City of Philadelphia
Department of Behavioral Health and Intellectual disAbility Services

Data Governance Framework Implementation Plan

Version 1.00
29 February 2016

Actionable Steps to Align Information Management Concepts, Practice and Context

DBHIDS
DEPARTMENT of BEHAVIORAL HEALTH and INTELLECTUAL disABILITY SERVICES

Arthur C. Evans Jr., Ph.D., Commissioner

CREDITS

Author: **Daniel J. Paolini, DBHIDS Chief Information Officer**

With input and guidance from the

DBHIDS Enterprise Data Management Team

Barbara Bunkle

Wulan Guo

Minky Kernacs

Jennifer Li

Members of the DBHIDS Data Governance Executive Board

Members of the DBHIDS Data Stewardship Council

and Jessica Costanzo, Consulting Data Architect

DBHIDS VISION

We envision a Philadelphia where all people can achieve health, wellness and self-determination through a comprehensive, holistic, community-based service delivery system.

ENTERPRISE DATA MANAGEMENT MISSION STATEMENT

In support of the DBHIDS population health mission we will define, cleanse and integrate the data needed to provide better clinical services, measure outcomes and make more informed policy decisions while reducing the cost of those activities. We will identify the core data in use across the department and establish a collaborative approach to managing it so as to provide information to those that need it while meeting the quality, format and content expectations of the organization.

"The society which scorns excellence in plumbing as a humble activity and tolerates shoddiness in philosophy because it is an exalted activity will have neither good plumbing nor good philosophy; neither its pipes nor its theories will hold water."

— John W. Gardner, *Excellence: can we be equal and excellent, too?* (1961)

TABLE OF CONTENTS

Image Credits

WWW.INFORMATION-MANAGEMENT.COM

WWW.MCKESSON.COM

WWW.THEDATACAVE.COM

WWW.SMARTDATACOLLECTIVE.COM

WWW.NORTHUMBERLAND.GOV.UK

WWW.APPLUSCONSULT.COM

WWW.MARSMETADATA.COM

WWW.COMPUTING.CO.UK

WWW.SANDYX.COM

AHABIBLEMOMENTS.WORDPRESS.COM

WWW.STUDY.COM

WWW.INFOSIGHTINC.COM

"You cannot improve what you cannot manage; you cannot manage what you cannot measure; you cannot measure what you cannot define; you cannot define what you do not govern."

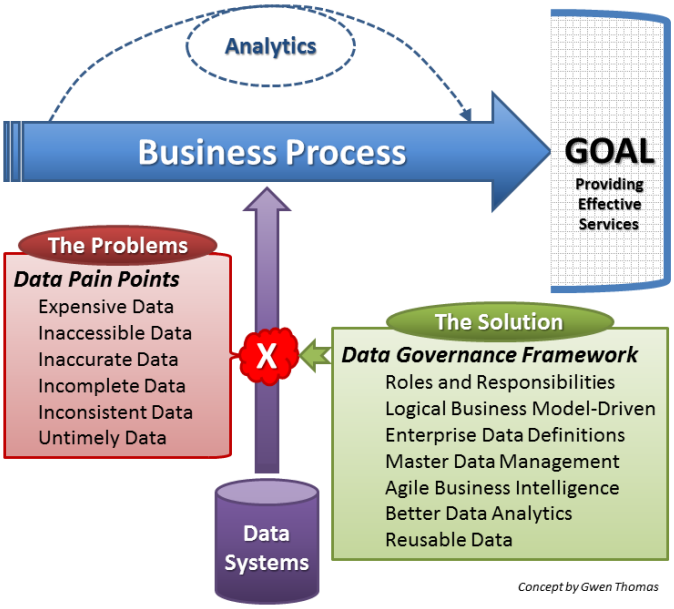
— **Daniel J. Paolini, *Introduction to Data Governance*** (2015)

DATA GOVERNANCE FRAMEWORK STRATEGIC PLAN EXECUTIVE SUMMARY

The Data Governance Framework Strategic Plan Executive Summary is provided here for context. This Implementation Plan is consistent with the Data Governance Framework Strategic Plan.

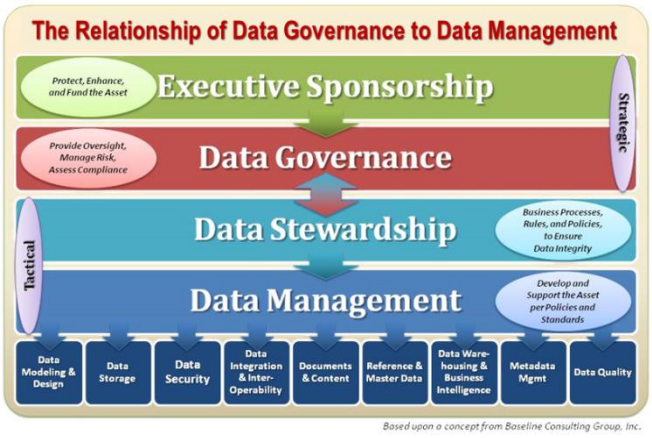
Information Technology is the Information, Applications and Infrastructure necessary to meet the needs of the Business. Of these, the most important resource is the information asset. It is this information that represents a model of the functioning business over time. Infrastructure and applications change to reflect changes in technology or process, but the data that records the transactions of key entities with the business must be logically defined and maintained so that it is consistent, persistent and useful.

Over time, the ability for data systems to support business processes in the pursuit of business goals degrades because data quality degrades. This results in more cost but less benefit. Attempting to fix these problems at the operational level will not work. Adding more data and data systems in an effort to "fix" these problems only makes them worse.



What will fix these problems is not more technology, more systems or more data, but Data Governance. Data Governance advances the goal of "Reusable Data"; data that is timelier, more accurate, more complete, more accessible, more useful and less costly.

Data Governance is not a technology function. It is driven by the business and forms a bridge between business management and technology providers. ***Executive Sponsorship*** comes from the business. ***Data Governance*** and ***Data Stewardship*** represent collaborations of business subject matter experts and information architecture staff. ***Data Management*** is provided by various technologists overseen by an enterprise information management unit.



The Data Governance Framework represents the desired future-state needed to address the problems that compromise our information asset today. It documents the Department's Information Architecture. Taking our cue from the Practice Guidelines, we will align information management concepts, practices and context.

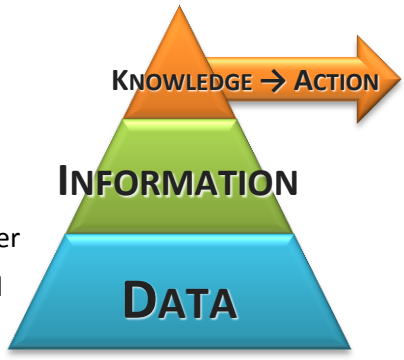
IMPLEMENTATION RECOMMENDATIONS SUMMARY

The implementation strategies have been grouped into three broad categories of recommendations.

1. Implement the Data Governance Framework
    a. Develop data management policies and standards for consideration by management
    b. Achieve compliance with those information architecture policies and standards
    c. Evangelize the Data Governance Framework principles to the department

2. Create the Office of Enterprise Data Management (OEDM) within the Office of the CIO to provide department-level focus and coordination for data management, architecture, governance, integration, collaboration and business intelligence.
    a. Establish a Data Governance Office within OEDM that will continue to evolve the Data Governance Framework through engagement with the Data Governance Executive Board and the Data Stewardship Council.
    b. Establish a Data Architecture Office within OEDM and staff it with an enterprise data architect. Acquire business data modeling software and deploying it to model and document enterprise data and structures.
    c. Establish a Data Integration Office within OEDM that will use a business model-driven approach to data integration to establish a DBHIDS Enterprise Data Warehouse. Acquire data warehouse automation tools as the foundation for data integration and master data management to provide agility, scalability and maintainability.
    d. Establish a Business Intelligence Office within OEDM that will deliver consumable data of known quality and provide self-service business intelligence reporting and visualization capability to all information consumers based upon their needs and abilities.
    e. Establish the role of Data Security Coordinator to work with OEDM and network security officials to protect the confidentiality and integrity of data while making it available to those that require it and are entitled to it.

3. Manage data as an organizational resource.
    a. Manage the enterprise business model and glossary
    b. Provide enterprise metadata management for users
    c. Manage master data management domains
    d. Manage the enterprise data integration and persistence layer
    e. Provision data for those that need it in the manner required



DBHIDS DATA GOVERNANCE FRAMEWORK IMPLEMENTATION PLAN

Actionable Steps to Align Information Management Concepts, Practice and Context

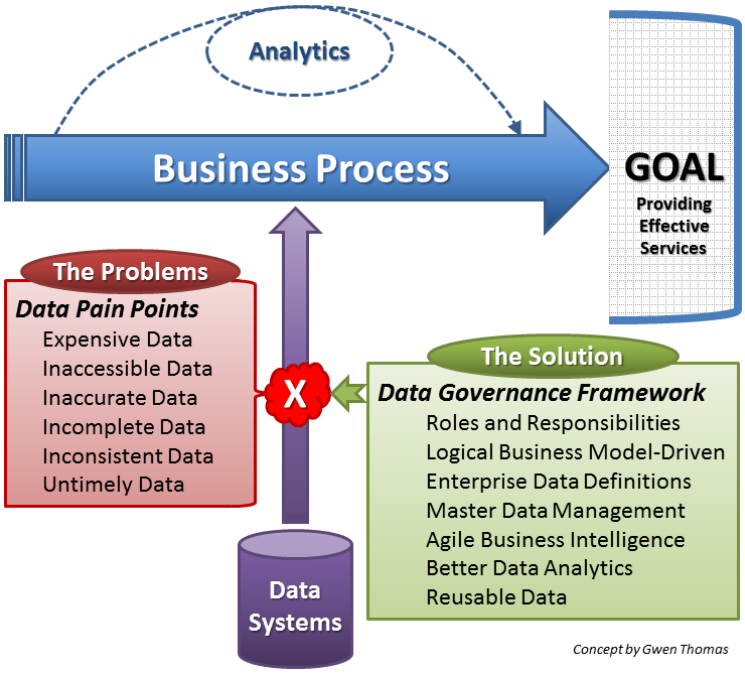
DATA MANAGEMENT STRENGTHS AND CHALLENGES

DBHIDS is a diverse organization with a decentralized approach to data management, even for data that is of use across organizational boundaries. This situation has been caused by a number of factors. Most data systems were developed "bottom-up" to meet business unit needs. Information management resources were deployed to meet the needs of those business units, and were funded by those units. Funding was often determined by various federal or state programs which placed either real or perceived restrictions on data management and use. Even when data was common between business units, it was addressed by units exchanging data from time to time as needed without attempting to agree on definition, structure or common management. These factors encouraged the development of data siloes and discouraged cooperation between units.

Developers and analysts are hard-working, earnest and intelligent but lack data integration tools, methodologies and training. Even without these, they are able to deliver meaningful value but are unable to keep up with an ever-increasing workload. In addition, as most analytical efforts are performed one-off without leveraging previous efforts, data quality problems are introduced into reports due to lack of consistent business rules and definitions. Of most concern, the data as a result of many of these efforts is maintained in collections of spreadsheets, often without documentation and unavailable for reuse.

While it is important to discuss the benefits to the organization that will be provided by an enterprise approach, it can be difficult to make these over-arching issues resonate with business management. While keeping the big picture of our enterprise information management and data governance landscape in perspective, we must demonstrate how the implementation plan activities address the pain points of the individual business units.

For each project, we must connect the enterprise activities required by this implementation plan to either the solution sought by the business or the avoidance of both seen and unforeseen problems; ideally, we will address both.



ORGANIZATION EXPECTATIONS FOR THE USE OF DATA WITHIN DBHIDS

The DBHIDS Data Governance Framework Strategic Plan reflects the expectations that stakeholders have for the use of data within the department. These expectations are framed in business terms and are informed by business needs. They are independent of any technology implementation but inform every technology solution or activity.

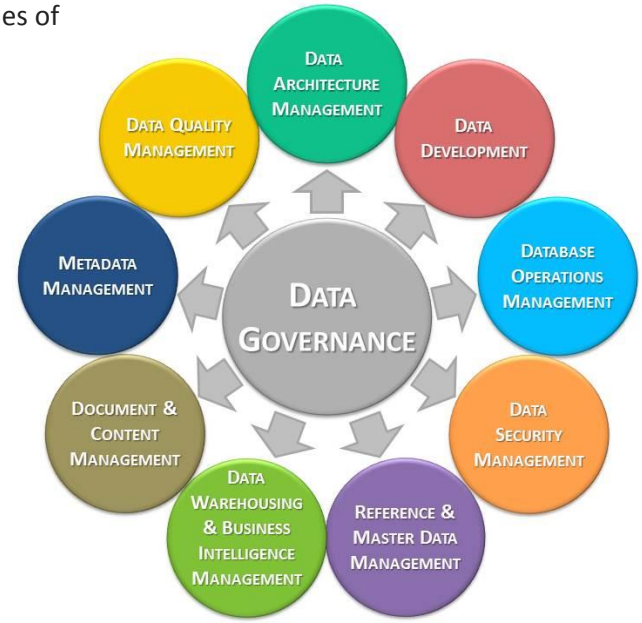
- Our systems are people-centric and provider-centric, not business unit-centric.
- Data that is important to our business is managed formally.
- We have a 360 degree view of our service recipients and our providers.
- We have access to data of known quality for policy evaluation and decision-making.
- We are agile when there is a need to analyze a new problem.
- We invest in ourselves wisely.

FROM THE DATA GOVERNANCE FRAMEWORK STRATEGIC PLAN

Strategic Objective: ***Establish an Enterprise Information Management focus under the CIO***

To accomplish this strategic objective, DBHIDS needs to create an Enterprise Data Management capability where one does not exist today for the express purpose of providing a defined, reusable data resource of known quality to those that need it in the format needed and to do in a responsive, agile manner. This will include the sub-disciplines of

- Data Governance
- Data Architecture
- Data Integration
- Master Data Management
- Metadata Management
- Data Quality Management
- Data Warehousing
- Data Development
- Database Operations
- Data Security



DBHIDS requires a fully-developed data governance approach to mandate enterprise data management. It requires an enterprise data architecture unit to design and document data structures, objects and their definitions. It requires an enterprise data integration unit to manage reusable data resources. It requires a business intelligence unit to make data available to knowledge workers. This will place Enterprise Data Management responsibilities into the hands of data management professionals rather than in the hands of technologists or analysts currently responsible, enabling those technologists and analysts to spend more time in their respective and equally important worlds.

This effort will be guided by the **DBHIDS Data Governance Framework Strategic Plan,** which includes

Ten Information Architecture Principles

1. Information architecture is the reflection of the business; it is not a just technology domain.
2. The identification and definition of data attributes must involve the business.
3. Data is an organizational asset and must be managed with an enterprise perspective.
4. Data that is common to more than one business unit must be defined through the consensus of representatives of those business units.
5. The value of data to the enterprise is in its fitness for reusability, not its exclusivity.
6. The value of data management staff is in its ability to build high quality, reusable data assets.
7. Different information use cases require different data management solutions.
8. In order to be sustainable, physical data stores must be governed by a logical understanding of the enterprise, captured in a logical business model.
9. The purpose of a data management organization is to produce a data product that meets the information needs of the business commensurate with the investment made by the business.
10. If it isn't documented it doesn't exist; if it can't be measured it has no value.

