



ArcGIS Data Reviewer: An Introduction

Chandan Banerjee

Jay Cary

Workshop Agenda

- **Importance of Data Quality**
- **What is ArcGIS Data Reviewer?**
- **Automated review**
- **Semi-automated review**
- **Managing errors and reporting data quality**
- **Summary/Resources**

Importance of data quality

Defining quality

A business perspective

- **Executive**

- Confidently make decisions
- Reduce financial risk
- Optimize organizational performance

- **Manager**

- Effective data stewardship
- Drive increased usage
- Maximize productivity

- **Knowledge worker**

- Increased efficiencies
- Confidence in GIS



Defining Quality

A Technical Perspective



ISO-19157:2013 Geographic information – Data quality

Data quality management

Capabilities of the ArcGIS Platform

Geodatabase integrity

- Schema constraints
- Geoprocessing tools
- Data load checks
- Versioning

Advanced data types

- Topologies
- Geometric networks

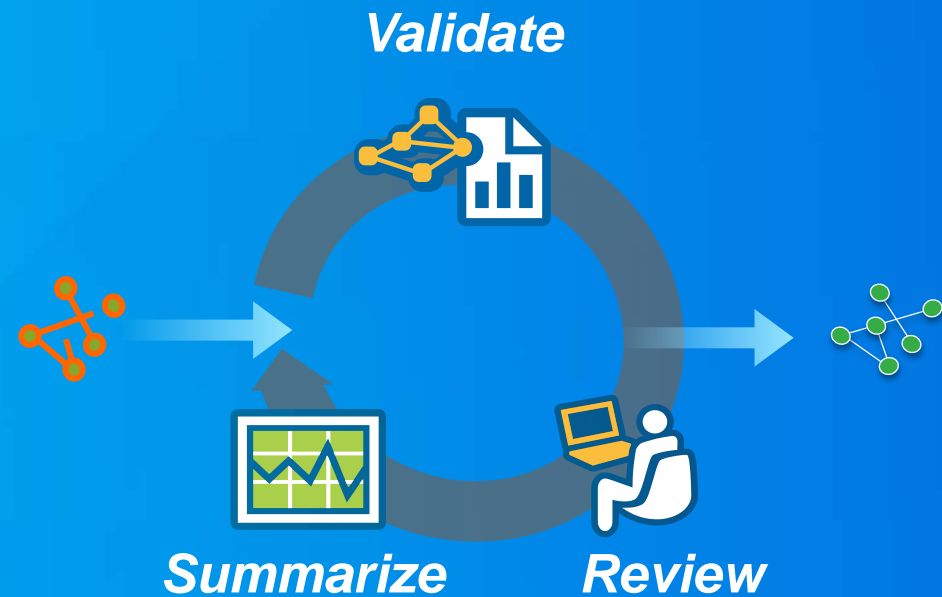
ArcGIS Data Reviewer

- Automated review
- Semi-automated review
- Error management
- Quality reporting

What is ArcGIS Data Reviewer?

