Data Warehouse and Analytics Discussion Paper

3/1/2013

Luna Rajbhandari, Director of Business Intelligence Solutions
Jason Schober, Director of Project Cafe

I.   Executive Summary

The investments made in the past decade in our enterprise business systems have improved many of our business processes and reporting capabilities. These systems continue to amass large amounts of data - data that could be further turned into valuable information and insight if integrated across our systems in an enterprise data warehouse.

Our Business Intelligence development priorities over the last few years were mainly driven by the urgency felt in enabling reporting within each enterprise business system. Each major business system now has a corresponding data mart which provides some level of reporting capability on the NU Business Intelligence (Cognos BI) platform. Please refer to Appendix 1 for a summary of current state of the data marts and the identified gaps based on known business requirements.

Despite the progress made in the last few years, there are still many unmet needs on campus as they relate to data access and integration. Our community continues to seek ease of access to analytical data and timely insights on our business enterprise without being limited to system-specific information silos. To address this growing need, this discussion document seeks to: a) provide an assessment of our current state, b) describe ongoing challenges in providing access to integrated data and c) suggest a model for addressing our needs moving forward.

In writing this paper, we are guided by the principle that our tools, processes and systems should enable a culture of analytics and data-driven decision making on campus by promoting ease of use, self-service and role-based data access.

Many recommendations are presented throughout the document around topics of staffing, self-service, data integration, data governance and information management, the highlights of which include:

1. **Refinement of the center's focus:** We recognize that a central BI team cannot meet all distributed reporting and analytics needs on campus. The center should therefore focus on enabling self-service usage of BI tools by providing the skills, processes and technology platform to drive adoption and usage of BI tools on campus. The focus of this team should be on providing highly specialized services such as data warehousing (data modeling, data loading, and data integration), advanced BI(reports, dashboards) development or cross system analytics, training the trainers, BI project lifecycle management, establishing BI standards and best practices, and coordinating and enabling a community of BI practitioners.

2. **Development of skills and resources outside of the BI team:** We need to expand the population of report developers skilled in creating Cognos BI reports and dashboards within each major enterprise function. We also need to continue to cultivate power users of self-service tools within the community, both in central units such as Institutional Research as well as analytics teams within all schools. Robust training and mentoring programs are required to support both of these populations, along with commitment across the institution to develop the analytics resources that are central to developing the culture of analysis and data-driven decision making.

3. **Data integration and access:** Much progress has been made in expanding the amount of data, and in improving the quality and consistency of data in the Northwestern data marts. However, there is still important work to be done in order to get to the type of cross-functional analytics our community has been asking for. Individual data marts need common data identifiers, elements, and definitions. Examples include the need for a consistent employee identifier and verification that the crosswalk approach to primary management units (PMU) will be successful.

    While some of the data marts were developed in-house, other data marts were purchased from external vendors and customized in-house (e.g., the financial data mart was purchased from Oracle, and the student data mart was purchased from Blackboard). Decisions on how best to approach integration and technical effort required in integrating these data marts will need to be prioritized against other needs.

    Similarly, disparate security rules in different data marts need to be matched and integrated. As the build out of individual data marts reaches maturity, and the focus turns to cross-functional analytics, having a group work through data access policies becomes increasingly important.

4. **Adoption of data as a service:** We expect that data access needs above and beyond those met via front-end tools such as Cognos BI will continue to exist and may even increase. We have an opportunity to implement a data as a service model, where another web portal or application could consume the information real time or nightly from the data warehouse through the NU Business Intelligence web service interface. Just as these services are being discussed for access to data in other enterprise systems, the same approach makes sense for providing the enterprise data that has been extracted and loaded into the BI environment. Here again, the need to have a group that develops clear data access policies, and insures that the guidelines are being followed, is very important

This paper is not a final strategy document; our hope is to provide necessary information and context to inform any future strategic decision making in the areas of data access, data integration and analytics. By sharing this now with our stakeholders across our partnering schools and central units, we hope to articulate and validate a shared vision and to capture what we have missed.

II.   Introduction

The topic of data access, analysis, and integration is consistently cited as a top area of need with regard to our information systems. Reporting and analytics has been treated as a high-priority area, with a great deal of time and effort invested since the selection and purchase of Cognos BI in 2006 and many efforts prior to that point within each of the University's enterprise systems. Yet, given this investment to-date, the areas of reporting and analytics remain one of the top areas of as-yet-unmet demand. We believe this demand can be summarized along two main themes: 1) preparing our institution to make use of the available features and capabilities and 2) making real investments, both in terms of effort and priority, into integrated reporting and data access capabilities.

As it relates to delivering enhanced reporting and data access capabilities from our enterprise information systems, we have accomplished quite a bit. The number of function-specific warehouses and marts that have been produced thus far and the volume of reporting that is conducted from these databases are significant. Even more capacities along these lines are being developed this fiscal year, both broadening and deepening the information available to the community. This development has enabled better decision-making, bringing speed and depth of analysis to processes and functional areas that previously required a great deal of manual effort to analyze. We now need to take what has been built and maximize its usage. This requires awareness of the available capabilities, the continued development of power users in the community, and an increase in our report developer population. Training and support are also required to sustain this resource base.

Yet, even with a substantial base of data and access tools available, we continue to see unmet demand. Users want greater ease of use and data to be made available in different formats. Users want access to data sets that are as yet unavailable. In many cases, these data sets require integration across system sources to unleash the true value of our institution information. Such integration requires both focus and effort (e.g., making these areas our priority) as well as work at a foundational level. The presence of common identifiers and the need for a shared approach to data security, including policies around the secure and appropriate sharing of data continue to pose hurdles that must be overcome.

We need a unified, strategic approach to advance beyond our current state as it relates to data access, analysis, and integration. If we do not pursue this as a priority, NU will continue to have greater data availability/depth in some areas, while gaps will persist in other areas. If such gaps persist, users will continue to pursue local solutions or spend their time compiling data sets instead of analyzing and making better decisions. Without the ability to readily access and analyze data within and across our enterprise systems, we are not reaping the benefits of our investments.