Nine Data Governance Strategic Objectives

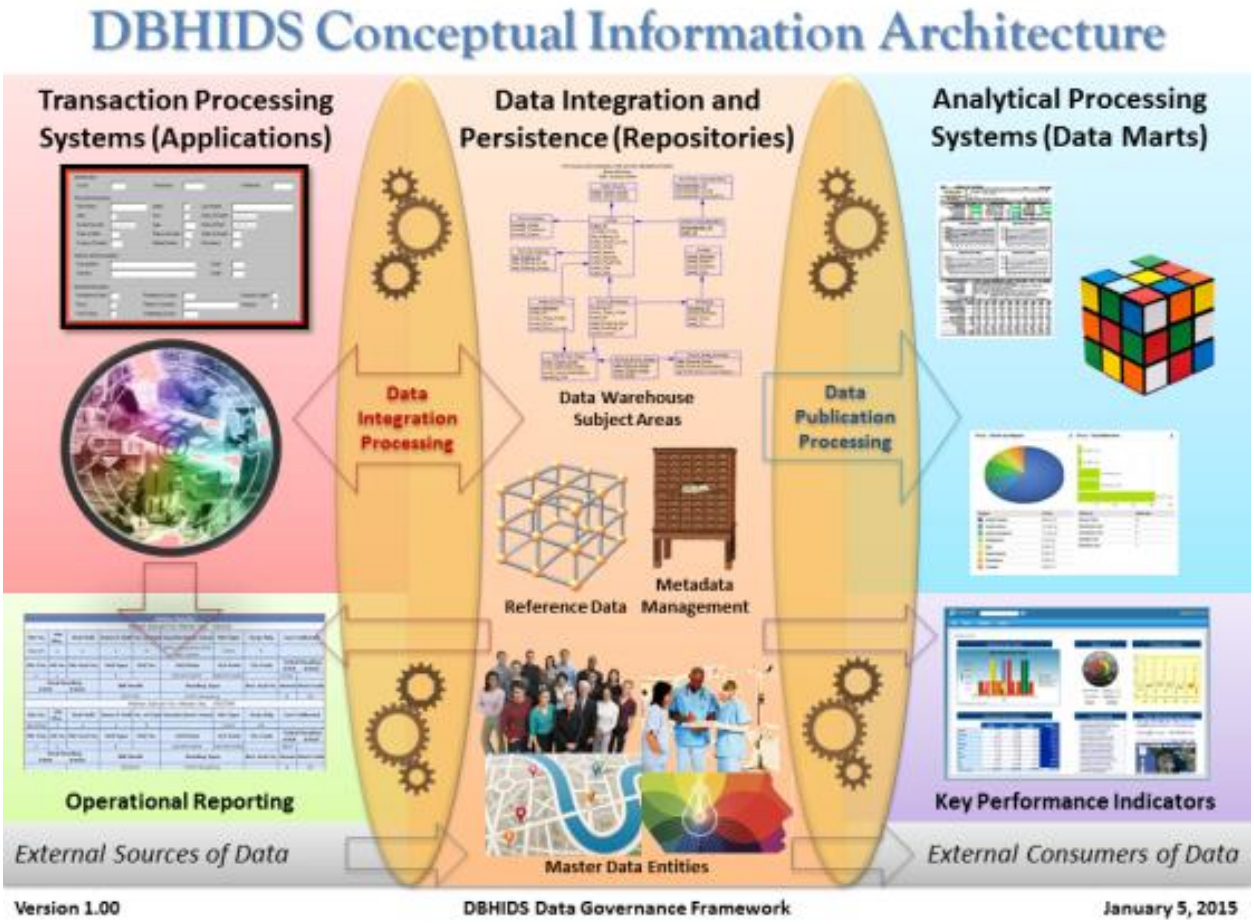
1. Create an information-centric and informed organizational culture.
2. Establish a data governance program to provide accountability for information assets.
3. Provide for effective and appropriate information security.
4. Improve the quality and usefulness of information by making it timelier, more accurate, more complete and more accessible.
5. Reduce the costs of managing information.
6. Share data through reusable processes; reuse data through shared processes.
7. Provide self-service business intelligence capabilities.
8. Develop enterprise-class data management staff.
9. Adopt enterprise-class data management tools.

Five Information Service Delivery Use Cases

1. Transactional Processing: "**To Do**"
2. Operational Reporting: "**To Know**"
3. Key Performance Indicators: "**To Measure**"
4. Analytical Reporting: "**To Learn**"
5. Data Integration & Persistence: "**To Remember**"

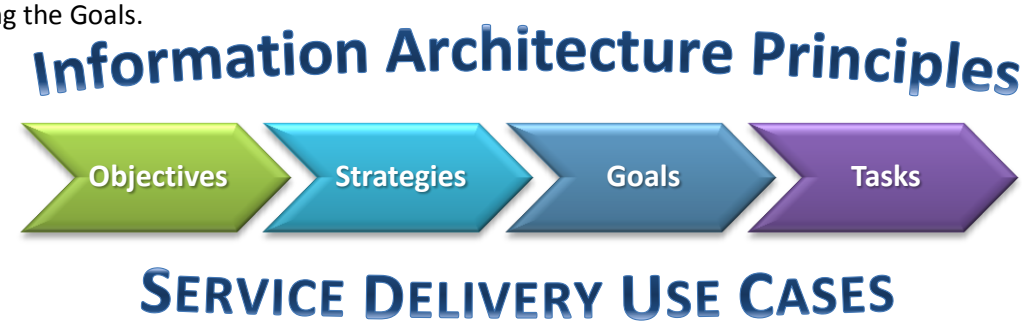
CONCEPTUAL INFORMATION ARCHITECTURE SCHEMA

This represents the manner in which the five different Information Service Delivery Use Cases are related to and interact with each other.



The DGF Implementation Plan addresses the DBHIDS Information Service Delivery Use Cases within the DBHIDS Conceptual Information Architecture Schema. The plan establishes Data Governance Strategies and Goals to fulfill the DBHIDS Data Governance Objectives while adhering to DBHIDS Information Architecture Principles.

The Strategies are described in the next section. Goals are described in Appendix C. The Tasks necessary to implement the Strategies and achieve the Goals will be developed by the units responsible for achieving the Goals.



DBHIDS DATA GOVERNANCE STRATEGIES

The following list of strategies will actualize the Data Governance Framework. The objectives that each supports are indicated in the table that follows.

1. Establish within the Office of the CIO a DBHIDS-wide data management unit consisting of
    a) DBHIDS Data Governance Office
    b) DBHIDS Data Architecture Office
    c) DBHIDS Data Integration Office
    d) DBHIDS Business Intelligence Office
2. Implement the Data Governance Framework
3. Achieve compliance with information architecture policies and standards
4. Develop new draft policies and standards for consideration by the DSC
5. Manage the enterprise logical data model and business glossary
6. Provide enterprise-level metadata management for business and technical users
7. Manage master data management domains
8. Manage the data integration and persistence layer for the department
9. Provision data for those that need it in the manner required
10. Evangelize the Data Governance Framework principles to the department

Relationship of Data Governance Strategies to Objectives

Data Governance Strategic Objectives / Data Governance Strategies		1 Information-centric culture	2 Data Governance Program	3 Information Security	4 Usefulness of Information	5 Reduce Information Costs	6 Share Data and Processes	7 Self-Service BI Reporting	8 Data Management Staff	9 Data Management Tools
1.	*Establish the Office Enterprise of Data Mgmt.*	√	√	√	√	√	√	√	√	√
2.	*Implement the Data Governance Framework*	√	√		√	√				
3.	*Achieve compliance with policies and standards*	√	√	√	√	√				
4.	*Develop new draft policies and standards*	√	√	√	√	√				
5.	*Manage enterprise LDM and business glossary*	√	√		√	√	√	√		
6.	*Provide enterprise metadata management*	√	√		√	√		√		
7.	*Manage master data management domains*	√	√		√	√	√			
8.	*Manage data integration & persistence layer*		√	√	√	√	√		√	√
9.	*Provision data for those that need as required*	√		√	√		√	√	√	
10.	*Evangelize the Data Governance Framework*	√	√						√	

The description of each strategy and its relevant objectives can be found in Appendix C.

DATA GOVERNANCE ROLES AND RESPONSIBILITIES

This section defines the roles and responsibilities of the Data Governance Framework and places them in one of three categories: Data Governance Bodies, Enterprise Data Management and Other Data-related.

DATA GOVERNANCE BODIES

Ultimately, data governance is a business responsibility. It must be endorsed by senior executive management. It must be led by executive management. It must require involvement by business unit representatives to make sure that data is properly identified and defined, made fit for purpose and used appropriately. While the Data Governance Office as the facilitator, enabler and implementer of data governance sits within the Office of the CIO, the responsibility for enterprise data governance resides with executive management. Enterprise Data Management follows from Enterprise Data Governance.

Executive Management Core

The Executive Management Core (EMC) team establishes the business vision, mission and priorities of the Department. It resolves data governance and data management priorities against other organizational priorities. Led by the Commissioner, it is the ultimate governance body.

Data Governance Executive Board

The Data Governance Executive Board (DGEB) is responsible for approving data governance and data management policies, resolving data governance issues presented by the Data Stewardship Council and prioritizing enterprise data management efforts and initiatives. The DGEB is led by the Commissioner and consists of members of the extended Executive Management team. The DGEB serves as the de facto executive sponsor of the enterprise data governance and management program.

Data Stewardship Council

The Data Stewardship Council (DSC) is responsible for defining enterprise data, developing data management policies, and guiding enterprise information architecture and data management efforts. Ultimately, its responsibility is to help the enterprise improve the competing dimensions of data quality, accessibility, security, and value. The DSC is led by the CIO/EDGO and consists of the Enterprise Data Architect, the Data Integration Director, the Business Intelligence Director, enterprise Chief Data Stewards, supervisors from the Department's various analytic and IT development units and business unit data stewards.

Enterprise Data Management Offices and Roles

The Office of Enterprise Data Management is the focal point for information architecture and data management for DBHIDS. Reporting to the CIO, it is comprised of offices for the enterprise activities of data management, architecture, governance, integration and business intelligence.

Data Governance Office

The Data Governance Office (DGO) is the unit responsible for executing data governance policies and supporting the Data Stewardship Council. The DGO is coterminous with the Office of the Chief Information Officer (OCIO), which provides staff support. All new data initiatives will be submitted for review by the DGO as well as Information Security and IT Strategic Planning. The review intensity will be based upon the initiative's risk, size/cost, cross-unit impact and potential to advance department objectives. The review will determine compliance with architecture and standards and identify any need for a waiver from them.

Enterprise Data Governance Officer

The DBHIDS Chief Information Officer (CIO) serves as the Enterprise Data Governance Officer (EDGO) for the Department. In this role, the EDGO is the individual responsible for ensuring that enterprise data governance policies and practices are followed. The EDGO is responsible for the proper application of the enterprise information and data architecture principals and the overall quality and usability of enterprise data assets. The EDGO leads the Data Stewardship Council.

Data Architecture Office

The Data Architecture Office (DAO) is the unit responsible for maintaining the Enterprise Logical Reference Model, reviewing application logical solution models for consistency, overseeing the maintenance of the enterprise business glossary and other metadata resources and for the proper application of enterprise data management policies, standards and practices. The DAO is also responsible for the design of data warehouse subject area models and Master Entity and Master Reference Data processes maintained by the Data Integration Office. The DAO is led by the Enterprise Data Architect (EDA).

Enterprise Data Architect

The Enterprise Data Architect (EDA) is the individual responsible for advancing the role of data architecture within the DBHIDS business and IT communities. The EDA shall guide the efforts of data architects and analysts within the Office of Enterprise Data Management and the business units in the development of conceptual and logical business models, and logical solution and physical data models. The EDA shall be responsible for overseeing the modeling of master data and the design of master data management processes that will be maintained by the Data Integration Office. The EDA reports to the CIO and is a member of the Data Stewardship Council.

DATA ARCHITECT

A Data Architect (DA) is an individual assigned to an IT project, business unit or subject area with the responsibility to develop a logical solution model, a business subject area model or a dimensional model consistent with the Enterprise Logical Reference Model. Logical solution models will be used to create physical data models for applications. Business subject area models will guide data integration in the Enterprise Data Warehouse. Dimensional models will define the structures required for reporting and analysis.

DATA INTEGRATION OFFICE

The Data Integration Office (DIO) is the unit responsible for integrating and maintaining DBHIDS and external data sources, supplying data to enterprise data marts, business unit data marts and external partners, and managing enterprise master entity data and master reference data. The DIO implements source system integration models provided by the Data Architecture Office and data reporting and analysis models provided by the Business intelligence Office.

DIRECTOR, DATA INTEGRATION

The Director of Data Integration (DDI) is the individual responsible for advancing a rational data integration approach within the DBHIDS business and IT communities for the purpose of developing reusable data resources. The DDI shall be responsible for the integrity of data integration processes and for the measurement of data quality against the expected baseline.

BUSINESS INTELLIGENCE OFFICE

The Business Intelligence Office (BIO) is the unit responsible for determining the information needs of knowledge workers and other data consumers, developing appropriate data reporting structures to meet those needs and providing them to the Data Integration Office, and implementing self-service reporting, analysis and visualization (business intelligence) tools so they are available throughout the organization.

DIRECTOR, BUSINESS INTELLIGENCE

The Business Intelligence Director (DBI) is the individual responsible for supporting a data-driven organization by identifying appropriate subject areas for reporting and evangelizing and implementing ubiquitous business intelligence capabilities that leverage those subject areas.

In addition to the above offices and directors, enterprise data management will also include:

DATA SECURITY COORDINATOR

The Data Security Coordinator is responsible for identifying data security policies that need to be followed throughout the organization, making implementation recommendations and advising the Data Governance Office. This role will be filled by the DBHIDS Security Policy Coordinator.

OTHER DATA-SPECIFIC ROLES

For additional information, see the **OWNER/STEWARD/CUSTODIAN MODEL** in Appendix B.

DATA OWNER

The Data Owner is the individual or organization that has statutory rights to the data, no matter who collects the data or who manages it. The rights can include copyright and intellectual property rights as well as the rights to exploit and/or destroy the data. The rights of the data owner apply even when the data is collected by a third party and/or combined with data owned by others.

The Data Owner can be an individual, for data such as personally identifiable information, or an organization, for data such as intellectual property owned by a company. For transactional and derived data collected or created for DBHIDS business purposes by DBHIDS units, the Data Owner is DBHIDS – not the individual business unit – unless the Data Owner has been designated as the City of Philadelphia or a state or federal agency.

DATA STEWARDSHIP ORGANIZATION

The Data Stewardship Organization (DSO) is the organization responsible for the definition and use of data collected or maintained within its functional areas. A data stewardship organization is responsible for developing policies specifically related to the use of the data, and for designating data stewards to implement and enforce those policies. The Data Stewardship Organization may also be the Data Owner or the Data Custodian, but not always (or even usually).

CHIEF DATA STEWARD

Each division or significant administrative unit will designate a Chief Data Steward (CDS). The CDS is a core member of the Data Stewardship Council. The CDS represents the data for which his or her unit is the Data Stewardship Organization. In addition, the CDS works with other data stewards to guide master data management efforts around data sets related to their data needs.

DATA STEWARD

The Data Steward is the individual within the Data Stewardship Organization identified as the point of contact for questions about the definition and use of data collected or maintained by the DSO within its functional areas and for the documentation of that data in a metadata registry. On behalf of the data stewardship organization, a data steward follows and/or implements policies, procedures, and guidelines that pertain to the data during the lifecycle of that data entrusted to his or her stewardship, and participates as a member of the Data Stewardship Council for enterprise data governance issues as well as representing his or her Data Stewardship Organization. The same data may have multiple data stewards; e.g. technical, business, collection, publication, etc.

DATA CUSTODIAN

A Data Custodian is the individual or organization granted authority to possess, use, and/or maintain data in accordance with requirements defined by its Data Steward. A data custodian is responsible to protect the rights of the data owner for the access, processing, maintenance, storage, protection,

and/or destruction of data and electronic records. Data custodians are responsible and accountable for the management and care of the data under their control.

DATA STAKEHOLDER

The Data Stakeholder is an individual or organization that is entitled to make use of data collected or maintained within a system or to derive benefits from its collection and/or publication.