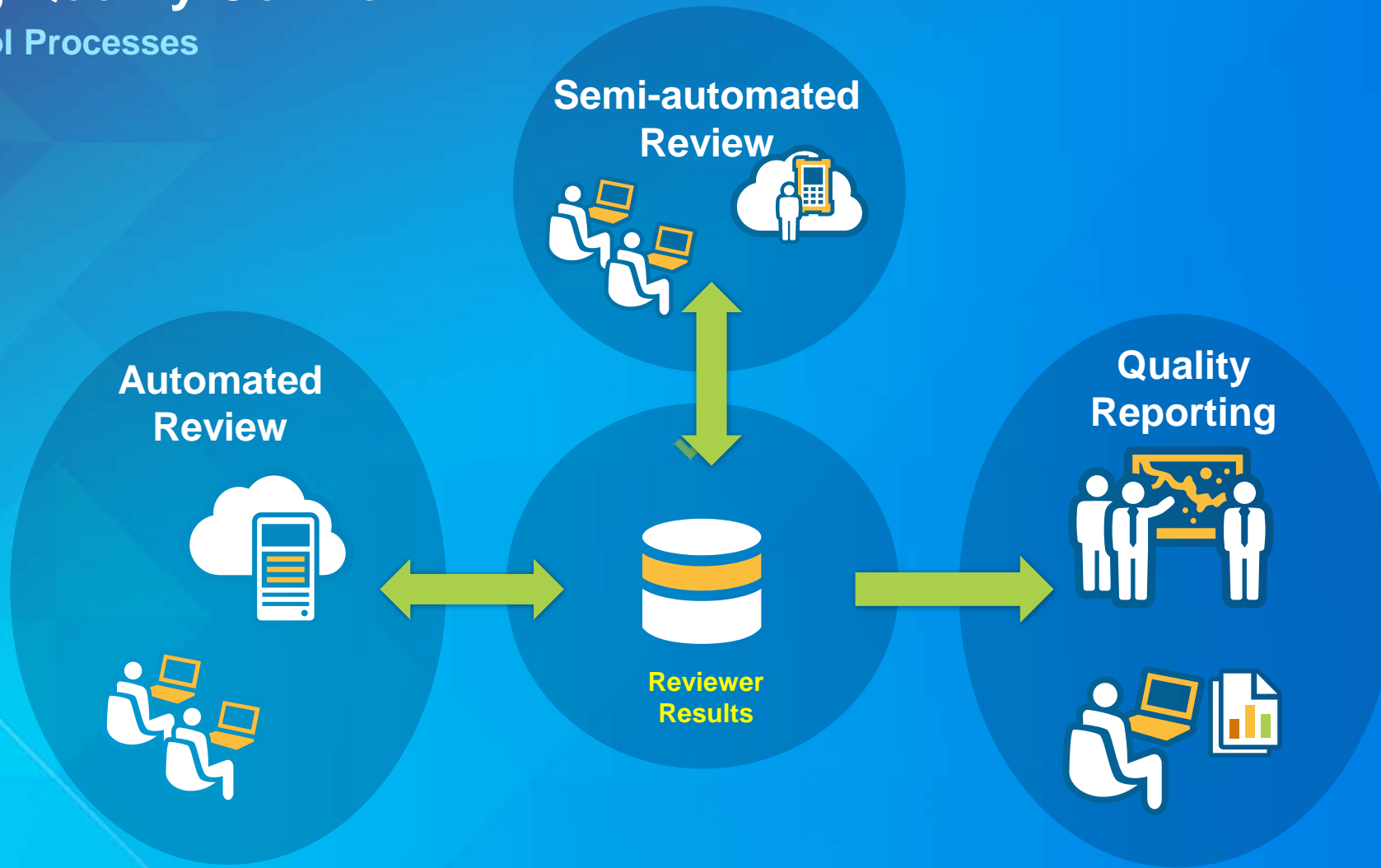
Data Quality Management for ArcGIS

- **Provides**
 - Rule-based validation
 - Interactive tools
 - Track errors
- **For individuals and enterprise**
 - Saves time/money
 - Less rework
- **For multiple domains**
 - Configurable
 - Extendable



Managing Quality Control

Quality Control Processes



What is ArcGIS Data Reviewer?

Data quality management in the ArcGIS platform

- **Data Reviewer for Desktop**

- ArcMap
- ArcGIS Pro

- **Data Reviewer for Server**

- Standard or higher

- **Data Reviewer API for**

- FLEX
- JavaScript



Automated Review

Types of quality control



Automated review

Fast

Consistent and repeatable

Objective

100% coverage

Defining Quality

Sources of data quality requirements

Industry standards / Specifications



Subject matter experts



Training and experience



Quality assurance plans



Automating Data Validation

- Implementing quality requirements
 - 40+ configurable checks
 - Attribute
 - Feature and table values
 - Spatial
 - Spatial relationships
 - Feature integrity
 - Collection rules
 - Metadata
 - Completeness/Content

Database Validation Checks

- Connectivity Rules**: Returns geometries for features that violate the geometric network connectivity rules.
- Domain**: Validates coded value and range domains to ensure that all values meet domain constraints.
- Relationships**: Searches for records that are orphans or have improper cardinality in a relationship class.
- Subtype**: Searches for feature classes with improper or null subtypes.

Table Checks

- Execute SQL**: Finds features based on a SQL query WHERE clause.
- Regular Expression**: Finds features with attribute values that violate the regular expression.
- Table to Table Attribute**: Returns rows whose attributes match those of a feature class or table and/or comply with a user-defined WHERE clause comparing the attributes between feature classes and/or tables.
- Unique ID**: Checks the values of a set of fields across a set of tables and feature classes for uniqueness within a given workspace.

Default Checks

- Invalid Geometry**: Finds features whose geometry is empty, nothing, or not simple, as well as features with empty envelopes.

Topology Checks

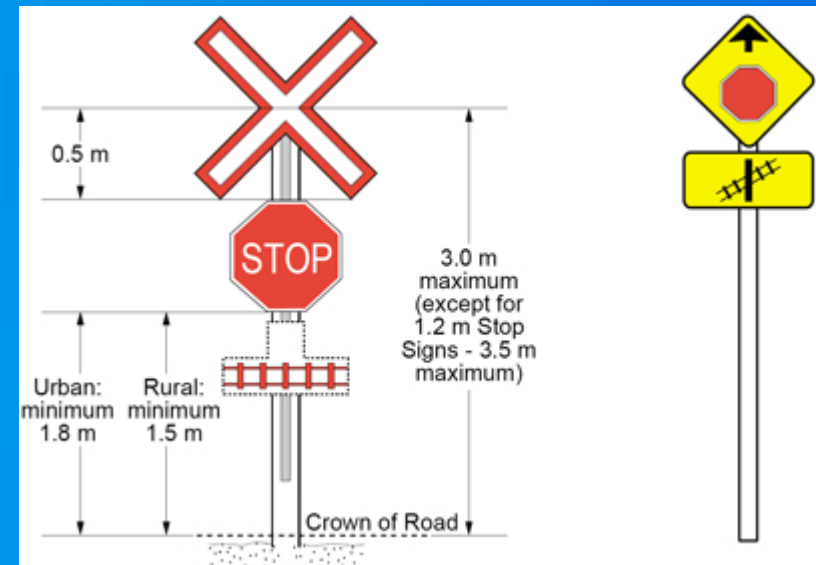
- Find Dangles**: Within a database topology, finds polyline features that have nodes that are within a tolerance but not connected to other features in the database topology. Comparison to Topology: Line must not have dangles.

<http://j.mp/DRCheckPoster>

Demo: Getting Started with Automated Review

Demo Scenario

- Department plans to use field staff with field data collection devices to map road-related features (signs, street furniture).
- **Goals**
 - Features should be positionally accurate based on siting criteria
 - Key attributes are populated and have the correct values
 - All features should be collected within the survey area



