



Design a Geodatabase

Rasu Muthurakku, Enterprise Solutions Architect / Consultant

An abstract 3D architectural graphic on the right side of the slide. It features a tall, blue, rectangular structure with white topographic contour lines. The structure is surrounded by various colored blocks and lines in shades of orange, red, green, and yellow, creating a complex, layered geometric composition.

**GIS
INSPIRING
WHAT'S
NEXT**

The "The Science of Where" logo is enclosed in a white square frame with a double-line border. The text "THE SCIENCE OF WHERE" is written in a bold, uppercase, sans-serif font, stacked vertically within the frame.

**THE
SCIENCE
OF
WHERE**

Agenda

Design a Geodatabase

- Overview
- Key Factors
- Geodatabase Design
- Recommendations

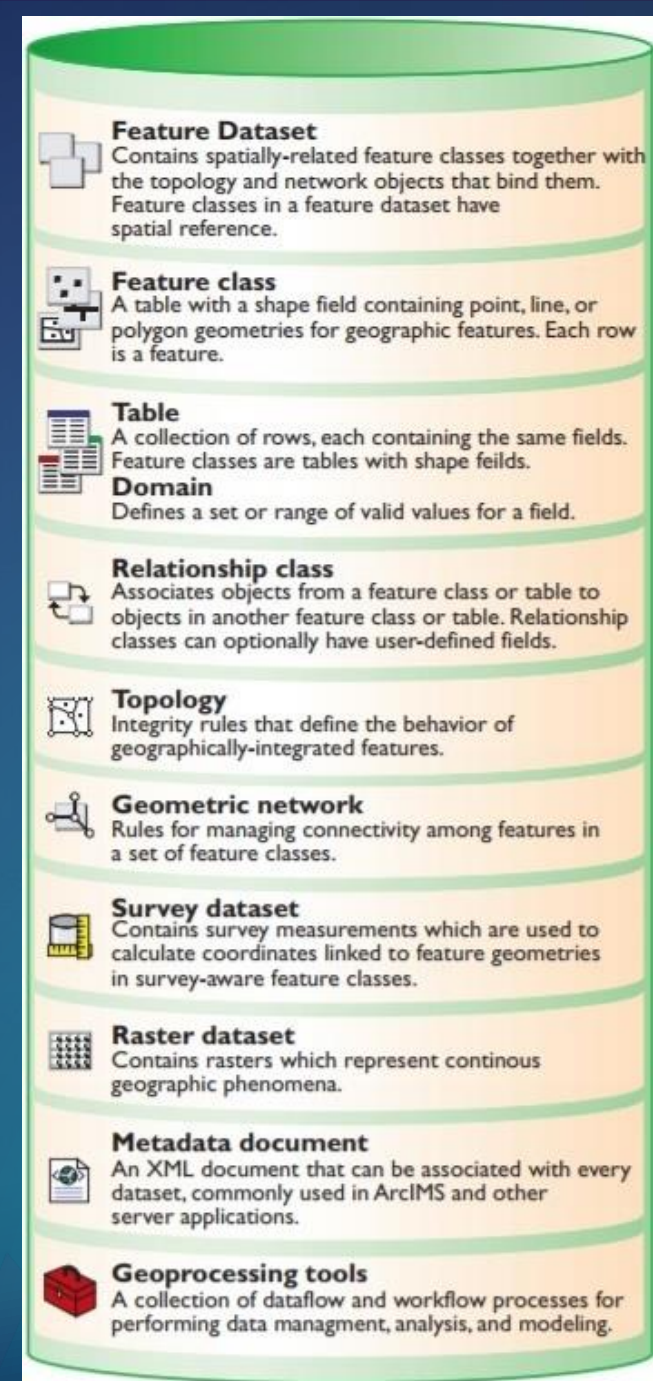
Key Considerations, Best Practices, Recommendations and Lessons Learned!

