Agenda

- Business Glossary
- Summary of the Meetings With Data Domain Trustees
- Sharing Information
- Administrative Data Management Policy
- Best Practices
Agenda

- **Business Glossary**
- Summary of the Meetings With Data Domain Trustees
- Sharing Information
- Administrative Data Management Policy
- Best Practices
The URL/web address to access the Business Glossary is: www.nyu.edu/business-glossary.

There is not a high volume of content yet, but the Glossary is in production.

[Live Demo]
Agenda

- Business Glossary
- Summary of the Meetings With Data Domain Trustees
- Sharing Information
- Administrative Data Management Policy
- Best Practices
Meetings with Data Domain Trustees

- Meetings took place over summer
- Some major themes emerged
  - Training
  - Data Accuracy and Cleanliness
  - Data Security
  - Systems Integration
Training

• Develop self-paced, online training modules
• Develop and publicize “tip sheets” for various systems
• Encourage *ad hoc* users to ask questions and get refresher training
• Instill a sense of ownership in all users for data in University systems
  – Encourage data users to report all instances of “bad” data
Data Accuracy and Cleanliness

• Create standardized reference tables across all systems
• Build in more data entry validation checks
• Take advantage of real-time validation possibilities
Data transfer can be secured by using NYU Box.
Systems Integration

- Develop automated interfaces between systems to replace manual processes
- Improve existing automated interfaces to more fully meet needs
Other Themes

- Encourage sharing of information by individuals who attend meetings as delegates from their schools/offices
- Investigate possibilities of moving to self-service applications
- Address issues relevant to the Global Network University as systems are implemented and maintained
Agenda

- Business Glossary
- Summary of the Meetings With Data Domain Trustees
  - **Sharing Information**
  - Administrative Data Management Policy
- Best Practices
We are dependent on one another for clean, accurate data, and for keeping up with system changes which affect multiple offices. Recognize that you may not always know all of the players who need to be informed of changes, so err on informing more offices rather than fewer.
The Administrative Data Management Policy has been approved as an official NYU policy.
Administrative Data Management Policy

• Statement of Purpose

All Administrative Data is owned by New York University and, as such, all members of the University community and affiliates are responsible for appropriately using and safeguarding that data.
Access and Confidentiality – only access data you need for your job, and keep that data confidential.

Training – make sure people are trained before being given access to data systems

Integrity, Validation, and Correction – clean and accurate data is everyone’s responsibility

Extraction, Manipulation, and Reporting – data, in whatever format, needs to be protected and used in an appropriate manner.
Administrative Data Management Policy

• Specifies roles and responsibilities
  – Data Trustee and Data Trustee Committee
  – Data Domain Trustee
  – Data Steward and Data Stewardship Advisory Group
  – Data Custodian
  – Data User
  – Office of Institutional Research

Link to policy can be found at www.nyu.edu/IR/DataGov
One of the thoughts when we began these quarterly Data Steward meetings was that we would have you, the Data Stewards, share some of the best practices you are following in your own office so that others could benefit from your experience. I reached out to one Data Steward who I know has done a wonderful job incorporating processes and procedures to ensure data integrity. I am very pleased to have Lisa Murphy, from the University Development and Alumni Affairs (UDAR) Office, share with you some of the things her office does to keep their data clean.
DATA INTEGRITY

NYU | University Development & Alumni Relations

University Data Steward Quarterly Meeting
Data is examined in two manners: on the way into Advance, and once in Advance.
Data arrives from the indicated data sources, and is loaded to a work table. The first set of clean-up scripts checks the data in the work table. Then, as data moves from the work tables to the database, there are additional checks performed.
These are some of the checks which are done on data coming from the University SIS and Graduate Stern AIS.

Salutation prefix is populated by Sex.

There are transformations on codes for academic information: School, Degree code, Major 1 and Major 2, Department, Grad Date, Honor 1 2 3 all go through conversion
These are just some of the automated checks performed during the Autoloader routine. MANY of these apply to manual entry as well....

<table>
<thead>
<tr>
<th>Dataloader Checks (subset)</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Validating TMS codes to ensure they exist and are active.</td>
</tr>
<tr>
<td>- Validating loaded data match the target column’s data type, e.g., it validates a datetime column contains valid date format.</td>
</tr>
<tr>
<td>- Validating start dates fall before stop dates.</td>
</tr>
<tr>
<td>- Validating users have the correct Application access.</td>
</tr>
<tr>
<td>- Validating email addresses are formatted correctly.</td>
</tr>
<tr>
<td>- Validating year falls within valid range (1700-2100).</td>
</tr>
<tr>
<td>- Validating phone is formatted correctly.</td>
</tr>
<tr>
<td>- Validating numeric values are numbers.</td>
</tr>
<tr>
<td>- Validating Entity IDs are valid.</td>
</tr>
<tr>
<td>- Validating check box/radio button values are provided (default values are inserted, if not provided).</td>
</tr>
<tr>
<td>- Validating partial dates contain correct values.</td>
</tr>
<tr>
<td>- Validating required fields have values.</td>
</tr>
<tr>
<td>- Validating min/max size (character fields).</td>
</tr>
<tr>
<td>- Validating min/max range (numeric fields).</td>
</tr>
</tbody>
</table>
Currently addresses are verified as added to Advance using software from QAS. We will be moving to a different vendor, Initium, in the future.
Conditional Validation

- Selection in one field determines available options in the next
- Fields invisible unless specific selection made
- Default values (based on user logged in)
- Pop-up validation messages
- Staff table lookup to auto-populate other fields

POP-UPS: When an Advance web input form is saved, the record's field values are temporarily stored in a transactional working upload ‘w’ table. A custom store procedure called a validation command can check those values, and/or perform other SQL checks to return to the user a warning (0 value), or a full stop (1 value) to prevent the save action.