Authoring Data Quality Rules

Batch Validation

Implementing Cumulative Review

- **Reviewer Batch Jobs**

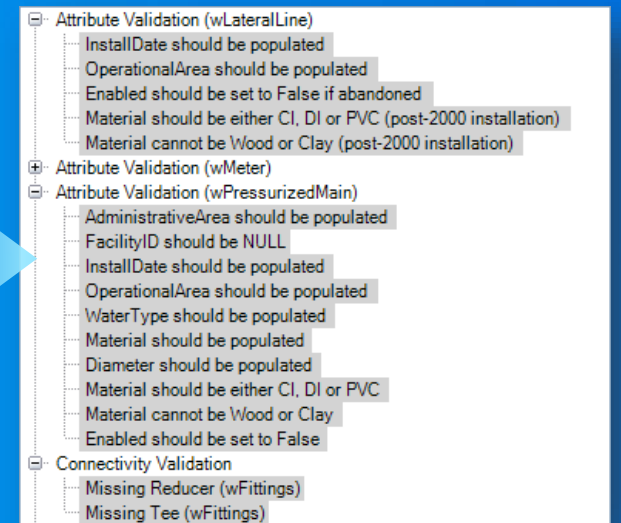
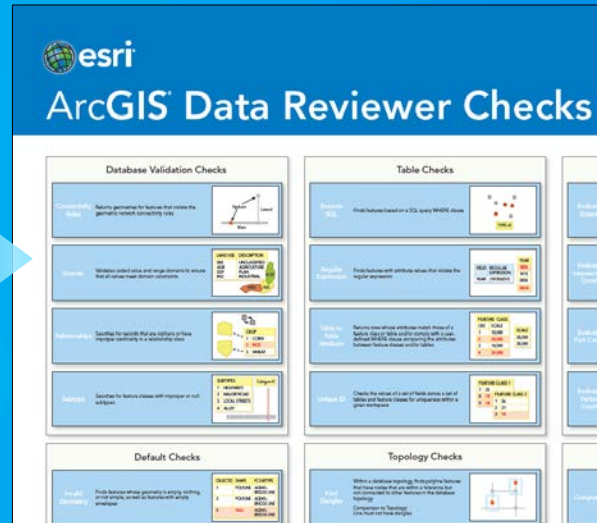
- Define and encapsulate quality requirements
- Captures domain knowledge
- Designed once and executed many times
 - ArcGIS for Desktop, ArcGIS for Server

Industry standards / Specifications

Subject matter experts

Training and experience

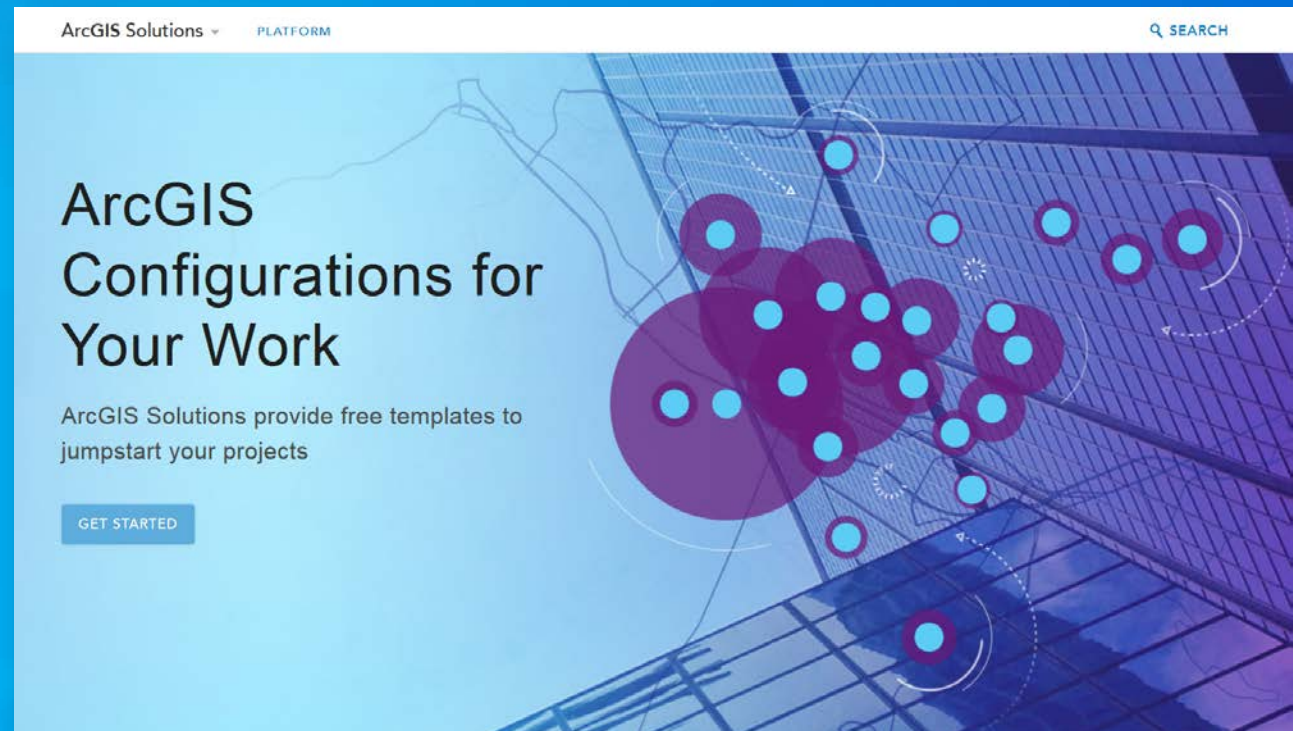
Quality assurance plans



Resources

Esri Solutions

- **Industry templates**
 - Address Management
 - Basemaps
 - Electrical Utilities
 - Gas Utilities
 - Roads & Highways
 - Tax Parcel Editing
 - Water Utilities
- **Based on Esri industry models**
- **Use as Starting point**