Overview

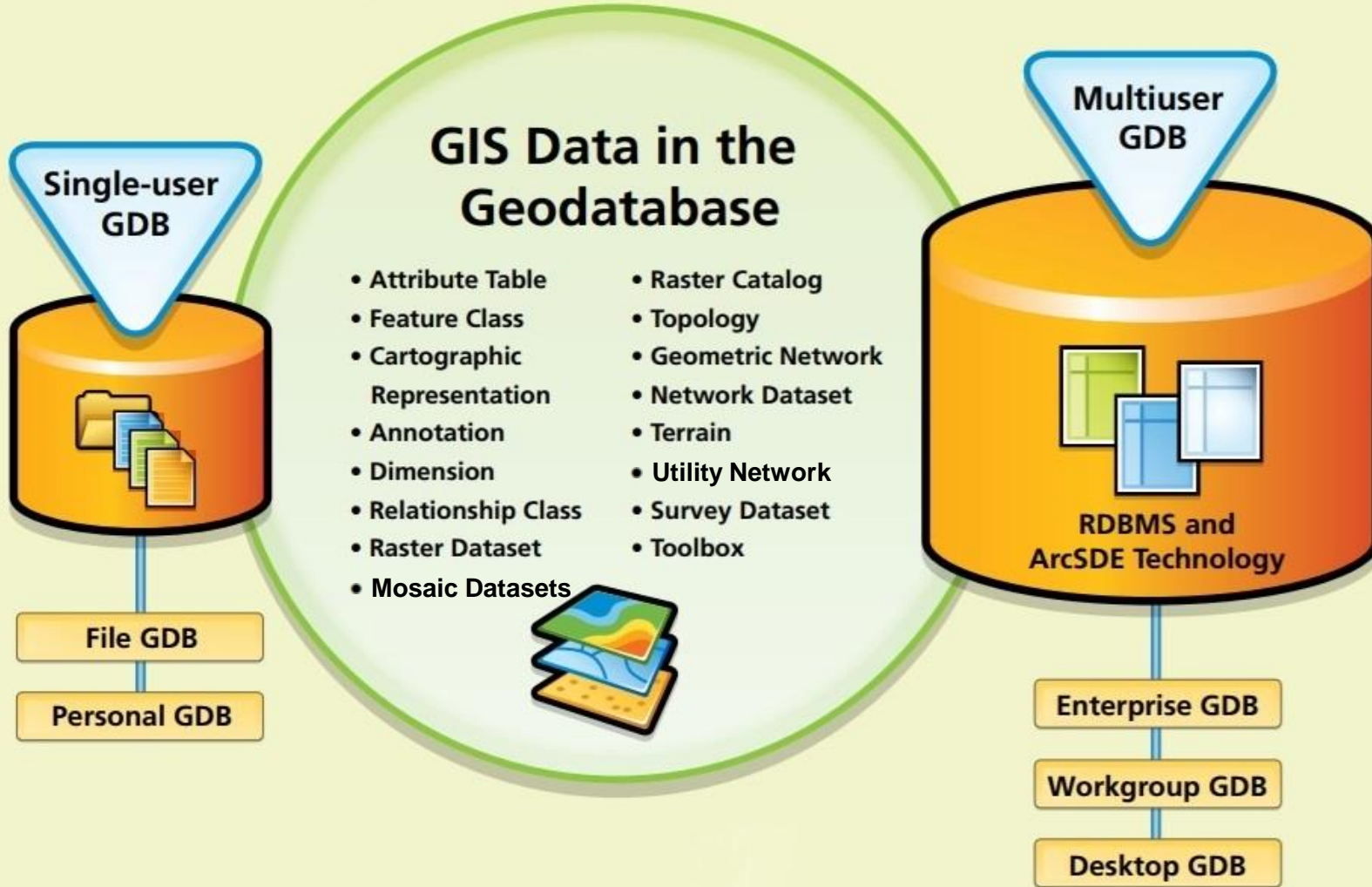
What is a Geodatabase (GDB)?

- Collection of Geographic Datasets of Various Types Stored in:
 - Common File System Folder
 - *Microsoft Access Database*
 - Multiuser RDBMS * / ArcSDE
 - Oracle
 - Microsoft SQL Server
 - PostgreSQL
 - IBM DB2, Informix
 - SAP Hana 2.0 SPS02 New
- Native Data Structure for ArcGIS
- Primary Data Format Used for Editing and Data Management
- Comprehensive Approach to Modeling and Managing Spatial Data

* RDBMS – Relational Database Management System

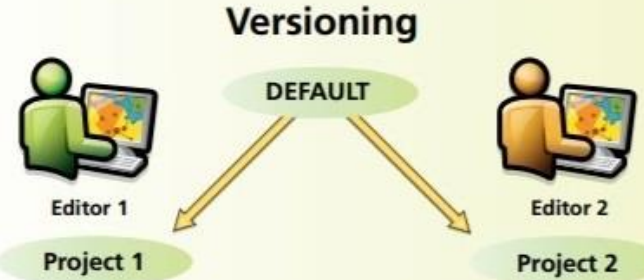


Types of Geodatabases



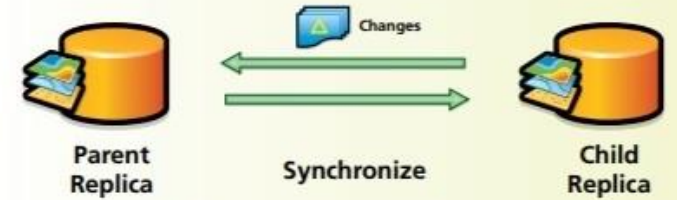
Choose per Workflow and Limitations!

Functionality



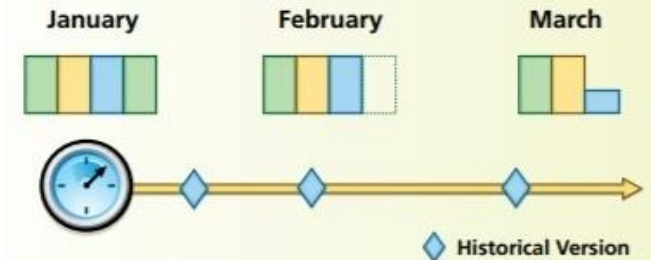
Versioning is the framework that enables multiple users to access and edit the same data simultaneously and provides long transaction (i.e., database changes that span long periods of time) support.

