DEVELOPING ANALYTICS CAPABILITY TO SUPPORT DATA INFORMED DECISION-MAKING

AIR Forum, Orlando, Florida

OCAIR-TAIR Panel Discussion, May 30, 2018

Henry Zheng, PhD, MBA
Lehigh University

About Lehigh University

FOUNDED:

1865 by Asa Packer, an industrial pioneer, entrepreneur and philanthropist

ACCREDITATION:

Regionally Accredited, Middle States Association of Colleges and Schools

STUDENT BODY:

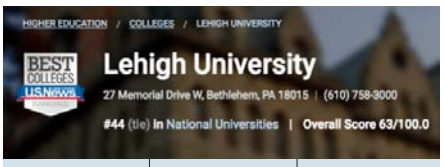
5,080 undergraduates (56% men, 44% women)
1,979 graduate students (55% men, 45% women)

4 COLLEGES:

College of Arts and Sciences, College of Business and Economics, P.C. Rossin College of Engineering and Applied Science and College of Education

LEHIGH ENDOWMENT:

$1.2 billion



29
AVERAGE CLASS SIZE
80% of classes have 35 or fewer students

10:1
STUDENT-TO-FACULTY RATIO

2,300+
COURSES OFFERED
100+ Undergraduate degree programs
13 Masters and Doctorate Degrees

Lenigh University is located in Bethlehem, PA., a vibrant and historic community that is a short drive from New York City and Philadelphia. Over 820,000 people live in the region.

SIZE OF CAMPUS: 2,358 acres. The Asa Packer Campus, Mountaintop Campus and Murray H. Goodman Campus are contiguous.

Strategic Analytics Plan

Strategic Analytics Planning Timeline

January 6, 2017
- Committee meets with executive sponsors for project kickoff
- Review of project charter
- Assignment of tasks and sub-committees

January - March, 2017
- Sub-Committees meet to work on assigned topics
- Update reports to full committee
- Recommendations on data inventory and governance, industry trends, best practices in higher education

Feb – June, 2017
- Gartner on-site interviews
- Sharing of draft reports to full planning committee
- Presentation of final recommendations to senior leadership

April 27, 2017
- Needs Assessment and Visioning discussion

June 5, 2017
- Full committee retreat on vision, goals
- Sub-group discussions of planning initiatives

July 10 , 2017
- Full committee review of the final draft of strategic analytics development plan

July 26, 2017
- Final plan presentation to Executive Sponsors
- Senior leadership endorsement of plan

Vision Statement

Lehigh University uses data analytics to support decision making, enhance operational effectiveness, and drive innovation.

Goals:

- Build a culture where data supports transparent and accountable decision-making
- Implement a robust data governance structure that supports and promotes the effective use of data assets
- Apply and enforce consistent data governance policies
- Identify and develop a common platform to enable collaborative analytics development

Guiding Principles

Data are strategic assets and will be at the foundation of all planning efforts and future operations. Data quality is ensured by following consistent standards and definitions. At all levels and units, data informed decision making will be evaluated based on how well we do the following cycle of activities:

- Systematically gather, report, and share data relevant to their area of operations
- Analyze and interpret data, so that it becomes "information"
- Use data intelligence to improve operations and track outcomes
- Gather and report data on the success of those actions

Short-term Outcomes: Leadership Commitment and Empowerment

- Lehigh leadership commits to supporting and promoting the effective use of data assets
- Lehigh leadership commits to building a culture of data-informed decision making
- Lehigh leadership empowers the development and communication of consistent data governance policies
- Lehigh adopts a common data source platform to enable collaborative analytics development
- Lehigh identifies general and specific requirements for new data analytics tools

Mid-term Outcomes: Organizing for Change and Enablement

- All units identify and develop critical and measurable metrics for their area of operations
- All units have data stewards to work with IR/BI specialists to receive training on the use of data analytics for practical assessment and tracking of their metrics
- Data stewards or designees are trained on use of data analytical and visualization tools and will in turn train other employees in their own unit
- First set of data reports and dashboards are accessible to authorized data users through the "Self Service Portal"

