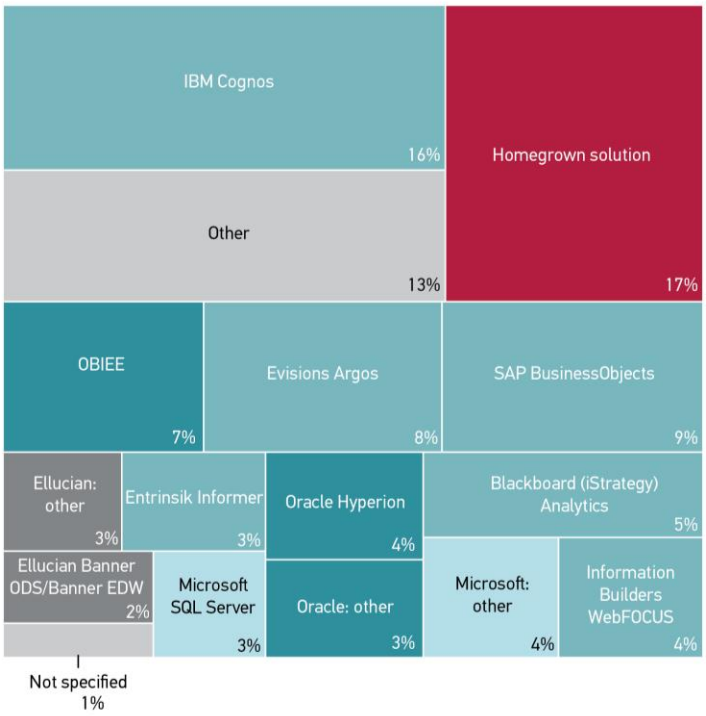

Figure 3. BI Reporting System Market Share

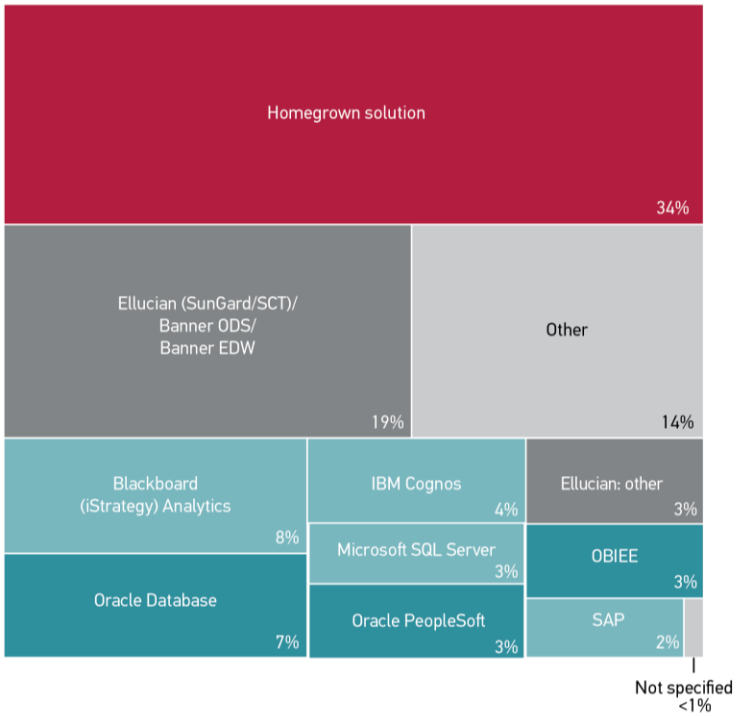

Figure 4. Data Warehouse System Market Share

Future Trends

BI reporting dashboards that visually monitor and display the status of processes and activities in the form of charts or metrics captured the top spot on the list of higher education's top-ten strategic technologies for 2014. The list's sixth-ranked strategic technology was administrative or business performance analytics to help target organizational resources to support organizational goals.[3] Institutions characterizing themselves as early adopters of technology will spend considerably more time than other institutions on five technologies; of these, three relate to the analytical portfolio in the form of emerging analytical techniques and data sources: predictive analysis, Hadoop, and big data.[4] As focus shifts to this application area, the following trends will impact the future of BI reporting and data warehouse systems in higher education:

- **Enhancements to the Current BI Reporting/Data Warehouse Platform:** Support for growing data-exploration and analysis requirements may necessitate enhancements to the current BI reporting/data warehouse platform—for example, greater data interoperability to create more possible data combinations (e.g., combining HR data with financial data or housing data with financial aid data) and the use of metadata to promote consistent understanding of data sources and definition among all data users, not just the data owners (e.g., metadata for GPA definitions to help users not employed in the registrar's office). In addition, the proliferation of smartphones and tablets feeds user expectations for anytime, anywhere data access, requiring mobile reporting and data-access capabilities (e.g., an alumni affairs associate downloading a prospective donor's gift-giving history before a fundraiser).
- **Emergence of the Analytical Portfolio:** Users' growing analytical expertise feeds the exploration of more complex questions and issues, which eventually may require dynamic and flexible blending of data as well as cutting-edge data analysis and visualization tools. Traditional BI reporting tools and data warehouses may not be robust enough to provide these capabilities, leading to the development of an expanding institutional portfolio of analytical tools and capabilities to address diverse information needs. For example, mature data warehouses and advanced BI systems can provide real-time and accessible reporting, dashboards, and data visualizations; systems that provide just-in-time advice and alerts can enable students and their advisors and instructors to take action to improve performance or enable administrators to optimize services and processes.[5] The deployment of analytics opens the door to the adoption of new supporting tools and platforms, as for example, predictive analysis.[6]
- **New Data Sources:** Institutions will likely mix new data sources into their analysis. Until now, the major type of institutional data has been structured data, such as is found in administrative and academic systems. In the future, higher education will most likely move to follow the commercial sector's use of unstructured data (e.g., social media such as Facebook and Twitter) to respond to customers' evolving requirements. For example, GPS and other geospatial data such as radio frequency ID, near-field communications, and Bluetooth low-energy proximity sensing can provide information on ways to enhance student services, though privacy issues may curtail some uses of unstructured data.

Tips for Successful Implementations

As the list of top-ten strategic technologies for 2014 illustrates, IT professionals recognize the importance of supporting higher education's burgeoning analytical culture. Whether IT organizations are building their first BI reporting tools and data warehouse systems or expanding current capabilities into an analytical portfolio, our research yielded several relevant recommendations to consider.

Assess Current Capabilities

Before moving forward, it is advantageous to understand the institution's current state in order to focus planning and investment optimally. Maturity models, such as the Oficina de Cooperación Universitaria BI Model[7] or ECAR's Analytics Maturity Model,[8] offer institutions a starting point to determine current strengths and weaknesses and guidance on how to proceed accordingly.

Create a Roadmap

Supporting an institution's analytical culture is not a project but an ongoing program, and the program's direction and approach should be driven by the institution's aims and questions to be answered. This requires a strategic plan that provides the big-picture aims, funding, and resources, as well as a detailed roadmap of the program's nuts and bolts—how the implementation will address the information needs of each constituency, as well as details on training and education in using the tools and understanding the data.

Create a Collaborative, Institutional Effort

There may be a tendency to focus on technology aspects because they are most familiar and tangible, but people and process are perhaps more important than the underlying infrastructure. Building analytical capability involves relationships, problem solving, prioritization, and governance of which problems to tackle—as well as an understanding of institutional culture. Successful collaborations require compelling need, a dedicated leader, and committed coalition members.[9] Focusing on the following elements will enable central IT to successfully navigate an institutional effort to build analytical capability:

- **Sponsorship:** Senior-level sponsorship is a vital component in any major IT initiative, but perhaps it is even more critical when supporting analytical capabilities, given their complexity and expense, as well as their direct impact on senior-leadership functions. The choice of institutional sponsor is equally critical; one needs a champion who understands thoroughly the ins and outs as well as the power of analytical tools, so that person can convey *knowledgeable* enthusiasm for the program's possibilities and benefits to colleagues and to the institution at large.
- **Balance between Central and Local:** Central IT may manage BI reporting and data warehouse systems at many institutions, but this approach should not preclude local involvement. The centralized unit may manage much of the infrastructure, data, and tools, but some analytical activities will address local (e.g., college), specific needs that may require the use of supplemental tools and a blend of local and institutionally certified data. The central and local units must work together to identify local needs and partnership opportunities.

- **Governance:** Coordinating these central and local interests requires strong governance, especially in more decentralized institutions. A representative group can prioritize program activities, but an institution can reap greater benefits from creating a higher-level steering committee that comprises leaders of the institution's data-producing and data-consuming areas. This helps align analytic activities with institutional strategic objectives. In addition, this steering committee can serve as a core decision-making body that guides future institutional investment and expansion of analytic capabilities.
- **Communication:** Grasping the potential and operation of analytic tools can be difficult, and thus change management and education are very important. An ongoing communication campaign, commencing at the program's inception, is required to demonstrate its capabilities and possibilities to staff members at every relevant level and area. The message has to be consistent and incessant to promote widespread awareness and understanding.

Prototype First

As with many complex IT projects, it may be behoove institutions to create a small prototype to show the value of the analytic solution and expand its capabilities over time. Such a strategy gives the program sponsor a quick win and establishes trust in future endeavors as the institution sees an operational example that demonstrates the possibilities and benefits.

Plan for the People

An implementation should take into account individuals in all roles—service owners, operators, end users, and others:

- **Staffing:** Staffing has both operational and strategic implications. There is a need for localized data experts or IT professionals with both functional knowledge and analytical skill to provide hands-on support to the dean or administrator who requires the information. For example, each college at Purdue University has a staff member focused on information, data, and analysis. However, one commonly expressed problem is the ongoing shortage of qualified individuals to fill these roles, and one potential solution is to draw on an institution's most valuable asset—its students—either in the form of internships or career-development opportunities. To support the strategic use of data and analysis, some institutions have hired a chief data officer to bring high-level leadership to these two areas and to coalesce central and local data stakeholders and activities into an enterprise endeavor.
- **Ease of Use:** Given the shortage of information-analysis professionals in higher education, it is imperative to fold ease of use into the program design. One means is to create common enterprise data sets whenever possible that are available for automated, institutional applications or that can be combined with local, customized applications. Again, this approach ties back to maintaining the central/local balance and grooming partnerships. In addition, some applications and tools, especially in the areas of data visualization and data discovery, can be complicated for the data novice to use. When creating or purchasing off-the-shelf tools, keep in mind the skill level of business or functional staff members who will be using them.