Building Business Intelligence solutions involves much more than simply pointing a business intelligence tool (such as Cognos) at a ready-made repository and creating reports to order. It is a full-scale, systems integration effort. Such efforts require project planning, resource staffing, business requirements analysis, architectural design, data modeling, hardware configuration, software development, integration testing, user training and system implementation. We have made a great deal of progress, yet have much still yet to do. This document outlines our vision for reporting and analytics, provides an assessment of our remaining needs and describes the implications of addressing these needs.

III.   Background and Strategy

As the University approached the replacement of its legacy financial systems several years ago, a workgroup was established to articulate our requirements and expectations of our future data access and reporting capabilities. At that time, like many of our peer institutions, the Northwestern community had historically met its reporting needs by deploying a variety of solutions. Developed by different units and users, many of these solutions addressed similar reporting needs and ultimately they offered a disjointed set of reporting alternatives. As a whole, these solutions were not integrated and thus not easy to use, requiring multiple log-ins and permissions, and were not providing all the functionality that business users needed.

The reporting workgroup recognized that our enterprise business systems are fundamentally designed and optimized to support business operations. They provide hundreds of screens that allow user to add, update, and retrieve records in support of essential business processes. Yet, when it comes to reporting, these applications typically provide relatively few "out-of-the-box" reporting capabilities.

As such, it was broadly agreed that NU would evaluate, select and implement a business intelligence (BI) tool to meet our reporting and data access needs. The reporting strategy workgroup in its January 2007 report titled "Recommendation for a Financial Reporting Strategy" recommended adherence to the following principles for the development of reporting solutions:

- Promote Ease of Use
- Provide Integrated Data
- Meet a Range of Reporting Needs
- Support Flexible Report Distribution, Scheduling and Printing
- Ensure Data Quality
- Deliver Training and User Support Services
- Maintain Compliance with Data Access Policies

The recalibration strategy outlined by the Enterprise Systems Executive Committee in its May 29, 2007 memorandum stated "While the focus of our administrative systems efforts over the next two years will remain on implementing the new financial system, we understand that robust data access and reporting solutions beyond the new financial system continue to be goals that are desired by both ESEC and the rest of the community. Resources will be reallocated to a small business intelligence team that will coordinate the engagement of the schools and central administration units in the development of a business intelligence deployment strategy."

As a result of the recalibration, a dedicated team within Project Café would continue to develop the data warehouse and reporting capabilities for NUFinancials, FAMIS and InfoEd while the new BI team would provide common infrastructure for enterprise BI capabilities and focus on warehousing the remaining administrative data (SES, FASIS, AIMS/CATracks and other research administration systems).

The central BI team was established in November 2007, and in a joint report on BI strategy presented by the Office of Change Management (OCM) and NUIT to the Board of Trustees in 2008 noted its goals as below:

1. Providing clear communications regarding the functionality and usage of the Cognos BI tool

2. Ensuring consistency in training, data security, data definitions, and data architecture
   a. Coordinating efforts with the reporting efforts underway within Project Café, HRIS, SES, and AIMS (now CATracks)
   b. Coordinating efforts with all existing enterprise systems' stakeholders
3. Deploying the Cognos BI tool for use in reporting.
4. Developing a consistent technical approach for data loading, data modeling, data security, and data storage
5. Developing a consistent technical approach for test, development, and production environments for the BI tool
6. Providing BI product expertise
7. Fostering sharing of vision and experience among a cross-university network of BI leaders by:
   a. Advising on philosophy and business objectives,
   b. Facilitating the sharing of developed expertise within units and school with other areas
      i. Creating a listserv
      ii. Creating an NUBI user group that will meet regularly
8. Working closely with the Office of Analytical Studies to address broader institutional reporting needs with those of the schools and individual administrative users
9. Facilitating the University's migration toward a more comprehensive solution serving the institution's business need for administrative data access.
10. Providing a central point of contact for the BI vendor and managing the relationships.

A BI Advisory Group formed subsequently in 2008 guided the strategies and priorities of the BI team. In order to deliver on the priorities, an incremental approach was taken by the BI team in building the data warehouse infrastructure, with projects with highest business value and breadth of usage determining the order in which data was populated in the data warehouse. A lot of attention was also given to building foundational elements of the data warehouse while concurrently meeting reporting needs on campus.

IV.   Our Platform and Tools

Since the formation, recalibration and establishment of the above strategies, the University has made significant progress in many fronts, both in deploying Cognos BI and in implementing the underlying data warehouse infrastructure or data marts for various subject areas. All the data marts today are accessed via Northwestern University's standard BI tool (Cognos BI). The data marts and the BI tool enable prompted reporting, ad-hoc querying/analyses, scheduling and other BI functionalities. Some of the reports support day to day operations while others provides aggregations, trends and other management level views of the data to support strategic decision-making.

The major functionalities of the Cognos BI tool include:

- Standard "authored" reports
- Ad Hoc query tool (Query Studio)
- Ad Hoc analysis tool (Analysis Studio)
- Advanced report authoring capabilities (Report Studio)

The BI System currently has approximately 2,600 active users on campus. The users represent all the schools and central units of the University. A majority of the users utilize standard prompted reports for

self service. Most users also have access to one or more ad hoc reporting or analysis tools. However, the usage of these more advanced self-service tools is limited to less than 200 users.

In addition, many of our enterprise systems still utilize other reporting tools such as Vista and Crystal Reports. Many of the Crystal Reports in SES have been migrated to the BI system, and some of the Vista reports for FASIS are being converted into BI reports. However, the independent reporting tools continue to serve unique needs.

It is also worth noting that the latest version of the BI tool (upgraded in 2013) also provides many other capabilities, most notably increased focus on self-service including users' capability to author or customize dashboards.

V.   Current State

The University has taken an evolutionary approach to building the data warehouse, meaning additional subject areas are incrementally built every fiscal year to address the highest priority business requirements. A great deal of reporting and data access capability is already in place. Still more is in progress, planned for delivery in FY13, or proposed and under consideration for delivery in FY14. Please refer to Appendix 1 for a more detailed data warehouse gap analysis

The last few years of development priorities have been in establishing many of the foundational data elements in the data warehouse (or data marts for student, financials, faculty and staff, development, research, etc.) to provide self-service reporting, whereas the newest set of priorities have evolved towards more integrated reporting capabilities. Some of the requested integrated reporting and analytics capabilities include:

- SES/FASIS integration for faculty course loads (expected vs. actuals)
- NUScholars data integration with FASIS
- NU Dashboard (School or University level) with institutional metrics from Student, FASIS, NUFinancials, Research
- Full Student Lifecycle analysis (applicants to alumni - Admissions, Records, Alumni/Development)
- Endowment Stewardship Reporting (Fundriver, NUFinancials)
- Alumni engagement data analysis (CATracks and EngageNU)
- Research: PI Portal (Financial – budget, expenditure; Research – protocols, proposals, awards; Student – tuition; HR – stipends)
- Research: PI Linkage Tool (ACUC, IRB, InfoEd)

In order to meet these growing needs on campus, a unified approach is needed to manage our information infrastructure and redefine our strategy for evolving the data warehouse.