Geodatabase Replication



Enables GIS data to be shared across two or more geodatabases. Data changes can be made in each geodatabase, then synchronized. Two-way, one-way, and checkout/check-in replication workflows are supported.

Geodatabase Archiving



When enabled on a dataset, archiving captures any and all changes made to the dataset in the DEFAULT version of the multiuser geodatabase.

Key Factors

Why?



esri®

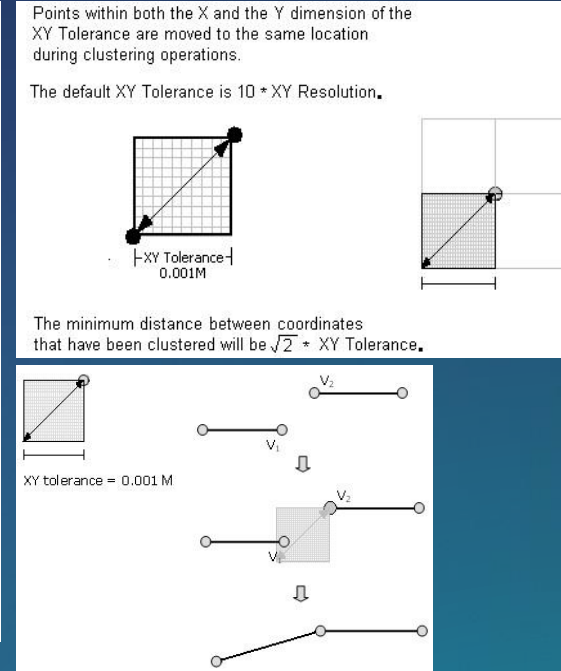
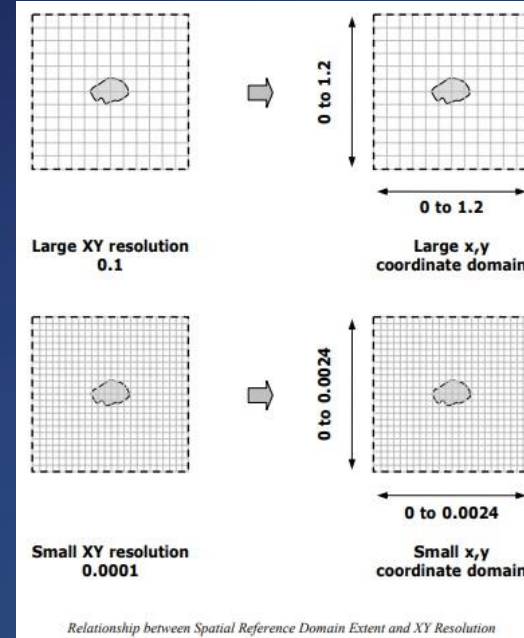
THE
SCIENCE
OF
WHERE™

Expensive Rework

A Stitch in Time Saves Nine!

- Data Model Designs
 - Change in XY Resolution / Tolerance
 - Multiple Projections
 - Objectid as Primary Key for Relationship Classes
 - etc.
- Data Conversion Specifications
 - Extra Vertices
 - Insufficient QA/QC
 - Large Feature Extent
 - Possibly by Data Conversion / Testing
 - Loading Static Raster Data Into GDB
 - etc.

Avoid Pitfalls!



Degraded Performance and Scalability

Risks



- **Workflow**

- Number of Outstanding Versions
- Versioning Levels
- Archiving
- etc.

- **Maintenance**

- Compact and Compress FGDB (File Geodatabase)
- Total Delta Table Records
- GDB Maintenance
- etc.

Versioning statistics

```
Number of versions: 18,055
Number of versions blocking DEFAULT: 5,966
Top 5 blocking versions...
GIS.Test_241625
RASU.WORK_ORDER1
ANDREW.ARCGIS_MONITOR1
ARCFM.SESSION_1254
ADMIN.Freeze_2018
Number of states: 20,139
Number of state lineages: 9,497,416
DEFAULT versions lineage length: 1,068
Last compress: JUN-18-2018
```

```
SQL > Exec DBMS_STATS.GATHER_SCHEMA_STATS ('ARCFM', estimate_percent=>100, DEGREE=> 15, CASCADE=>TRUE, No_Invalidate=>>false);
PL/SQL procedure successfully completed.

---ArcSDE A Table Rows
SQL> select sum(num_rows) as Total_A_Table_Rows from dba_tables where table_name in (select 'A'||registration_id from sde.table_registry);

TOTAL_A_TABLE_ROWS
-----
12,154,746

---ArcSDE D Table Rows
SQL> select sum(num_rows) as Total_D_Table_Rows from dba_tables where table_name in (select 'D'||registration_id from sde.table_registry);

TOTAL_D_TABLE_ROWS
-----
9,338,738
```

Separate Myths / Emotions from Evidence Based Facts!

Requirements and Workflows

Drives the Selection and Number of GDBs!

