ArcGIS Data Reviewer
An Introduction
Jay Cary & Michelle Johnson
Workshop Overview

- Importance of Data Quality

- What is ArcGIS Data Reviewer
  - Automated Data Review
  - Batch Data Review
  - Visual Data Review

- Managing Errors and Reporting Data Quality

- Summary/Resources
Ensuring Data Quality with ArcGIS Data Reviewer
Data Quality Matters

• Cost of poor quality can be extreme
  - $$$ to fix
  - Impacts customers
  - Reputation

• Consider
  - Do you publish a doc without spell checking?
  - Do you operate a PC without a virus scanner?
Defining Quality

- Business Perspective
  - Intended purpose and use
  - Client requirements
  - Technical specifications
  - Industry standards

- Technical Perspective
  - Positional Accuracy
  - Attribute Accuracy
  - Completeness
  - Logical Consistency
  - Lineage
Defining Quality Control

• Quality Assurance Plan
  - Defines desired level of data quality
  - Specifies acceptable and unacceptable levels of error
  - Identifies how quality control will be performed

• Quality Control
  - The tasks required to ensure data quality
  - Methods to assess and evaluate data
What is ArcGIS Data Reviewer?
Data quality management for ArcGIS

• Provides
  - Rule based workflows
  - Interactive tools
  - Track errors

• For individuals and enterprise
  - Saves time/money
  - Less rework

• Standard extension to ArcGIS
  - Desktop extension since 8.0
  - Server extension at 10.1

Trusted data through improved quality management
Managing Quality Control
Scalable framework for managing QC lifecycle

1. REVIEW
   - Find & Record Errors
2. CORRECT
   - Perform Edits or Note Exceptions
3. VERIFY
   - Acceptable or Unacceptable
An Introduction to Data Reviewer Components

The Quality Control Process with Data Reviewer

- Heart of the Data Reviewer
  - Tracks all anomalies/errors from review
- Review Data & Identify Errors
  - Interactive table
- Automated Checks
  - Similar to an edit session
  - Tracks all anomalies/errors
- Reviewer Session
  - Allows data interaction for QC purposes
- Reviewer Workspace
  - Stores spatial & attribute information about errors
  - Personal, File, or Enterprise Geodatabase
  - Feature class & tables created automatically
- Report Quality Information
  - Summarize quality control results
  - Interpret data error
  - Report data accuracy
Ensuring Data Quality with ArcGIS Data Reviewer
Performing Quality Control
Automated Data Review
Types of Quality Control

Automated Quality Control
- Fast
- Consistent and repeatable
- Objective
- 100% coverage

Visual Review
- As subjective as needed
- Better for finding patterns and missing elements
Understanding Data Reviewer Checks

- Over 40 data checks
- Grouped into 11 categories
- Configurable
- Additive/Scriptable

www.esri.com/datareviewer
Table Checks

Execute SQL Check

Mains installed after 1 January 2000 should only be Cast Iron, Ductile Iron, or PVC
Feature on Feature Checks

Geometry on Geometry Check

Geometry on Geometry

- Searches for features from two different feature classes or within the same feature class that spatially interact (e.g., intersect) or are within a tolerance of each other.

  Comparison to Topology:
  - Line: Must not overlap, intersect, or overlap with.
  - Polygon: Must not overlap, contain point, or overlap with.

Intersection on Geometry

- Returns geometries for features in Feature Class 1 that intersect with the intersections from features from Feature Class 2 and 3.

Polygon Overlap/Gap

- Returns overlap/gap geometries between polygon features from two feature classes that have a thinness ratio beneath a user-specified threshold; optionally requires that the overlap/gap polygons be beneath a maximum area threshold.

  Comparison to Topology:
  - Polygon must not have gaps.
Combine Multiple Rules for Batch Check

- Encapsulate workflow / process
- Designed once and executed many times
- Complete spec check

Subject Matter Experts
Industry Standards / Specifications
Training & Experience
Quality Assurance Plans

Configure as Rule-based Checks
Build Batch Job

- Attribute Validation (wLateralLine)
  - InstallDate should be populated
  - OperationalArea should be populated
  - Enabled should be set to False if abandoned
- Material should be either CI, DI or PVC (post-2000 installation)
  - Material cannot be Wood or Clay (post-2000 installation)
- Attribute Validation (wMeter)
  - AdministrativeArea should be populated
  - InstallDate should be populated
  - OperationalArea should be populated
  - WaterType should be populated
  - Material should be populated
  - Diameter should be populated
- Material should be either CI, DI or PVC
  - Material cannot be Wood or Clay
  - Enabled should be set to False

- Connectivity Validation
  - Missing Reducer (wFittings)
  - Missing Tee (wFittings)
- Disconnected Features
  - Disconnected hydrants
Batch Validation

- Run data validation in several ways:
  - Within ArcMap
  - From ArcToolbox
  - Within a model/Python script
  - ArcGIS Workflow Manager
  - ArcGIS Server service
New Capabilities for 10.1
Server-based Data Validation

- Automated quality control for the enterprise
  - Scheduled and on-demand data validation
  - Scalable data validation capability based on ArcGIS Server
- Frees data editors from running checks

ArcGIS Server

Batch Validation

Data Validation GP Service

Error Results
- Spatial extent
- Properties

Reviewer Results Workspace

Production Workspace

Business Rules
Enhancements for 10.1
Automated Data Validation

- Check Enhancements
  - Geometry on Geometry
  - Intersection on Geometry

- Performance improvements
  - Increase batch validation performance

<table>
<thead>
<tr>
<th>Version</th>
<th>Records Written</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.0</td>
<td>19,000</td>
<td>15:37</td>
</tr>
<tr>
<td>10.1</td>
<td>19,000</td>
<td>1:54</td>
</tr>
</tbody>
</table>
Automated Data Review
Performing Quality Control
Visual Data Review
Types of Quality Control

