Data Steward Meeting

December 2, 2014

NYU Institutional Research And Data Integrity
Good afternoon. I’m Fred Cohen, and I’m the Associate Director of Data Governance and Institutional Research. In this position, I also carry out the role of Chief Administrative Data Management Officer for NYU.

Today, we’re going to focus on two main topics:
1. Reviewing what you told us at our kick-off meeting about the data-related problems you face.
2. What are some of the upcoming data management projects which are planned or in the works

And then we’ll have some fun with an exercise that will help you in the future with your data steward work.
We had a good turnout at the first meeting, and hopefully it gave everyone idea of what the data management project is all about.

We ended that session with an exercise, where we asked: What data-related problems are you currently facing in working with NYU administrative data systems? The most common responses were....
10 – **Greater integration among all administrative systems.** This will most likely occur as we develop new systems; it may also happen as need arises.

9 – **Data profiling to find bad data.** We can produce standard exception reports to identify problems which need to fixed.

8 – **Data lineage available so know where data came from.** The goal is to know what element on what screen affects which data in which reports or tables for *ad hoc* reporting.
7 – Let “downstream” users know of changes which will affect them. There is a need for better communication regarding changes to processes, data values, etc.

6 – Enhance training regarding information available, how data captured. what information is on which report; what information is in which reporting environment; what are the business processes which create data

5 – Clean up duplicate records and prevent new ones from being created.
4 - **Add data integrity constraints in administrative systems.** We want to keep data clean at point of entry by making appropriate fields mandatory; checking for valid values or combination of values.

3 - **Use data for intended purpose.** For example, don’t use housing billing dates to control door access in residence halls, as they are not the same.

2 - **Streamline method for requesting/approving access across systems.** Make the process easy for someone new to university by having the ability to request access in one place.
1 – Create a data dictionary/business glossary so that there is a common understanding of the meaning of terms. It’s basic: If we’re not speaking the same language – can’t work together.

As an example: What does “Secure the building” mean?
To the Navy it means to turn off the lights and lock the doors.
To the Army it would mean to occupy the building so no one could enter
For the Marines it means assault the building, capture it, and defend it with suppressive fire and close combat.

One phrase, but very different meanings to different people.
Agenda

- Review of Responses to Data Problems/Solutions from last meeting
- **Upcoming Projects**
- Interactive Exercise
The first quote may not be the exact quote, but it summarizes the gist of the conversation between Alice and the Cheshire Cat in *Alice in Wonderland*.

That’s not the approach for us.

Our destination is improved data for improved operations and decision making. But if we don’t know where we are starting from, we can’t plot a path to reach our destination.

Kitty Bridges from NYUIT will discuss the first of our upcoming projects, which will deal with getting our bearings on where we are.
Data Management Maturity Model

• Maturity Models
  – Where we are
  – Where we’re headed
  – What we need to do to get there

• Carnegie Mellon Maturity Model
  – Adjusting for NYU

• Surveys in February? To determine our maturity in a number of areas
Two Parts for Data Management

• Departmental Process, e.g.
  – When, how do we check on data accuracy?
  – Roles assigned for data management activities?
  – Business Glossary for functional terminology in place?

• Systems, e.g.
  – What checking is done in an application for data integrity?
  – Roles
  – Technical Data Dictionary (data element x is in table y and has values of a, b, or c)

• Both, e.g.
  – Technical Lineage in place (data element x gets added to data element y when it is transformed for UDW+)
<table>
<thead>
<tr>
<th>AREA/THEM</th>
<th>Office/ Data Area 1</th>
<th>Office/ Data Area 2</th>
<th>Office/ Data Area 3</th>
<th>System 1</th>
<th>System 2</th>
<th>System 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Governance</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Process &amp; Roles</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Data Integrity</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>1.75</td>
<td>2.4</td>
</tr>
<tr>
<td>Tools</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Architecture</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Upcoming Projects

Meetings with Data Domain Trustees
• Identify major data-related problem areas
  – Data received from another office
  – Data created and used in your office
  – Data sent to other offices
• Begin discussions about Service Level Agreement (SLA) for correcting bad data
Upcoming Projects
Data Definitions

Types of Definitions:

1. Business Process Definition

Procurement: i-Buy

i-Buy is an electronic marketplace that integrates with Procurement’s e-Req/i-Buy System. You can search for items, create a shopping cart and order items directly with the supplier. You can create one cart with many suppliers and the system will automatically create separate purchase requisitions for each supplier. Once the purchase requisition is approved by your departmental administrator, a purchase order will be immediately transmitted directly to the supplier, without review/handling by Procurement.
Upcoming Projects
Data Definitions

Types of Definitions:

2. **Business Term Glossary**

*Payment Handling Code* is a combination of characters used to define the routing of a payment to a vendor. Examples of valid values are M (mail); E (pull enclosures to send with the mailed check); and H (hold payment for pick up).
Upcoming Projects
Data Definitions

Types of Definitions:

3. Physical Data Dictionary

Environment: EPMPROD
System: PeopleSoft Campus Solutions
Table: SYSADM.PS_F_SRVRC_INDCTRS
Column: PLACED_METHOD

(example; not true entry)
Now that you’ve seen the four types of definitions, let’s focus in on the second type – the Business Term Glossary. As you saw earlier – this is the number one request from our exercise last meeting.

We’ve heard you, and the Decision Support Group has begun working on a pilot project. It is still in development, but I’d like Steve Vassallo to give you a brief overview of the shape of things to come. The examples may not be exactly what will be used, but will give you an idea of the information that will be available.
Introduction

• **August 2014:** administrative data stewards and trustees met to discuss data-related issues with NYU administrative systems.