Some types of conditional validation are:

**Field Influence Field**
When an end-user makes a choice in a drop down field, depending on what value has been selected, the options for other fields on the form will be limited based upon that choice.

**Field Invisible Based on Other**
When an end user chooses a value in a drop down field, other fields either appear or disappear.

**Default Values / Assignment**
Fields auto-populate with default / specific values.

**Pop Ups / Validation**
When certain conditions are not met, a pop-up warning message will display. In some instances the save will not be permitted, while in others, it will go through.

**Staff Table Lookup Reference**
This enhancement allows the end user to choose a dropdown and before it is committed to the database, more data will be pulled based upon that initial dropdown selection (referencing the Staff table).
These are some of the scripts which are run overnight to ensure data integrity.

Special Handling – the code indicating “Special Handling” is automatically removed on the scheduled end date.

Appeals – the processing of donor appeals is inactivated on the appeal stop date.

Prospects – the stage (progress) of a prospect is adjusted based on the status of Proposals.

Proposals/Assignments – if there is a gift received or some other appropriate status change, there is an appropriate change to the proposals and assignments data.
Here are some of the data integrity jobs run by UDAR, along with the frequency with which they are executed.

<table>
<thead>
<tr>
<th>DESCRIPTION</th>
<th>FREQUENCY</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;ACTIVE&quot; records without an Active &quot;PREFERRED&quot; address</td>
<td>Weekly</td>
</tr>
<tr>
<td>&quot;FOREIGN&quot; address that does not contain &quot;FOREIGN CITY ZIP&quot; information</td>
<td>Weekly</td>
</tr>
<tr>
<td>&quot;LOST&quot; records that contain &quot;ACTIVE&quot; and &quot;PREFERRED&quot; address</td>
<td>Bi-weekly</td>
</tr>
<tr>
<td>&quot;SEASONAL ADDRESSES&quot; that do not contain a start date</td>
<td>Monthly</td>
</tr>
<tr>
<td>Active Address with a past Stop Date</td>
<td>Weekly</td>
</tr>
<tr>
<td>Alumni Entities that do not have an Alum record type</td>
<td>Weekly</td>
</tr>
<tr>
<td>Deceased with an active Telephone number or Email</td>
<td>Weekly</td>
</tr>
<tr>
<td>Domestic Addresses without a City and State</td>
<td>Weekly</td>
</tr>
<tr>
<td>Entity records that are &quot;ACTIVE&quot; but the individual is &quot;DECEASED&quot;</td>
<td>Monthly</td>
</tr>
<tr>
<td>Entity records that have a foreign Address where City and State are populated</td>
<td>Weekly</td>
</tr>
<tr>
<td>Entities that are Deceased but contain &quot;active&quot; Affiliations</td>
<td>Weekly</td>
</tr>
<tr>
<td>Entity records that contain active Appeals where the Appeals are inactive</td>
<td>Monthly</td>
</tr>
<tr>
<td>Foreign Addresses that contain no information in Addr Line 1</td>
<td>Weekly</td>
</tr>
<tr>
<td>No address info and preferred address listed is an email address</td>
<td>Weekly</td>
</tr>
<tr>
<td>Non-alpha characters in any Name field</td>
<td>Weekly</td>
</tr>
<tr>
<td>Parents not linked to Students and &quot;Parents&quot; linked to Child(ren) who are not Students</td>
<td>Every Semester</td>
</tr>
<tr>
<td>Record type of Alumni, Faculty/Staff, or Student and do not contain a Net ID, N# or Email address</td>
<td>Weekly</td>
</tr>
<tr>
<td>Record type of Faculty/Staff and do not contain a Net ID, N# or Business email address</td>
<td>Weekly</td>
</tr>
<tr>
<td>Records with multiple or duplicate 'Home' &amp; 'Business' address types.</td>
<td>Weekly</td>
</tr>
<tr>
<td>ETC...</td>
<td></td>
</tr>
</tbody>
</table>
On a quarterly basis we invoke two batch processes for addresses. The first uses the Postal Service’s National Change of Address database to update addresses based on change of address information received by the Postal Service. We also process all records noted as “Lost,” i.e. that mail has been returned or that we have somehow lost contact with the alum, through a screening database. We continually find new, valid addresses through both of these processes.
We are not done. We are looking to hire our first Data Analyst, who will be working on data integrity issues. We are designing an “import hub” which should ease future updating to new systems or data structures. We are increasing our outreach activities to our alumni.
Thank you!

Lisa Murphy
Associate Director, Information Systems
University Development & Alumni Relations
Information Technology Services (UDAR ITS)

lisa.murphy@nyu.edu
212.992.7611
Thank You!