Automated Quality Control
- Fast
- Consistent and repeatable
- Objective
- 100% coverage

Visual Review
- As subjective as needed
- Better for finding patterns and missing elements
An Introduction to Visual Data Review

- Visual inspection of the data
  - Level of judgment/discretion
  - Compare between data sources
  - Requires strong knowledge of the data & ability to make consistent decisions

- Simplify visual review with Data Reviewer
  - Organize and structure the visual review process
  - Manage error information in the Data Reviewer table
Working with the Reviewer Overview Window

- Structure review by geographic extent
- Navigate to cells & track review progress
Sampling Data

- Determine a subset of features
  - Fixed number of features
  - Percentage of features
  - Statistically significant sample
  - By polygon grid
Submitting Results during Visual Review

- **Commit to Reviewer Table**
  - Features selected interactively
  - Results of automated processes

- **Flag Missing Features/Notepad**
  - Capture x,y location or geometry
  - Identify type of missing feature
Quality Feedback via the Web

Expanding the dialog on quality data management

- Extending quality control workflows into other communities
  - QC review across ArcGIS platform
  - Simple to use tools for error identification
  - Deployed via intranet or Internet
Visual Data Review
Measuring and Reporting Data Quality
Working with Reviewer Table Records

- Interact with features
  - Pan/Zoom & Select
- Enter correction and verification status
Summarizing Information

- Sort and group by table fields
- Summarize quality control information
  - Identify most common data errors
  - View progress of quality control
Reporting

- Automated reporting of quality control results

- Available Reports
  - Automated Check by Origin Table
  - Automated Check by Subtype
  - Automated Check by Group
  - Total Record Count
  - Sampling
Dashboard Reporting
Enabling transparency in data quality

- Better decision making by communicating data quality across stakeholders
  - Open quality reporting
  - Shared across ArcGIS system
  - New tools and methods to communicate quality
Measuring and Reporting Data Quality
ArcGIS Data Reviewer
Automate, Simplify, and Improve your Quality Control Process

- Extension to Desktop/Server
- Supports complete QC process
- Provides
  - Rule based workflows
  - Interactive tools
  - Track errors
- For individuals and enterprise
  - Saves time/money
  - Less rework
Resources

- Product Information  www.esri.com\datareviewer
  - 60-day Free Evaluation
  - ArcGIS Data Reviewer Checks poster

- Training  www.esri.com\training
  - Instructor Led Training: Data QC Using GIS Data Reviewer
  - Virtual Campus:
    - Introduction to GIS Data Reviewer
    - Using ArcGIS Data Reviewer to Assess Data Quality

- Data Reviewer Resource Center
  - resources.arcgis.com

- Email questions and comments to DataReviewer@esri.com
# Other Data Reviewer Sessions

<table>
<thead>
<tr>
<th>Day and Description</th>
<th>Type</th>
<th>Time</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tuesday July 24</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Validating Data on the Web Using ArcGIS Data Reviewer</td>
<td>Demo Theater</td>
<td>11:00 – 12:00 AM</td>
<td>Web &amp; Server GIS Theater</td>
</tr>
<tr>
<td>Using ArcGIS Data Reviewer to validate the quality of your Community Maps content</td>
<td>Demo Theater</td>
<td>1:30 – 2:00 PM</td>
<td>Online GIS Theater</td>
</tr>
<tr>
<td>Extending the ArcGIS Data Reviewer validation framework</td>
<td>Demo Theater</td>
<td>3:00 – 3:30 PM</td>
<td>Geodatabase Mgmt Theater</td>
</tr>
<tr>
<td><strong>Wednesday July 25</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Using ArcGIS Data Reviewer to validate the quality of your Community Maps content</td>
<td>Demo Theater</td>
<td>9:30 – 10:00 AM</td>
<td>Online GIS Theater</td>
</tr>
<tr>
<td>Assessing positional accuracy using ArcGIS Data Reviewer</td>
<td>Demo Theater</td>
<td>10:00 – 10:30 AM</td>
<td>Geodatabase Mgmt Theater</td>
</tr>
<tr>
<td>Validating Data on the Web Using ArcGIS Data Reviewer</td>
<td>Demo Theater</td>
<td>12:00 – 1:00 PM</td>
<td>Web &amp; Server GIS Theater</td>
</tr>
<tr>
<td>Validating item metadata with ArcGIS Data Reviewer</td>
<td>Demo Theater</td>
<td>2:00 – 2:30 PM</td>
<td>Geodatabase Mgmt Theater</td>
</tr>
<tr>
<td>Integrating ArcGIS Data Reviewer &amp; ArcGIS Workflow Manager to automate data quality</td>
<td>Demo Theater</td>
<td>2:30 – 3:00 PM</td>
<td>Geodatabase Mgmt Theater</td>
</tr>
<tr>
<td><strong>Thursday July 14</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leveraging Geoprocessing for data validation within ArcGIS Data Reviewer</td>
<td>Demo Theater</td>
<td>1:00 – 1:30 PM</td>
<td>Analysis and GP Theater</td>
</tr>
<tr>
<td>ArcGIS Data Reviewer: An Introduction</td>
<td>Technical Workshop</td>
<td>1:30 – 2:45 PM</td>
<td>Ballroom 6C</td>
</tr>
</tbody>
</table>
Steps to evaluate UC sessions

- My UC Homepage > “Evaluate Sessions”

- Choose session from planner
  OR

- Search for session

www.esri.com/ucsessionssurveys
• Thank you for attending
• Have fun at UC2012
• Open for Questions

• Please fill out the evaluation:

  www.esri.com/ucsessionssurveys

  First Offering ID: 602
  Second Offering ID: 859