DATA CONSUMER

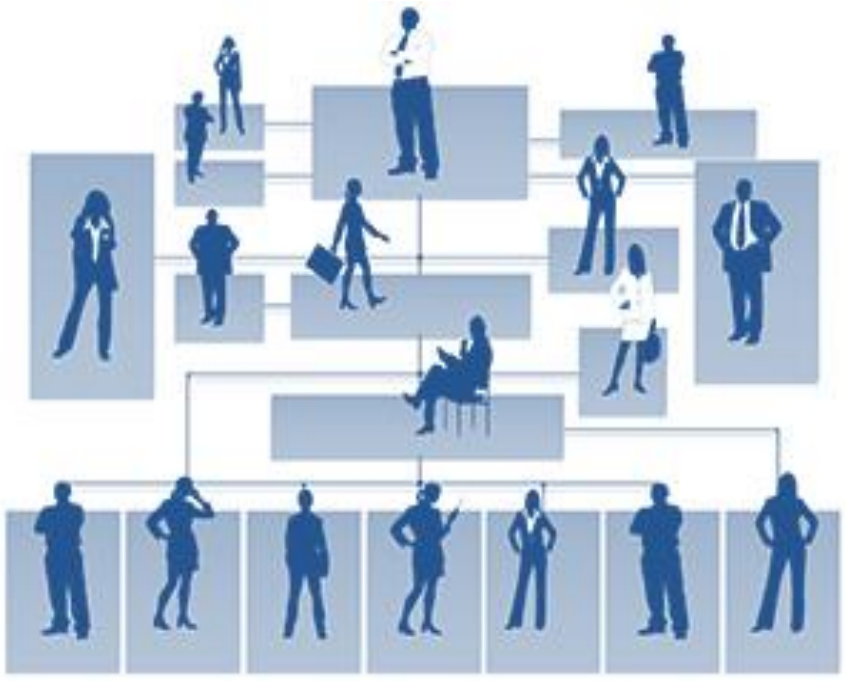
The Data Consumer is an individual or organization that receives data on a regular basis, either in the form of a report or as a data extract file. Typically, this report or data is in support of a predefined, standardized process. A Data Consumer does not typically make ad hoc one-time requests for data. A Data Consumer may be a Knowledge Worker but does not have to be.

KNOWLEDGE WORKER

The Knowledge Worker is an individual that uses enterprise data to make decisions, measure effectiveness or develop policy. A Knowledge Worker will assist in identifying the types of reporting and analytic solutions that are required. A Knowledge Worker may be a Subject Matter Expert but does not have to be.

SUBJECT MATTER EXPERT

The Subject Matter Expert is an individual within a business unit or the analyst assigned to the business unit that has the most complete understanding of a particular subject area or process. The Subject Matter Expert is engaged to make sure that the right data is being used appropriately to meet business needs for both transaction processing and reporting purposes. A Subject Matter Expert may also be a Data Steward and/or a Knowledge Worker, but does not have to be.



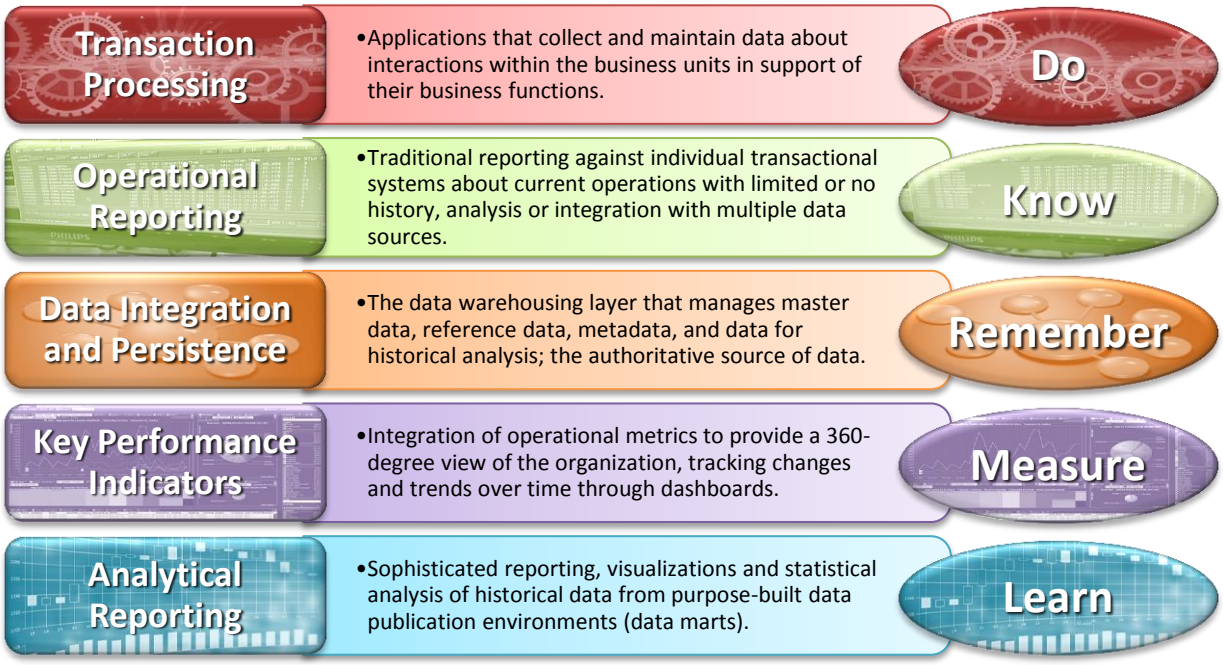
MAPPING OF RESPONSIBILITIES TO ROLES

The following table indicates for each of the six areas whether it is responsible to **GUIDE, LEAD, PERFORM OR SUPPORT** one of ten data management activities. Lead can be inclusive of Perform, and Perform can be inclusive of Support. The six areas include the four OEDM offices as well as a category representing other IT units and a final category representing business unit participation.

Responsible Area / Category	Data Governance	Data Architecture	Data Integration	Business Intelligence	Other IT Units	DBHIDS Business
Requirements Gathering	GUIDE	LEAD	SUPPORT	PERFORM	NOT APPLICABLE	GUIDE
Modeling & Design	GUIDE	LEAD	SUPPORT	SUPPORT	NOT APPLICABLE	GUIDE
Metadata & Glossary Management	LEAD	PERFORM	SUPPORT	SUPPORT	NOT APPLICABLE	PERFORM
Data Warehouse Development	GUIDE	SUPPORT	LEAD	SUPPORT	NOT APPLICABLE	NOT APPLICABLE
Data Mart & Dashboard Development	GUIDE	SUPPORT	SUPPORT	LEAD	NOT APPLICABLE	NOT APPLICABLE
Business Intelligence Data Delivery	GUIDE	GUIDE	SUPPORT	LEAD	NOT APPLICABLE	GUIDE
Data Quality Assurance	SUPPORT	LEAD	SUPPORT	SUPPORT	NOT APPLICABLE	SUPPORT
Platform & Tool Administration	NOT APPLICABLE	NOT APPLICABLE	SUPPORT	SUPPORT	LEAD	NOT APPLICABLE
Business Intelligence Help Desk	NOT APPLICABLE	NOT APPLICABLE	NOT APPLICABLE	LEAD	PERFORM	NOT APPLICABLE
Master Data Management & Administration	GUIDE	LEAD	PERFORM	NOT APPLICABLE	NOT APPLICABLE	SUPPORT

Mapping Information Management Use Cases

Information Service Delivery Use Cases define the design patterns to meet the competing information needs of the organization.

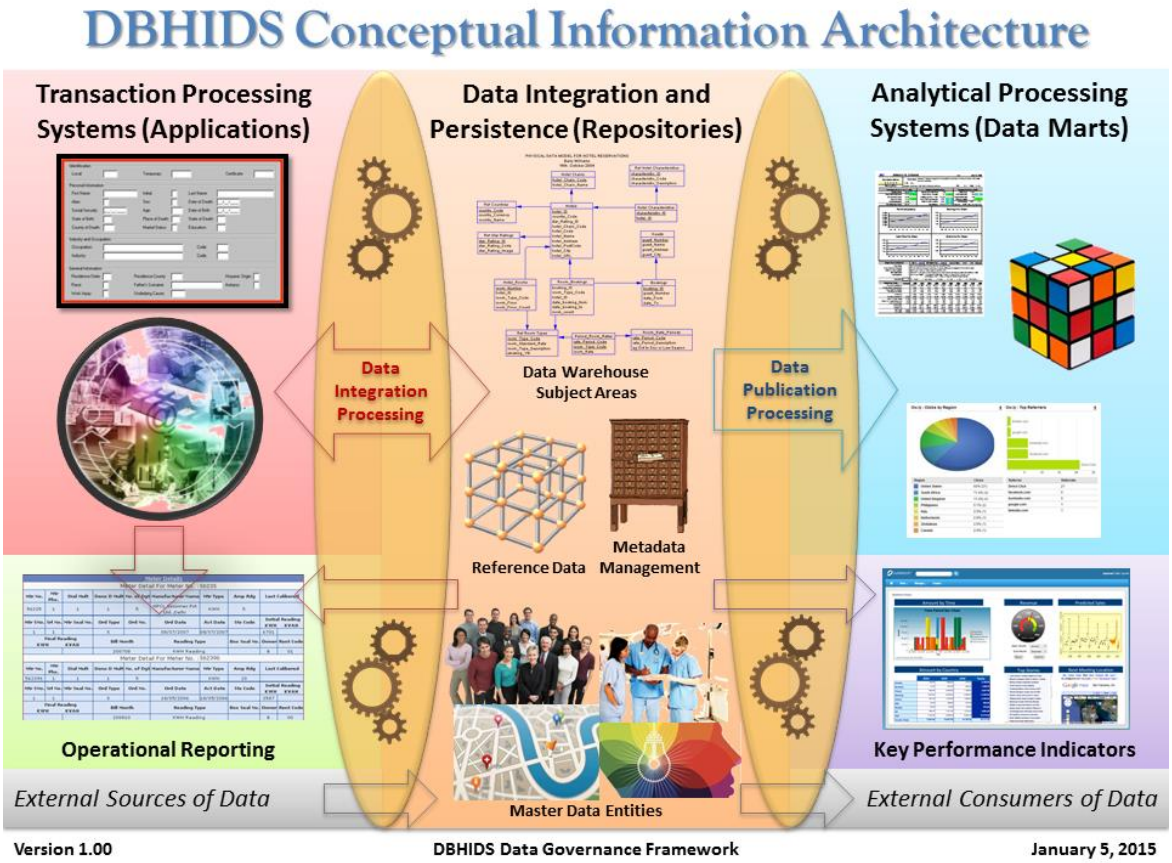


The following table maps the various roles in the data management landscape to the five use cases and categorizes each role as one of: **Guide, Lead, Perform, Support or Not Applicable.**

Role	Do	Know	Measure	Learn	Remember
Application Development	Perform	Support	Support	Perform	Support
Business Intelligence	Not Applicable	Perform	Perform	Lead	Support
Business Management	Lead	Lead	Lead	Guide	Not Applicable
Data Architecture	Support	Support	Support	Support	Lead
Data Governance	Guide	Guide	Guide	Guide	Guide
Data Integration	Not Applicable	Not Applicable	Support	Support	Perform
Data Stewards	Guide	Guide	Guide	Guide	Guide
Executive Management	Guide	Guide	Guide	Guide	Guide
Other Technologists	Support	Support	Support	Support	Support
Subject Matter Experts	Guide	Guide	Guide	Guide	Guide

The following table maps the various components of the conceptual information architecture to the five use cases and categorizes each component as one of: **Essential, Preferred, Optional, Sub-Optimal, Anti-Pattern or Not Applicable**. An anti-pattern is a component that is at cross-purposes to the use case.

Components	Do	Know	Measure	Learn	Remember
Dashboards	Not Applicable	Optional	Essential	Preferred	Not Applicable
Data Marts	Not Applicable	Optional	Preferred	Essential	Sub-Optimal
Enterprise Data Warehouse	Not Applicable	Optional	Preferred	Preferred	Essential
Master Data Management	Preferred	Optional	Optional	Preferred	Essential
Metadata Repository	Preferred	Preferred	Preferred	Preferred	Essential
Operational Reporting	Sub-Optimal	Essential	Sub-Optimal	Anti-Pattern	Anti-Pattern
Transactional Systems	Essential	Optional	Sub-Optimal	Anti-Pattern	Anti-Pattern
Visualization Tools	Not Applicable	Optional	Essential	Preferred	Not Applicable



APPENDIX A: GLOSSARY OF INFORMATION ARCHITECTURE TERMS

Where possible, the following terms are defined using commonly accepted definitions within the data management profession, but the definitions may be adjusted to represent the way in which these terms are used within the context of the DBHIDS Data Governance Framework and Common Information Architecture.

A B C D E I K L M O P R S T U W

Term	Definition
ANALYTICAL REPORTING	Analytical reporting is the discovery and communication of meaningful patterns in data. Especially valuable in areas rich with recorded information, analytical reporting relies on the simultaneous application of statistics, computer programming and operations research to quantify performance. Analytical reporting often favors data visualization to communicate insight. Analytical reporting is a multidimensional discipline. There is extensive use of mathematics and statistics, the use of descriptive techniques and predictive models to gain valuable knowledge from data and data analysis. The insights from data are used to recommend action or to guide decision making rooted in business context. Thus, analytical reporting is not so much concerned with individual analyses or analysis steps, but with the entire methodology. Since most, though not necessarily all, analytical reporting requires processing to prepare and summarize data the term is often treated as synonymous with online analytical processing (OLAP). [Home]
ARCHITECTURE	Architecture in the context of information technology is both the process and the product of planning and designing technology-based business solutions and/or their components so as to meet expected requirements and either fit into the existing business and technology environment or improve it in a considered and intended way. See also Data Architecture, Enterprise Architecture, Enterprise Data Architecture, Information Architecture, Service-Oriented Architecture and Solution Architecture. [Home]
ATTRIBUTE	An attribute is a specification that defines the logical property of an entity. It may also refer to or set the specific value for a given instance of such. When instantiated in a physical data model, an attribute becomes a column, field or tag depending on the technology. [Home]
BIG DATA	Big data is a broad term for data sets so large or complex that traditional data processing applications are inadequate. Challenges include analysis, capture, data curation, information privacy, search,

Term	Definition
	sharing, storage, transfer, visualization and querying. The term often includes the assumption of the use of predictive analytics or certain other advanced methods to extract value from data, regardless of the size of the data. Big data can be defined by the "V dimensions": Volume, Variety, Velocity, Variability, Veracity, Visualization and Value. [Home]
BUSINESS GOAL	A business goal describes a specific outcome an organization expects to achieve over a specific period of time, whereas business objectives are broader and longer in scope. Organizations usually outline their objectives and goals in their strategic plans. [Home]
BUSINESS INTELLIGENCE OFFICE (BIO)	The Business Intelligence Office (BIO) is the unit responsible for determining the information needs of knowledge workers and other data consumers, developing appropriate data reporting structures to meet those needs and providing them to the Data Integration Office, and implementing self-service reporting, analysis and visualization (business intelligence) tools so they are available throughout the organization. [Home]
BUSINESS INTELLIGENCE PUBLISHING TOOLS (BIP TOOLS)	Traditionally, business intelligence referred to the tools used to access, query, analyze, or publish integrated data in different types of data warehouses. More and more, the term is being used by vendors to describe the entire data integration and publishing process, which includes the creation of various types of data warehouses. For this reason, our conceptual information architecture refers to the traditional category of tools as Business Intelligence Publishing (BIP) tools. These are query, analysis, reporting and visualization tools that provide rapid development of reports and dashboards (graphical representations of key performance indicators) by business people, not technologists. These are the tools that provide meaningful information to data consumers. [Home]
BUSINESS MODEL	Business models fall into two general categories, each with two perspectives. The categories are Process Models and Data Models. Business process models document the functions and transactions important to an organization and the business rules that guide them. Business data models document the information that describes things important to the organization and the business rules that govern that information. Together the models completely describe the business in a way that can be implemented in a technology solution. Each category of model can be represented in either a conceptual or