solutions.arcgis.com

Demo: Authoring Data Quality Rules

Authoring Batch Jobs

Tips and Tricks

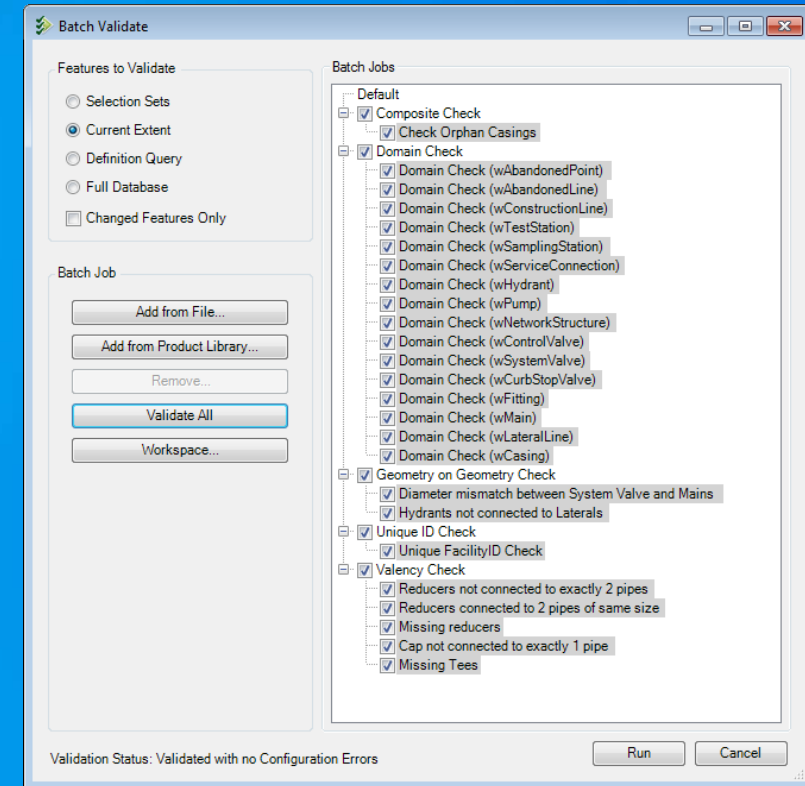
-  Create sample data for testing
-  Configure and execute from Reviewer Toolbar
-  Leverage Mini-Browser to assess results
-  Validate a pilot area with known issues