• While a number of problems were identified, there was overwhelming consensus on one particular issue:

  *There needs to be a *business glossary* / *data dictionary* to ensure there is a common understanding of the meaning of terms.*
Proposed Solution

• Utilize ServiceLink Knowledge Base to create Wikipedia-like knowledge structure.

• Define terms **but also** contextualize and connect them to related concepts.

• Allow for seamless navigation to related terms without searching.

• Connect business terms and concepts to other NYU resources (e.g. FinanceLink (website), iLearn (website), other web content, forms).
Navigating to the ServiceLink Knowledge Base

- Go to https://nyu.service-now.com/servicelink/
- Enter any search term or phrase.

About Service Link

Use ServiceLink to get information about services, and assistance from a variety of NYU service centers and service providers. You can find information and get help from:

- Decision Support Group (Financial/HR Reporting)
- Finance Operations & Treasury
- Information Technology (ITS, QTS, Stern, Silver Libraries IT, NYU Abu Dhabi, and NYU Shanghai)
- PeopleLink (Human Resources, Payroll, and Benefits)
- Sponsored Programs and Accounting
- Student Services (Bursar, Financial Aid, Registrar and Residential Life & Housing)
Example: Purchase Order

Procurement: Purchase Order
Article: 180076873

A purchase order is a document sent to a supplier or vendor, authorizing shipment of a product to the customer at a specified price and terms. The creation of a purchase order creates a legally binding contract which cannot be changed without the consent of both parties.

There are different processes to create a purchase order, including:
1) using iBuy
2) using a Req for a non-iBuy vendor

The purchase order is shown on the UDW+ budget reports as an unnumbered. When the purchase order is fulfilled the associated amount is de onumbered and becomes an actual expense.

Click here for Buying and Paying Information from NYU Financial link

source: draft article in ServiceLink Knowledge Base

NYU Institutional Research
And Data Integrity
Example: **Purchase Order**

**Procurement: Purchase Order**

A purchase order is a document sent to a supplier or vendor, authorizing shipment of a product to the customer at a specified price and terms. The creation of a purchase order creates a legally binding contract which cannot be changed without the consent of both parties.

There are different processes to create a purchase order, including:

1. **Using iBuy:**
   - Link to “how to”
   - The purchase order is shown on the UDW+ budget reports as an account. When the purchase order is fulfilled, the associated amount is de-encumbered and becomes an actual expense.

2. **Link to an additional term**
   - Link to external knowledge source

---

**Source:** draft article in ServiceLink Knowledge Base

---

[Logo: NYU Institutional Research and Data Integrity]
Example: click on “encumbrance”

Procurement: Purchase Order

A purchase order is a document sent to a supplier or vendor, authorizing shipment of a product to the customer at a specified price and terms. The creation of a purchase order involves a legally binding contract which cannot be canceled without the consent of both parties.

There are different processes to create a purchase order, including:
1) using RUP
2) using e-Req for a non-Buy vendor

The purchase order is shown on the UNH+ budget reports as an encumbrance. When the purchase order is fulfilled the associated amount is re-encumbered and becomes an actual expense.

Click here for Buying and Paying information from NYU FinancialLink

Encumbrance

An encumbrance is a method to set aside or reserve a budget for a specific purpose.

This is to ensure that the department has sufficient funds reserved to for that purpose. An encumbrance is not an expense. An encumbrance is a higher level of commitment than a budgeted expense, because monies are reserved for a specific purpose. However, the transaction has not occurred yet so it is not an actual expense.

source: draft article in ServiceLink Knowledge Base

NYU Institutional Research and Data Integrity
Next Steps

To create a fully developed and integrated knowledge structure, we need:

– Work with data stewards to continue developing glossary terms and other content.

– Create a “sub-style” guide within the current Kbase style guide:
  
  • create ~ 3 standardized templates for the different types of articles we want to include: example, pure definition, systems article, etc.

– We need to setup a formal process of “sign-off” and maintenance of articles.
The first definition of Customer ID is a tautology – the definition is circular and doesn’t offer new information.

The first definition of spoon is incomplete. It allows for other eating utensils such as a fork or chop sticks.
Data Definitions

What makes a good definition?
A definition containing at least 2 of these 3 components:
Broad Term – “is a”
Distinguishing Characteristic – “has a”
Function Qualifier – “used for”

Is a = general class to which the term belongs
  A student is an individual
  A college is an organization

Has a = what distinguishes this term from the broader term
  A student has enrolled classes
  A college has an academic program

Used for = how the term is being used, usually involves a verb
  A student is being educated
  A college is used for educating students

A student is an individual who is enrolled in classes to gain an education.
A college is a learning organization that has one or more academic programs used for educating students.
A student is an individual who is enrolled in classes to gain an education.

A college is a learning organization that has one or more academic programs used for educating students.

Both of these examples have all three parts.

Broad term – is a
Distinguishing characteristics – has a
Function qualifier – used for

OK – now it’s your turn.

We’ll start with a group exercise, and then move to individual exercises.

At your tables, work together to come up with the best definition for these two terms:
  faculty
  donor

After x minutes, I’ll ask you to switch to an individual exercise, where I’d like each of you to define one term you use in your office. We’ll give you just a few minutes.
Interactive Exercise

As a Group

Define:

- Faculty
- Donor

Individually

Define one term you use in your office. Try and choose a term which might be ambiguous if not clearly defined.
Thank You!