Term	Definition
	logical perspective. A conceptual model documents the key processes or data and their basic relationships. A logical model documents the detailed descriptions of the process logic or data attributes. [Home]
BUSINESS PROCESS	A business process is a function or transaction performed by a business unit as part of its organizational mission. [Home]
BUSINESS RULE	A statement that defines or constrains some aspect of the business as it is implemented in the data model. Data-related business rules are statements, phrased in absolute terms, about data (i.e., a telephone number must have 10 digits) and about relationships between data (i.e., if a phone number is entered, the phone type must be entered). [Home]
CHIEF DATA STEWARD (CDS)	Each division or significant administrative unit will designate a Chief Data Steward (CDS). The CDS is a core member of the Data Stewardship Council. The CDS represents the data for which his or her unit is the data stewardship organization. In addition, the CDS works with other data stewards to guide master data management efforts around data sets related to their data needs. [Home]
CHIEF INFORMATION OFFICER (CIO)	The DBHIDS Chief Information Officer (CIO) provides strategic vision, executive leadership and coordination for the agency's information management, technology and initiatives. [Home]
CHIEF INFORMATION OFFICER, OFFICE OF THE (OCIO)	The Office of the Chief Information Officer (OCIO) is the DBHIDS business unit led by the Chief Information Officer and containing the Office of Enterprise Data Management as well as the individuals responsible for Web 2.0 Coordination, Information Security and Information Technology Planning and Coordination. [Home]
CONCEPTUAL DATA MODEL (CDM)	A conceptual data model (also referred to as a Conceptual Schema) is a high-level description of a business's informational needs. It typically includes only the main concepts and the main relationships among them. This is a first-cut, more abstract model, with insufficient detail to build an actual database. A conceptual data model is a map of data concepts and their relationships. This describes the semantics of an organization and represents a series of assertions about its nature. Specifically, it describes the things of significance to an organization (entities/entity-types), about which it is inclined to collect information, and characteristics of (attributes) and associations between pairs of those things of significance (relationships). [Home]

Term	Definition
CONCEPTUAL INFORMATION ARCHITECTURE (CIA)	The DBHIDS Conceptual Information Architecture (CIA) represents the interrelationship of all of the information management components within the department. It categorizes these components by the information service delivery use cases each addresses. [Home]
DAMA	The Data Management Association, an organization of professionals working in data architecture, data modeling, data management and related fields. [Home]
DATA	A discrete recorded fact or value. Data is the raw material of a system supplied by data producers and used by information producers. Data can be of multiple types, i.e. tabular, spatial (location-based), video, audio. [Home]
DATA ACQUISITION	The process of obtaining a set of related information about a given subject area. For example, the extraction of source data from a legacy system for the purpose of creating a new data store relating that data to other data (perhaps extracted from another legacy system) is data acquisition. Developing a new application to capture and store information is another form of data acquisition. [Home]
DATA ANALYTICS	See Analytical Reporting. [Home]
DATA ARCHITECT (DA)	A Data Architect (DA) is an individual assigned to an IT project, business unit or subject area with the responsibility to develop a logical solution model, a business subject area model or a dimensional model consistent with the enterprise reference data model. Logical solution models will be used to create physical data models for applications. Business subject area models will guide data integration in the enterprise data warehouse. Dimensional models will define the structures required for reporting and analysis. [Home]
DATA ARCHITECTURE	An orderly arrangement of enterprise resources to achieve: (1) a common understanding of data resources available; (2) a planned approach to data acquisition, storage and retrieval to achieve a high degree of responsiveness to user demands; and (3) a high degree of data sharing and data mobility to reduce program delivery costs. [Home]
DATA ARCHITECTURE OFFICE (DAO)	The Data Architecture Office (DAO) is responsible for maintaining the enterprise reference data model, reviewing application logical solution models for consistency, overseeing the maintenance of the enterprise business glossary and other metadata resources and for the proper application of enterprise data management policies, standards and

Term	Definition
	practices. The DAO is also responsible for the design of data warehouse subject area models and master entity data and master reference data processes maintained by the Data Integration Office. The DAO is led by the Enterprise Data Architect. [Home]
DATA CONSUMER	A data consumer is an individual or organization that receives data on a regular basis, either in the form of a report or as a data extract file. Typically, this report or data is in support of a predefined, standardized process. A data consumer does not typically make ad hoc one-time requests for data. A data consumer may be a Knowledge Worker but does not have to be. [Home]
DATA CUSTODIAN	A data custodian is the individual or organization granted authority to possess, use, and/or maintain data in accordance with requirements defined by its data steward. A data custodian is responsible to protect the rights of the data owner for the access, processing, maintenance, storage, protection, and/or destruction of data and electronic records. Data custodians are responsible and accountable for the management and care of the data under their control. [Home]
DATA DICTIONARY	A data dictionary represents information about physical data and database structures. It functions as a catalog of all data objects and elements within a database schema, containing names, structures and information about their usage. [Home]
DATA ECOSYSTEM	See Information Architecture as a deliverable. [Home]
DATA ELEMENT	A data element is an atomic unit of data that has precise meaning or precise semantics. Each data element has a name, a definition and one or more representation terms. Optionally, a data element may have a defined value domain and/or a list of synonyms. Data elements are represented as attributes in a logical model, or as columns, fields or tags in a physical model. [Home]
DATA GOVERNANCE	Data governance is the exercise of decision-making and authority for data-related matters. Formal data governance is a system that provides rules and policies proactively to enable ongoing efficient service delivery while providing mechanisms to address data quality issues as they are identified. Data governance also refers to the organizational bodies, rules, decision rights, and accountabilities of people and information systems as they perform information-related processes. In other words, data governance is the identification of those with decision-making

Term	Definition
	responsibility for data management. Data governance is not a technology function. It is driven by the business and forms a bridge between business management and technology providers. Executive sponsorship comes from the business. Data governance and data stewardship represent collaborations of business subject matter experts and information architecture staff. Data management is provided by various technologists overseen by an enterprise information management unit. Where information architecture encompasses data management (the what, when and where), data governance describes roles and responsibilities (the who, how and why). [Home]
DATA GOVERNANCE FRAMEWORK (DGF)	The DBHIDS Data Governance Framework (DGF) provides guidance for the department's information architecture and data management activities. It defines the process by which enterprise data is defined and verified as fit for purpose. The DGF also addresses the quality, design, modeling, transport, transformation, storage, access and maintenance of data. It provides a framework for the development and enforcement of policies, standards and guidelines needed to manage data as an enterprise resource with the ultimate goal to deliver effective business intelligence to the user community. [Home]
DATA GOVERNANCE OFFICE (DGO)	The Data Governance Office (DGO) is the unit responsible for executing data governance policies and supporting the Data Stewardship Council. The DGO is coterminous with the Office of the Chief Information Officer (OCIO), which provides staff support. All new data initiatives will be submitted to the DGO for review. The Enterprise Data Architect and the DBHIDS IT Planning Coordinator will also participate in the review. The review intensity will be based upon the initiative's risk, size/cost, cross-unit impact and potential to advance department objectives. The review will determine compliance with architecture and standards and identify any need for a waiver. [Home]
DATA INTEGRATION	Data integration involves combining data residing in different sources by following agreed upon business rules and definitions to provide users with a unified view of the data. Data integration becomes more important as the volume and the need to share existing data explodes. Data integration can be performed in bulk with persistent storage using data warehousing technology. Data integration can also be performed in real-time on individual records using service-oriented architecture technology. [Home]
DATA INTEGRATION OFFICE (DIO)	The Data Integration Office (DIO) is the unit responsible for integrating and maintaining DBHIDS and external data sources, supplying data to

Term	Definition
	enterprise data marts, business unit data marts and external partners, and managing enterprise master entity data and master reference data. The DIO implements source system integration models provided by the Data Architecture Office and data reporting and analysis models provided by the Business intelligence Office. [Home]
DATA INTEGRATION TIER	While all data is an enterprise asset, the governance of that data is the responsibility of multiple business units. To better define data stewardship (governance) responsibilities, the Data Governance Framework classifies data into one of four integration tiers, 0 to 3. See also Universal Data, Enterprise Data, Line-of-Business Data and Programmatic Data. [Home]
DATA INTEGRITY	The state that exists when data is handled as intended and is not exposed to accidental or malicious modification, destruction or disclosure. Also, the preservation of data for its intended use. [Home]
DATA MANAGEMENT	The technical function of acquiring, defining, certifying, organizing, protecting and delivering data and the metadata that describes it. It focuses technical planning and data operations to meet program delivery objectives and business goals. [Home]
DATA MART	A data mart is a data store sourced from a data warehouse that contains smaller subsets of data and focuses on a particular business subject area. In the DBHIDS information architecture it is a pre-defined and pre-formatted subset of data from the enterprise data warehouse (EDW) or an operational data store (ODS) that has been created to address questions that need to be answered by the report community. Data marts are built for the needs of the specific report community, so the same data may exist in many ways and many combinations in different data marts. A data mart may be logical, consisting of views of EDW data, or physical, consisting of extracts of EDW data. Data marts should always receive data from a consistent, integrated source – never directly from individual operational systems – so the answer to the same question from any data mart is always the same. The DBHIDS Common Information Architecture (CIA) supports the development of dependent data marts (sourced from the EDW environment or an ODS) using conforming dimensions (common reference data used by multiple data marts). When business needs require rapid deployment of a reporting solution, the CIA permits the development of an independent data mart provided that the data to source that mart is first brought into the EDW and staged for future

Term	Definition
	integration. [Home]
DATA MODEL	A data model provides a pictorial view of data, groupings of data, relationships between data groupings, and the organization of data groupings by dependencies. There are several perspectives of data models: conceptual, logical, and physical. These models can have different contexts: enterprise, subject area or solution. A conceptual model does not include all of the detailed attributes of an entity. A logical model is fully attributed view that documents both relationships and unique identifiers. A logical model, however, does not reference the characteristics of a particular database system or the physical storage of data. Those would be reflected in a physical model. A solution model is specific to a particular system. A subject area model defines an area of the business without regard to organizational boundaries. An enterprise model serves as a reference that describes data of interest to all or part of an entire enterprise. See also Enterprise Reference Data Model, Conceptual Data Model, Logical Data Model and Physical Data Model. [Home]
DATA OWNER	The data owner is the individual or organization that has statutory rights to the data, no matter who collects the data or who manages it. The rights can include copyright and intellectual property rights as well as the rights to exploit and/or destroy the data. The rights of the data owner apply even when the data is collected by a third party and/or combined with data owned by others. The data owner can be an individual, for data such as personally identifiable information, or an organization, for data such as intellectual property owned by a company. For transaction and derived data collected or created for DBHIDS business purposes by DBHIDS units, the data owner is DBHIDS – not the individual business unit – unless the data owner has been designated as the City of Philadelphia or a state or federal agency. [Home]
DATA PERSISTENCE	In computer science, persistence refers to the characteristic of state that outlives the process that created it. Data persistence is the characteristic of data to represent a point in time when it was valid regardless of what changes may have occurred prior or subsequent to that point in time. Transaction system data is volatile and dynamic so data persistence in best provided in a data warehousing environment. [Home]
DATA QUALITY	Data quality is the summary characteristic of data that represents three broad dimensions: 1) fitness for intended use; 2) fidelity to the

Term	Definition
	real-world construct to which it refers; and, 3) internal consistency. Applying ISO 9000:2015, data quality can also be defined as the degree to which a set of characteristics of data fulfills requirements such as: completeness, validity, accuracy, consistency, availability and timeliness. Requirements can be implied or obligatory. [Home]
DATA REPOSITORY	Any data container - typically, a database. See also Data Store. [Home]
DATA REUSABILITY	Data reusability is the practice of leveraging investments in data cleansing and integration so that future consumers of the data benefit, from initial efforts. Data reusability achieves data integration through standardization of reference and master data and rationalization of definitions and formats for data used across the department. This is more effective than traditional data sharing, where data is moved between units and both the definitions and the transformations of that data are inconsistent. [Home]
DATA SECURITY COORDINATOR	The Data Security Coordinator is responsible for identifying data security policies that need to be followed throughout the organization, making implementation recommendations and advising the Data Governance Office. [Home]
DATA STAKEHOLDER	A data stakeholder is an individual or organization that is entitled to make use of data collected or maintained within a system or to derive benefits from its collection and/or publication. [Home]
DATA STEWARD	A Data Steward is the individual within the Data Stewardship Organization (DSO) identified as the point of contact for questions about the definition and use of data collected or maintained by the DSO within its functional areas and for the documentation of that data in a metadata registry. On behalf of the DSO, a data steward follows and/or implements policies, procedures, and guidelines that pertain to the data during the lifecycle of that data entrusted to his or her stewardship, and participates as a member of the Data Stewardship Council for enterprise data governance issues as well as representing his or her DSO. The data steward establishes business rules, defines data elements, identifies valid data values, establishes certification standards and determines completeness and availability of the data. [Home]
DATA STEWARDSHIP	Data stewardship is the practice of creating, implementing and following policies, procedures, and guidelines that pertain to specific

Term	Definition
	data during its lifecycle. This includes the establishment of business rules, definition of data elements, identification of valid data values and the establishment of certification standards for the completeness and availability of the data. [Home]
DATA STEWARDSHIP ORGANIZATION (DSO)	The data stewardship organization (DSO) is the organization responsible for the definition and use of data collected or maintained within its functional areas. A DSO is responsible for developing policies specifically related to the use of the data, and for designating data stewards to implement and enforce those policies. A DSO may also be the Data Owner or the Data Custodian, but not always (or even usually). [Home]
DATA STORE	While the term database is used to indicate just about any formal data storage, it is necessary to categorize databases by their function. The DBHIDS common information architecture provides different roles, both logical and physical, to address these functions. A data store is a database implementation to fulfill one of these specific roles. [Home]
DATA WAREHOUSE	Generically, a data warehouse is any informational database, or collection of databases, used to store shareable data for reporting and analysis outside of the system(s) that collected it. The warehouse is usually created through data extracts from operational databases. The warehouse typically allows users to tap into an organization's vast store of operational data to track and respond to business trends and to facilitate forecasting and planning efforts. Multiples types of data stores can all be referred to as "data warehouses", which leads to confusion. The DBHIDS Data Governance Framework does not use the stand-alone term "data warehouse" for that reason. The Gartner IT Glossary differentiates data warehouses from data marts in this way. A data warehouse contains data arranged into abstracted subject areas with time-variant versions of the same records, with an appropriate level of data grain or detail to make it useful across two or more different types of analyses most often deployed with tendencies to third normal form. A data mart contains similarly time-variant and subject-oriented data, but with relationships implying dimensional use of data wherein facts are distinctly separate from dimension data, thus making them more appropriate for single categories of analysis. See also Enterprise Data Warehouse and Data Mart. [Home]
DATA WAREHOUSING	Data warehousing encompasses the practices of identifying business data integration requirements, analyzing the source systems that can meet those requirements, modeling data structures to store the

Term	Definition
	integrated data and designing and building routines to extract, transform and load the source data into a data warehouse structure. See also Enterprise Data Warehousing. [Home]
DATABASE	A collection of related data organized to serve one or more applications. In the broader sense, it describes any organized collection of data regardless of the physical storage method. This is often used erroneously as synonym for either the data in the collection or the software product that manages the collection. [Home]
DIRECTOR OF BUSINESS INTELLIGENCE (DBI)	The Director of Business Intelligence (DBI) is the individual responsible for supporting a data-driven organization by identifying appropriate subject areas for reporting and evangelizing and implementing ubiquitous business intelligence capabilities that leverage those subject areas. [Home]
DIRECTOR OF DATA INTEGRATION (DDI)	The Director of Data Integration (DDI) is the individual responsible for advancing a rational data integration approach within the DBHIDS business and IT communities for the purpose of developing reusable data resources. The DDI shall be responsible for the integrity of data integration processes and for the measurement of data quality against the expected baseline. [Home]
DMBOK	DMBOK is the Data Management Book of Knowledge maintained by DAMA, the Data Management Association. [Home]
DOMAIN	A domain is a formal category. A data domain refers to all the values which a data element may contain. The rule for determining the domain boundary may be as simple as a data type with an enumerated list of values. An architecture domain refers to one of the three components of enterprise architecture: business process, information or technology. A data management knowledge domain refers to one of the ten knowledge areas defined by the Data Management Association as comprising the Data Management Book of Knowledge (DMBOK) Wheel, in which Data Governance is at the center of nine other knowledge domains. [Home]
ENTERPRISE	Generically, the enterprise is that which falls under central direction to meet a common mission. Enterprises are hierarchical and each level of an enterprise, such as a business unit, department or jurisdiction can think of itself as an enterprise while still being a component of a larger