VI.   Future Vision and Strategy

Before providing specific recommendations, we would like to describe a future state we aspire to reach in which our tools, processes and systems promote ease of use, self-service and role-based data access in order to enable a culture of analytics and fact-based decision making across campus.

We would like to articulate the following future vision for the use of BI in the University.

- Users will have access to the data they need to support their job function in the University.
- Users will not be limited to system silos in accessing the data they will need. Institutional metrics and analytics will span multiple systems.
- Users will have access to local power BI users within their schools or units who can facilitate their use of self-service ad hoc analysis and ad hoc reporting tools.
- Users will have access to training courses that can ramp them up on the analytical tools needed to excel at their jobs. A "train the trainers" approach will develop specialized, distributed trainers with a training center coordinating user centric training efforts.
- There will be institution level data governance with an established data access policy that supports an analytics culture, treats all enterprise data as the University's asset and enables access to the information in line with job duties and responsibilities of individuals on campus.
- The multiple data marts that form the foundation of future BI needs are robust enough and rich enough to facilitate the analysis that users will need to perform within a functional area.
- There will be a distributed set of advanced reporting skills that facilitate the delivery of advanced reports and dashboards within a functional area. The distributed reporting resources will conform to University BI standards and best practices established by a central unit.
- All data marts will share common identifiers, dimensions and data definitions to support integrated reporting and analytics.
- There will be a central unit empowered to integrate data and deliver BI applications that provide cross system analytics. The central team will possess advanced skills and provide specialized services such as data warehousing, advanced reporting, BI project management, training the trainers and facilitating a community of BI practitioners.

We recognize that a future data warehouse strategy should consider both organizational and architectural implications.

i.   Organizational Strategy

The organizational strategy should address efficient usage of resources across the organization ensuring effective balance of central staff, distributed advanced reporting resources and a larger community of power users across the organization. A combination of central and distributed BI resources enables self-service, agility and efficiency. The model we are suggesting here utilizes a central team to develop a common set of platform and BI capabilities, integrate data and coordinate distributed efforts. Dedicated resources with advanced skill levels within functional areas or enterprise system teams meet ongoing reporting needs within the functions. Distributed analytics staff or power users in schools and central units provide local expertise. A broader community of faculty and staff consume information as reports or dashboards developed by one of the groups within the circle.

These concepts are illustrated in the diagram below and described more fully on pages 15 – 17:
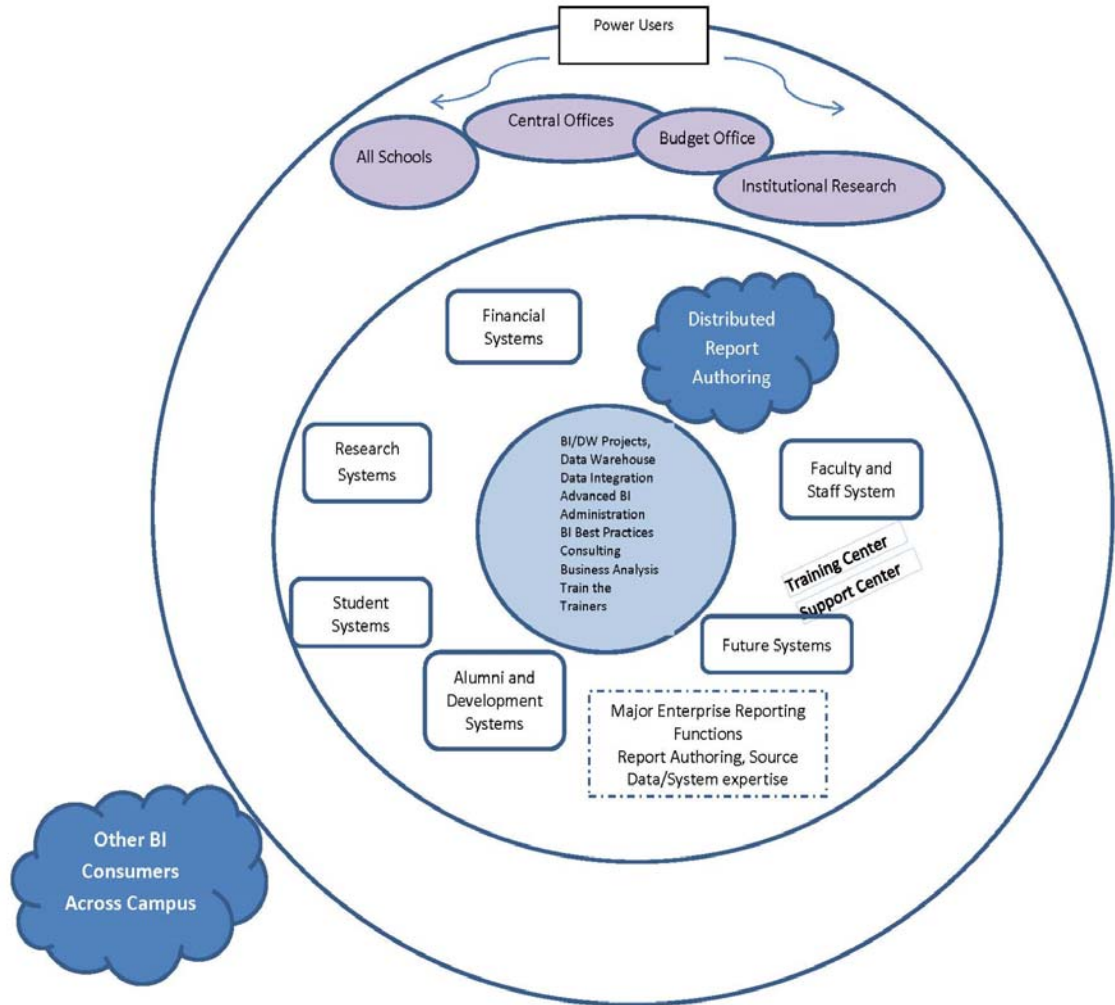
Diagram 1: Central and Distributed Resources

- **Central Resources:**
  This group provides a unified data warehouse and BI infrastructure, along with expertise to integrate disparate data sets and implement BI solution projects. This group should service the broader BI user community with the goal of promoting data access and integration, self-service, standards and best practices to ensure effective adoption of tools and solutions.