Long-term Outcomes: Development of Analytics Culture and Maturity

- Administration will use data to support strategy execution and operational improvement
  - Use data to reduce overall costs to students and improve the efficiency and effectiveness of the institution
  - Use data to support budget, space, personnel, technology, or other resource requests with relevant data
- All units will use data to create targeted interventions to improve student success
  - Faculty will use data to assess what students are learning and how they are progressing
  - Departments will use data to determine optimal schedule offerings
  - Advisors will use data to monitor student engagement and satisfaction
  - Admissions and outreach units will use data to assess efforts at recruitment, retention, and success
- Data collection will be coordinated and follow standard protocols to enable the organic linking of different data assets to portrait the entire engagement cycle of our students before, during, and after their education experience at Lehigh University

Key Plan Initiatives and Milestones

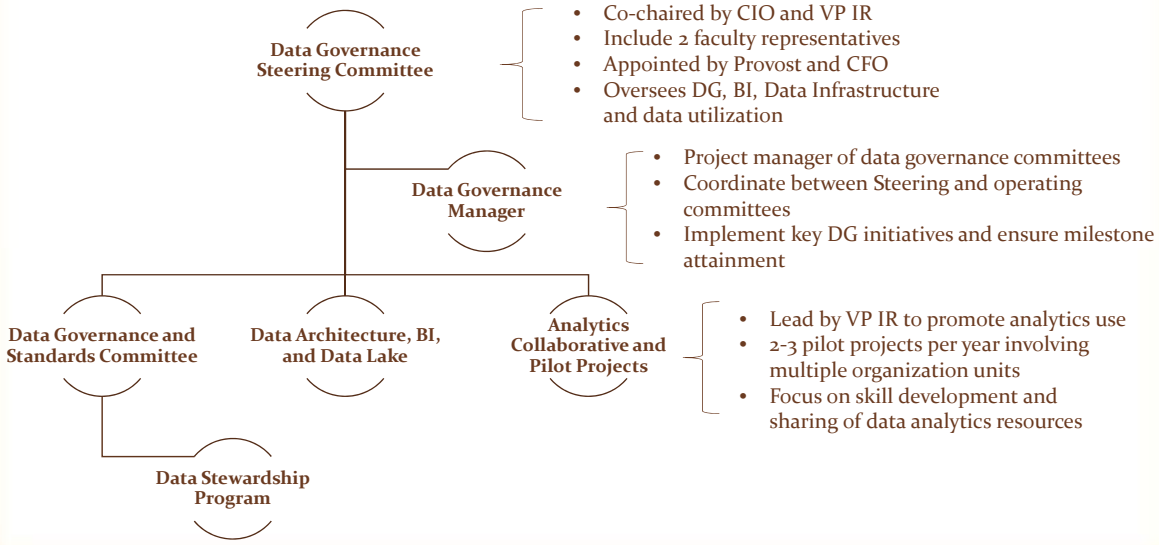
- Data Quality, Data Governance, and Data Stewardship
- Data Architecture, Business Intelligence, and Reporting Platform
- Data-informed Decision Making and an Analytics Culture