Term	Definition
	enterprise. [Home]
ENTERPRISE ARCHITECTURE	Enterprise architecture (EA) is the well-defined practice for conducting enterprise analysis, design, planning, and implementation, using a holistic approach at all times, for the successful development and execution of strategy. Enterprise architecture applies architecture principles and practices to guide organizations through the business, information, process, and technology changes necessary to execute their strategies. These practices utilize the various aspects of an enterprise to identify, motivate, and achieve these changes. In the context of information technology, enterprise architecture consists of three primary sub-architectures: Business Process, Information and Technology. [Home]
ENTERPRISE DATA (TIER 1 DATA)	Enterprise data is federal, state-wide or city-wide data with a common format used by multiple jurisdictions and agencies or to serve common needs, such as geospatial data. It is governed by an external authority, jurisdiction or organization. [Home]
ENTERPRISE DATA ARCHITECT (EDA)	The Enterprise Data Architect (EDA) is the individual responsible for advancing the role of data architecture within the DBHIDS business and IT communities. The EDA shall guide the efforts of data architects and analysts within the Office of Enterprise Data Management and the business units in the development of conceptual and logical business models, and logical solution and physical data models. The EDA shall be responsible for overseeing the modeling of master data and the design of master data management processes that will be maintained by the Data Integration Office. The EDA reports to the CIO and is a member of the Data Stewardship Council. [Home]
ENTERPRISE DATA GOVERNANCE OFFICER (EDGO)	The Enterprise Data Governance Officer (EDGO) is the individual responsible for ensuring that enterprise data governance policies and practices are followed. The EDGO is responsible for the proper application of the enterprise information and data architecture principals and the overall quality and usability of enterprise data assets. The EDGO leads the Data Stewardship Council. The DBHIDS Chief Information Officer (CIO) serves as the DBHIDS EDGO. [Home]
ENTERPRISE DATA MANAGEMENT, OFFICE OF (OEDM)	The Office of Enterprise Data Management (OEDM) within the Office of the CIO is responsible for managing the enterprise data and supporting infrastructure. With guidance from the Data Governance Executive Board and the Data Stewardship Council, OEDM manages this enterprise data environment and conducts efforts specific to enterprise projects or activities. [Home]

Term	Definition
ENTERPRISE DATA WAREHOUSE (EDW)	From the William Inmon definition, a data warehouse is a purpose-built data store of shareable data in an integrated, subject-oriented, non-volatile and time-variant structure intended to meet operational and analytical requirements to support the organization. From the Bin Jiang, Ph.D. definition, a data warehouse is an infrastructure based on the information technology for an organization to integrate, collect, and prepare data on a regular basis for easing analysis In the DBHIDS common information architecture an enterprise data warehouse adheres to a single enterprise data model to ensure consistency of decision-support data across the enterprise. The DBHIDS EDW is the central integration and storage environment for enterprise data that is gathered from a variety of sources to support data analysis. An enterprise data warehouse is the single version of the truth that supplies historical data to data reusability partners, as well as to analysis areas called data marts. It is not a single database, but a consistent data integration environment that consists of multiple subject areas, staging, archiving and persistent storage and multiple physical databases. It is rarely accessed directly by end-users. The DBHIDS EDW environment can accommodate solutions for a subject area, a business unit or the department as a whole. [Home]
ENTERPRISE DATA WAREHOUSING	Enterprise Data Warehousing is the application of data warehousing practices using a centralized approach to create a single source of reusable data for the enterprise. [Home]
ENTERPRISE REFERENCE DATA MODEL (ERDM)	The ERDM is the enterprise logical data model. It defines and standardizes data used to conduct business operations across business units. By documenting the natural relationships between different groups of data, it serves as a starting blueprint for database design activities. The ERDM gives a graphical view of the universal, enterprise and line-of-business data tiers – information that is common to all business units or shared between them. The ERDM supports the management of the overall data assets to achieve optimal integration, sharing, access, and utilization of technology resources and infrastructure. The ERDM is based upon existing and emerging federal reference models and standard industry data models to the greatest extent possible. [Home]
ENTITY	An entity is a thing capable of an independent existence that can be uniquely identified. An entity is an abstraction from the complexities of a domain. It is some aspect of the real world that can be

Term	Definition
	distinguished from other aspects of the real world. Entities can be thought of as nouns. In data modeling, an entity (as used in entity-relationship modeling) is more accurately an "entity-type"; the inclusive category of all entity instances. An entity, strictly speaking, is an instance of a given entity-type. There are usually many instances of an entity-type. Because the term entity-type is somewhat cumbersome, entity has become an accepted synonym for this term. [Home]
INFORMATION	A commodity derived from data through analysis or by the orderly presentation of data for human interpretation. [Home]
INFORMATION ARCHITECTURE	As a practice, information architecture is the structural design of shared information environments; the art and science of organizing and defining data to support reusability and find-ability. Typically, it involves a model or concept of information that is used and applied to activities which require explicit details of complex information systems. As a deliverable, information architecture represents the formal definition of and mapping between information system components in an enterprise. This indicates which component is responsible for which functionality and the relationship each has with the others. Information architecture can be expressed from a conceptual, logical or physical perspective. The conceptual perspective aligns the information goals of the organization. The logical perspective defines the responsibilities for each component. The physical perspective represents the topology of all of the information systems. See Common Information Architecture. [Home]
INFORMATION SECURITY (INFOSEC)	Information security, sometimes shortened to InfoSec, is the practice of defending information from unauthorized access, use, disclosure, disruption, modification, perusal, inspection, recording or destruction. It is a general term that can be used regardless of the form the data may take (e.g. electronic, physical). It encompasses both IT security and information assurance. IT security, sometimes referred to as computer security, is information security applied to technology (most often some form of computer system). In this context, a computer is any device with a processor and some memory. Such devices can range from non-networked standalone devices as simple as calculators, to networked mobile computing devices such as smartphones and tablet computers.

Term	Definition
	Information assurance ensures that data is not lost when critical issues arise. These issues include, but are not limited to: natural disasters, computer/server malfunction, physical theft, or any other instance where data has the potential of being lost. [Home]
INFORMATION TECHNOLOGY (IT)	Information technology (IT) is the application of computers and telecommunications equipment to store, retrieve, transmit and manipulate data, often in the context of a business or other enterprise. The term is commonly used as a synonym for computers and computer networks, but it also encompasses other information distribution technologies such as television and telephones. [Home]
IT INFRASTRUCTURE	IT infrastructure is shown, as defined in the Information Technology Infrastructure Library (ITIL) standard, version 3, is a combined set of hardware, software, networks, facilities, etc. (including all of the information technology), necessary to develop, test, deliver, monitor, control or support IT services. In practice, infrastructure is used to describe the technology environment within which an application is hosted and/or functions. It can be synonymous with information technology, when information technology is used to describe technologies and not system functionality. [Home]
KEY PERFORMANCE INDICATOR (KPI)	A key performance indicator (KPI) is a type of performance measurement. KPIs evaluate the success of an organization or of a particular activity in which it engages. [Home]
KNOWLEDGE WORKER	A knowledge worker is an individual that uses data to make decisions, measure effectiveness or develop policy. A knowledge worker will assist in identifying the types of reporting and analytic solutions that are required. A knowledge worker may be a subject matter expert but does not have to be. [Home]
LINE-OF-BUSINESS DATA (TIER 2 DATA)	Data that is common to multiple business units in the same line of business, such as Homelessness. [Home]
LOGICAL DATA MODEL (LDM)	A logical data model (LDM) is a tool used to provide a pictorial view of data, groupings of data, relationships between data groupings and the organization of data groupings by dependencies. A logical data model is a view that does not reference the characteristics of a computerized system or of the physical storage of data. An LDM can be for the enterprise, a business subject area or a particular solution. It is a fully attributed view that documents both relationships and unique identifiers. It is created in a fully normalized (non-redundant, logically related) way. It does not reference the

Term	Definition
	characteristics of a particular database system or the physical storage of data. An LDM is derived from and is consistent with the ERDM. It provides documentation of new data structures to the ERDM. An LDM is a prerequisite for any new systems development. It is used to produce a Physical Data Model (PDM). Any changes needed in that PDM should be first captured in the LDM to maintain consistency. [Home]
MASTER DATA	Master data represents the business data objects which are agreed on and shared across the enterprise. Master data can take two forms; either master entity data or master reference data. It can also address two purposes; either operational data quality or analytical reporting. Master entity data describes the classes of people, places or things that are of interest to the organization. Examples are Service Recipient, Provider and Location. Master reference data describes the classes of standard code sets that are used throughout the organization. Examples are Program Activity Codes, Diagnostic Codes and Zip Codes. [Home] Master data is used to improve operational data quality by providing standard reference data for classification and by integrating standard entities to avoid creating multiple data versions of the same entity. Master Reference Data addresses analytical reporting by providing standardized coding across systems. Master entity data is the glue that connects data from disparate systems. [Home]
MASTER DATA MANAGEMENT (MDM)	Master data management (MDM) comprises the processes, governance, policies, standards and tools that define and manage the critical entity and reference data of an organization to provide a single point of reference in the organization to ensure consistency and control in the ongoing maintenance, application use and reporting and analysis of this information. [Home]
MASTER ENTITY DATA	See Master Data. [Home]
METADATA	Metadata is often describes as data about data. It can be conceptual, logical, physical or any combination. Conceptual metadata is in the form of business glossary. Logical metadata is documented in a logical data model. Physical metadata is captured in a data dictionary. Metadata includes attributes such as data name, length, domain of valid values and definition. It also includes business requirements, business rules, referential integrity, models and other forms of system documentation. Metadata describes how, when and by whom a particular set of data

Term	Definition
	was collected and how the data is formatted. In a data warehouse, Metadata can include elements such as the original source of the data, timestamps, data conversion routines, data transformations, volatility of the data, refresh periods, data reliability indicators and relationships between data from multiple sources. [Home]
METADATA MANAGEMENT	Metadata management involves managing data about other data, called metadata, whereas this other data is generally referred to as content data. Forms of metadata include data catalogs, data dictionaries, business glossaries and taxonomies. Metadata management is the end-to-end process and governance framework for creating, controlling, enhancing, attributing, defining and managing a metadata schema, model or other structured aggregation system, either independently or within a repository and the associated supporting processes. [Home]
METADATA REPOSITORY	A metadata repository is a database of information describing the characteristics (metadata) of data. Typically, the repository stores a broad range of descriptive information, including business rules, data models and process models that help to elaborate on the usage of data in various systems. Repositories can store metadata for the purpose of identifying and retrieving sets of actual data. Metadata that describes a map is an example. [Home]
MODEL-DRIVEN DEVELOPMENT (MDD)	Model-driven development (MDD) is the practice of using conceptual, logical and physical data models to create information system databases consistent with the enterprise definition of data. The DBHIDS information architecture requires a model-driven approach to development. This process begins with the development of a high-level conceptual data model (CDM) to capture the key information needs of the business. After creation of the CDM, the logical data model (LDM) process begins capturing user requirements and business rules to produce a fully normalized and attributed LDM. The LDM is used to produce a physical data model (PDM) for the solution. It is in this PDM that any changes to data structures to address performance, security, or development issues are made. The LDM remains fully normalized and the physical changes are mapped from it. The approved PDM is used to generate the data description language (DDL) needed to create the actual database structures required. When changes later need to be made to the application those changes should first be made in the LDM. The changes are then progressed through the PDM to the actual database. In this way, the documentation remains accurate and synchronized, and the impact of

Term	Definition
	changes on data integrity is fully understood. In a MDD approach, the conceptual and logical data models are guided by existing overarching models. These models can be in the form of the enterprise reference data model, subject area models specific to a business area or reference models from external authoritative source. [Home]
Online Analytical Processing (OLAP)	See Analytical Reporting. [Home]
Online Transaction Processing (OLTP)	See Transaction Processing. [Home]
Operational Data Store (ODS)	An Operational Data Store (ODS) is a central repository of current operational data gathered from a variety of existing transaction systems to present a single rational view of operational data for a single subject area or business unit, or for an entire agency or line-of-business group. The ODS can serve as a transitional facility while an organization's systems undergo a systematic re-write, as it insulates reports and interfaces from changes in the underlying transaction systems. As legacy systems are re-written, new systems can write directly to the ODS. History should not be stored in the ODS. Reporting can occur directly against an ODS or data can also be replicated into operational reporting areas called operational data marts (opera marts). [Home]
Operational Reporting	Operational reporting is reporting about operational details that reflects current activity. Operational reporting is intended to support the day-to-day activities of the organization. Operational reporting is typically performed by the same business unit that is responsible for the transaction processing of the data being reported. Operational reporting is typically produced out of the same transaction processing system or from a copy of it made specifically for reporting purposes. When reporting requires data from other sources, significant transformation of the data or historical data not otherwise required by the transaction processing system, the reporting should be moved into an analytical reporting system. [Home]
Physical Data Model (PDM)	A physical data model (PDM) is what most developers and many business people think of when they hear the term "data model". It may be a relational model or a dimensional model. A PDM is derived

Term	Definition
	from and is consistent with the LDM. It will document variations from the fully normalized LDM that are necessary for the physical implementation. The PDM is the design of the physical database structures for a system. It is used to produce the physical database. Any changes needed in that database should be first captured in the LDM, and pushed out to the PDM to maintain consistency. [Home]
PREDICTIVE ANALYTICS	See Analytical Reporting. [Home]
PROGRAMMATIC DATA (TIER 3 DATA)	Data owned and managed by a specific business unit for a specific program, and which does not normally need to be shared. [Home]
REFERENCE DATA	See Master Data. [Home]
REFERENCE DATA MANAGEMENT	See Master Data Management. [Home]
REPORTING, ANALYTICAL	See Analytical Reporting. [Home]
REPORTING, OPERATIONAL	See Operational Reporting. [Home]
SCHEMA	A schema is a description of an information model typically expressed in terms of data structure, domain values, relationships and other constraints. Constraints are expressed using some combination of grammatical rules governing the order of elements, Boolean predicates that the content must satisfy, data types governing the content of elements and attributes, and more specialized rules such as uniqueness and referential integrity constraints. A schema can be expressed in a conceptual, logical or physical perspective. A conceptual schema is represented as a logical data model. A logical schema is represented as a physical data model. A physical schema is represented by the database implementation itself. [Home]
SERVICE-ORIENTED ARCHITECTURE (SOA)	A service-oriented architecture (SOA) is an architectural pattern in software design in which application components provide services to other components via a communications protocol, typically over a network. The principles of service-orientation are independent of any vendor, product or technology. As an architectural style it supports service-orientation. Service-orientation is a way of thinking in terms of services and service-based development and the outcomes of services. A service is a self-contained unit of functionality, an operation that may be discretely invoked. A service is a logical representation of a

Term	Definition
	repeatable business activity that has a specified outcome. It is self-contained. A service may be composed of other services but it will appear as a "black box" to consumers of the service. [Home]
SOLUTION ARCHITECTURE	As an activity, solution architecture is the practice of defining and describing the design of a system delivered in context of a specific solution and as such it may encompass description of an entire system or only its specific parts. As a deliverable, solution architecture is the description of a discrete and focused business operation or activity and how information technology supports that operation. An individual solution architecture typically applies to a single project or project release, assisting in the translation of requirements into a solution vision, high-level business and/or IT system specifications, and a portfolio of implementation tasks. Solution architecture activities and deliverables combine guidance from different enterprise architecture viewpoints (business process, information and technical). [Home]
SOURCE SYSTEM OF RECORD	A source system of record is the authoritative system that provides data for integration, aggregation, reporting and analysis. It is typically a transaction processing system but may be an internal or external reference system. [Home]
SOURCE-TO-TARGET MAPPING	Source-to-target mapping is the process of evaluating each field in a target environment to determine if there is a corresponding field in one or more source environments that is suitable for populating the target field. This process is generally performed during a gap analysis to surface problems resulting from inadequate source system data. The expected result of a source to target mapping exercise is the identification of the authoritative source field(s) to provide the necessary data for each target field. [Home]
SPATIAL DATA	Information that identifies the geographic location and characteristics of natural or constructed features and boundaries on the earth. This information may be derived from sources such as remote sensing, mapping and surveying technologies. It includes both attributes (text) as well as spatial (map) data. [Home]
SUBJECT AREA	A subject area is the category of knowledge around a specific business function. Subject areas often represent functions that cross organizational boundaries and in that roles become the focus for data integration or master data management efforts. [Home]