Executing Automated Validation

Methods for executing data validation

Execute data validation using

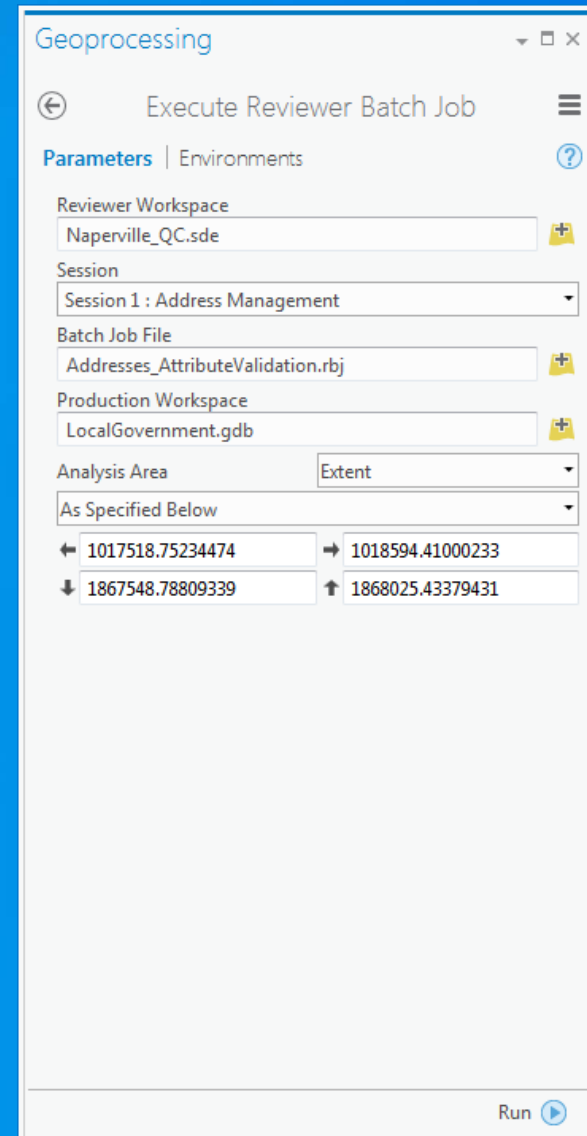
- ArcMap



Methods for executing data validation

Execute data validation using

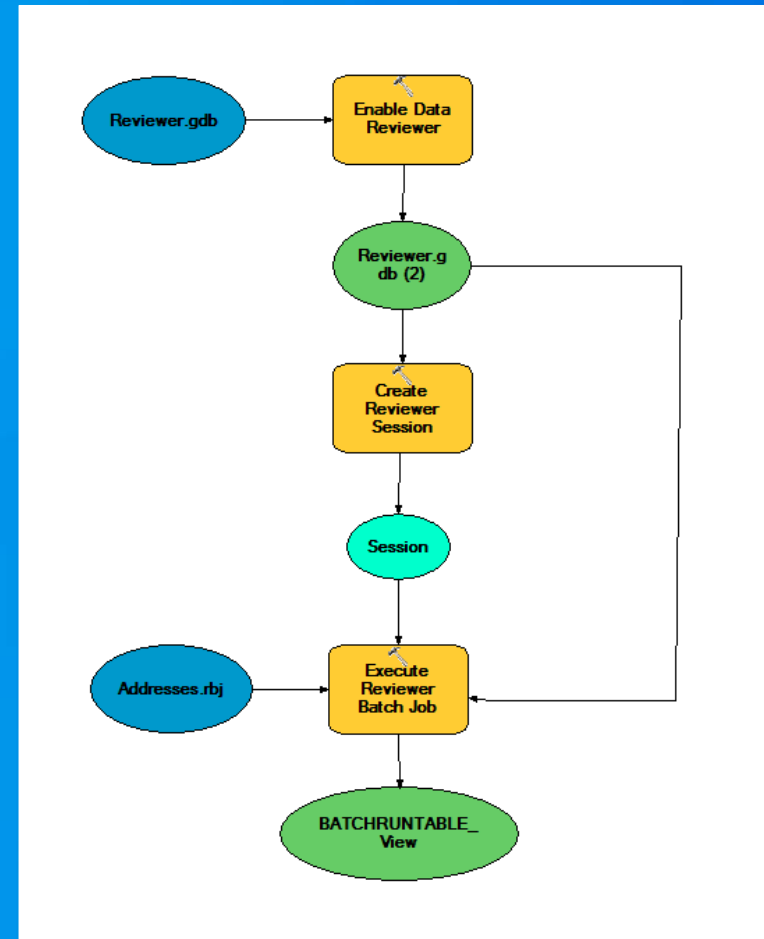
- ArcMap
- ArcGIS Pro



Methods for executing data validation

Execute data validation using

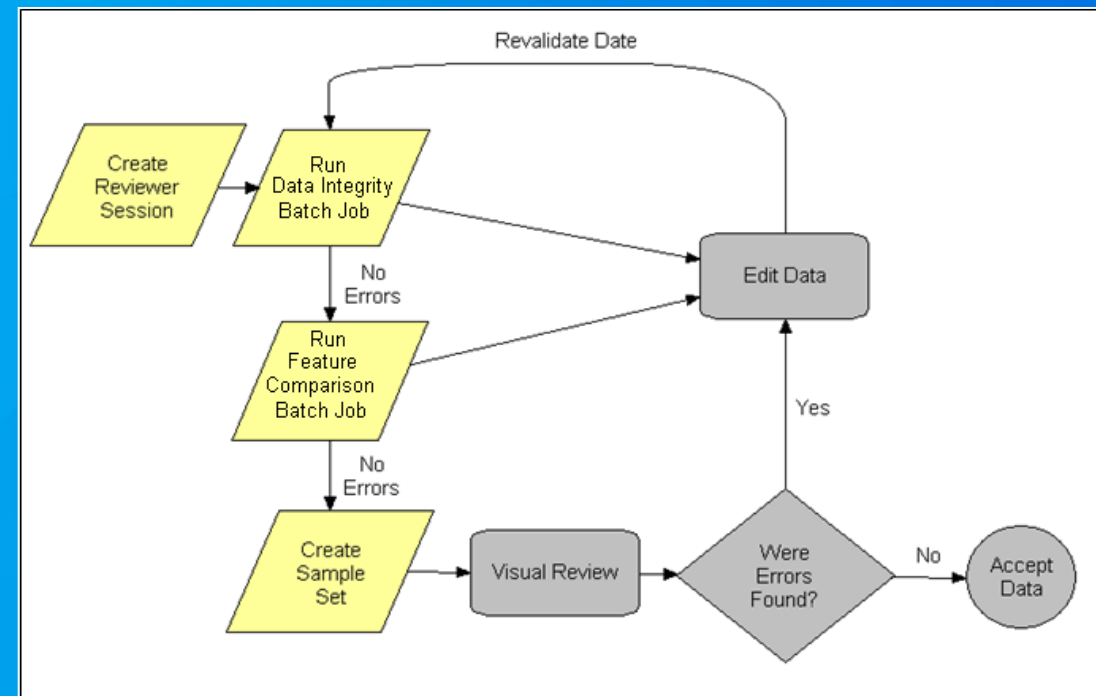
- ArcMap
- ArcGIS Pro
- Model/Python script



Methods for executing data validation

Execute data validation using

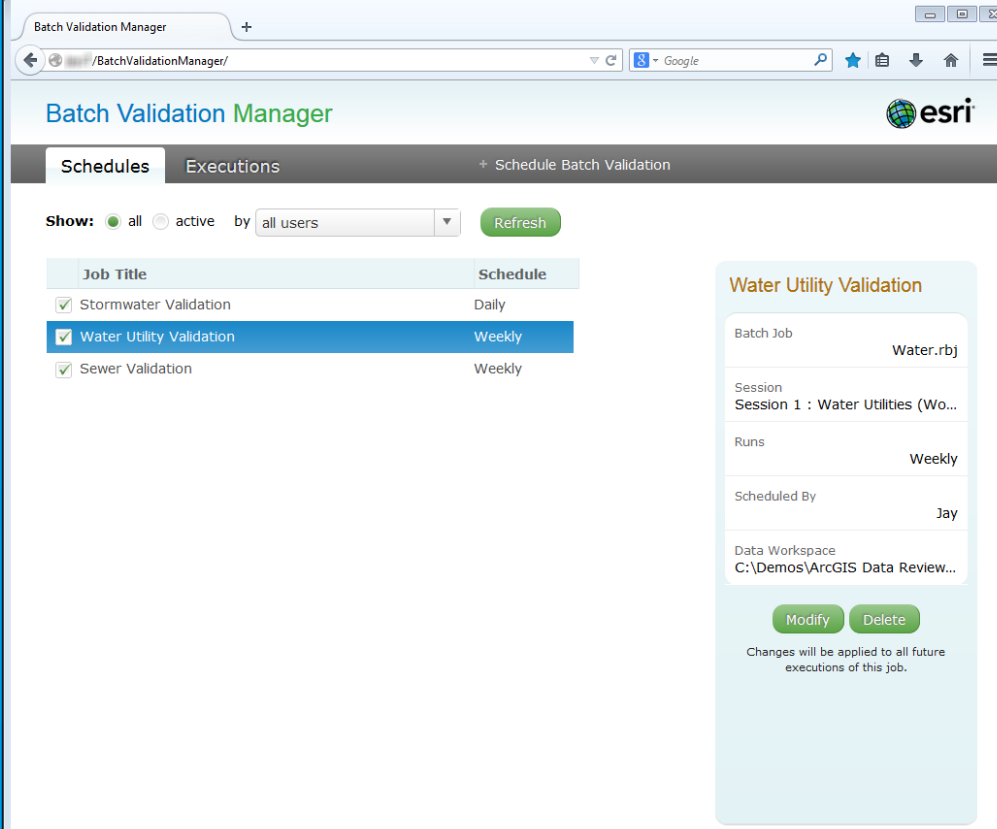
- ArcMap
- ArcGIS Pro
- Model/Python script
- ArcGIS Workflow Manager