Look to Peers

The peer benchmarking data included in CDS can help inform decision making throughout an implementation effort. Knowing which solutions peer institutions are using, as well as the extent to which their systems are customized and whether they plan to replace them, can provide additional insight for planning and roadmaps. Collaborating with peers through networks such as the Higher Education Data Warehousing Forum[10] and shared services can help uncover best practices and opportunities for operational efficiency.

Where to Learn More

- Chapple, Michael. "Effective Data Governance Practices." Presentation at the 2013 Security Professionals Conference, St. Louis, Missouri, April 17, 2013. http://www.educause.edu/events/security-professionals-conference/2013/2013/effective-data-governance-practices.
- Collins, Marcus. *2014 Planning Guide for Data Management*. Research report. Stamford, CT: Gartner, December 9, 2013. http://www.educause.edu/library/resources/2014-planning-guide-data-management.
- Higher Education Data Warehousing Forum. http://www.hedw.org/.
- Lang, Leah. *2013 CDS Executive Summary Report*. Research report. Louisville, CO: EDUCAUSE, February 2014. Available from http://www.educause.edu/coredata.
- Grajek, Susan. *Higher Education's Top-Ten Strategic Technologies in 2014*. Research report. Louisville, CO: ECAR, February 2014. Available from http://www.educause.edu/ecar.
- Grajek, Susan, Leah Lang, and David Trevvett. *The 2011 Enterprise Application Market in Higher Education*. Research report. Louisville, CO: ECAR, June 2012. Available from http://www.educause.edu/ecar.
- *HBR Blog Network*, "Are You Ready for a Chief Data Officer?" blog entry by Thomas C. Redman, October 30, 2013. http://blogs.hbr.org/2013/10/are-you-ready-for-a-chief-data-officer/.

Acknowledgments

The authors thank Theodore M. Bross, Associate Director, Administrative Systems, Princeton University; Ora Fish, Executive Director of the Data Warehouse and Business Intelligence, New York University; Beth Ladd, Director, Business Intelligence, Illinois State University; and Aaron Walz, Director, Business Intelligence Competency Center, Purdue University, for sharing their professional experiences and insights about business intelligence and data warehouses in higher education applications.

About the Authors

Leah Lang (llang@educause.edu) is Manager of the Core Data Service at EDUCAUSE. Judith A. Pirani (Judith.pirani@gmail.com) is an EDUCAUSE consultant and President of Sheep Pond Associates.

Citation for This Work

Lang, Leah, and Judith A. Pirani. *BI Reporting, Data Warehouse Systems, and Beyond*. Research bulletin. Louisville, CO: ECAR, April 23, 2014. Available from http://www.educause.edu/ecar.

Notes

1. Rankings for figure 1 are out of 20 system areas covered in Module 8 of CDS 2013. Rate-of-change ranking is based on age of system, number of institutions implementing the system, and number of institutions replacing the system. This ranking is an indication of how ripe the market is for vendors in this space (1 = high change; 20 = low change).
2. Susan Grajek, *Higher Education's Top-Ten Strategic Technologies in 2014,* research report (Louisville, CO: ECAR, February 2014), available from http://www.educause.edu/ecar.
3. The survey was distributed to 10,393 EDUCAUSE members as part of the Top-Ten IT Issues Survey in November 2013; 444 members responded and indicated, for each technology, the attention their institution was planning to devote to 78 technologies in 2014. The survey defined *strategic technologies* as "relatively new technologies institutions will be spending the most time implementing, planning, and tracking in 2014." For more information, see Grajek, *Higher Education's Top-Ten*.
4. Ibid., 19.
5. Susan Grajek and the 2012–2013 IT Issues Panel, "Top-Ten IT Issues, 2013: Welcome to the Connected Age," *EDUCAUSE Review* 48, no. 3 (May/June 2013), http://www.educause.edu/ero/article/top-ten-it-issues-2013-welcome-connected-age.
6. For more information about analytics, see Jacqueline Bichsel, *Analytics in Higher Education: Benefits, Barriers, Progress, and Recommendations*, research report (Louisville, CO: ECAR, August 2012), available from http://www.educause.edu/ecar.
7. OCU BI Maturity Model, Jisc infoNet, http://www.jiscinfonet.ac.uk/infokits/business-intelligence/measuring-success/ocu-maturity-model/.
8. ECAR, "Analytics Maturity Index," https://www.surveygizmo.com/s3/1020013/Analytics-Maturity-Index.
9. Jack McCredie and Judith A. Pirani, *A Dozen Gurus Describe IT Collaborations That Work*, research bulletin (Louisville, CO: ECAR, May 22, 2012), available from http://www.educause.edu/ecar.
10. See http://www.hedw.org/.