Stony Brook University

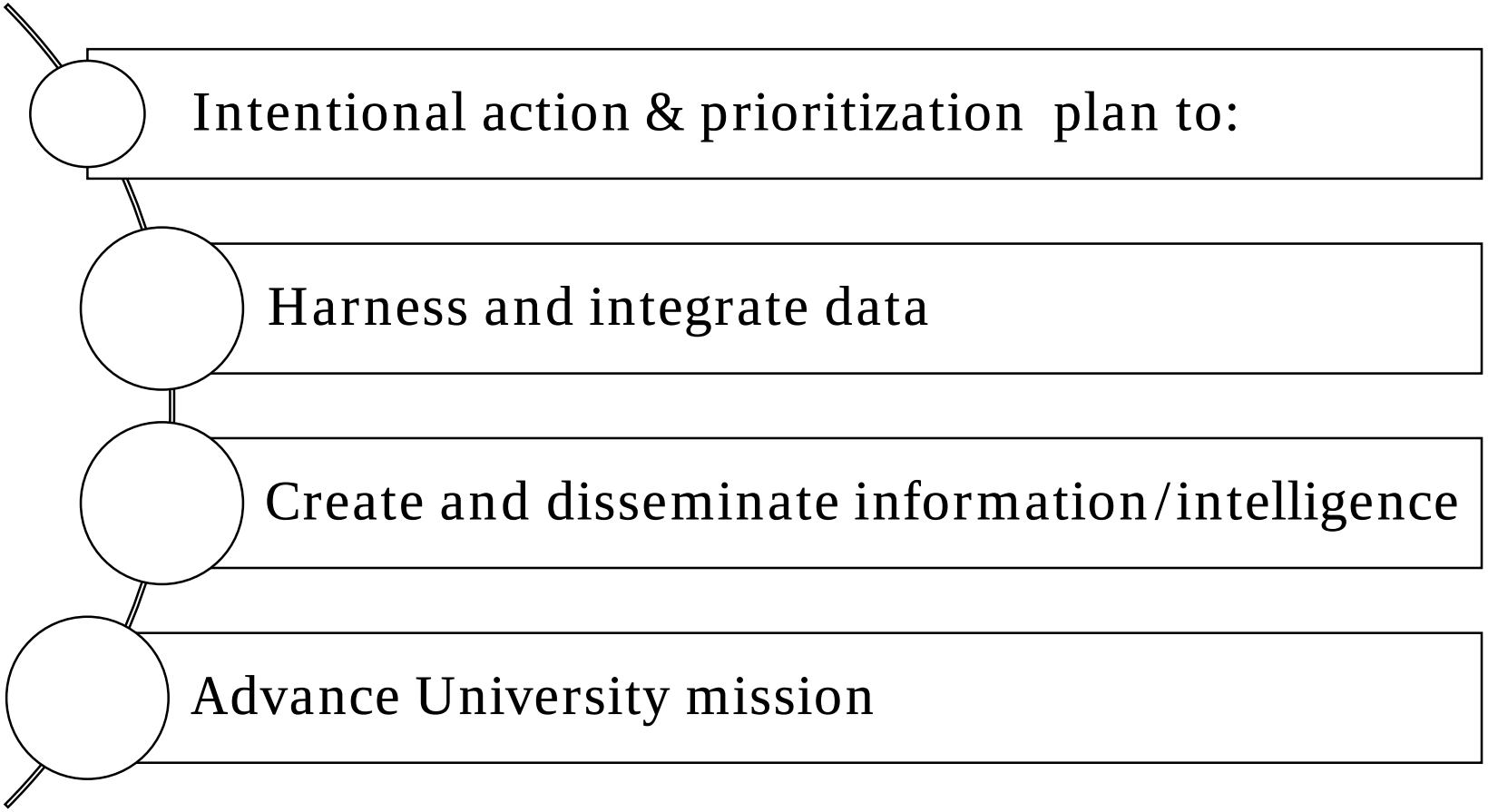
FAR BEYOND

Stony Brook University Data Strategy

Presented to the Data Governance Council
June 8, 2017

Stony Brook University

What is a data strategy?



Why do we need a data strategy?

