Section 3. ETA Database Domain

The Database Domain describes the technical components of the software systems that support storage and retrieval of data and the types of database software that will support applications. It includes the two topics of Database and Other Data Access Methods, and Data Management. Database and Other Data Access Methods addresses the components Hierarchical, Networked, Relational, and Object-oriented databases, and Other Data Access Methods. Data Management addresses the components Data Recovery and Backup, Data Dictionary, Database Administration, Enterprise Information Integration (EII), Database Design (Standards and Tools), and Data Modeling components.

Domain-wide Requirements

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

DB-R-01 **Security, Confidentiality and Privacy Policies.** Production databases shall be implemented in adherence with all security, confidentiality and privacy policies and applicable statutes.

DB-R-02 **Support Tools Version/Release Levels.** The version/release levels of all databases and related tools used to develop or support Commonwealth and/or agency "*mission critical applications*" shall have vendor or equivalent level support.

DB-R-03 **Assess Business Recovery Requirements.** An assessment of business recovery requirements is mandatory when acquiring, developing, enhancing or outsourcing database solutions. Based on that assessment, appropriate disaster recovery and business continuity planning, design and testing shall take place.

DB-R-04 **Restrict Free-Form Data Entry/Update.** Data entry and update to production databases using direct database access shall be restricted, logged and reported to business owners or other appropriate staff. Production database owners shall provide written delegated authority for this type of access.

Database and Other Data Access Methods

A database is a collection of information organized in such a way that a computer program can quickly select (access) desired pieces of data. A database management system (DBMS) is a software application providing management, administration, performance, and analysis tools for databases. The Database and Other Data Access Methods topic has Hierarchical, Networked, Relational, and Object-oriented (Object) components.

DB-R-05 **Minimize DBMS Number/Version.** Agencies shall minimize the number and versions of database management systems utilized.

DB-R-06 **Support Connectivity.** Newly deployed database technologies shall support Java Database Connectivity (JDBC) and Microsoft connectivity technology

(such as Open Database Connectivity (ODBC) or Object Linking and Embedding Database [OLEDB]).

Hierarchical Database

A hierarchical database is a kind of database management system that links records together in a tree data structure such that each record type has only one owner, e.g. an order is owned by only one customer. Hierarchical structures were widely used in the first mainframe database management systems. However, due to their restrictions, they often cannot be used to relate structures that exist in the real world. See DB-S-01 Table below for component standards.

Networked Database

A networked database is a database model conceived as a more flexible alternative to the hierarchical model. Where the hierarchical model structures data as a tree of records, with each record having one parent record and many children, the network model allows each record to have multiple parent and child records, forming a lattice structure. See DB-S-01 Table below for component standards.

Relational Database

A relational database is a database model in which the database is organized and accessed according to the relationships between data items without the need for any consideration of physical orientation and relationship. Relationships between data items are expressed by means of tables.

DB-R-07 **Relational DBMS for New Applications/Solutions.** A Relational DBMS shall be used as the "Database and Other Data Access Method" for newly developed or acquired applications/solutions.

DB-R-08 **Support Security Using Database Access Controls.** The SQL implementation and relational database products shall support database security using the following database access controls: GRANT and REVOKE privilege facilities, the VIEW definition capabilities, and some Discretionary Access Control (DAC) mechanisms.

Object-oriented (Object) Database

An object database (more correctly referred to as ODBMS or OODBMS for Object DBMS or Object Oriented DBMS, respectively) is a DBMS that stores objects as opposed to tuples (one row of a database table…one record) or records in a RDBMS (Relational Database Management System) or record-based DBMS. As data is stored as objects it can be interpreted only using the methods specified by its class. The relationship between similar objects is preserved (inheritance) as are references between objects. See DB-S-01 Table below for component standards.

Other Data Access Methods

Indexed **S**equential **A**ccess **M**ethod (ISAM) is a common disk access method that stores data sequentially while maintaining an index of key fields to all the records in the file for direct access. The sequential order would be the one most commonly used for batch processing and printing (account number, name, etc.).

Virtual **S**torage **A**ccess **M**ethod (VSAM) is an IBM access method for storing data, widely used in IBM mainframes. It uses the B+tree method for organizing data.

See DB-S-01 Table below for component standards.

The following table provides strategic direction for agencies that are acquiring database and other data access method products.

Table DB-S-01: Database and Other Data Access Methods Technology Component Standard *(Updated April 1, 2009)*
Strategic:
Microsoft SQL Server 2005,Microsoft SQL Server 2000 Oracle 9i or 10g DB2 Version 8.x MySQL (shall have vendor or equivalent quality level support if used for *Mission Critical Applications*)
Emerging:
EnterpriseDB CACHÉ Other Object-oriented (Object), Multidimensional, and Real Time Databases
Transitional/Contained:
IMS VSAM Adabas MAPPER, BIS, Cool Ice
Obsolescent/Rejected:
Desktop database products (Such as Microsoft Access, Lotus Approach, or Paradox, are considered desktop productivity tools. ***They shall not be used for multi-user applications***.) All Networked Databases All Hierarchical Databases not categorized as "Transitional/Contained" All versions/release levels of Database and Other Data Access Methods that do not have vendor or equivalent level quality support All other non-specified Database and Other Data Access Methods
Exception History:
02/12/2009: CIO approved adding the MS SQL 2008 DBMS product as a strategic technology *Notes:* • *This table is under review and will be updated in the 2nd quarter of 2009* • *Oracle version 9.2 extended support ends 7/2010* • *DB2 version 8.1 support ends on 9/08/2009*

Data Management

Data Management defines the set of capabilities that support the usage, processing and general administration of unstructured information. The Data Management topic has Data Recovery and Backup, Data Dictionary, Database Administration, Enterprise Information Integration (EII), Database Design (Standards and Tools), and Data Modeling components. Other than the Domain-wide requirements identified above, no specific requirements are identified for the Database Design (Standards and Tools) component.