Methods for executing data validation

Execute data validation using

- ArcMap
- ArcGIS Pro
- Model/Python script
- ArcGIS Workflow Manager
- ArcGIS for Server



The screenshot displays the Batch Validation Manager web interface. The browser address bar shows the URL `/BatchValidationManager/`. The page title is "Batch Validation Manager" and the Esri logo is visible in the top right corner. The interface has two tabs: "Schedules" (selected) and "Executions". A "+ Schedule Batch Validation" link is present in the top right of the main content area.

Below the tabs, there is a "Show:" section with radio buttons for "all" (selected) and "active", and a "by" dropdown menu set to "all users". A green "Refresh" button is located to the right of the dropdown.

The main content area contains a table with the following data:

Job Title	Schedule
<input checked="" type="checkbox"/> Stormwater Validation	Daily
<input checked="" type="checkbox"/> Water Utility Validation	Weekly
<input checked="" type="checkbox"/> Sewer Validation	Weekly

To the right of the table is a detailed view for the "Water Utility Validation" job. It includes the following information:

- Batch Job: Water.rbj
- Session: Session 1 : Water Utilities (Wo...)
- Runs: Weekly
- Scheduled By: Jay
- Data Workspace: C:\Demos\ArcGIS Data Review...

At the bottom of this view are "Modify" and "Delete" buttons. A note below the buttons states: "Changes will be applied to all future executions of this job."

Demo: Executing Automated Validation

Batch validation

Events when validation may be required

- ✓ Before data loading
- ✓ During editing
- ✓ Prior to sharing/publishing maps
- ✓ Prior to rolling-out new applications

Semi-Automated Review

Types of quality control



Automated review

Fast

Consistent and repeatable

Objective

100% coverage



Semi-automated review

Guided Workflows

Streamlined processes

Subjective

Sampling

Semi-automated review methods

Visual review

- Redlining
- Systematic review
- Sampling

Data inspection

- Feature counts
- Attribute review
- Version differences

Positional accuracy

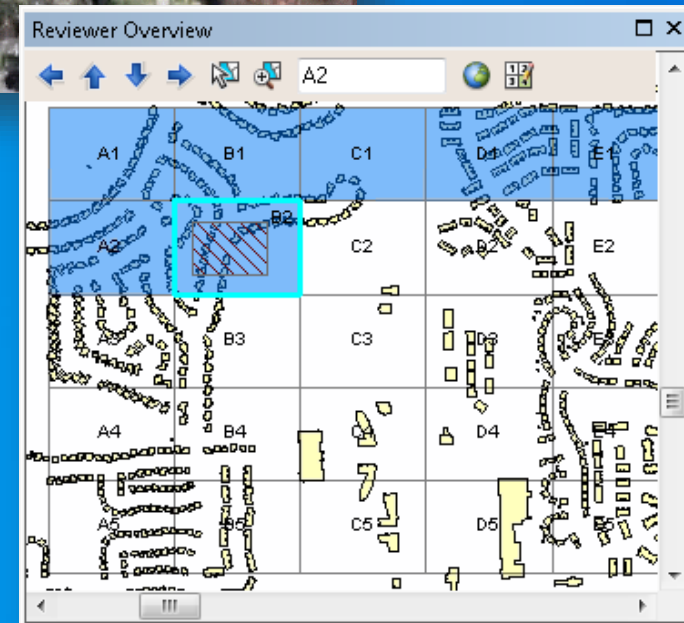
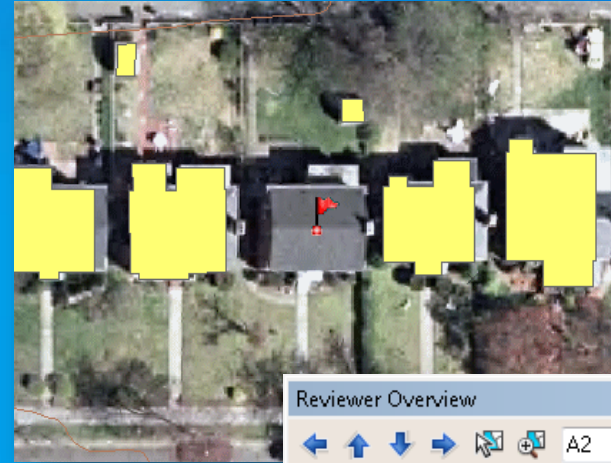
- Assessment tool