- Number of Users and Types of Users
- Workflows
 - Multi User Editing - *Enterprise / Workgroup GDBs*
 - Single User Editing – *FGDB*
 - Replication - *EGDB → FGDB / EGDB*
 - Read Only / Publication – *FGDB / EGDB*
 - etc.
- Generally more than one Geodatabase is required!

Enterprise Geodatabase Vs File Geodatabase		
Key Characteristics	Enterprise Geodatabase	File Geodatabase
Description	A collection of various types of GIS datasets held as tables in a relational database.	A collection of various types of GIS datasets held in a file system folder.
Number of users	Multiuser: many readers and many writers	Single editor and can support multiple readers.
Storage format	Oracle	Each dataset is a separate file on disk. All the datasets that belong to one geodatabase are contained in a single folder.
	Microsoft SQL Server	
	IBM DB2, Informix	
	PostgreSQL	
	SAP HANA	
Size limits	Size is controlled by the DBMS	By default, each dataset can grow to one TB. The 1 TB limit can be raised to 4 or 256 TB for extremely large image datasets. Each feature class can scale up to hundreds of millions of vector features per dataset.
Versioning support	Fully supported across all DBMSs Note:SAP HANA geodatabases do not support the traditional versioning type, but only Branch Versioning.	Does not support versioning workflows
Platforms	The DBMS are supported on multiple operating systems. Consult the system requirements for the full list.	Cross-platform.
Security and permissions	Managed by the DBMS.	Managed by the operating system.

Avoid Limitations!

Geodatabase Design



esri®

THE
SCIENCE
OF
WHERE™

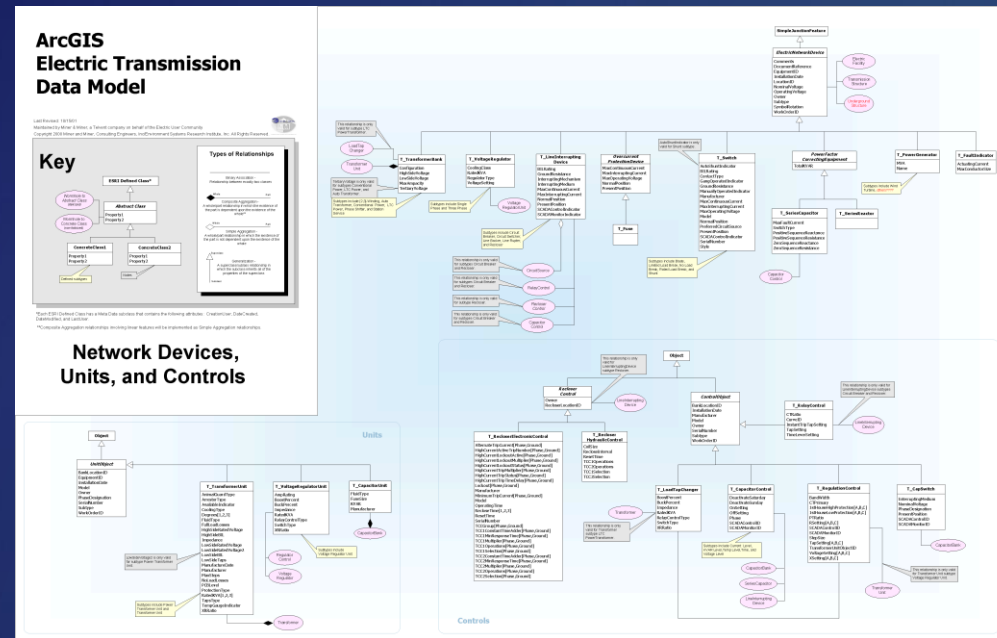
Geodatabase Design – Data Modeling

• Conceptual Design

- Identify Business Requirements
- Identify Thematic Layers
- Identify Required Applications
- Leverage Data Model Templates
- Document

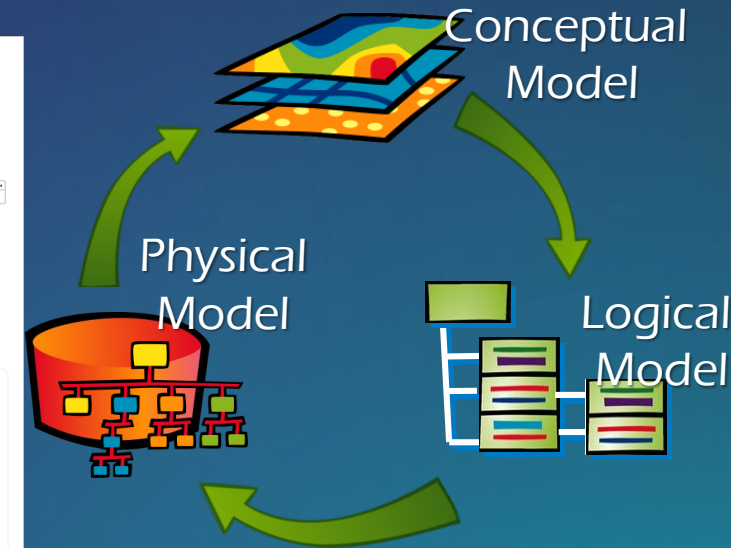
• Logical Design

- Define Tabular Database Structure
- Define Relationships
- Determine Spatial Properties
- Document



