

---

## Section 4. ETA Information Domain

The Information Domain assists business and technical leaders in making sound decisions related to data warehouse design and acquisition of data warehouse, business intelligence, and other reporting tools and products. This domain also provides a framework for defining responsibility for data integrity and distribution. An effective Information Domain will enable the Commonwealth to leverage the most value from its data assets. This domain addresses the Reporting, Data Management, Business Intelligence, and Knowledge Management topics. Other than the Information Domain-wide requirements identified below, no specific requirements are identified for the Knowledge Management topic.

### *Domain-wide Requirements*

The following domain-wide requirements pertain to all topics and components in the Information Domain.

- INF-R-01 Security and Privacy** – All Information Domain IT systems, as listed above, shall be implemented in adherence with all security, confidentiality and privacy policies and applicable statutes.

Agencies should ensure that all of their mission critical applications are or can be adequately supported by the vendors of all hardware and software products used to support those applications. Software needs to be at a current and supportable release levels and support must be available from a reputable supplier that can be purchased as needed. Non-complying software tools and hardware require an approved Commonwealth Enterprise Technical Architecture Exception.

- INF-R-02 Software Tools Version/Release Support** – All software used to support Mission Critical Information/ Business Intelligence Applications shall be on version/ release levels that are fully supported by the vendor or third party and have traditional paid-for support available.

Currently no central repository exists in the Commonwealth that captures software tool use for the types of components found within the Information Domain. Maintaining an ‘as-is’ inventory is critical in moving the enterprise to a desirable ‘to-be’ state.

- INF-R-03 Maintain Software Tools Inventory** – The Commonwealth shall collect data on agency use of software tools, maintain an up-to-date inventory, and perform research in order to create a more effective and efficient environment in support of the Information Domain.

Many Information Domain activities involve the storage of artifacts by electronic means. These files, source code listings, reports, models, etc., need to be accessible and available today and in the future. In order to ensure that these artifacts remain accessible, they must either be in an open format or, at a minimum, the software tool which manages the artifact needs to have the ability to easily export the item to a format that is accessible and can be transported to other tools as needed.

- 
- INF-R-04**    **Artifact Accessibility** – All electronic repositories of Information/Business Intelligence source code, metadata, development artifacts, models, documentation, etc. shall have their contents accessible either by an export facility or by a direct access method. This ability is required to allow the repository contents to be transferred from one methodology or tool to another as needed.

### ***Reporting***

Requirements have been identified within the *Reporting* topic for the Ad Hoc End-User Reporting, Standardized/Canned Reporting, and Online Analytical Processing (OLAP) components. Other than the Information Domain-wide requirements identified above, no specific requirements are identified for the Standardized/Canned Reporting component.

#### **Ad Hoc End-User Reporting**

Ad hoc query provides the business analyst with the ability to pose specific questions to produce a result without needing the programming of a report by IT. The ad hoc nature of these queries implies a short shelf life where some situation is being researched or a new opportunity is being explored.

Following are requirements that all newly acquired Information/Business Intelligence ad-hoc reporting software tools must support.

- INF-R-05**    **Ability to share queries** – An ad hoc end-user tool shall be able to share an ad hoc query with others. This enables the reuse and efficient utilization of agency resources
- INF-R-06**    **Intuitive interface** – The tool used to build a report shall have an intuitive interface, with “point and click” features for adding elements, filtering data, and sorting the results, with no programming knowledge required.

#### **Online Analytical Processing (OLAP)**

OLAP tools view information in the form of cubes, or multiple dimensions and allow the user to drill down to lower levels of detail, and slice across different dimensions such as time or commodity. These tools are generally used by the business analyst in conducting research to answer business questions as part of the decision making process.

Following are requirements that all newly acquired Information/Business Intelligence OLAP software tools must support.

- INF-R-07**    **Drill-down capability** – OLAP tools shall have the ability to drill into the details of a cell in an OLAP cube by going to the source database.
- INF-R-08**    **OLAP Export** – OLAP tools shall have the ability to export the results to a standard spreadsheet format such as .csv or .xls.
- INF-R-09**    **Easy cube manipulation** – The interface to manipulate data in the cube shall have “point and click” and “drag and drop” features for analyzing the available data.