- **Distributed Reporting Resources:**
  Each functional area should have designated report author(s) who can utilize advanced reporting features in Cognos BI. These resources may reside in respective enterprise teams or may be designated resources in a central business unit such as the Registrar's Office. In addition to report development, these resources provide BI expertise within their functional area.

- **Power Users in Schools and Central Units:**
  The power users in schools and central units are the analytics resources who use data directly for decision making or conduct analysis and create reports to support decision makers. These resources need access to ad hoc reporting and analysis tools so that they can analyze data and create reports without having to rely on IT staff. These resources should be able to provide local support and expertise pertaining to appropriate usage of data and tools.

ii. **Architectural Strategy**

The architectural strategy should address how the individual data marts developed over the last few years can be further developed and integrated to support the various types of information needs on campus. In addition, it should allow for new types of data to be added to the data warehouse.

The data marts in production today are built to meet specific BI needs within functional areas. These are the foundational data marts (shown as data marts layer in diagram 2 on page 10). Although all the data marts are accessed via the standard Cognos BI interface, different tools and technologies have been used over the past to develop the underlying data marts.

An integration layer built on top of the data marts can facilitate integrated reporting and analytics. As illustrated in the chart below, the integration layer can be built as either physical databases or virtual BI models using different integration techniques. This approach ensures that the investments made to establish the various data marts are minimally disrupted. Architectural best practices such as common dimensions or 'conformed dimensions' should be applied to the foundation layer. However, additional work is needed to make the data marts conform to each other so that they can be linked through common dimensions such as person, organization and time. Any new data mart should conform to this best practice so that new data marts do not exist as silos of information.

Dashboards   Analysis   Reporting   Ad-hoc Reports   Data as a Service

BI Portal (Cognos)

Research Space Metrics

Student Lifecycle

Future Analytics

This layer may be physical or virtual.

Data Integration Layer

Finance   Space   Research   Faculty/Staff   Alumni/Development   Student   Future Marts

Each Data Mart contains multiple subject areas. All Data Marts should share Common Dimensions (Employees, Protocols, Time, PMU etc.)

Data Marts Layer

NUFinancials   FAMIS   InfoEd   FASIS   CATracks   SES   Other Future DBs

SIMS   Granite   NUCore   ISIS   EngageNU   IRB&ACUC
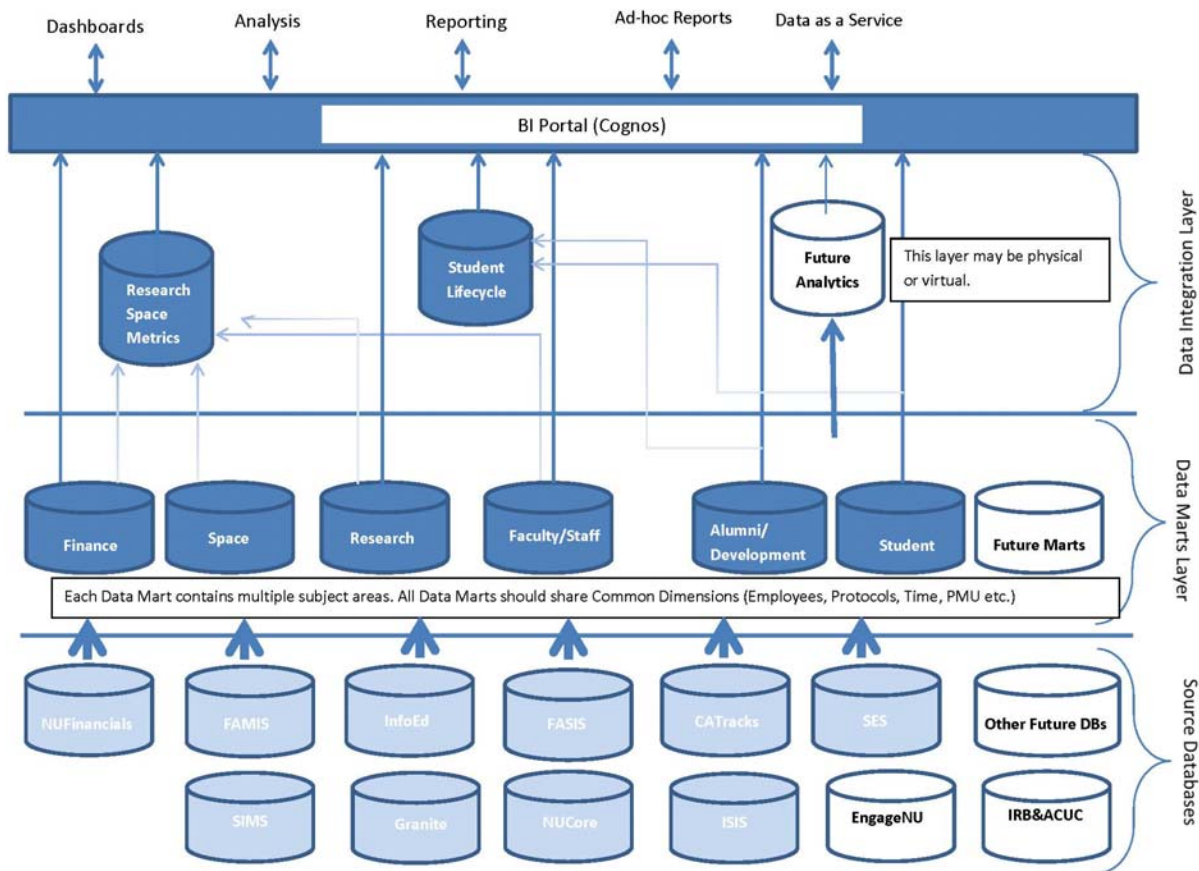
Source Databases

Diagram 2: Data Warehouse Building Blocks and Analytical Data Integration for Reporting and Analytics- Conceptual Architecture

Existing reporting capabilities and future needs within a functional area can continue to exist. All data marts should share common dimensions as an architectural best practice to enable data integration. The "Future Marts" and "Future Analytics" in the above diagram highlight the ongoing need for both extensibility and integration.