Semi-automated review

Leveraging ArcGIS for Desktop

Tools supporting

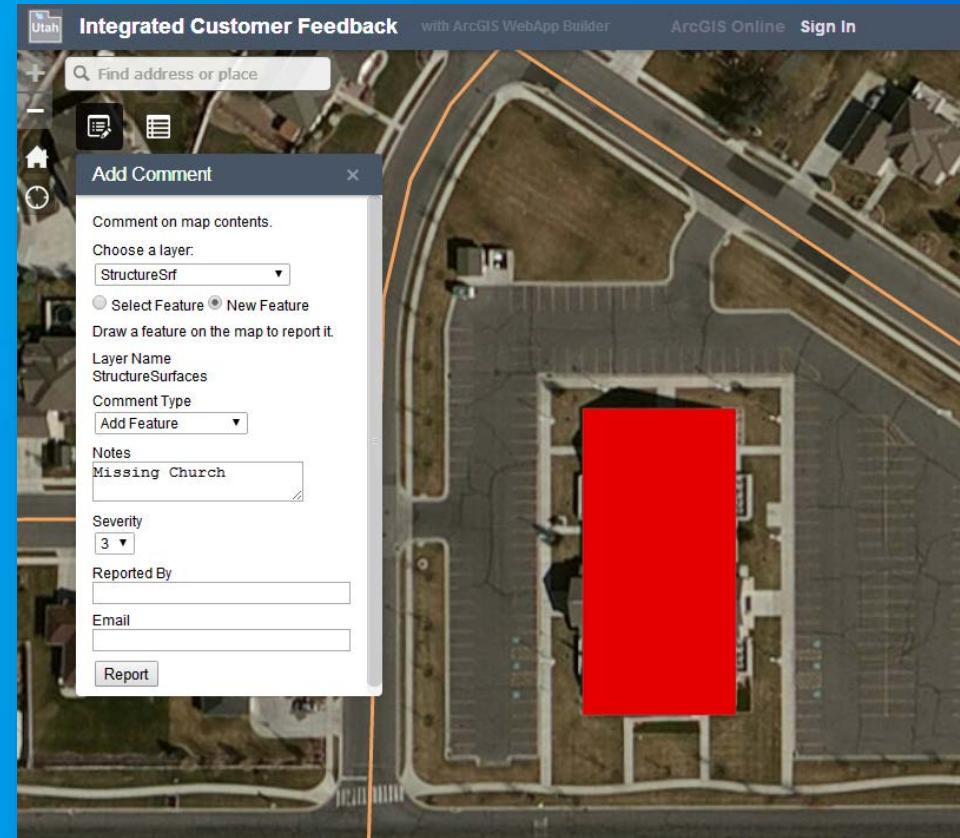
- Selecting/browsing features
- Redlining missing features
- Flagging existing features in error
- Generating random samples
- Assessing positional accuracy
- Comparing geodatabase versions



Semi-automated review

Leveraging ArcGIS for Server

- Extending quality control workflows into other communities
 - QC review across ArcGIS platform
 - Simple to use tools for error identification
 - Manual QC workflow “automation”



Demo: Visual Data Review

Reporting Quality

Data Quality Reporting

ArcGIS for Desktop

- Automated reporting of quality control results
- Adhoc result summary tools
- Pre-configured reports
 - Automated Check (Origin Table, Subtype, Check Group)
 - Total Record Count
 - Sampling

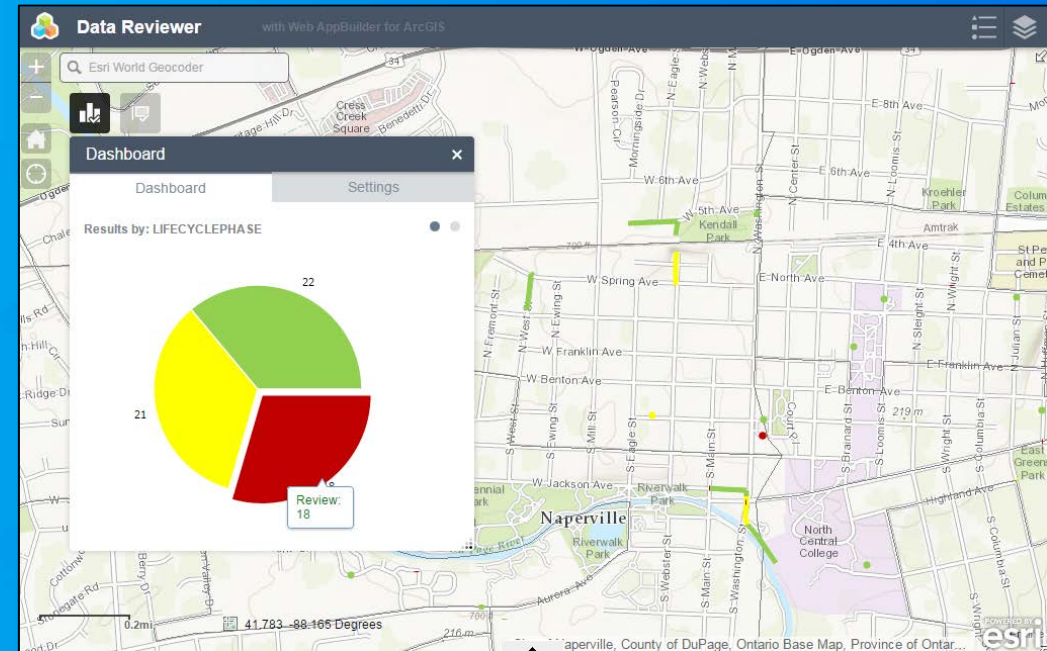
Automated Check Report By Origin Table											
Report generated on 7/20/2015 8:45 PM by xxxxxx											Reviewe
Workspace Location: C:\Demos\ArcGIS Data Reviewer\DataReviewerAddressesfor10.2\MapsandGeodatabase\QC_Results.gdb											
Session(s): Session 2 : Address Validation (4/21)											
Time	Origin Table	Check Type	Check Title	Severity	Records Validated	Total Results	Total Percent Accuracy	Distinct Results	Distinct Percent Accuracy	Batch Job Name	Date/T
	RouteValidation.rbj	986	81	91.78%	10	98.99%	C:\Demos\ArcGIS Data Reviewer\DataReviewerAddressesfor10.2\ApplicationAddresses_Attrib				
		2	0	100.00%	0	100.00%	4/21/2015 5:30:52 PM				
							AddressEntrancePoint				
							Domain Check				
	Address Point Must	3	2	0	100.00%		Address Entrance Point				
							Have valid Domains				
	100.00%						AddressPoint		50	0	100.00%
							Domain Check				0
							Address Point Must Have Valid Domains		3	50	0
	100.00%						FacilitySite		1	0	100.00%
							FacilitySitePoint				0
	Domain Check						Facility Site Must Have Valid Domains		3	1	0
											100.00%
											0
	100.00%										100.00%
	Domain Check						Facility Site Point Must Have				0