---

## ***Data Management***

The *Data Management* topic is concerned with the components that affect the quality, management, meta-management, accessibility, and recovery of electronic data resources.

Requirements have been identified within the *Data Management* topic for the Data Standards and Data Classification (security and access) components.

### **Data Standards**

It is important to address the issues of data and data quality through the use of data standards. Data standards are important in the quest for data integration and consist of a framework used to classify or define data. These standards may include Data Element Naming, Database Object Naming, Metadata Requirements, Data Modeling, and Geo-Spatial Requirements.

Following are requirements that all newly acquired Information/Business Intelligence software tools must support.

- INF-R-10**     **Standard file formats** – Agencies shall ensure that all software tools or packages that create files or data stores do so in a format that is based on an underlying open or de facto standard or provides the capability to export to such a format.

### **Data Classification (security and access)**

Data must be classified according to its degree of sensitivity in a universally understandable manner. The degree of sensitivity can be determined by applying the appropriate state, local or federal laws or regulations to the data. Sensitivity levels are determined by the type of information that is in an automated system. The information that has the least amount of sensitivity might include things such as summary revenue and expense data for the Commonwealth. Data that is made generally available without specific custodian approval and that has not been explicitly and authoritatively classified as confidential is not considered sensitive. Highly sensitive information would include information that must be protected to meet state and federal Privacy Act requirements including data such as social security numbers, credit card numbers, criminal and medical histories, etc. It is also data whose loss, corruption, or unauthorized disclosure would be a violation of state and federal statutes, mandates and regulations. The term "in a universally understandable manner" implies there should be standard definitions for the different sensitivity classifications. In addition, the data needs to maintain its security classification as it traverses any physical or logical boundary such as an agency, computer-related device, network, or software application system.

- INF-R-11**     **Sensitivity classification** – Data that is sensitive shall be classified by the agency according to its degree of sensitivity in a universally understandable manner.
- INF-R-12**     **Security classification** – Data that requires a security classification shall maintain its security classification as it traverses any physical or logical boundary such as an agency, computer-related device, network, or software application system.

---

## ***Business Intelligence***

Business intelligence (BI) is a broad category of application programs and technologies for gathering, storing, analyzing, and providing access to data to help enterprise users make better business decisions.

Requirements have been identified within the *Business Intelligence* topic for the Data Warehouse/ Data Marts component. Other than the Information Domain-wide requirements identified above, no specific requirements are identified for the other Business Intelligence components: Operational Data Stores, Extraction, Transformation and Loading (ETL), Data Storage Structures, Data Mining, Demand Forecasting and Management, Balanced Scorecard, Decision Support and Planning, Business Analytics Suites, and Dashboards.

The phrase business intelligence (BI) may refer to

- 1) a set of business processes,
- 2) the technology used in these processes, or
- 3) the information obtained from these processes.

## **Data Warehouse / Data Marts**

A data warehouse is a database designed to support decision-making in an organization or enterprise. It is refreshed, or batch updated, and can contain massive amounts of data. When the database is organized for one department or function, it is often called a "data mart" rather than a data warehouse. The data in a data warehouse is typically historical and static in nature.

**INF-R-13     Read-only Data Warehouse** – Access shall be restricted to read-only for end users of the data warehouse.

**INF-R-14     Database Standard** – Data warehouses and data marts that use relational databases shall conform to all of the Requirements and Technology Product Standards for databases as defined above in Section 3: ETA Database Domain.

To ensure that data warehouse and data mart implementations are built to meet the current and future business needs of an agency, executive sponsorship and representation by the business community on the project is required. Without this leadership, business intelligence (BI) projects run the risk of not providing the anticipated rewards or even failing altogether.

**INF-R-15     Business community representation** – A representative of the business community shall be involved in the entire development life cycle of all BI projects.

**INF-R-16     Executive sponsorship** – Project sponsorship shall be obtained from one or more executives within the upper management of the related organization prior to initiating any Data Mart or Data Warehouse project.