Data Quality, Data Governance, and Data Stewardship

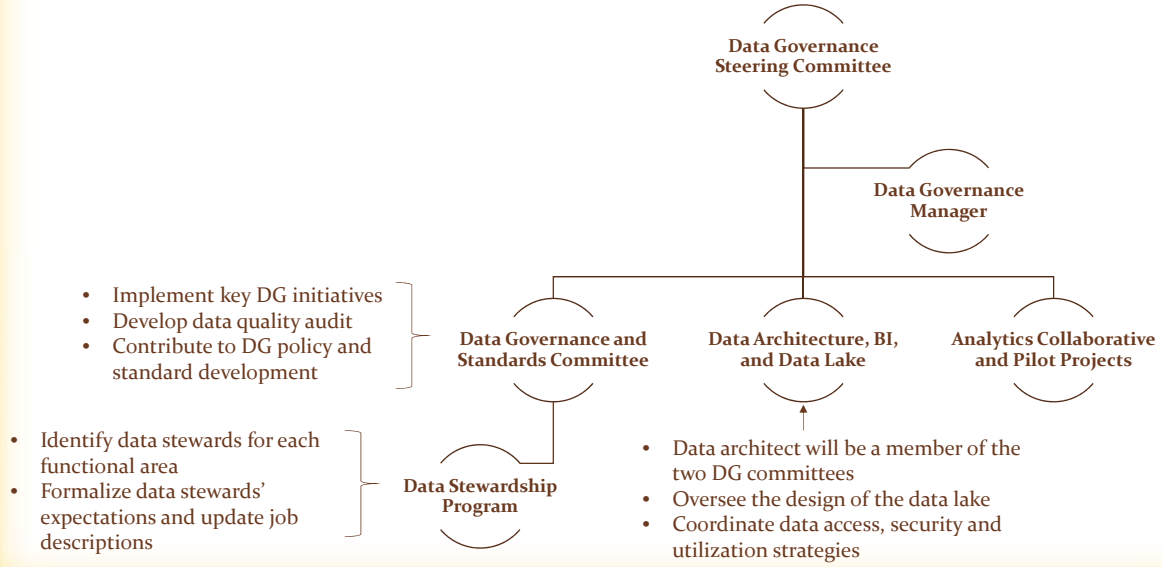
Year 1	Year 2	Year 3

• Recast data governance structure
  o Create a steering group co-chaired by CIO and VP IR and include key functional area leaders.
  o Include faculty involvement in data governance
  o Appoint a data governance manager
  o Facilitate DG activities, lead some of the efforts
  o Responsible for stewardship training design/development
• Restructure current Data Governance and Standards Committee to include chief data stewards from major functional areas
• Review and update data dictionary (start with an inventory and review of what prior Data Advisory Council had completed)
• Start inventory of data assets – Phase I (e.g. registration data)
  o Identify highest priority/areas of most integration
• Develop data stewardship programs
  o Identify data steward for each area,
  o Update expectations and data steward's job descriptions
  o empower with training
	• Review, update, and approve data governance policy and procedures handbook
	• Publish data governance documents and standards online to be accessible to all campus users
	• Initiate data quality audit pilots through the data stewardship program
	• Complete Phase II of data assets inventory to include non-priority areas
	• Develop data quality and data stewardship training for on-boarding of new employees
	• Create an active directory of data stewards as well as their areas of expertise and responsibilities
		• Complete first round of data quality audit across the campus for all administrative offices and identify key gaps and process for improvement
		• Fully develop the data governance web-site to host all data definitions, glossary, data maps, and data asset inventory.
		• Data quality and stewardship training for all staff who have touch points in the data generation process

New Structure for Data Strategy Coordination



New Structure for Data Strategy Coordination



Data Architecture, Business Intelligence, and Reporting Platform

Year 1	Year 2	Year 3

• Hire Data Architect
• Implement Data Laundry Initiative and Initiate the "Data Lake" development:
  o Survey and/or conduct focus group to determine where data resides in campus systems
  o Extract data from systems of record, re-code data to preserve context and conform to new formats/standards
  o Normalize data to new context and as inputs for models
  o Propose a workflow process for adding data assets to "data lake" with proper documentations
• Determine Data Access and Security needs, implement classification scheme
	• Develop a timeline for integrating key university data assets into a coordinated platform (data lake) using a federated data system
	• Uploading structured or unstructured data to "Data Lake"
	• Identify and acquire the right tool for data extraction and analytic reporting
	• Hold extensive internal reviews of data uses for legal and policy compliance
		• Gradual implementation of the Connected Campus philosophy – making it possible to organically connect different data sets to understand students' learning, engagement, and other factors to promote student success
		• Enable Self-Service portal to empower campus data users to access data reports
		• Functional areas will publish performance indicator reports/dashboards to promote an accountable culture

Data-informed Decision Making and an Analytics Culture

Year 1	Year 2	Year 3
• Present this comprehensive plan (from this committee) to senior leadership for buy in • Memo from senior leadership to clearly state university commitment and support • Appoint VPIR to lead analytics program development, focusing on developing collaborative analytics projects • Call for ideas and inputs from campus stakeholders to identify 2-3 high-value projects as Phase I of efforts to foster analytics collaboration (Analytics Open Lab) • Develop training workshops to improve analytical skills and promote knowledge sharing • Develop initial framework on university dashboard to use data for performance improvement	• Build an analytics community by empowering more people with easier access to data and greater level of transparency • Create a shared service model for providing analytical support to key projects on campus • Implement next phase of 2-3 key collaborative analytics projects • Build analytics skills by developing and delivering a series of training workshops on data governance and analytics for both new employees onboarding and current employees for skill enhancement • Move BI team from IT to Institutional Research and rename office to Office of Institutional Research and Strategic Analytics	• Improved outcomes and decision making by incrementally developing a system of scorecards or dashboards to track and evaluate organizational performance to enhance accountability • Data scientist will provide active shared services and expertise to functional areas to enable local level analytics projects • Implement "connected campus" model for student success