• Physical Design

- Create and Implement Model Design
- Generate Physical Schema in the RDBMS / FGDB
- Testing and Validation
- Document

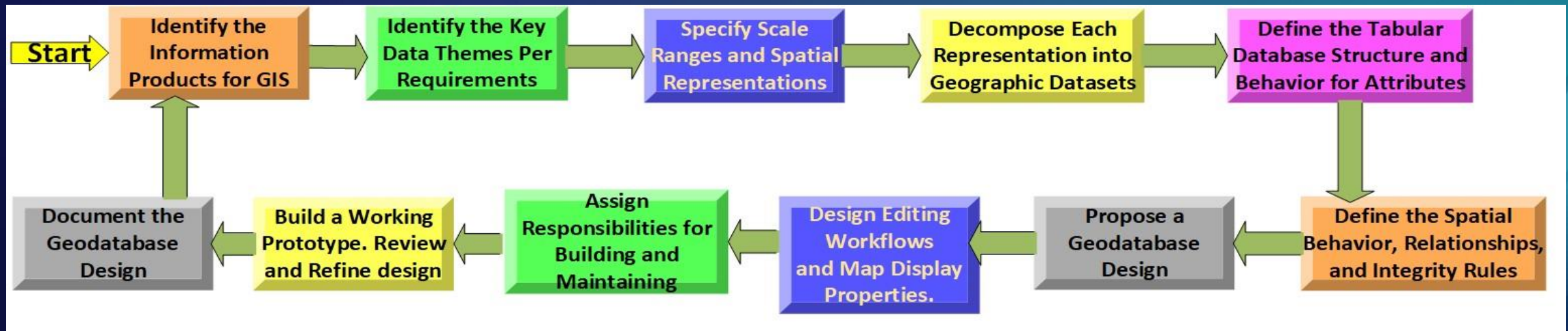
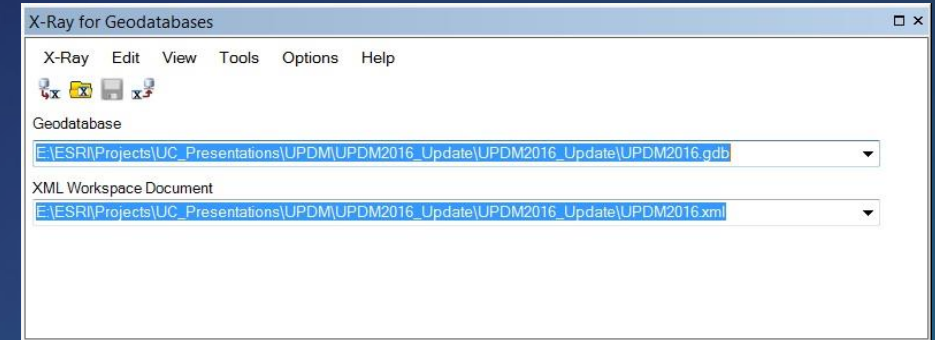


Geodatabase Design – Process

- Tools

- X-Ray Add-In
- Geodatabase Diagrammer
- Sparx Systems' Enterprise Architect
- Geometric Network Configuration Manager









- 11 Technical Steps!



Geodatabase Design – ArcGIS Data Models



Leverage hundreds of ready-to-use maps, apps, and platform configurations to meet your needs, delivering innovation and best practices for your organization.

- 
Local Government
- 
State Government
- 
Emergency
- 
Water
- 
Electric
- 
Gas
- 
Defense
- 
Telecommunications
- 
Parks and Gardens



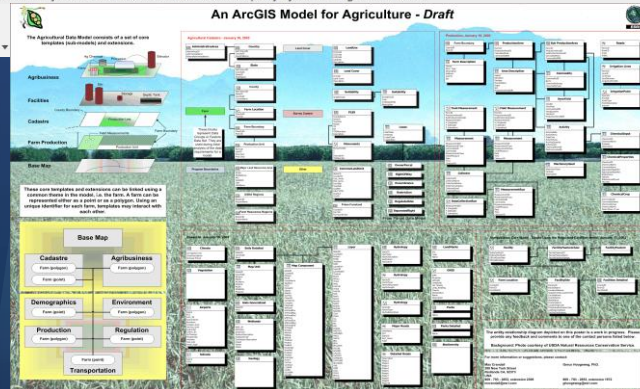
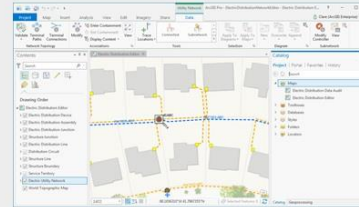
Overview

The Electric Utility Network Configuration models a utility network to behave like an electric network and provides maps, services, and workflows to facilitate data management and analysis. This configuration includes both electric distribution, electric transmission, and a combined electric distribution and transmission utility network.