VII. Challenges and Implications:

The BI Solutions and other enterprise applications teams have been steadily at work on a prioritized list of initiatives aimed at improving data access. While a lot of the foundational data elements are in place as illustrated by the list of completed and ongoing projects, there are still a few key gaps that, if filled, could facilitate the majority of the reporting requirements from our core enterprise systems and accelerate building specialized applications that integrate data across multiple areas. Even if we are to remain on our current trajectory (using our existing approaches and resources), there are several key hurdles that lie in our immediate path. If we aspire to accelerate our pace or increase our ambitions, additional challenges must be tackled. These challenges are described below along with some of the

implications foreseen in working through these topics. The challenges can be categorized into the following two broad categories: i) Information Management and Data Governance and ii) Human Resources.

i. Information Management and Data Governance

Currently, data is governed at the individual source system level by very separate security and access policies. This is equally true for the separate BI data marts (Student, HR, CATracks, Finance, etc.) we have built over the last few years. BI solutions needed to support decision making on campus will span across these multiple functions. This presents an opportunity for the University to develop an institution level data governance strategy. Not only do we need common policies and procedures to enable necessary and responsible use of data, but also a common strategy and perspective to help shape the data warehouse infrastructure as a central enterprise repository of analytical data on campus.

a. Data Quality and Consistency: Common Data Identifiers, Data Elements and Definitions

A key component of integrated reporting is the existence of common data elements across multiple source systems. Data quality and consistency across all enterprise systems accelerate data integration. For example, in order to successfully track multiple aspects of a faculty across all enterprise systems through the data warehouse (e.g., finance, HR, student), the same employee identifier should exist for all faculty members across the three systems. However, the employee identifier is not consistent in some cases in our enterprise systems.

Also of particular note are the disparate organizational unit values/rollup hierarchies present in many of our enterprise systems. The primary management unit (PMU) project has made some solid headway towards using a crosswalk approach to this potential data integration challenge. It is unclear at this stage, however, if we can state with confidence that this approach will enable the cross-system analyses that the institution wishes to perform, or, if we need to plan to approach a standardized organizational data dimension in another fashion (e.g., change the organizational unit data structures at their source in the enterprise systems).

This type of data quality and consistency issue in our enterprise systems should be a very high priority item to address. Other such common identifiers would include time periods and financial Chart of Accounts values. Once common data elements exist in source enterprise systems, they can then become common dimensions in the data warehouse that can be utilized to link the multiple data marts to answer business questions.

Implications

- Work will need to be invested to analyze our options related to those dimensions requiring common identifiers. Building common identifiers may need capacity to be dedicated that would otherwise be available for business analysis, data modeling, etc. The portfolio of projects approved may be thus affected.

- Action plans will need to be developed to implement the chosen solutions. These will be projects in their own sense, posing a potential opportunity cost to build out this "infrastructure". All enterprise systems that utilize the identifiers or the business processes that generate the common identifiers could be impacted. Therefore, addressing this would require sound enterprise architecture design, and broad support and commitment across the various data stewards, systems owners and business owners.

b. Security and Data Access Policy

Gaining access to the University's enterprise systems generally requires the submission of a security form and each system has its own security form. On these forms are a variety of roles that users may choose between. The forms require the user to obtain varying levels of approvals. These forms are then submitted to the appropriate security administrator who performs the necessary setup tasks in each system. Most of the University's enterprise systems also utilize some form of "row level" security, meaning data are locked down from a viewing standpoint to some extent. Typically, this access is granted by the equivalent of department or school, although it varies by system. The appropriate "row level" security is requested as part of the security form and is set up manually by each security administrator. If a user has uniform access to the same data across all systems, it is the result of independent setup in each system.

Cognos BI can admit or deny users access to the reporting website based on their logon credentials. It can also determine "who can see what" to some extent based on their respective user role. Each role, for example, can determine which reports users may access and view (e.g., limiting access to senior level administrators such as deans, VPs, and their immediate staff for some reporting applications while granting broader access to reports such as budget statements).

The University has no single data governance policy by which to guide security design efforts at a global level. In some cases, this omission has already posed hurdles to data integration efforts that had to be cleared in order to implement solutions that support University operations. As the University pursues additional initiatives aimed at providing access to integrated data, it is likely that a revisiting of our current "row level" security practices will be necessary. Else, reporting solutions may need to be either 1) restricted to access by a select few users that are authorized to "view all data," 2) constructed using highly complex security logic that combines the row level security rules from each source system, or 3) developed and maintained using a separate (or multiple) security paradigms altogether.

If we could develop a security framework for data access that incorporates either role based or primary management unit (PMU) based security (or both) across our enterprise data, it would significantly improve efficiency in developing and maintaining the BI applications. Although this approach would incur a larger initial cost and work would still be required to link each BI application to this data security framework, the return on investment over the longer term in time savings/ efficiency could be significant. The key point is that an approach of custom security provisioning incurs added cost for each new application that is built in the BI tool (and is even more costly if multiple approaches are taken within an application) and limits the way the University can produce reporting/analytics on its data.

Implications

- We need to embrace a unified strategy that can be employed across system and data silos. While we may be able to build upon data classification standards work done in the past, this is likely to be a much-debated topic.
- Once resolved, we might expect that some (perhaps substantial) design and build effort be required to implement the unified security approach. Existing solutions may need to be retrofitted to embrace this approach.
- Again, broad support and commitment across the various data stewards, systems owners and business owners are needed but this effort will result in better capability to integrate and analyze data, and reduced time in developing and maintaining unique security rules.

c. Integration of disparate data marts

Many of the foundational data marts were built independently of each other to meet reporting requirements within the functional areas. Some data marts were purchased and customized in house, such as Oracle EPM warehouse for Financials and Blackboard Analytics for Student data. The other data marts such as for Faculty and Staff (FASIS) and Alumni and Development (CATracks) are built in-house. Although data elements needed to meet the integrated reporting requirements may exist in these data marts, data integration across disparately governed and managed data marts will require effort.

