• How did we get here?
• What are we doing?
• What does this mean for users?
• Helpful URLs
HOW DID WE GET HERE?
IN THE WORLD...

1997
- Oracle 8 Released
- UMD Data Warehouse

2002
- Tableau Founded
- Oracle 10g Released
- UMD Profiles Goes Live

2007
- Oracle Purchases Hyperion
- Last “DIB” Published
- Profiles “Drill Down” Expands

2012
- Amazon Redshift Released
- Current WH Expands to 7 Yrs

2017
- UMD Practices for Data Sharing

AT UMD...

1997
- UMD Data Warehouse

2002
- Oracle 10g Released

2007
- Oracle Purchases Hyperion

2012
- Amazon Redshift Released
- Profiles “Drill Down” Expands

2017
- UMD Practices for Data Sharing
TAKING INVENTORY

Oracle Tables for System Integration

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

Generally Nightly

Census Dates

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

Shadow Databases
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• **Design is Outdated**: Warehouse was designed using outdated standards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
<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Performant, Outdated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>
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
WHAT ARE WE DOING?
# PROJECT OVERVIEW

<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><strong>Description</strong></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><strong>Legacy Systems</strong></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><strong>Risk Mitigation</strong></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><strong>Enhancements</strong></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><strong>Solution Space Systems</strong></td>
<td>Tableau, AWS</td>
<td>SQL, Tableau, Jasper, SAS, ...?</td>
<td>Amazon Redshift</td>
<td>Data Docs</td>
</tr>
</tbody>
</table>

### Notes:
- AWS DW transformation
- Ensuring all data are documented in DataDocs prior to moving to the new DW
- Field level definitions - DAD Replacement Report Documentation
- Table Documentation
- Data Docs
A SHIFT IN PHILOSOPHY

Individuals Approved

Data Sets / Elements Approved

Registration Data Set
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
WHAT DOES THIS MEAN?
CHANGES FOR USERS

• Finding data
  • New location, hopefully consolidated and easier to find
  • New look and feel, hopefully more intuitive

• Generating data
  • New tools, hopefully with more stability and features
  • Some pain in learning, hopefully with community support

• Documenting data
  • New tools, hopefully with collaboration and support
  • New responsibility, hopefully with payoff
HOW CAN YOU PREPARE

• Identify:
  • Questions / decisions
  • Data sets that answer these

• Start learning:
  • Install Desktop Tableau
  • Explore resources

• Prep your data:
  • Data documentation
  • Data cleanup