Term	Definition
SUBJECT MATTER EXPERT (SME)	A subject matter expert (SME) is an individual within a business unit or the analyst assigned to the business unit that has the most complete understanding of a particular subject area or process. A SME is engaged to make sure that the right data is being used appropriately to meet business needs for both transaction processing and reporting purposes. A SME may also be a data steward and/or a knowledge worker, but does not have to be. [Home]
TIER 0 DATA	See Universal Data. [Home]
TIER 1 DATA	See Enterprise Data. [Home]
TIER 2 DATA	See Line-of-Business Data. [Home]
TIER 3 DATA	See Programmatic Data. [Home]
TRANSACTION	A transaction symbolizes a unit of work performed within a database management system against a database, and treated in a coherent and reliable way independent of other transactions. A transaction generally represents any change in the database. By definition, it must be atomic, consistent, isolated and durable. These properties of database transactions are collectively referred to by the acronym ACID. Transactions have two main purposes to support consistency: 1) to provide reliable units of work that allow proper recovery from failures; and 2) to provide isolation between programs accessing a database concurrently. If this isolation is not provided, the programs' outcomes can be inconsistent and/or inaccurate. [Home]
TRANSACTION PROCESSING	Transaction processing is information processing that is divided into individual, indivisible operations called transactions. Each transaction must succeed or fail as a complete unit; it can never be only partially complete. Transaction processing is designed to maintain database integrity in a known, consistent state, by ensuring that interdependent operations on the data are either all completed successfully or all canceled successfully. Since most, though not necessarily all, transaction processing today is interactive the term is often treated as synonymous with online transaction processing (OLTP). [Home]
TRANSACTION PROCESSING SYSTEMS	These data stores are where the results of business transactions with the organization are stored. They can be in relational, hierarchical, or file-based database management systems. They can be on a

Term	Definition
	mainframe or on a distributed (network) server. They can be batch processing systems, on-line transaction processing (OLTP) systems, or a hybrid. These systems are typically built to perform a specific business function. An organization can have hundreds, or even thousands, of different transaction processing systems. Because many of these systems were built in isolation without concern for future data integration, they are referred to as silo or stovepipe applications. [Home]
UNIVERSAL DATA (TIER 0 DATA)	Common data that is used by most, if not all business units, such as people places and things, but which describes attributes that exist independently of interactions with the business systems, such as an individual's name or date of birth. [Home]
UNSTRUCTURED DATA	Unstructured data refers to information that either does not have a pre-defined data model or is not organized in a pre-defined manner. Unstructured information is typically text-heavy, but may contain data such as dates, numbers, and facts as well. This results in irregularities and ambiguities that make it difficult to understand using traditional programs as compared to data stored either in highly-structured form in databases or annotated (semantically tagged) in documents. The term is somewhat imprecise insofar as this type of data generally has some form of structure, just not the formally defined structure of databases. Examples include emails as well as text documents accompanied by metadata. In this context, unstructured means that whatever structure exists is insufficient for direct, automated processing of the data within the structures. It may also be called semi-structured data or non-tabular data.[Home]
USE CASE	A use case is a list of actions or event steps, typically defining the interactions between a role and a system to achieve a goal. The role can be a human, an external system, or time. In systems engineering, use cases are used at a higher level than within software engineering, often representing missions or stakeholder goals. [Home]
WEB SERVICE	A Web service is a service offered by an electronic device to another electronic device, communicating with each other via the World wide web. In a web service, web technology such as the hyper-text transport protocol (HTTP) is utilized for transferring machine readable file formats such as XML and JSON. [Home]

APPENDIX B: THE OWNER/STEWARD/CUSTODIAN MODEL

The Data Owner/Steward/Custodian Model is the relationship between data owner, steward, and custodian is often confusing, as the roles overlap depending on the data. A transaction record often has composite ownership, with the details of the transaction occurrence being owned by the transacting organization but the details about people, places or things involved in the transaction owned by others. The following example illustrates the application of the various roles.

Organization A must collect the following information from Individual X to process a transaction:

- Individual X Name at time of transaction
- Individual X Date of Birth
- Individual X Social Security Number
- Individual X Address at time of transaction
- Transaction Date
- Transaction Amount
- Scanned copy of Individual X Birth Certificate provided by Organization B

As a result of the transaction, Organization A issues a document with the following information:

- Individual X Name
- Individual X Date of Birth
- Individual X Address at time of transaction
- Document Identifier
- Document Expiration Date

As a result of the transaction, as permitted by law or regulation, Organization A makes the following information available to Organization C:

- Individual X Name
- Individual X Date of Birth
- Individual X Address at time of transaction
- Document Identifier
- Document X Expiration Date

The following table indicates who serves in the indicated role for specific data in this transaction.

DATA OWNER/STEWARD/CUSTODIAN MODEL

Data Object	Data Owner	Enterprise Data Steward	Business Data Steward	Data Custodian
X Name	Individual X for the name itself	Organization A for the transaction, Organization B for birth certificate	Organization C for the reused data	Organizations A, B & C
X Date of Birth	Individual X for the date of birth itself	Organization A for the transaction, Organization B for birth certificate	Organization C for the reused data	Organizations A, B & C
X Social Security #	Individual X for the social security number itself	US Social Security Administration	Organization A for the transaction	Organization A
X Address	City of Philadelphia	Master Address Organization	Organization A for the transaction	Organizations A & C
X Birth Certificate	Individual X	Organization B	Organization A for the transaction	Organization A
Transaction Amount	DBHIDS	Organization A	Organization A	Organizations A
Document Identifier *	Individual X or DBHIDS *	Organization A	Organization A	Organizations A & C
Document Expiration	DBHIDS	Organization A	Organization A	Organizations A & C
Transaction Record *(this is the transaction as a whole)*	DBHIDS *(subject to personal privacy considerations)*	Organization A	Organization A	Organization A
Transaction Audit Log *(any log created of the transaction)*	DBHIDS *(subject to personal privacy considerations)*	Organization A	Organization A	Organization A

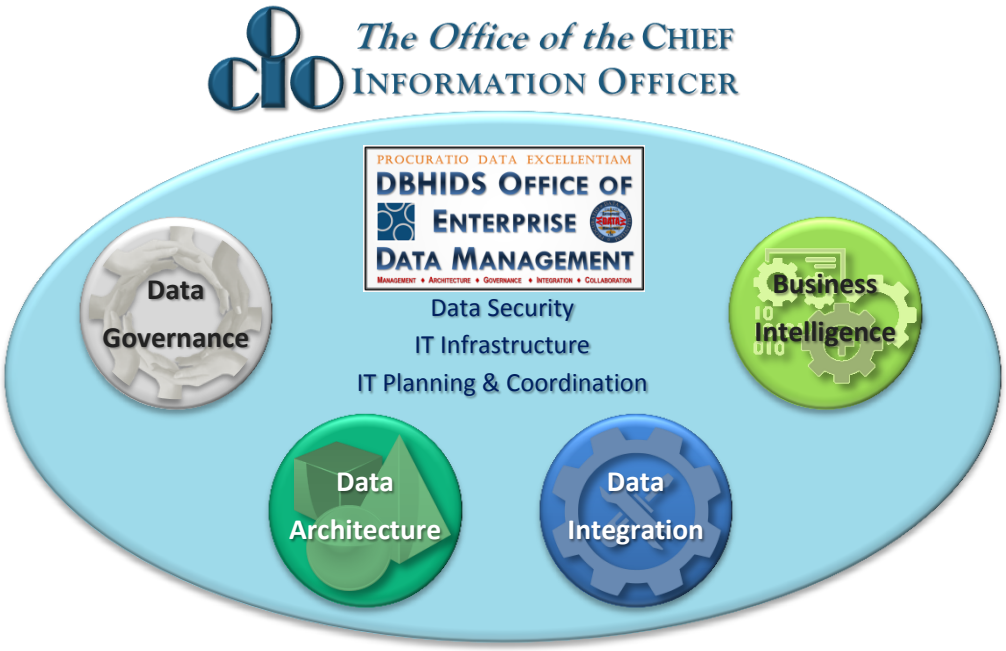
* Where a document identifier, such as a driver's license number or social security number is considered to be personally identifiable information (PII), the owner is the individual to which it refers. Otherwise, it is DBHIDS.

Note that where the data owner is not an external party, the data owner is always DBHIDS, not a specific unit.

APPENDIX C: IMPLEMENTATION PLAN STRATEGIES AND GOALS

STRATEGY #1: ESTABLISH THE DBHIDS OFFICE OF ENTERPRISE DATA MANAGEMENT (OEDM)

We must recognize that to be a data-driven organization we must make decisions using data of known quality and that known data quality requires an enterprise approach to data governance and management. While the Data Stewardship Council provides high-level guidance, there needs to be a focal point to operationalize data governance. Individual units today perform many aspects of data management, but DBHIDS-wide objectives require a DBHIDS-wide unit to implement. Individual business units have their own respective priorities upon which they must focus. Reporting to the DBHIDS Chief Information Officer (CIO) the OEDM will be the focal point for and enabler of enterprise data management, architecture, governance, integration, collaboration and business intelligence.



The Office will include department-level offices to serve as enterprise focal points for coordinated data governance and data management. Each Office will be responsible for developing best practices in its respective domain and ensuring that they are used throughout the department.

- Data Governance – led by the Chief Information Officer
- Data Architecture – led by the Enterprise Data Architect
- Data Integration – led by the Enterprise Data Warehouse Manager
- Business Intelligence – led by the Business Intelligence Director

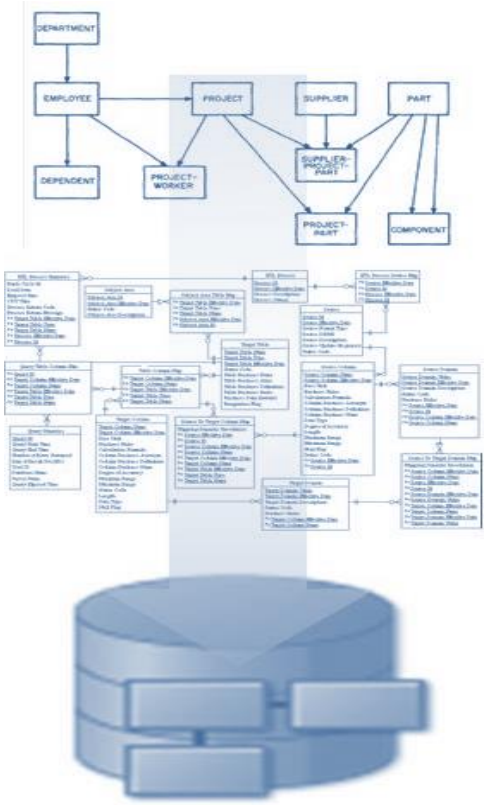
In addition, the DBHIDS Security Policy Coordinator will also coordinate Data Security.

From this strategy, all others follow. This strategy supports each of the Data Governance Objectives.

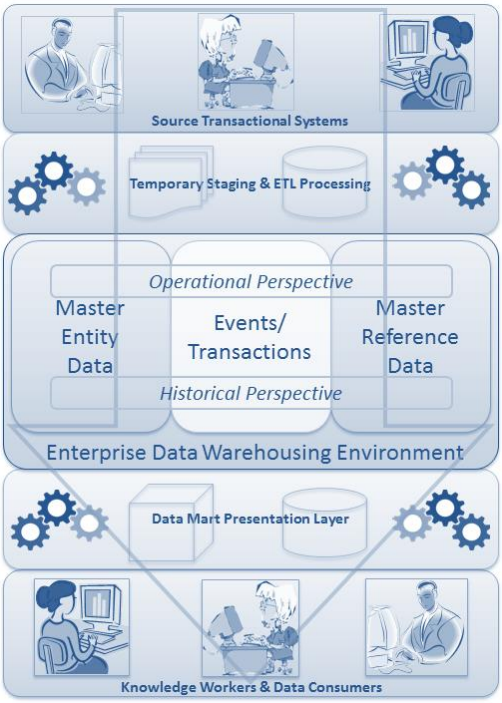
These are the essential goals of this strategy.

1. Under the **Data Governance Office**
    a. Create and use a data architecture review process and a data governance process to synchronize business definitions and database definitions and make sure that all new transactional and analytical work is consistent with the enterprise standards and business model.
    b. Coordinate these data reviews with an overall IT review process.
    c. Maintain a Business Glossary with authoritative definitions of business terms and make the glossary available to both business and technical users.

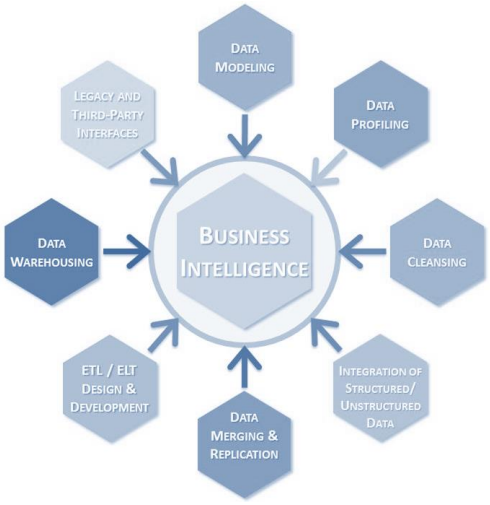
2. Under the **Data Architecture Office**
    a. Use model-driven development techniques to create business models that map to all existing operational data and use those models for all new development.
    b. Define, create and manage authoritative Master Reference Data as well as the Master Entity Data domains of service recipients, providers, facilities, services and programs.
    c. Document and maintain data dictionaries for all transactional and analytical systems, map their terms to the enterprise business model and glossary and make them available to both technical and business users.



3. Under the **Data Integration Office**
    a. Create and manage an enterprise data warehousing environment that supports persistent, normalized integration of transactional data.
    b. Make the enterprise data warehouse data available for reuse to supply data marts, provide data for external partners and support ad hoc and exploratory data querying (data mining) to identify possible future analytical models.
    c. Integrate and maintain master entity and reference data history within the enterprise data warehousing environment to support historical analysis.



4. Under the **Business Intelligence Office**
    a. Design and deliver optimized analytical data solutions in the form of data marts and analytical cubes.
    b. Couple analytical data solutions with self-service business intelligence and data visualization tools.
    c. Empower business users to become knowledge workers by providing access to appropriate data of known quality when and in the format needed to achieve organization objectives.



5. Working with the **Security Policy Coordinator**
    a. Recommend data security policies to the CIO and Data Governance Executive Board.
    b. Develop best practices for implementation of data security policies.
    c. Provide guidance to data management and technology staff, knowledge workers and data consumers on best practices usage.
    d. Monitor data security policy compliance.

STRATEGY #2: IMPLEMENT THE DATA GOVERNANCE FRAMEWORK

We must implement the Data Governance Framework by embracing the philosophy that data of any value to the Department is an enterprise asset that must be managed as such. This will help to change our culture from one that values data of poor quality provided quickly into one that values quality data and not only accepts but expects rigor in the preparation and documentation necessary to achieve it.

Besides implementing the Data Governance Framework, the Office of Enterprise Data Management must continue to evolve it through engagement with the Data Governance Executive Board and the Data Stewardship Council.

This strategy supports the following strategic objectives:

Objective 1	Create an information-centric and informed organizational culture.
Objective 2	Establish a data governance program to provide accountability for information assets.
Objective 4	Improve the quality and usefulness of information by making it timelier, more accurate, more complete and more accessible.
Objective 5	Reduce the costs of managing information.

These are the essential goals of this strategy and the unit(s) responsible.