Implications

- This effort should be given a high priority so that architectural best practices and coordination with data governance group noted above can enable sound design and sustainable infrastructure for integrated reporting and analytics.
- The scope of the work may be significant, and knowledge and time of staff from different system groups are needed.

d. Data as a Service

As the data warehouse becomes the central repository of integrated administrative information with common definitions and institutional data access policies, it can be viewed as the central repository of validated and reliable source of institutional data that other applications may want to consume. Rather than developing one-off data feeds to disparate areas within campus or administering a multitude of data base accounts, we envision a data as a service model, where another web portal or application could consume the information real time or nightly from the data warehouse. Because the IBM architecture for Cognos BI is based on Service Oriented Architecture (SOA), we anticipate applications to connect with the data warehouse via web service protocols.

Implications

- An enterprise level SOA system with common registry is beneficial in management of the service layer.
- Capabilities and skills may need to be developed within distributed application development teams to consume information via web service calls.
- Data access and data sharing policy should ensure proper usage of data so that it does not encourage proliferation of 'multiple versions of truth' and duplication of effort on campus.

ii. **Human Resources and Prioritization:**

As we continue to add more applications (data marts, reporting packages, ad hoc tools) to the production system, the demand for enhancements, support and maintenance continues to increase. At the same time, we are taking on more new projects annually. The demand for more analytics/reporting is certainly more than our capacity to deliver as is evident by the list of outstanding requests. Balancing these two demands within the constraint of fixed number of BI resources is a challenge that will become more evident in the coming fiscal years as we add new solutions to the production environment.

A sustainable staffing strategy that takes advantage of both central expertise and distributed staffing to achieve efficiency and agility is recommended. In addition the following challenges and associated implications also need to be taken into account:

a. **Pace of Development and Staffing of BI Projects**

Although the current BI prioritization process selects projects with highest business impact and breadth of usage, there are many requests that simply do not make it to the short list for implementation. Larger projects selected for implementation have to be broken down into multiple deliverables over multiple fiscal years because at the current staffing level, it is impossible to allocate more than three FTEs in the BI team annually to work on these major enterprise system projects (examples are CATracks and FASIS). If the consensus is that the current staffing level is not keeping up with the demand, we should consider scaling up the effort to sustain the demand, at least in the next few fiscal years until the subject areas that facilitate at least 80% of reporting requirements on campus are met.

Because of staffing limitations and the resulting prioritization process, projects that benefit just a school or a unit are rarely selected for implementation. An example is Kellogg's Career Management BI solution proposal. Since the career management system is unique to Kellogg, this type of request is never likely to bubble up high enough on BI priority list to be implemented. Implementing each data repository takes time and specialized skills; therefore, providing tool access and basic training to Kellogg IT staff is not adequate to implement this type of solution. A key question is how we can better leverage a combination of central and school resources to accelerate development to meet outstanding business needs that are high priority to an individual school or department.

We need to creatively explore resourcing and funding models to support these types of initiatives. Otherwise we run the risk of units pursuing one off solutions inconsistent with the University's BI strategy.

Implications

- A University-wide prioritization may be needed when considering whether to warehouse the core components of our major enterprise systems (to support 80% of anticipated analyses) or build more targeted, specialized BI applications in the near term. A hybrid of both options is also a potential option. In all cases, the pace at which we proceed towards our desired outcomes and the associated staffing needed to keep up with the pace are factors to consider. Creative funding and staffing models (such as use of consultants) may allow for unique requests or a school/unit based request to be added to the project portfolio. Regardless of projects selected, long term support and staffing implications should be taken into account during prioritization.
- Although it makes sense for new projects that require developing data marts and integration layers to be centrally led, subject matter experts or other flexible resources from enterprise teams or broader University business units should be available as dedicated project resources to ensure most effective and efficient delivery of proposed solutions.

b. Report Authorship

Although ad hoc, self-service capabilities continue to expand, the need for standardized, authored reports continues to exist. We have learned that the work involved to create data repositories to support data access via the Cognos BI tool is substantial. Because the central BI resources are expected to provide report development and training, we are already, in effect, constrained in our capacity to tackle additional projects without additional resources in some of these highly specialized areas. Also, there remain limited resources within the FASIS, SES, and CATracks teams trained in and devoted to using Cognos Report Studio, raising questions about who will construct, implement, and maintain these reporting capabilities during the next 1-2 years and beyond. Opportunity may exist to broaden the population of staff trained in skills such as report writing to alleviate, at least in part, this resource constraint. For example, report development may be a task that could be undertaken by many of the more data-savvy business analysts on the various enterprise system teams.

Implications

- Not all functional areas have developed report authors with Cognos Report Studio skills. In aggregate, we need more of the skilled resources than we currently have, across all our functions. Via training, conversion of position vacancies as they occur, or other means, staff will need to be trained in the Cognos suite, particularly Report Studio and proper usage of the underlying data to produce standardized reports.
- Where these report authors reside may warrant a standardized approach. Currently there is no consistency, and this function may reside in either an enterprise system team or in a business office.
- Developing report authors from existing staff positions will bring with it an opportunity cost; those resources that will now be available to author BI reports may no longer be available to perform their prior job duties. In some cases, this

may be satisfactory given a phasing-out of the legacy tools. Such resources might traditionally be viewed as "technical", and while the skills that they will acquire are indeed specialized, some of our "functional" resources may be strong candidates for these roles, as these same resources, often, are already well-versed in other report writing tools and SQL query.

