PPMO-0053: Faculty Interactions with Corporations Data Analysis Summary

Business Case

In many instances across the University of Illinois System, the same record is stored in multiple systems and maintained by separate and different procedures in multiple departments. Of immediate interest due to the Blackbaud and Kuali Coeus implementations is the corporate record. The corporate record is used to manage grants to Sponsored Research, endowment contributions and gifts, and university procurement. Currently, tracking interactions with corporations and foundations is handled by each organization separately. Sponsored Research, the UI Foundation, and other units interacting with corporations and foundations have their own procedures for capturing data, their own naming conventions, and their own systems and databases. There are currently at least three independent systems storing corporate records: Banner, Blackbaud, and InfoEd/Kuali Coeus. This duplication of data adds complexity and confusion that manifests in lost opportunity. No single system can give a complete picture of the relationship between a corporation or foundation and the University of Illinois.

Analysis

The team reviewed discrepancies in the naming convention guidelines for Banner, Foundation, and Sponsored Research systems, compared the data structure of these three systems, and assessed the degree of organization name-matching that occurred among the systems in a sample of the top 100 corporations by expenditure from FY15. The team also performed an impact analysis to determine how much work is required to pursue naming conventions.

To gain a basic understanding of the situation we performed some simple comparisons of names utilized for organizations in Banner, Blackbaud and Kuali Coeus. It should be noted that these comparisons were based on data received through queries and requests to individuals that were not reviewed by an additional error correction process. As a result, each dataset may have many assumptions in the underlying query that delivers the data to us. An example of such a case was where we eliminated records from Banner where an organization was an individual, which we assumed, were university employee payments.

Statistics for the data set were as follows:

265,718 rows of data from Blackbaud

15,602 rows of data from Kuali Coeus

108,862 rows of data from Banner

With this data we performed direct name comparisons with uppercase conversion between the appropriate fields. Our analysis returned only a 15% or less rate of agreement between the datasets. Matching statistics are available at the end of the document.

Additionally, the name (address1, sponsor name, Spriden Last Name) fields in each system were of different lengths and name fields were also being used to track additional business needs. These reasons led us to the conclusion that the business cost of transitioning to a single naming standard would be very high.

Final Recommendation

Organization (corporations, vendors, sponsors, etc.) names and identifying codes are stored in many systems across the university. Based on the business process a system supports, the name of the organization and its attributes stored in that system may be different from other systems. However, a common identifier among nearly all organizations is their Federal Employee Identification Number (FEIN). We recommend that a central data store be generated and maintained including the Federal Employee Identification Number and the Preferred Name for all organizations having a relationship with the university. Any system storing organization-related information would store the FEIN while managing other vendor attributes as needed internally. Any requests to assemble a comprehensive dataset of an organization's interaction with the university can then combine data from all systems based on the common FEIN and show the Preferred Name if desired. Existing products like the LexisNexis® suite of tools may provide this functionality already and should be explored before a local solution is developed.

Appendix

Matching results by direct name comparison with uppercase conversion for both fields were:

1,080 matches between Blackbaud "Address1" and Kuali Coeus "Sponsor Name"

2,341 matches between Banner "Spriden Last Name" and Kuali Coeus "Sponsor Name"

9,974 matches between Banner "Spriden Last Name" and Blackbaud "Address1"