The Electric Utility Network Configuration includes:

- Configuration instructions
- Schema for common electric components
- Rules to define network behavior
- Project for configuring and publishing the utility network services
- Maps and apps to support common electric utility workflows
- Sample data

The Electric Utility Network Configuration is a starting place to implement a utility network for electric. It is not designed to provide an all-encompassing solution for all workflows and assets managed by an electric utility. It is intended to be extended to meet any additional asset, workflow, or 3rd party system integration needs.



FAQ: Does Esri have industry-specific Data Models?

Question

Does Esri have industry-specific Data Models?

Answer

With the ArcGIS platform, the Esri vision is to build many industry-specific data models. Our basic goals are to simplify the process of implementing projects, and to promote and support standards that exist in our user communities.

Academic and industry leaders collaborate with Esri to create and design data model templates that can be used with one GIS platform. The result is this set of data model structures that can be implemented for each of the industries and scientific disciplines that Esri serves.

Read about the goals and process in the Introduction to ArcGIS Data Models, and visit the Data Model tips and tricks page designed as a resource for common methods and best practices.

Visit the Esri datamodel discussion forum to share your ideas, thoughts, and questions with other users.

Use this list to explore the industry areas. Some have available for download case studies, some have design templates, some have tools, and some have two or more of these available.

Note:
Data models available for download here are not applicable to all the product versions listed. Please note the relevant product and version listed for each data model.

Download Links

Below is a list of data models that are available for download.

Related Information

- Address
- Agriculture
- Atmospheric
- Basemap
- Biodiversity
- BroadbandStat
- Building Interior Space
- Carbon Footprint

Leverage the Existing Data Models!

Distribution

Leverage geographic information to maintain comprehensive electric utility records, plan and conduct inspections, communicate with customers, and more.

Transmission

Organize your network assets, map critical infrastructure, collect system data in the field, and more.

Generation

Create facility maps, analyze data, provide information to your entire organization, and more.

Land Records

Manage an accurate inventory of real property, describe who owns the land, analyze land value and associated improvements, and maximize property tax revenue.

Water Utilities

Maintain comprehensive water distribution, sewer, and stormwater records; and improve the operation of utility networks that provide clean drinking water.

Public Works

Maintain right-of-way assets; meet transportation and environmental quality needs; coordinate and plan capital projects and operate parks, and government facilities in a safe and effective way.

Recommendations



esri®

THE
SCIENCE
OF
WHERE™

Geodatabase Design – Key Considerations

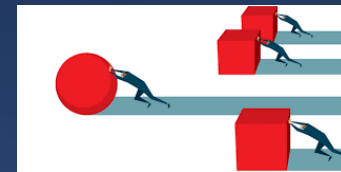
Poor Design = Slow Performance and Bugs



- Empty Feature Classes / Columns
- Missing / In-Correct Domains and Aliases for Fields
- Column / Domain Names and Field Lengths
 - >10 Characters in Field Names
 - Length of Text/NCLOB – 256 or 1,073,741,822
 - Choice of Field Type
 - Selection of Precision and Scale
 - Define Not Null Fields



Field Properties	
Alias	CustomerName
Allow NULL values	Yes
Default Value	
Domain	
Length	1073741822



Geodatabase Design – Best Practices

Poor Design = Slow Performance and Bugs

- **No Attributed Relationship Classes for Empty Tables!**
- **Use Many to Many Relationship Classes Only When Necessary**
- **Don't Use Objectid as Primary Key for Relationship Classes**
 - Unexpected Replication Behavior
 - Additional Processing During Synchronization

Relationship Class Properties

General Rules

Name: ELECDIST.SwiFac_DynProtectDev

Type: Simple

Cardinality: 1 - M

Notification: None (no messages propagated)