- A more distributed staffing structure for operational reporting will increase report development activities and will certainly put pressures on central resources supporting and coordinating the distributed efforts in order to ensure common standards and best practices are adopted. Capacity for administration, meta-data model (business views of data) or data mart changes (ETL development), training, support, mentoring and coordination of distributed efforts need to be planned appropriately.
- It needs to be underscored that taking a 5 day "Report Studio" training is not adequate to convert a SQL expert or a crystal report developer to Cognos report developer. Ongoing mentoring and co-development efforts are needed to ensure long term success. New projects where distributed resources are included as project team members may provide such opportunities.
- The advantage however to the proposed approach is that the backlog for new operational report requests can be lessened more quickly and the central team can focus on continuous innovations in the areas such as self-service BI capabilities, advanced analytics (predictive, text, other future trends etc.), dashboards, mobility and development of integrated reporting applications, also integral to our long term BI strategy.
- A reporting governance team or a user group should also be established with participation from all "report authoring" units. Without coordination, established standards and governance we run the risk of inconsistencies in reporting, information silos or uncontrolled growth in reporting inventory, thus diminishing usability for end users.

c. **Training**

Training of End Users

As each solution is deployed, users expect to receive basic training in the various tools they will need for reporting. Furthermore, they expect quick access to online resources and documentation to provide help as necessary. The reporting community also expects the solution to include some level of help desk support. At a minimum, the help desk should assist users in trouble-shooting any problems they encounter. Beyond this, reporting users (especially analysts and report authors) would profit tremendously from guidance provided by a core team of reporting tool specialists and data experts. To meet this overall expectation, this solution must include a dedicated group of resources who will provide the administrative, training and help desk services required to support this reporting solution.

Most BI users benefit most from one-on-one or small group training sessions, especially those catered more specifically to users' individual business needs. Training is time intensive and the central BI team does not have dedicated training staff to support this. Although the team has taken on training responsibilities for many of the project rollouts, it would be ideal to train local trainers who are closer to the business units they serve.

Building power users

The rollout of the Cognos BI suite within the past 2-3 years represents a substantial change that not all users have acclimated to at this point. The various tools (e.g., canned reports, ad-hoc reporting via Query Studio, Analysis Studio, etc.) are very powerful, but, thus also require training and, for many, a learning curve. We must do more to enable users at all levels to maximize the benefits of these tools. This may need to follow a tiered approach, where the focus on the casual user is helping them locate and run commonly used reports. A middle tier of users can likely become quite adept at using some of the ad-hoc query and analysis tools; particularly with assistance such as training on the tools and data structures and a forum for the sharing of best practices amongst users, so as to help users see what is possible and translate it to a means to solve their own data access/analysis needs. Finally, the more that can be done to enable advanced users to write their own reports (including across multiple data sources or data sources that have been made "integrated" via data warehousing techniques), the more powerful our tools can be in supporting data-driven business decision making.

Implications

- Training may need capacity to be dedicated that would otherwise be available for business analysis, etc. The portfolio of projects approved may be thus affected.
- In order to develop local power BI users within schools or units who can facilitate their use of self-service ad hoc analysis and ad hoc reporting tools we need:
  - Management support in units
  - Data driven management and decision making processes
  - Commitment to develop or hire the skillsets
  - Appropriate training (data, business rules and tools)
- The training of end users/power users in the use of canned reports and tools such as Analysis Studio and Query Studio (or their equivalent features in the next version) should be handed off to groups that are most familiar with the data and business processes being analyzed. A "train the trainer" approach is likely the most sustainable format, with a central team owning responsibility for the core curriculum and a distributed group of trainers conveying the knowledge in context to users of that subject area.
- As the number of power users increases, the level of support expected of central teams may fluctuate or increase over time depending on adoption rates of the new tool sets. This can directly impact future staffing needs.

VIII. Planning for the Future

a. Data Warehouse Where Appropriate

We recognize that all reports will <u>not</u> be delivered by the warehouse. Direct query access and the use of delivered reports within our enterprise business systems will continue to serve practical purposes. Analyses that rely on highly complex business rules or joining or aggregations of data across many sources are well suited to being delivered via a BI solution. There is a cost to migrate existing reports to the BI tool. There is also a cost of not migrating that should be considered (e.g.,

maintaining licenses for multiple tools, requiring staff with diverse skills sets, forcing users to learn and use multiple tools, etc.)

A key attribute of some current reporting tools are that they are embedded into our enterprise business systems, enabling users to access reports and data "in line" with their other business activities. This is indeed a powerful capability and we should pursue opportunities to embed Cognos BI reports within our enterprise business systems and related web sites in a similar capacity. In this way, these same reports can be accessed via a variety of points – via the BI tool, within the application, or via an informational web site or form.

As described in the "data as a service model" in this document, we recognize that there are business needs to access the data within the warehouse in formats other than on-screen, print-ready, or flat-file reports. To the extent that system-to-system access is merited, this should be done via a service call rather than a read-only database account. Coupled with sound data management practices, this approach will help to promote reusability in our solutions and to safeguard against misinterpretation or misuse of data in the data warehouse.

b. Expect Ongoing Change

Regardless of the approach taken in building a warehouse (customizing an off the shelf solution, upfront development of an enterprise warehouse or iterative approach based on business needs), once built, it is inevitable that the data warehouse will continue to evolve with time. Changes in the University's business operations/business processes, adoption of new services or systems, new regulatory/compliance requirements, and other new University priorities or strategies will continue to impact how we analyze our data, measure performance and predict future outcomes.

There are also many emerging desires in this area, including the development of management dashboards, creating advanced financial reporting formats (e.g., school/department income statements), and implementing planning and forecasting capabilities and the ability to integrate planning and forecasting data with actual activity incurred in a time period. Access to data as a feed or service is also sought after; if implemented, an enterprise Service Oriented Architecture could enable data as a service to other applications on campus. Our data access strategy should provide a roadmap for these future capabilities.

Our users will continue to expect easy access to reliable data that are tailored to meet the specific business needs of our users. Many trends in higher education and in the analytics space will continue to drive change. A robust data warehouse strategy that is based on sound architecture, robust infrastructure, skilled staff and data governance will provide long term value.