Data Quality Reporting

ArcGIS for Server

Enabling transparency in data quality

- Better decision making by communicating data quality across stakeholders
- Open quality reporting
- Shared across ArcGIS system
- Tools and methods to communicate quality



Demo Scenario

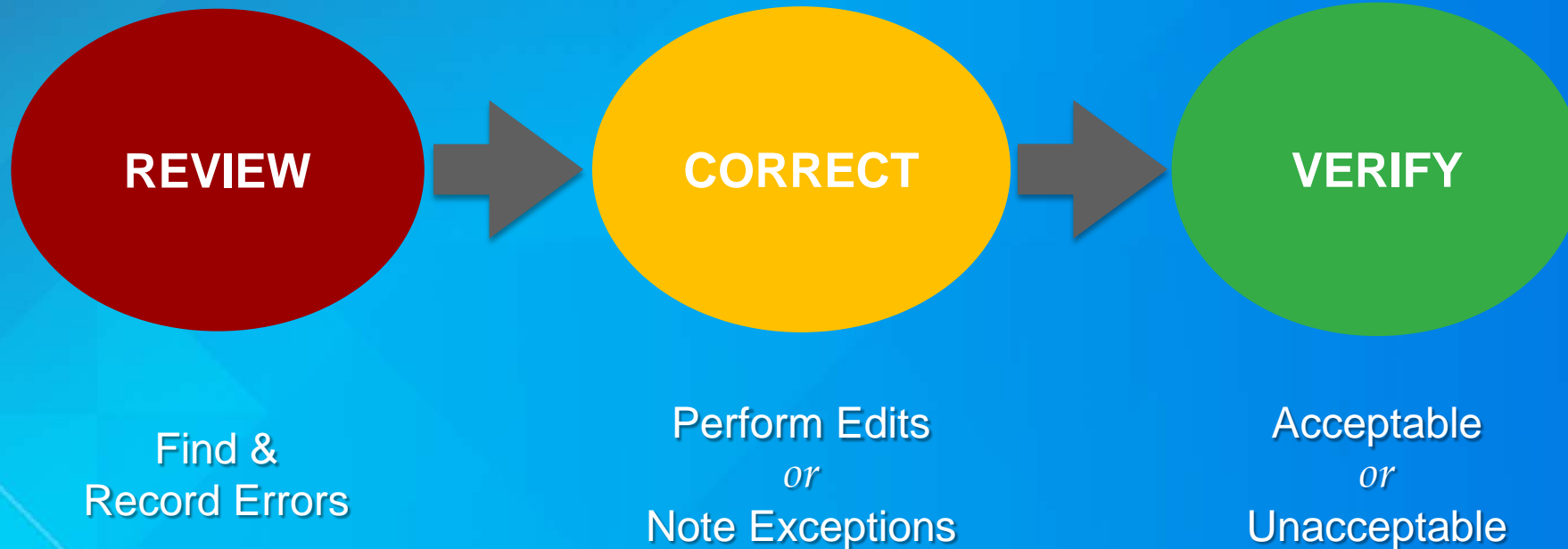
- **As a GIS manager, I need to manage the quality of the GIS database to ensure that it can support existing and new staff operations.**
- **Business Requirements**
 - Features should be positionally accurate based on siting criteria (95% accuracy)
 - Key attributes are populated and have the correct values (95% accuracy)

Demo: Reporting Quality

Managing Quality Results

Managing Quality Control

QC lifecycle management



Demo:

Managing Quality Results

Workshop Review

- Importance of data quality
- Forms of data quality control
 - Automated review
 - Semi-automated review
- ArcGIS Data Reviewer
 - Automated validation
 - Semi-automated tools
 - Error lifecycle management
 - Reporting

The image displays the ArcGIS Data Reviewer software interface. At the top, the Esri logo and the title "ArcGIS Data Reviewer Checks" are visible. Below the title, there are several panels for different types of checks:

- Database Validation Checks:** Includes checks for nullable parameters, null values, and table relationships.
- Table Checks:** Includes checks for table relationships, table names, and table statistics.
- Spatial Parameter Evaluation Checks:** Includes checks for spatial parameters, spatial relationships, and spatial statistics.
- Default Checks:** Includes checks for default values and default relationships.
- Topology Checks:** Includes checks for topology rules and topology relationships.
- Advanced Checks:** Includes checks for advanced rules and advanced relationships.

The main window shows a map view of a street grid with various check results overlaid. A table at the bottom right displays the results of the checks, including columns for Status, ID, Name, Check Type, and Severity. The table shows several errors related to address management, such as "Invalid Address Point" and "Invalid Address Point Type".

Want to learn more?

- **Documentation**

- [Resource Center](#)

- **Training**

- [Assessing Data Quality using ArcGIS Data Reviewer \(Seminar\)](#)
- [Data QC with ArcGIS: Automating Validation \(Web Course\)](#)
- [Data QC with ArcGIS: Visual Review \(Web Course\)](#)
- [Quality Control Using ArcGIS Data Reviewer for Desktop \(Instructor Led\)](#)

- **Questions/Comments**

- GeoNet ([Data Reviewer](#) place)
- Email (datareviewer@esri.com)





Understanding our world.