Major Milestones

Year 1	Year 2	Year 3

• **Data Governance / Stewardship**
  • Appoint DG Executive Committee
  • Hire/assign a DG manager
  • Restructure DG and Standards Committee
  • Set up Data Stewardship Prog.
• **Data Architecture / BI / Reporting**
  • Hire data architect
  • Allocate resources for "Data Lake" development
• **DBDM and Analytics Culture**
  • Senior leadership memo to authorize the Strategic Analytics Plan
  • Assign VPIR to lead campus-wide collaborative analytics development
  • 2-3 high value collaborative analytics pilot projects
	• **Data Governance / Stewardship**
	  • Update all DG documents
	  • DG web-site with data maps and dictionaries for major data assets
	  • Start data quality audit pilots and implement staff training
	• **Data Architecture / BI / Reporting**
	  • Work with key stakeholders to design and create "Data Lake"
	  • Create a timeline for data assets inclusion in "Data Lake"
	  • Recommend decision on enterprise-wide BI/ reporting platform
	• **DBDM and Analytics Culture**
	  • Hire a data scientist
	  • Formalize data analytics user group and start sharing of data analysis practices and skills
	  • Next phase of analytics pilots
		• **Data Governance / Stewardship**
		  • 1st round of data quality audit done across campus
		  • DG web-site fully done to host all data definitions, glossary, and data maps
		  • Data stewardship training and on-going skill update
		• **Data Architecture / BI / Reporting**
		  • "Data Lake" fully functional
		  • Single platform for BI / data reporting
		• **DBDM and Analytics Culture**
		  • Develop Executive Dashboards
		  • Work with campus stakeholders to implement Lehigh's version of "connected campus"
		  • Use of predictive analytics to support admissions / student success

Backup Slide

Key Terms: Data Warehouse vs. Data Lake

DATA WAREHOUSE	vs.	DATA LAKE
structured, processed	DATA	structured / semi-structured / unstructured, raw
schema-on-write	PROCESSING	schema-on-read
expensive for large data volumes	STORAGE	designed for low-cost storage
less agile, fixed configuration	AGILITY	highly agile, configure and reconfigure as needed
mature	SECURITY	maturing
business professionals	USERS	data scientists et. al.

A data warehouse is a hierarchical, structured data repository of integrated data from multiple sources, organized for creating analytical reports. They often use multiple databases for different types of data storage, such as ingestion, staging, and transformation, and for processing, such as online analytical processing or online transaction processing.

A data lake is a storage repository that holds a vast amount of raw data in its native format and stores it unprocessed until it is needed. A data lake uses a flat architecture to store data so the data is completely unstructured and left in the format in which it was originally ingested, although each data element in the lake is assigned a unique identifier and tagged with a set of extended metadata tags. That way when a query is run, it can be run against a smaller set of data with the specific tags rather than processing all of the data in the lake.