Support objectives to:

- Promote operational effectiveness, excellence & efficiency
- Retain and grow revenue
- Reduce risk
- Drive innovation

Proliferation of data assets

Increasing organizational size and complexity

Advances in analytical tools

Selected Stony Brook data assets

And other faculty activity data sources

Assessment Data
Help Desk Tickets
Card Swipes
Surveys

Stony Brook's mission

The university has a five-part mission to provide and carry-out:

- Highest quality comprehensive education
- Highest quality research and intellectual endeavors
- Leadership for economic growth, technology, and culture
- State-of-the-art innovative health care, with service to region and traditionally underserved populations
- Diversity and positioning Stony Brook in global community

Elements of Stony Brook's data strategy


Data acquisition


Data governance


Data quality


Data access

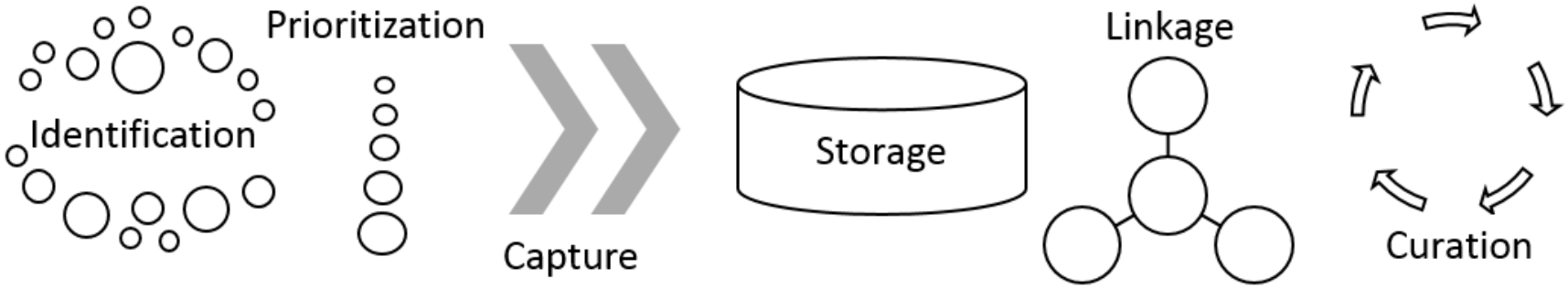

Data usage & literacy


Data extraction & reporting


Data analytics

Data acquisition

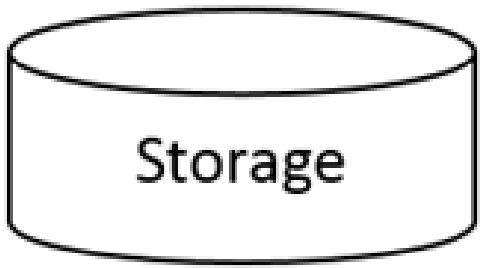
- Data acquisition involves identification, prioritization, capture, storage, linkage, and curation of data assets most valuable to the enterprise



Data acquisition
Identification & prioritization

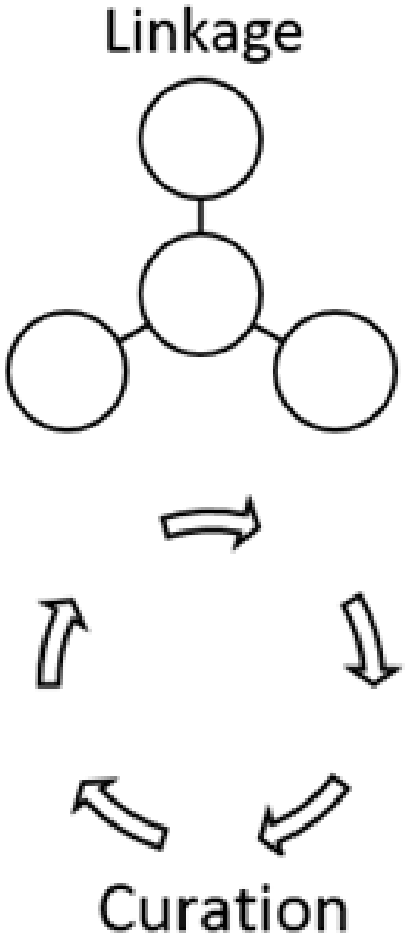
- Establish and maintain an inventory of data assets and assess acquisition maturity
- Establish a process to prioritize integration into data infrastructure

Data Acquisition
Capture & storage



- For each data asset identify current and optimal capture procedures
- For each data asset identify current and optimal storage areas

Data Acquisition
Linkage & curation



- For each data asset identify current and optimal procedures to link to other data sources
- For each data asset identify how data will be updated and maintained to preserve value

Data governance

- Data governance formalizes behavior around how data are defined, produced, used, stored, and destroyed to enable and enhance organizational effectiveness.