Appendix 1:				
Data Warehouse Gap Analysis		✓	**Available in Data Warehouse**	
		●	**In Progress**	
		✗	**Not Available**	
Area	**Type of Information/Subject Areas**	**Available**	**Source System**	**Sample Metrics or Attributes***
Student Lifecycle				
	Student	✓	SES	ethnicity, IPEDS ethnicity, citizenship, career, age bands, admit type, cohort type, visa type
	Student Plans	✓	SES	plan count (program, plan type)
	Registrations	✓	SES	majors, minors, units taken, units earned, terms, enrollment count, drop count, subject, career, advisor
	Degrees Awarded	✓	SES	degree, degree honors, years to graduate, first term at institution, prior major/minor/program, credits attempted, credits earned, credits audited
	Performance	✓	SES	GPAs, grade distribution (faculty or class)
	Course Offerings/Teaching Assignments	✓	SES	courses by term, section count, section capacity, primary, secondary instructors and TA s.
	Student Term	✓	SES	eligible to enroll, academic standing, enrollment, retention
	Course Schedules/Class Utilization	✓	SES	class meeting pattern, section capacity, utilization%
	Student Financials	✓	SES	item type, fund, activity, project, account, program, fiscal period, term, amount, student
	Alumni	✓	CATracks	contact information, giving to NU (see alumni/development)
	Admissions - Undergraduate(Ev) and Kellogg	●	SES	test scores, high school performance, academic level, yield, demographics
	Admissions - TGS, Law, FSM, SCS, NUQ	✗	Multiple Admissions Systems	test scores, high school performance, academic level, yield, demographics
	Student Activities	✗	Multiple Systems	club memberships, fraternity, sorority, athletic participation
	Career Management	✗	School Systems	student interactions, job placements, salary post graduation, employer
	Surveys	✗	External	TBD
	Financial Aid	✗	SES	TBD (no outstanding request)
Faculty Activities				
	Teaching/Instruction(Actual)	✓	SES	class size, grade distribution, secondary instructors and TAs, number of sections/courses taught
	Course Loads (Expected)	✗	FASIS	expected course loads
	Research (Grants and Awards)	✓	InfoEd/ NUFinancials	(see Proposals and Awards)
	NUScholars	●	NUScholars	publications, collaborations, affiliations,hindex
	CTEC (Course/Teacher Evaluation)	✓	SES	response %, ratings, questions, responses
	TCEs (Kellogg)	✗	SES	response %, ratings, questions, responses
Human Resources: Faculty and Staff				
	Employee Profile	✓	FASIS	exempt or non, full time/part-time, demographics, benefit eligibility, leaves of absence
	Faculty Profile	✓	FASIS	appointments, tenure status, rank, title, demographics, career path, position category
	Postdocs	✓	FASIS	see employees
	Compensation	✓	FASIS	salary, salary grades, funding type, merit increases, merit %
	Turnover	✓	FASIS	terminations, transfers, promotions, demotions, other job transactions
	Payroll Expense Distribution	●	FASIS	funding source, payroll details, benefit amounts, encumbrances, stipends
	Census	●		see Faculty Profile
	Turnover Metrics	●	FASIS	
	Faculty Course Loads	●	FASIS/SES	see faculty activities
	GL data analysis	●	FASIS	base pay, additional pay, funding type
	Benefits analysis	✗	FASIS	
	Recruiting	✗	e-recruit	

Area	Type of Information/Subject Areas	Available	Source System	Sample Metrics or Attributes*
Space				
	Assignment	✓	SIMS	assigned department, productivity department, usage function, square footage, usage %, assigned %,
	Facilities Management	✓	FAMIS	includes capital projects, job cost, service requests, parts and equipment inventory
Alumni and Development				
	Donors (and Alumni)	✓	CATracks	biographic information, addresses, spouse info, overall giving summaries, degrees, school specific giving
	Tracking Gifts and Pledges	✓	CATracks	cash, new gifts and commitments, outright gift, pledge payment, matching gift, soft credit, appeals, allocation
	Prospects	✓	CATracks	prospect manager assignment, number of contacts, number of visits, giving capacity
	Proposal Management	✓	CATracks	ask amount, anticipated FY cash, anticipated commitment, proposal amount bands, proposal assignments
	Gift Clubs	✓	CATracks	memberships - leadership circles, planned giving societies
	Special Handling	✓	CATracks	donor contact preferences, restrictions
	Appeals Tracking	✓	CATracks	appeals response rates, target audience
	Events	✗	CATracks	event attendance, registrations
	Alumni Engagement	✗	EngageNU	online behaviors, social network relationships, social activities,email marketing campaign successes or bouncebacks
Alumni and Development				
	Mailing Information	✗	CATracks	privacy preferences, opt-ins, opt-outs,
	Student Activities	✗	TBD	activities (committees, athletics), see also: student lifecycle
Research Administration				
	Animal Care (Facilities and Census)	✓	Granite	census counts, locations, species, housing types, number of days, protocol, cost centers
	NU Core Facilities	✓	NUCore	billing, payments, facility name, services, instruments, account
	Research Safety and Compliance	✓	ISIS	safety training, certifications, hazardous materials, laser safety, lab annual review
	Proposal and Awards	✓	InfoEd/ NUFinancials	Pre-award: Proposal activity and trends, award activity and trends, activity by unit/sponsor, negotiation status, research safety Post-award: Expenditure detail and summary, project attributes, milestones/deliverables, revenue, billing/AR, F&A
	Human Subject Protocols	✗	eIRB	protocol details
	Animal Care Protocols	✗	ACUC	protocol details
Financials				
	Chart of accounts	✓	NUFinancials	expenditure, revenue, budget, purchasing/AP
	User security/workflow	✓	NUFinancials	
	Budgets - original, current, changes	✓	NUFinancials	
	Financial - summary and detail	✓	NUFinancials	includes fiscal year end reporting; also specialty reporting by fund type (gift, endowment, agency, plant, etc.)
	Purchasing	✓	NUFinancials	includes purchasing of capital equipment, open encumbrances, trends; includes reconciliation of Voyager purchases
	Expense reports	✓	NUFinancials	
	Vouchers	✓	NUFinancials	includes online voucher; includes match exceptions
	Vendors	✓	NUFinancials	
	Assets	✓	NUFinancials	
	Financial reconciliation	✓	NUFinancials	
Endowment Stewardship				
	Endowment Information	●	Fundriver	endowed fund performance

Area	Type of Information/Subject Areas	Available	Source System	Sample Metrics or Attributes*
	Endowment Stewardship	●	Fundriver, CATracks, NUFinancials, NUScholars, SES, FASIS	contributed, amount spent, usage, account manager,scholarship receipient (see student), endowed faculty (see faculty activities), donor details (see alumni and development)

Area	Type of Information/Subject Areas	Available	Source System	Sample Metrics or Attributes*
			Fundriver, CATracks, NUFinancials, NUScholars, SES, FASIS	