Backup Slide

Key Terms: BI and Analytics tools

tableau SOFTWARE

Qlik

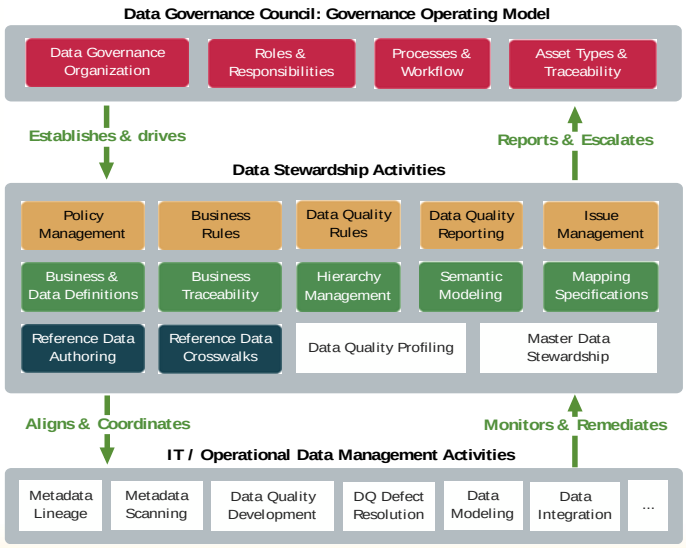
Power BI

salesforce

IBM COGNOS

Backup Slide

Key terms: Data Governance and Stewardship



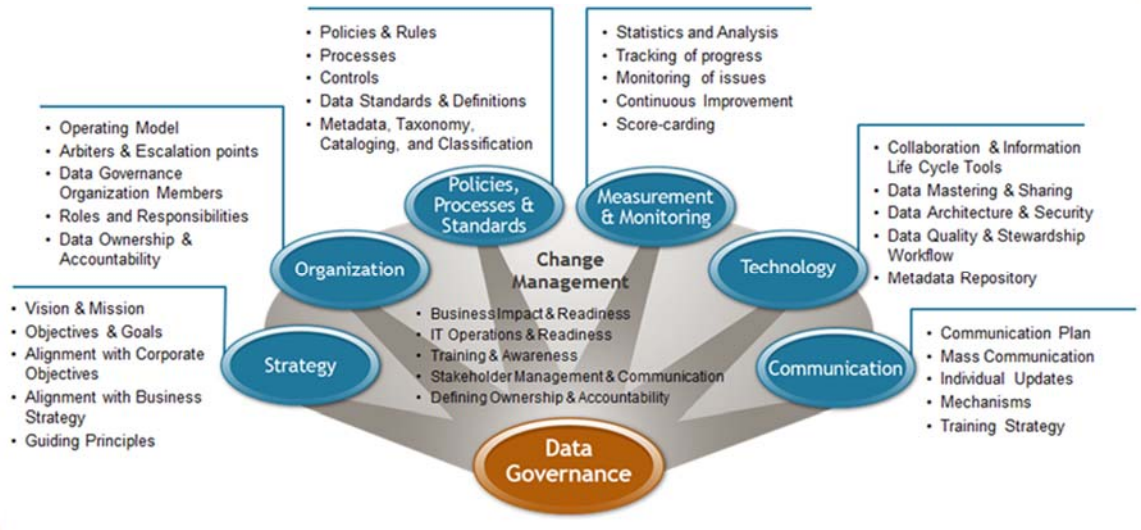
Data Governance Executive Committee – functional area leaders, DG manager, data architect

DG manager, data architect, appointed data stewards, power users

Data architect, IT systems, BI team, and data security

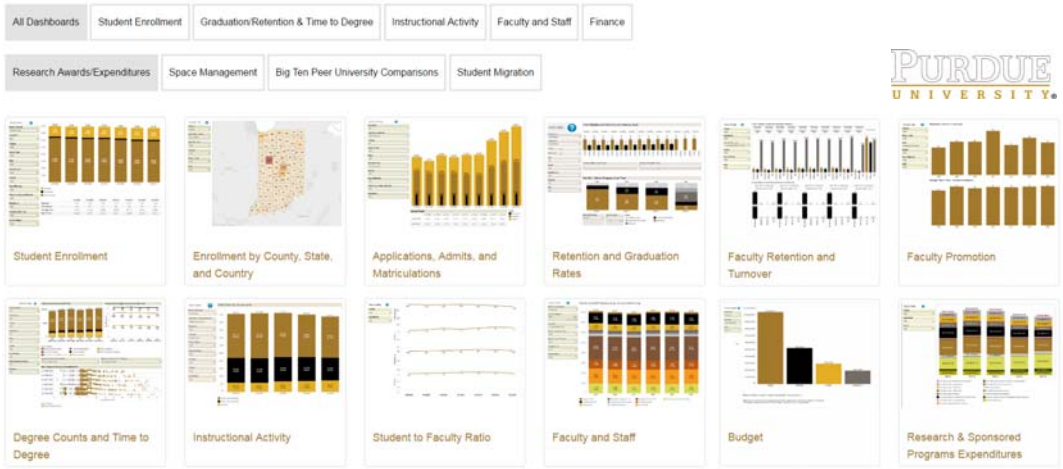
Backup Slide

Key terms: Data Governance and Stewardship



Backup Slide

Key terms: Collaborative BI Portal Development



Source: Purdue University IR Web Site