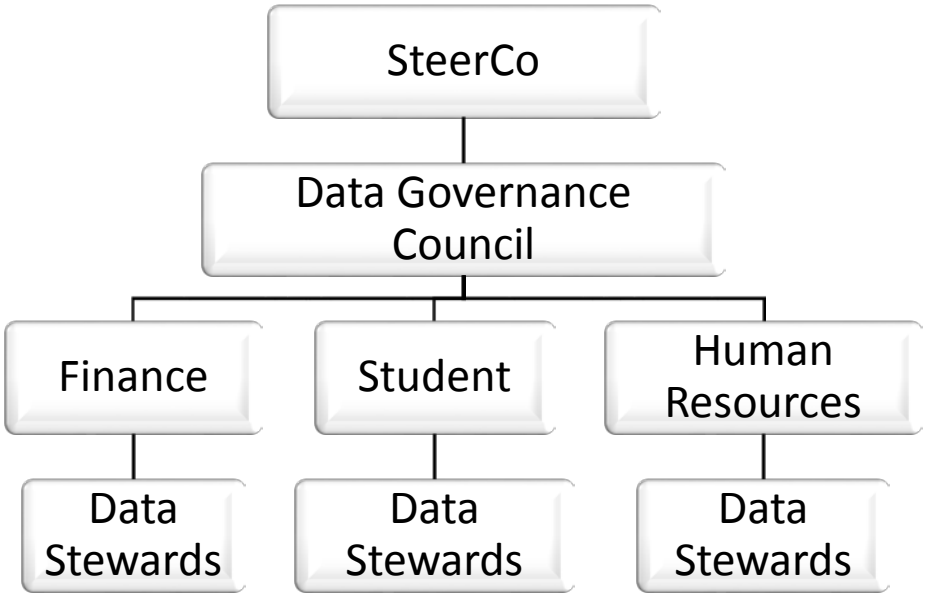
PeopleSoft and the Data Warehouse are governed by the University Data Governance Council

Establish expectations for all other data assets to have formal data governance

Data governance Requirements

- Designated decision-making body
- Formal data dictionaries and descriptions of architecture
- Individuals designated to provide stewardship
- May opt to be governed through the Stony Brook Data Governance Council

Stony Brook
Data Governance Framework*



*Applies to PeopleSoft and the Data Warehouse (as of 9/26/16)

Data Quality

- Data quality is the state of completeness, validity, consistency, timeliness and accuracy that makes data appropriate for a specific use.

The Data Governance Council is charged with improving data quality for PeopleSoft and the Data Warehouse. A roadmap to achieve this has been developed

For each data asset, develop and execute a plan to maintain and improve data quality; automate when justified by ROI.

Data access

- Data access ensures authorized individuals can obtain and use data when and where they are needed and protects privacy and sensitive information by preventing unauthorized use.

Accessibility | Authorization | Security

Data usage and literacy

- Data usage and literacy entail people regularly obtaining data; understanding them; and using them to improve operational effectiveness .

Establish for all data assets:

Usage metrics

Effectiveness metrics

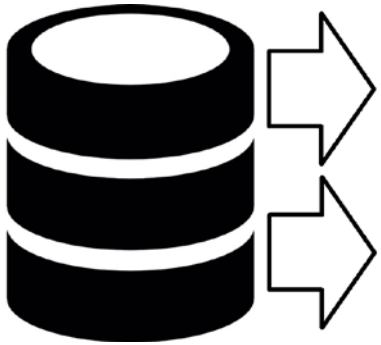
Training inventory

Data User Responsibilities

1. Recognize data complexities; understand data meanings and limitations
2. Cite sources; assume broad audiences
3. Respect privacy
4. Secure data and reports
5. Report data quality issues

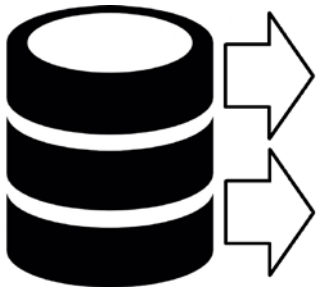
Data extraction and reporting

- Data extraction and reporting represent the ways that data are queried and retrieved from storage and then delivered to users through regularized and ad hoc reporting to support day-to-day operations.



Extraction | Reporting

Data extraction and reporting



Extraction

Reporting

- Methods for querying and extracting data from storage should be identified, including user types associated with each extraction method
- Reports should be linked to operational objectives
- Report inventories should be maintained in an accessible area.
- Reports should be automated depending on ROI
- Reports should include effectiveness metrics

Data analytics

- Analytics deliver dynamic and visual analysis of data, internal & external benchmarking, exploratory and causal analysis, and predictive/forecasting capacity

Requirements

Maturity in data acquisition, governance, quality, access, usage, & extraction

Tools capable of performing analyses and communicating effectively

Speed and ease of use

Data asset strategy document compiled for each data asset

Data Asset Strategy Doc
e.g. IPEDS

Description & use

Data acquisition

Priority

	Current	Plan	Date
Capture			
Storage			
Linkage			
Curation			

Data governance plan

Data quality protocols

Data access plan
- Accessibility
- Authorization
- Security

Data usage and literacy

Data extraction/reporting

Data analytics