1. Identify and define Data Governance roles and responsibilities. (CIO, DSC)
2. Identify individuals to fill the required roles. (CIO, DSC, Mgmt)
3. Establish lines of authority and a communications plan to enable individuals to work on their respective responsibilities in a collaborative manner. (CIO, DSC)
4. Assess the effectiveness of the Data Governance Framework through efforts of the Data Governance Executive Board, Data Stewardship Council and the Data Governance Office and adjust it as necessary to meet evolving business needs or to address challenges. (CIO)

STRATEGY #3: ACHIEVE COMPLIANCE WITH INFORMATION ARCHITECTURE POLICIES AND STANDARDS

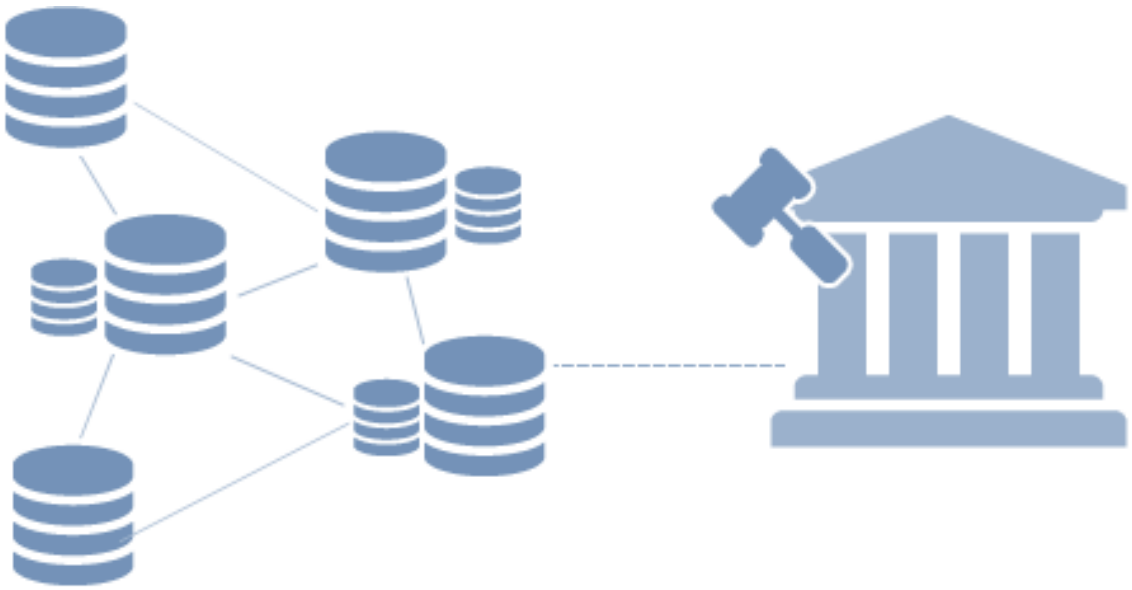
Standards are necessary to achieve consistency and reach enterprise goals. In order for a standard to be enforceable, it must draw authority from a policy. Approved policies must carry the weight of executive management behind them. Enforcement of policies and standards must first be seen as an opportunity for better education. When that is not sufficient, there must be a defined process for escalating an issue.

This strategy supports the following strategic objectives:

Objective 1	Create an information-centric and informed organizational culture.
Objective 2	Establish a data governance program to provide accountability for information assets.
Objective 3	Provide for effective and appropriate information security.
Objective 4	Improve the quality and usefulness of information by making it timelier, more accurate, more complete and more accessible.
Objective 5	Reduce the costs of managing information.

These are the essential goals of this strategy and the unit(s) responsible.

1. Publish approved policies and standards and disseminate information and/or training as needed. (DGO)
2. Monitor approved policies and standards for compliance. (DGO)
3. Attempt to gain compliance where there are gaps. If unsuccessful, bring the issue(s) to the attention of the Data Stewardship Council and, if needed, the Data Governance Executive Board. (DGO)



STRATEGY #4: DEVELOP NEW DRAFT POLICIES AND STANDARDS FOR CONSIDERATION BY THE DSC

Business needs, technologies and capabilities evolve over time. Policies and standards must evolve concurrently in order to maximize value and relevance. The development of policies and standards must be through a repeatable and rational process, and approved policies must carry the weight of executive management behind them.

This strategy supports the following strategic objectives:

Objective 1	Create an information-centric and informed organizational culture.
Objective 2	Establish a data governance program to provide accountability for information assets.
Objective 3	Provide for effective and appropriate information security.
Objective 4	Improve the quality and usefulness of information by making it timelier, more accurate, more complete and more accessible.
Objective 5	Reduce the costs of managing information.

These are the essential goals of this strategy and the unit(s) responsible.

1. Identify topics for data-related policies and standards. (Mgmt, DGEB, DSC, OEDM, IT)
2. Develop policy and standard recommendations. (DSC, DGO)
3. Present recommended policies and standards to the Data Governance Executive Board for approval. (DGO, DGEB)
4. Publish and socialize approved policies and standards and provide additional information and/or training as needed. (DGO)

STRATEGY #5: MANAGE THE ENTERPRISE REFERENCE DATA MODEL AND BUSINESS GLOSSARY

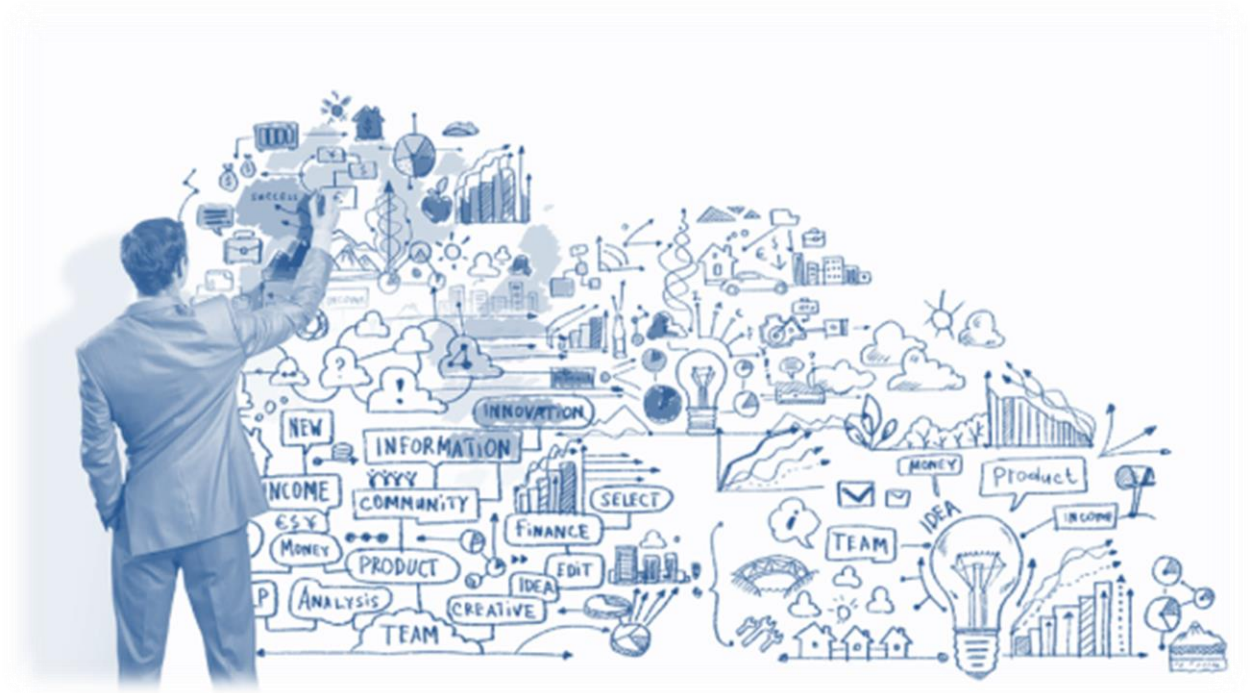
Traditionally, data design and documentation has followed a bottom-up approach. The focus has been on documenting already-built physical systems in data dictionaries. While necessary, this is not the most important documentation for an effective enterprise data management approach. We must guide all data collection, integration and publishing efforts with an agreed-upon logical data model of the organization. This Enterprise Reference Data Model must in turn be guided by and consistent with authoritative business terms, rules and definitions maintained in the Enterprise Business Glossary.

This strategy supports the following strategic objectives:

Objective 1	Create an information-centric and informed organizational culture.
Objective 2	Establish a data governance program to provide accountability for information assets.
Objective 4	Improve the quality and usefulness of information by making it timelier, more accurate, more complete and more accessible.
Objective 5	Reduce the costs of managing information.
Objective 6	Share data through reusable processes; reuse data through shared processes.
Objective 7	Provide self-service business intelligence capabilities.

These are the essential goals of this strategy and the unit(s) responsible.

1. Create, maintain and publish a Business Glossary that defines terms and related business rules used within the business and represented within information systems. (DAO, DGO)
2. Create, maintain and publish an Enterprise Reference Data Model that represents the logical, canonical, normalized, integrated representation of data used across the department. (DAO)



STRATEGY #6: PROVIDE METADATA MANAGEMENT FOR BUSINESS AND TECHNICAL USERS

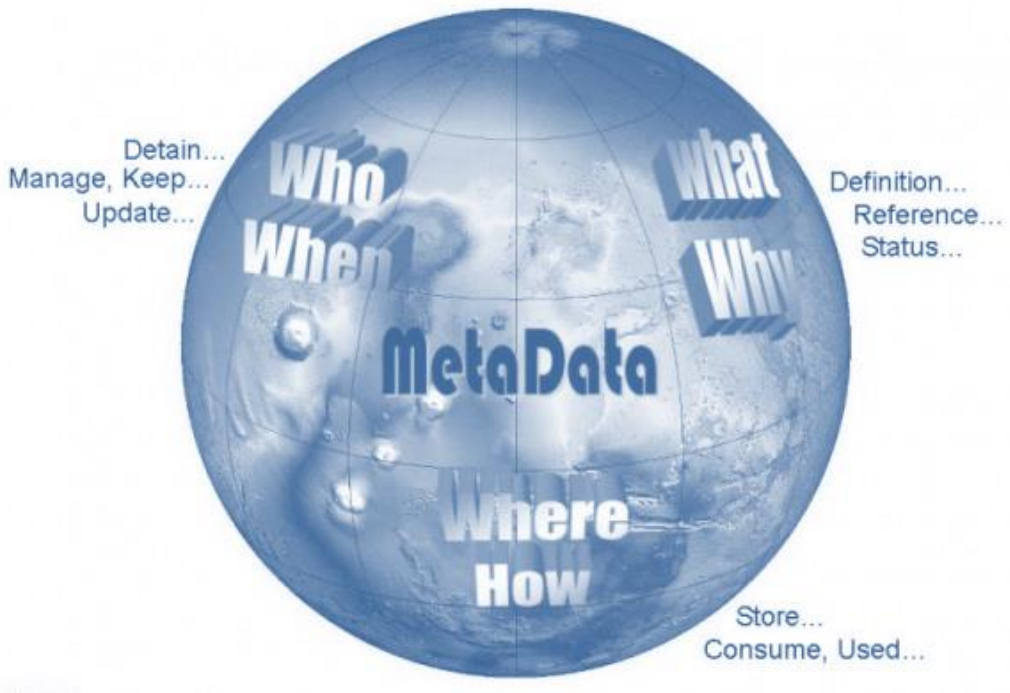
Metadata is the information resource that helps us achieve data quality goals. It includes documentation of definitions, formats and business rules essential to ensure that data is properly collected, stored and published and is fit for the intended use. Often, this documentation does not exist. Even when business and/or technical metadata exists in some form or fashion within the organization, it is not coordinated, consistent or consumable.

This strategy supports the following strategic objectives:

Objective 1	Create an information-centric and informed organizational culture.
Objective 2	Establish a data governance program to provide accountability for information assets.
Objective 4	Improve the quality and usefulness of information by making it timelier, more accurate, more complete and more accessible.
Objective 5	Reduce the costs of managing information.
Objective 7	Provide self-service business intelligence capabilities.

These are the essential goals of this strategy and the unit(s) responsible.

1. Create, maintain and publish a Data Dictionary that documents physical database system technical information in a consistent way. (DAO, IT)
2. Map Business Glossary terms and Data Dictionary data objects to the Enterprise Reference Data Model so that business and technical users will be able to trace any term or object to its authoritative definition and representation, and understand how terms and data are related across different systems. (DAO)



STRATEGY #7: MANAGE MASTER DATA MANAGEMENT DOMAINS

Master data is the foundation for data integration, data warehousing and business intelligence. Master entity data ensures that the data that defines unique instances of people, places and things for which we collect and report are de-duplicated and well-defined. Master reference data ensures that the data used to categorize data is consistent so as to improve data quality for both collection and publishing. By definition, master data has enterprise value and must be managed with an enterprise approach.

This strategy supports the following strategic objectives:

Objective 1	Create an information-centric and informed organizational culture.
Objective 2	Establish a data governance program to provide accountability for information assets.
Objective 4	Improve the quality and usefulness of information by making it timelier, more accurate, more complete and more accessible.
Objective 5	Reduce the costs of managing information.
Objective 6	Share data through reusable processes; reuse data through shared processes.

These are the essential goals of this strategy and the unit(s) responsible.

1. Identify, define and model the key enterprise master entity domains. (DAO, DSC)
2. Establish data governance for each master entity domain. (DGO, DSC)
3. Implement a master data management process for each master entity domain. (DAO, DIO)
4. Identify, aggregate and publish authoritative reference data for reuse. (DAO, DIO)

STRATEGY #8: MANAGE THE DATA INTEGRATION AND PERSISTENCE LAYER FOR THE DEPARTMENT

In order to achieve the goals of the DBHIDS Data Governance Framework, it is essential that there is a single reusable, non-redundant, governed and defined collection of enterprise data. By definition, this collection must be managed at the enterprise level. The data integration and persistence layer represents the "Remember" information service delivery use case. Guides by the Enterprise Reference Data Model, it includes Master Entity Data, Master Reference Data, the time-normalized enterprise data warehouse data and the tools to support the integration and management of this data.

This strategy supports the following strategic objectives:

Objective 2	Establish a data governance program to provide accountability for information assets.
Objective 3	Provide for effective and appropriate information security.
Objective 4	Improve the quality and usefulness of information by making it timelier, more accurate, more complete and more accessible.
Objective 5	Reduce the costs of managing information.
Objective 6	Share data through reusable processes; reuse data through shared processes.
Objective 8	Develop enterprise-class data management staff.
Objective 9	Adopt enterprise-class data management tools.

These are the essential goals of this strategy and the unit(s) responsible.

1. Identify, implement and maintain an enterprise data warehouse infrastructure. (DIO)
2. Identify, implement and maintain a master data management infrastructure. (DIO)
3. Establish, collect and monitor data quality and process quality metrics. (DIO, DAO)
4. Optimize data warehouse data resources for reusability. (DAO, DIO)

STRATEGY #9: PROVISION DATA FOR THOSE THAT NEED IT IN THE MANNER REQUIRED

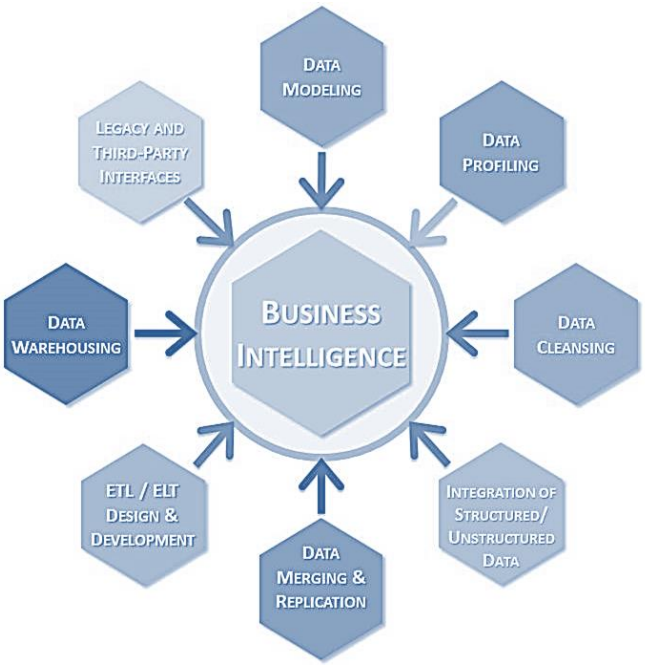
Fundamentally, this is the definition of Business Intelligence. While we have always attempted to provide data to those that need it, it was not always performed consistently or efficiently. Often, data consumers did not receive data in the appropriate format because they were expected to use the format designed for a different need. This is not the reusability that the Data Governance Framework seeks. Most importantly, these efforts were not done in a governed fashion, were not driven by an enterprise model and had inconsistent definitions and business rules applied. An enterprise business intelligence approach addresses each of these shortcomings.