Origin Table/Feature Class

Name: ELECDIST.SwitchingFacility

Primary Key: OBJECTID

Foreign Key: StructureObjectID

Destination Table/Feature Class

Name: ELECDIST.DynamicProtectiveDevice

Primary Key: OBJECTID

Foreign Key: DeviceObjectID

Labels

Forward: Dynamic Protective Device

Backward: Switching Facility

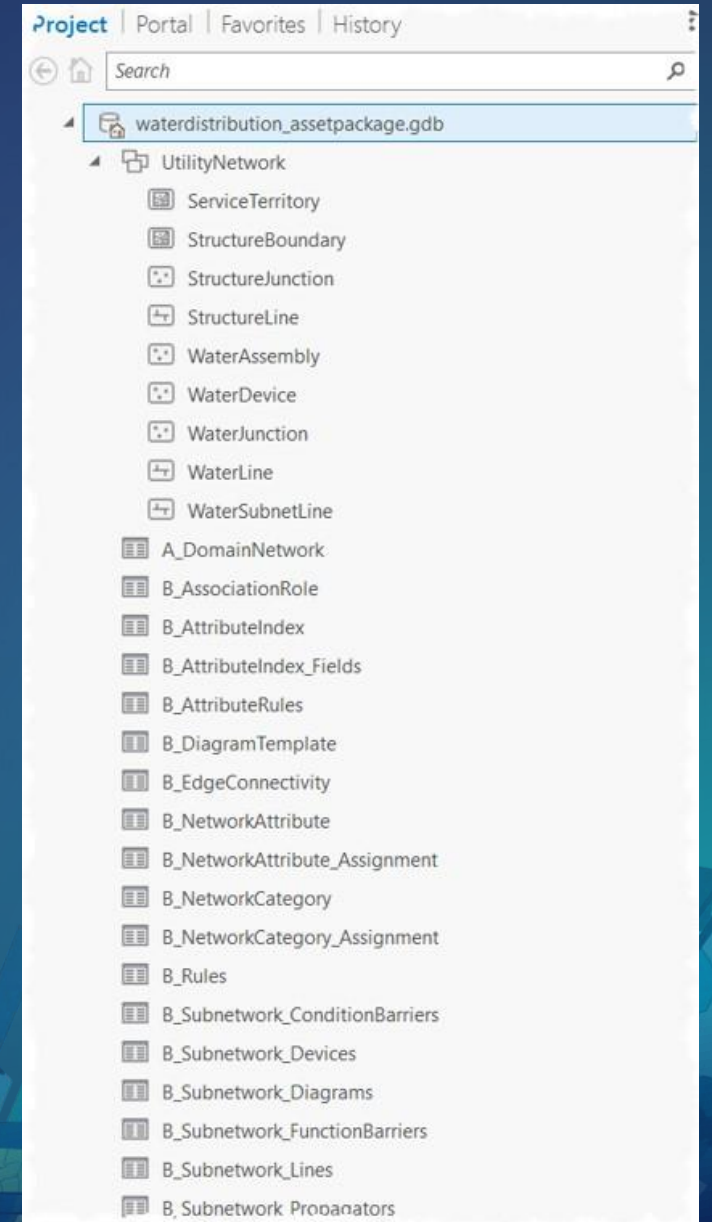
OK Cancel Apply

Geodatabase Design – Best Practices

Prevention is Better Than Cure!

- **Leverage the Existing ArcGIS Data Models**
 - Drop Redundant Feature Datasets / Classes, Columns, etc.
 - Stand Alone Feature Classes are Fine!
 - Possibly Split the Feature Classes Per Scale Levels
 - Less Complex and Attributed Relationship Classes
 - Test, Refine and Tune the Data Models
- **Integrate Related Feature Classes using Topology**
- **Deploy Necessary Information Models**
 - Geometric Network Vs Utility Network *(New!)*

Justify Every Single Geodatabase Element!



Geodatabase Design – Best Practices

Prevention is Better Than Cure!

- **Select Single Coordinate System**
 - On the Fly Projection is Expensive
 - Geometric Network Editing Does Not Support “On the Fly Projection”
- **Create Feature Datasets or Databases for Each LOB (Line of Businesses)**
 - Depends on Size, Access, Usage and Maintenance
 - E.g. Landbase, Gas, Electric, Water GDBs, etc.

QA / QC - Recommendations

Data Integrity and Validation!

- Design and Implement QA / QC Workflows

- Data Requirements for Software Functions
- Accurate Data for Business
- Maintain Data Integrity

- Tools

- Domains, Subtypes, Topology, etc.
- Attribute Assistant Add-In
- ArcGIS Data Reviewer
- ArcGIS Workflow Manager
- Business Partner Products
- Customization

Capture, Load and Maintain Data Accurately!

ArcGIS Data Reviewer Checks

The image displays a comprehensive set of 24 ArcGIS Data Reviewer Checks, organized into 8 categories. Each check card includes a title, a brief description of the check, and a small illustrative diagram.

- Database Validation Checks:** Includes checks for consistency (verifying generalization functions), domain (verifying nested value and range domains), subtypes (searching for records that are not in a subtype), and subtypes (searching for records that are not in a subtype).
- Table Checks:** Includes checks for consistency (verifying that a table is a table), integrity (verifying that a table is a table), table (verifying that a table is a table), and table (verifying that a table is a table).
- Spatial Parameter Evaluation Checks:** Includes checks for consistency (verifying that a table is a table), table (verifying that a table is a table), and table (verifying that a table is a table).
- Default Checks:** Includes checks for consistency (verifying that a table is a table), table (verifying that a table is a table), and table (verifying that a table is a table).
- Topology Checks:** Includes checks for consistency (verifying that a table is a table), table (verifying that a table is a table), and table (verifying that a table is a table).
- Advanced Checks:** Includes checks for consistency (verifying that a table is a table), table (verifying that a table is a table), and table (verifying that a table is a table).
- Polygon Checks:** Includes checks for consistency (verifying that a table is a table), table (verifying that a table is a table), and table (verifying that a table is a table).
- Z-Value Checks:** Includes checks for consistency (verifying that a table is a table), table (verifying that a table is a table), and table (verifying that a table is a table).
- Feature on Feature Checks:** Includes checks for consistency (verifying that a table is a table), table (verifying that a table is a table), and table (verifying that a table is a table).
- Polyline Checks:** Includes checks for consistency (verifying that a table is a table), table (verifying that a table is a table), and table (verifying that a table is a table).
- Duplicate Geometry Checks:** Includes checks for consistency (verifying that a table is a table), table (verifying that a table is a table), and table (verifying that a table is a table).