Data Recovery and Backup

Data Recovery and Backup defines the set of capabilities that support the restoration and stabilization of data sets to a consistent, desired state.

DB-R-09 **Test Production Databases.** Production databases shall be periodically tested for recoverability according to requirements for their use and preservation.

DB-R-10 **Business/Recovery Strategies Shall Address Business Requirements.** All backup and recovery strategies shall address the business requirements of the data regarding availability, accuracy, and timeliness of data.

DB-R-11 **Backup Metadata.** Metadata (database schemas, structures, data definitions, etc.) shall be backed up along with the data.

DB-R-12 **Recover to Point-In-Time and Point-Of-Failure.** Production databases supporting mission critical applications shall be recoverable to a point-in-time and point-of-failure.

DB-R-13 **Define High Availability Strategy.** Databases requiring 24 x 7 availability shall have a high availability strategy such as failover, mirroring, and/or the use of online backups.

DB-R-14 **Production Databases.** Production databases shall be on different physical machines than the test and development databases.

Data Dictionary

A Data Dictionary is a database about data and databases. It holds the name, type, range of values, source, and authorization for access for each data element in the organization's files and databases. It also indicates which application programs use that data so that when a change in a data structure is contemplated, a list of affected programs can be generated. The data dictionary may be a stand-alone system or an integral part of the DBMS.

DB-R-15 **Implement a Data Dictionary.** A Data Dictionary is required for any development that results in new databases and any enhancement activities that result in new tables being added to existing databases.

Database Administration

Database administration is the process of establishing computerized databases and insuring their recoverability, integrity, security, availability, reliability, and performance.

DB-R-16 **Assign DBA (Database Administrator) Responsibilities.** Agencies shall formally assign the responsibilities for database administration.

DB-R-17 **Limit DBA Permissions.** Database permissions shall be granted at the minimum level required. Limit the members of the System or Database Administrators role to trusted DBAs. Create custom database roles, if required, for better control over permissions. Business data manipulation by DBAs shall not be permitted.

DB-R-18 **Control Application Access and Passwords. Reset Default Access.** Production application programs or interfaces shall never be given System or Database Administration authority. Default accounts shall be changed.

Production passwords shall be changed from test and development environments.

DB-R-19 **Limit Query/Reporting Database Access to Read-Only.** Direct production database access for ad-hoc queries and end-user reporting shall be read-only.

DB-R-20 **Evaluate and Apply Patches.** DBAs shall evaluate the latest service packs and security patches released by DBMS vendors. When the DBMS is utilized by a 3rd party application, all patches shall be certified by that application vendor before being applied. Security patches shall be applied and the other service packs and patches should be applied according to DBMS and related 3rd party application vendor recommendations as needed.

DB-R-21 **Monitor Databases for Planning and Availability.** Databases for mission critical applications shall be monitored proactively for capacity planning purposes and to maintain high availability.

Enterprise Information Integration (EII)

EII is the industry acronym for **Enterprise Information Integration**. It describes the process of using data abstraction to address the data access challenges associated with data heterogeneity and data contextualization. Data is the foundation upon which the "Information Age" and critical components such as the burgeoning Web 2.0 and a future Semantic Web are being built. Uniform data access and uniform information representation are critical aspects of this journey.

An EII product offers virtualization of heterogeneous data where data takes the form of SQL, Extensible Markup Language (XML), Data-returning Web services, and other Universal Resource Identifier (URI) resources that may be referenced. Such SQL data is typically accessible via Open Database Connectivity (ODBC, Java Database Connectivity (JDBC), Active X Data Objects (ADO.NET), Object Linking and Embedding Database (OLEDB) APIs. XML is generally URI based, and is thus accessible via (Web-based Distributed Authoring and Versioning) WebDAV.

EII products enable loose coupling between homogenous-data consuming client applications and services and heterogeneous-data stores. Such client applications and services include desktop productivity tools (spreadsheets, word processors, presentation software, etc.), development environments and frameworks (J2EE, .NET, Mono, Simple Object Access Protocol [SOAP] or RESTian [Representational State Transfer] web services, etc.), Business Intelligence (BI), Business Activity Monitoring (BAM), Enterprise Resource Planning (ERP), Customer Relationship Management (CRM), Business Process Management (BPM) and/or Business Process Execution Language (BPEL), and Web Content Management.

DB-R-22 **Enterprise Information Integration (EII) Tool.** Agencies (excluding institutions of higher education) shall not purchase an EII tool without VITA approval.

Data Modeling

Using modeling tools to describe (usually graphically) the attributes and tables (fields and records) of the organization of a database; it is often created as an entity relationship diagram. In many tools, the SQL code that defines the data structure (schema) in the database is automatically created from the visual representation.

DB-R-23 **Implement a Data Modeling Tool.** Agencies shall select and implement a

consistent data modeling tool.