These are the essential goals of this strategy and the unit(s) responsible.

Objective 1	Create an information-centric and informed organizational culture.
Objective 3	Provide for effective and appropriate information security.
Objective 4	Improve the quality and usefulness of information by making it timelier, more accurate, more complete and more accessible.
Objective 6	Share data through reusable processes; reuse data through shared processes.
Objective 7	Provide self-service business intelligence capabilities.
Objective 8	Develop enterprise-class data management staff.

These are the essential goals of this strategy.

1. Eliminate “Management by Spreadsheet” processes. (BIO, DGO, DSC)
2. Identify and deploy business intelligence and data visualization tools. (BIO, IT)
3. Identify reporting and analytics needs and deploy purpose-built data marts. (BIO, DAO, DIO)
4. Establish best practices for data reporting, visualization and analytics. (BIO, DGO, DSC)
5. Provide continuous assessment and optimization of reporting environments. (BIO, IT)



STRATEGY #10: EVANGELIZE THE DATA GOVERNANCE FRAMEWORK PRINCIPLES TO THE DEPARTMENT

For the Data Governance Framework to be effective and the Common Information Architecture to be realized, all stakeholders must be informed of the principles that underlie them. Data practitioners and consumers must also be educated about the definitions, practices, methodologies and expectations. In addition, the Data Governance Framework must be a living document. This requires that all stakeholders not only have awareness of it but are able to contribute to its evolution.

This strategy supports the following strategic objectives:

Objective 1 Create an information-centric and informed organizational culture.

Objective 2 Establish a data governance program to provide accountability for information assets.

Objective 8 Develop enterprise-class data management staff.

These are the essential goals of this strategy and the unit(s) responsible.

1. The Data Governance Executive Board will be kept informed of the activities of the Data Stewardship Council and the Office of Enterprise Data Management.
2. Data practitioners will meet regularly to share techniques and discuss common challenges.
3. Data governance-related presentations will be made available to DBHIDS staff on a regular basis.
4. The Data Governance Framework Strategic and Implementation plans will be made available for reference and appropriate sections cited in RFPs, project proposals and any other description of a data-related initiative.

APPENDIX D: DBHIDS DATA ARCHITECTURE PROGRAM TACTICAL PLAN

The following Data Architecture Program Tactical Plan was developed by consultant Jessica Costanzo of Resilient Consulting and was submitted December 11, 2015. It is copyrighted © 2015 to Ms. Costanzo and Resilient Consulting and licensed to DBHIDS without restriction.

PURPOSE OF THE TACTICAL PLAN:

The purpose of this tactical plan is to identify the specific actions to be taken in order to launch an Enterprise Data Architecture Program for DBHIDS.

WHERE YOU ARE NOW

An assessment of existing gaps around the use of data:

Data Management is splintered throughout DBHIDS. There are multiple units that collect data but each operates in a silo. Much of the same data is used in each silo and there is minimal coordination or unity between the units where the data is concerned. This results in much wasted effort and poor data quality. Furthermore, the data contained in the siloes are used to create independent data marts. While independent data marts can report pertinent information from the data, they may not be the most efficient design for data usage. Independent data marts typically do not support:

- Data Reusability
- Data Integration
- Data Standardization

Often times when each disparate unit creates its own independent data mart using the same data, the result is different answers to the same questions! There are multiple reasons for this:

- The data elements in each silo are used differently, even if they are from the same source.
- The data elements are defined differently in each silo or not at all causing them to be reported differently.
- The data stored in each silo are saved from different points in time, or not saved at all, which results in reports being published from different perspectives.

Additionally, DBHIDS lacks the appropriate tools:

- To collect essential information about the data it collects. This essential information about the data is called metadata (or data about data.)
- To create standardized logical and physical models.
- To capture business rules and ETL data mapping for use by the developers of a data warehouse or data mart solution.

Lastly, DBHIDS requires training to establish a skillset in a model-driven development methodology.

- Training in Conceptual, Logical, Physical and Dimensional Modeling.
- Creating standardized, reusable templates for Business Rule and ETL mapping documentation.

- Creation of a metadata repository to capture business, logical and technical information about the data.

WHAT YOU CAN DO

Develop a Data Architecture Program for creating an Enterprise Data Warehouse and managing Master Data using a model-driven approach:

The development of a Data Architecture Program will address these challenges by using a model-driven methodology to capture valuable information about DBHIDS data. This valuable information is documented and stored, using tools and techniques to create standardization, maintain consistency and manage all data and information about data centrally. A model-driven approach to data architecture will enable DBHIDS to document and illustrate information about the data on at least three levels: conceptual, logical and physical.

Gathering information on the conceptual, logical and physical levels is manifested in two different scenarios within each discipline. Consider the conceptual:

- Conceptual Business Glossary – This includes gathering metadata information pertaining to common business terms, business definitions and business owners.
- Conceptual Modeling – This is a high level model, which is preliminary in structure, possibly abstract in content and sparse in attributes, that is intended to represent a business area. It is preliminary in structure because it may contain many-to-many relationships.

There is an implicit relationship between the Conceptual Business Glossary and a Conceptual Model based around similar data domains. The Conceptual Business Glossary encapsulates information about the Conceptual Model and vice versa.

When considering the logical model in model-driven data architecture; a logical model is a type of model showing a detailed representation of an organization's data, independent of any particular technology. A logical model is fully attributed with entities and attributes named logically and consistently. Metadata for the Logical Business Glossary (or Lexicon), including the logical names and definitions of entities and attributes is captured and documented. Similar to the conceptual version, there is an implicit relationship between the Logical Model and Logical Business Glossary.

Within model-driven data architecture, the logical model is then used to create a physical model. A physical model (or database design) is a representation of a data design which takes into account the facilities and constraints of a given database management system. The physical model and physical metadata (data dictionary) about the tables, columns, business rules and ETL mapping is captured and documented.

A central repository of data in an Enterprise Data Warehouse enables the creation and implementation of environments for reporting using dependent data marts. Dependent data marts draw data from an enterprise data warehouse that has already been created. Pulling data from an enterprise data warehouse better ensures that the use of the data elements for reporting purposes is consistent. What is reported on and published will **not** result in different

reporting results from siloes of data. Instead, the information is reported from a consistent, standardized and defined view of the world. Furthermore, once a report is created, it can be easily reproduced without having to start from scratch. Data is reused consistently, because data is defined consistently.

Thus, a model-driven approach to developing a Data Architecture Program and Enterprise Data Warehouse will produce following:

- A consistently defined conceptual illustration of the business-level information assets.
- A standardized logical representation of all data that resides in the centralized enterprise data warehouse.
- A standard for designing the logical representation of new data that is brought into the data warehouse.
- The defining of data elements consistently at the logical level, including definitions, data types and relationships.
- A standard around the naming of the physical containers (tables and columns) of data.
- The ability to create physical database structures from the logical representation instead of building from scratch at the database level.
- A reporting environment that publishes information consistently from a single centralized source of data – the enterprise data warehouse!

WHERE YOU CAN BEGIN

A toolkit of guides to get you started:

- ERwin - "How To" Quick Reference Guides:
  - Installation
  - Administration
  - Usage
- Training Material Entity-Relationship Diagram Data Modeling
  - Conceptual Modeling Concepts and Exercises
  - Logical Modeling Concepts and Exercises
  - Physical Modeling Concepts and Exercises
- ETL Development Hand-Off
  - Creating Source to Target Mapping Documents
  - Creating Business Rule Documents
- Metadata
  - Business Glossary Template
  - Logical Glossary Template
  - Physical Metadata Template

SUGGESTED FUTURE TRAINING

TDWI courses

https://tdwi.org/pages/education/onsite/course-list.aspx

Data Analysis and Design

- TDWI Data Modeling: Data Analysis and Design for BI and Data Warehousing Systems
- TDWI Advanced Data Modeling Techniques
- TDWI Dimensional Data Modeling Primer: From Requirements to Business Analysis
- Advanced Dimensional Modeling: Techniques for Practitioners
- Dimensional Modeling Beyond the Basics: Intermediate and Advanced Techniques

In the end, you will have a strong Data Architecture framework that will enable data reusability, standardization and consistency.

WORKS CITED

FEA Consolidated Reference Model Document (whitehouse.gov) Oct 2007. Web.
https://en.wikipedia.org/wiki/Physical_data_model. 1 Oct. 2015

Haughey, Tom. "The Conceptual Data Model." Erwin.com/community. Industry Expert Blogs, 4 Mar. 2011. Web. erwin.com/community/industry-expert-blogs/the-conceptual-data-model. 1 Oct. 2015.

Matthew West and Julian Fowler (1999). Developing High Quality Data Models. The European Process Industries STEP Technical Liaison Executive (EPISTLE). Web.
https://en.wikipedia.org/wiki/Logical_data_model. 1 Oct. 2015.

Oracle8*i* Data Warehousing Guide Release 2 (8.1.6)A76994-01 (1999). Web.
https://docs.oracle.com/cd/A81042_01/DOC/server.816/a76994/marts.htm

APPENDIX E: MISSION AND BRANDING OF THE ENTERPRISE UNIT

WE SEEK DATA ARETE

Arete (“Ahr-reh-tee “; Greek: ἀρετή), in its basic sense, means "excellence of any kind". The term may also mean "moral virtue". In Greek literature, this notion of excellence was ultimately bound up with the notion of the fulfillment of purpose or function: the act of living up to one's full potential. Data Arete means using data of excellent quality to its full potential.

Arete is frequently associated with bravery, but more often with effectiveness. Data Arete therefore is data of the highest effectiveness. The concept implies a data-centered universe in which data management is of paramount importance. It recognizes that the world is a place of conflict and difficulty, and data value and meaning is how this conflict and difficulty is managed and overcome.

Data Arete is explicitly linked with human knowledge; "knowledge is virtue " and "Data Arete is knowledge". The highest data potential is knowledge and all other capabilities are derived from this central capacity. If Data Arete is knowledge, the highest data knowledge is knowledge about knowledge itself; metadata – the ultimate product of Data Architecture.

(**adapted from Wikipedia**)

Our Mission

In support of the DBHIDS population health mission we will define, cleanse and integrate the data needed to provide better clinical services, measure outcomes and make more informed policy decisions while reducing the cost of those activities. We will identify the core data in use across the department and establish a collaborative approach to managing it so as to provide information to those that need it while meeting the quality, format and content expectations of the organization.

OPERATIONAL BRANDING

Since the role of the Office of Data Arete would not be understood by most people, we instead will operate collectively as the Office of Enterprise Data Management (EDM).

Our motto shall be

PROCURATIO DATA EXCELLENTIAM

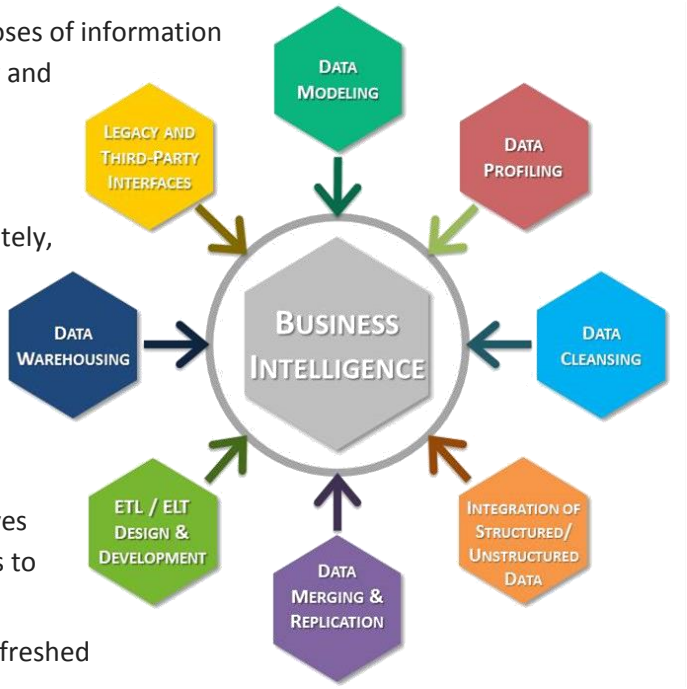
(Excellence in Data Governance)

APPENDIX F: DBHIDS IT SERVICES REALIGNMENT

Concurrent with the creation of the Enterprise Data Management organization, the overall information technology units of the department will also be realigned.

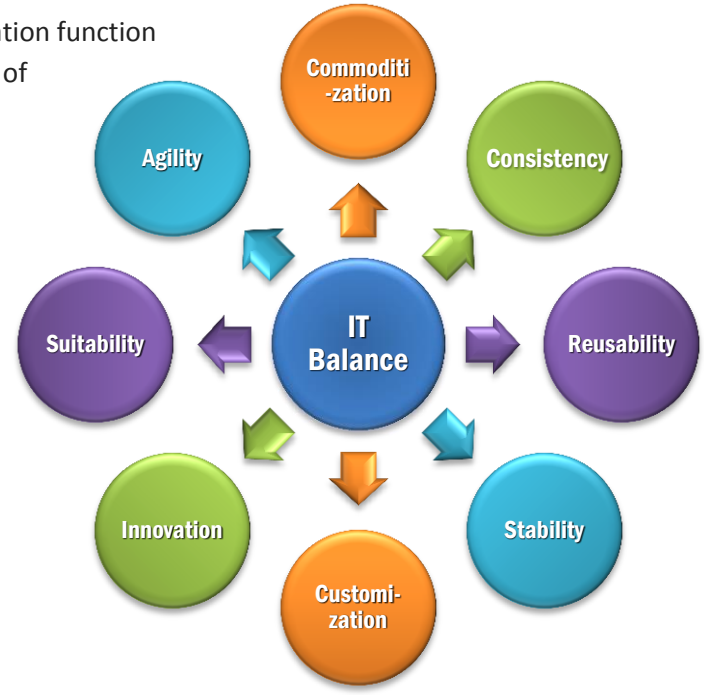
STRATEGIC GOALS

- Embrace the philosophy that the purposes of information technology are to maximize the quality and usefulness of Department data, to optimize information delivery to staff, service delivery partners, service recipients and constituents and, ultimately, deliver effective business intelligence to the organization.

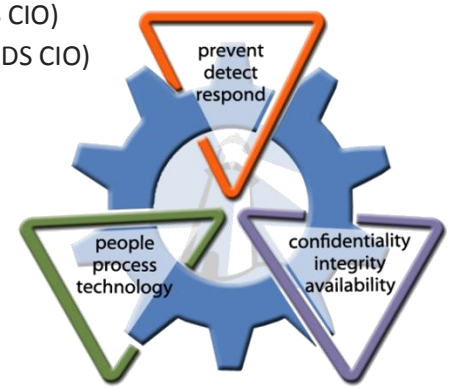


- Develop an organization culture where both business unit and IT management collectively
  - Identify organizational objectives and make technology decisions to meet those objectives
  - Choose to invest in new and refreshed technology to meet organizational objectives and needs
  - Adjust existing technology priorities when new technology projects or priorities are identified

- Develop an IT Planning and Coordination function to balance the competing objectives of
  - **Commoditization** *versus* **Customization**
  - **Consistency** *versus* **Innovation**
  - **Reusability** *versus* **Suitability**
  - **Stability** *versus* **Agility**



- Establish a department-level IT Planning & Coordination Committee (ITPCC) to serve as the enterprise focal point for IT planning and coordination. The Committee will be responsible for developing IT and networking best practices and ensuring that they are used throughout the department. The ITSPC will consist of the following:
    - DBHIDS Chief Information Officer
    - CBH Chief Information Officer
    - DBHIDS RIM/IT Director
    - BHSI/OAS IT Manager
    - DBHIDS Web 2.0 Coordinator (office of the DBHIDS CIO)
    - DBHIDS IT Planning Coordinator (office of the DBHIDS CIO)
    - DBHIDS Security Policy Coordinator
      (will also serve as Data Security Coordinator)



DBHIDS IT STRATEGIC PLANNING & COORDINATION TEAM
EFFECTIVE 2/1/2016