Questions and Answers

Contact Info:

Rasu Muthurakku

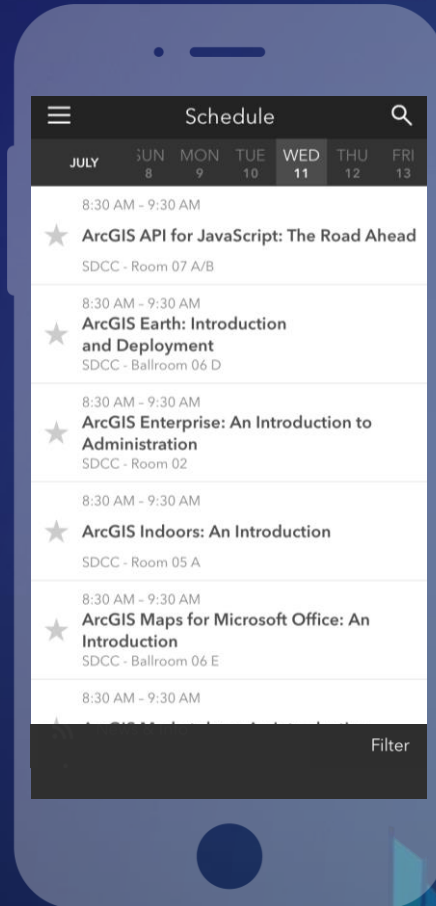
rasu@esri.com

Please Take Our Survey on the App

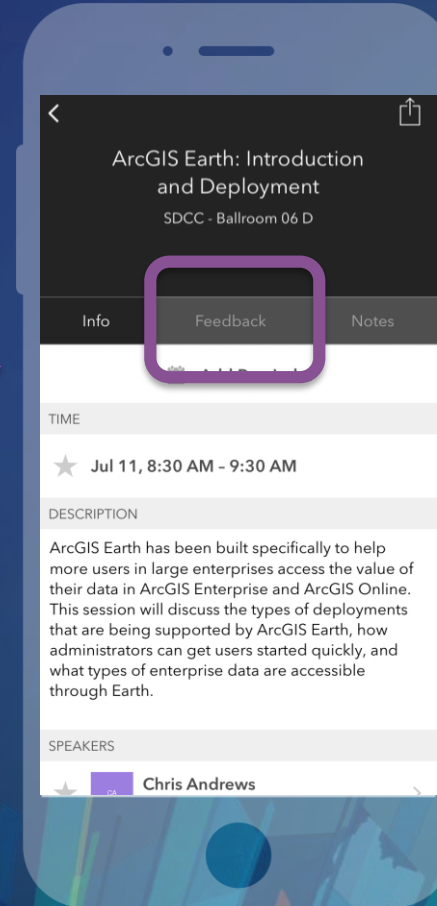
Download the Esri Events app and find your event



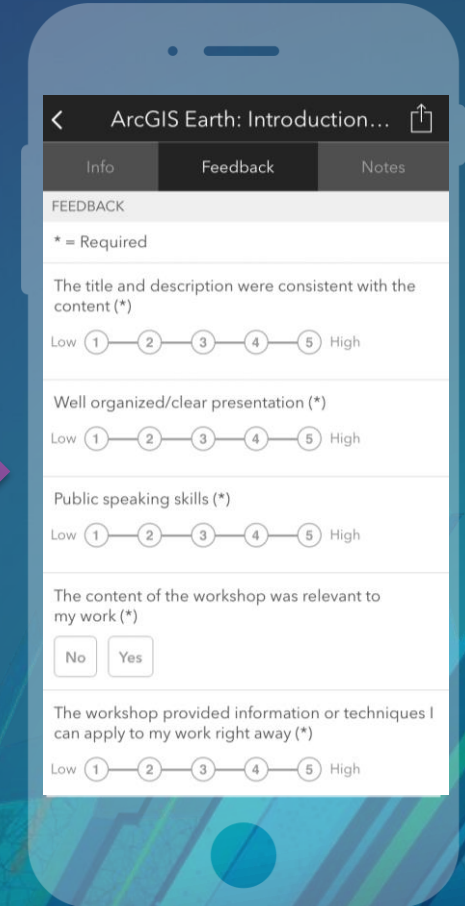
Select the session you attended



Scroll down to find the feedback section



Complete answers and select "Submit"



See Us Here

WORKSHOP	LOCATION	TIME FRAME
<ul style="list-style-type: none">• Esri Best Practices: Implementing an Enterprise Geodatabase	<ul style="list-style-type: none">• SDCC - Room 17 B	<ul style="list-style-type: none">• Wednesday 7/11/2018 08:30 AM - 09:30 AM
<ul style="list-style-type: none">• Esri Best Practices: Implementing an Enterprise Geodatabase	<ul style="list-style-type: none">• SDCC - Room 05 B	<ul style="list-style-type: none">• Thursday 7/12/2018 10:00 AM - 11:00 AM

Thanks!



esri

THE
SCIENCE
OF
WHERE