AGENDA

• History of UMD’s Data Ecosystem
• Project Overview
• Data Policy and Practice
• Role of Data Stewards and Managers
• DIT Data Administration and Partnership
A SELECTED TIMELINE

IN THE WORLD...

- Oracle 8 Released
- Oracle 10g Released
- Tableau Founded
- Cold Fusion 9 Released
- Oracle Purchases Hyperion
- Amazon Redshift Released
- UMD Profiles Goes Live
- Last “DIB” Published
- Profiles “Drill Down” Expands
- Current WH Expands to 7 Yrs
- UMD Practices for Data Sharing

AT UMD...

- UMD Data Warehouse
- 1997
- 2002
- 2007
- 2012
- 2017
WHY MODERNIZE?

- **Software Out of Support**: Manufacturer no longer provides patches or technical support
- **Software Incompatible**: Software is incompatible with new versions of operating systems or browsers
  
  **Risks**: Data Security, System Stability, User Frustration

- **Software Is No Longer Performant**: Product does not provide experience users expect – response time, features, etc.

- **Design is Outdated**: Warehouse was designed using outdated standards
  
  **Risks**: User Frustration
TAKING INVENTORY

Transaction Systems

KFS Data Warehouse
Managed by Admin. Affairs

Data Store
Managed by DivIT

Current Warehouse
Managed by DivIT

Frozen Warehouse
Managed by DivIT & IRPA

Oracle Tables for System Integration

Generally Nightly

Census Dates

Shadow Databases

Data Documentation via “Gunpowder” Site
Uneven coverage across tables & fields
Inconsistent updates
Inconsistent definitions

Gaps in data availability

Inconsistent reporting

Processing data is brittle and slow

Inconsistent application of data policy

Limited unit-record and reporting flexibility
<table>
<thead>
<tr>
<th>System</th>
<th>Issue</th>
<th>Risk Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Profiles</td>
<td>Out of Support, Performant</td>
<td>High</td>
</tr>
<tr>
<td>WOW</td>
<td>Incompatible, Out of Support</td>
<td>High</td>
</tr>
<tr>
<td>Desktop Hyperion</td>
<td>Support ends 2018?</td>
<td>Medium</td>
</tr>
<tr>
<td>Warehouse</td>
<td>Performant, Outdated</td>
<td>Medium</td>
</tr>
</tbody>
</table>
PROJECT GOALS

• Replace Profiles before Cold Fusion goes out of support
  Provide more useful information faster and with less effort.
  Add data sources (internal and external).

• Involve decision makers and decision supporters

• Implement a data architecture that supports Practices for Data Sharing

• Implement modern analytic tools
<table>
<thead>
<tr>
<th>Workstream</th>
<th>Executive Reporting (ER)</th>
<th>Data Extraction + Ad Hocs (DE)</th>
<th>Data Modernization (DM)</th>
<th>Data Documentation (DD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>Any user on campus should be able to access predigested information with context to aid in their understanding</td>
<td>Power users on campus generate reports to do their day-to-day work and need the ability to run ad hoc queries</td>
<td>AWS DW transformation</td>
<td>Ensuring all data are documented in DataDocs prior to moving to the new DW</td>
</tr>
<tr>
<td>Legacy Systems</td>
<td>Profiles, Frozen Warehouse, WOW</td>
<td>Current Warehouse, WOW, Advise on the Web, Hyperion Studio</td>
<td>Oracle, Mainframe</td>
<td>DAD site, Oracle Tables</td>
</tr>
<tr>
<td>Risk Mitigation</td>
<td>Profiles Replacement WOW - Partial Replacement</td>
<td>WOW Replacement</td>
<td>Informatica implementation SIS Replication / Data Store</td>
<td>Field level definitions - DAD Replacement Report Documentation</td>
</tr>
<tr>
<td>Enhancements</td>
<td>New Executive Reports New Content Areas</td>
<td>Hyperion/Studio Replacement Advise on the Web Replacement</td>
<td>Warehouse redesign New data sources Implementation of &quot;Data are Free&quot; Policy</td>
<td>Table Documentation</td>
</tr>
<tr>
<td>Solution Space Systems</td>
<td>Tableau, AWS</td>
<td>SQL, Tableau, Jasper, SAS, ...?</td>
<td>Amazon Redshift</td>
<td>Data Docs</td>
</tr>
</tbody>
</table>
HOW MUCH CENTRAL TUITION IS GENERATED?
TUITION BY PROGRAM - EXISTING
TUITION BY PROGRAM - TABLEAU

Central Undergraduate Tuition

BSOS Majors generated the most tuition revenue. However, BSOS does not have the largest enrollment. Out-of-state residency appears to account for this difference.

Tuition Term

<table>
<thead>
<tr>
<th>(All)</th>
<th>Fall 2014</th>
<th>Spring 2015</th>
</tr>
</thead>
</table>

Tuition is defined as charges allocated to Central Undergraduate Tuition. Enrollment is defined as number of Primary Majors. Due to differences in data systems, tuition data were not available for all enrolled undergraduates.
WORKING DEFINITION OF DONE

- All data elements used in a report have approved DEFINITIONS
- Report specifications and appropriate usage has been DOCUMENTED
- Report has functional PARITY with Profiles, WOW, etc.
- SECURITY is implemented using standard campus systems
- Data managers confirm that the report is ACCURATE
- Report has been TESTED and VETTED with stakeholders
- Report design/display is appropriate across multiple DEVICES
- Report evaluated for ACCESSIBILITY
DATA POLICY

UMD Practices for Data Sharing
A SHIFT IN PHILOSOPHY

Individuals Approved

Data Sets / Elements Approved

Registration Data Set

Course

Term
FOUR TYPES OF ACCESS

• **Public**: aggregate, summary, basic data

• **Campus-wide**: aggregate, summary, decision support

• **Protected-use**: de-identified individual records for research or legitimate business use

• **Full-use**: individual records identified for use in research, legitimate business use, and applications
DATA MANAGEMENT RESPONSIBILITIES

- Data Stewards

- Data Managers
  
  Assure compliance with data security standards
  Define data elements
  Train and assist in data interpretation

http://it.umd.edu/dataadmin/ManageStructure/roles.html
IT PARTNERSHIP

• Collaboration – Critical Success Factor
  • New Technology & Data Infrastructure
    • Modernized Data Models & Data Warehouse
  • Enhanced Data Administration Services
    • Hiring new Data Manager (refill – in progress)
    • Data Documentation is critical for:
      • Meta-data to improve ease-of-use with data
      • Master Data Management (bridge, unit codes, SCAT, etc.)
      • Data quality and automated checks
WHAT’S NEXT?

- Stakeholder groups
- New Data Docs tool
- Priorities for Executive Reports
- Corresponding data documentation
- Vetting and roll-out of new documentation and reporting
- Gradual movement into new data architecture and tools
QUESTIONS?