2013 Esri International User Conference
July 8–12, 2013 | San Diego, California

Technical Workshop

# ArcGIS Data Reviewer: An Introduction

Jay Cary

Michelle Johnson

# Workshop Agenda

- Importance of Data Quality
- What is ArcGIS Data Reviewer
- Automated Review
- Visual 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
  - Maximize productivity
  - Drive increased usage
- Knowledge Worker
  - Increased efficiencies
  - Confidence in GIS

# Defining Quality

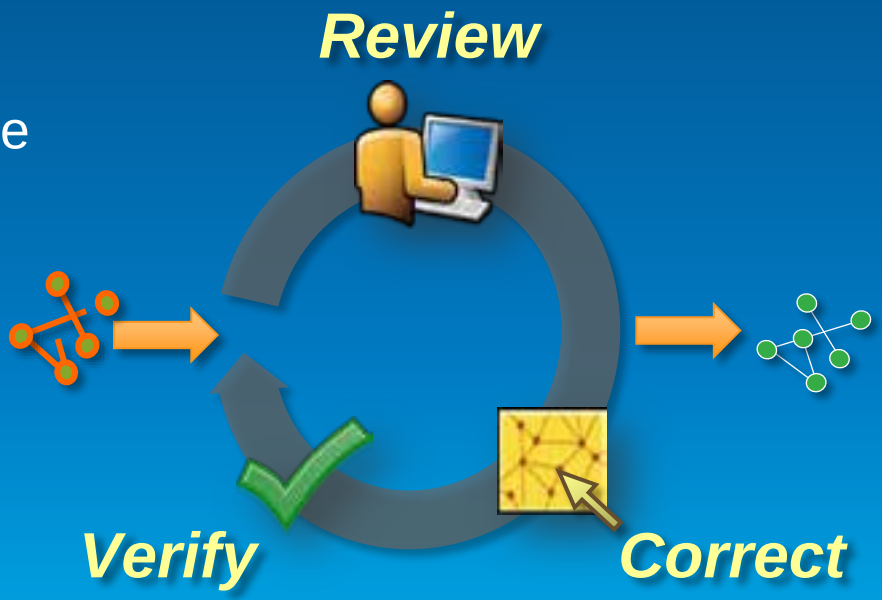
## A Technical Perspective

ISO-19157 (DRAFT), 2010

# What is ArcGIS Data Reviewer?

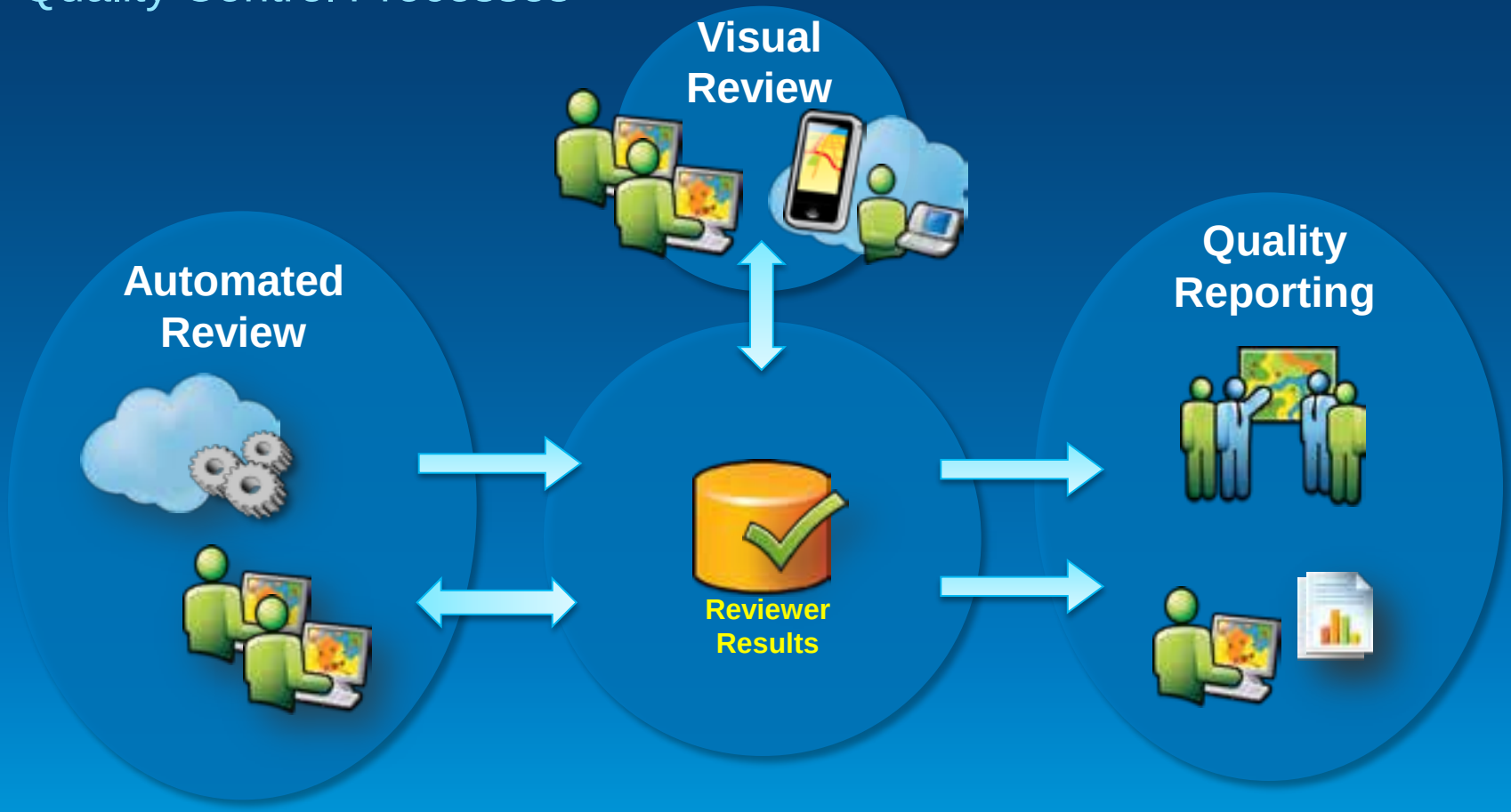
## Data Quality Management for ArcGIS

- Provides
  - Rule-based validation
  - Interactive tools
  - Track errors
- For individuals and enterprise
  - Saves time/money
  - Less rework
- Standard extension
  - ArcGIS for Desktop
  - ArcGIS for Server

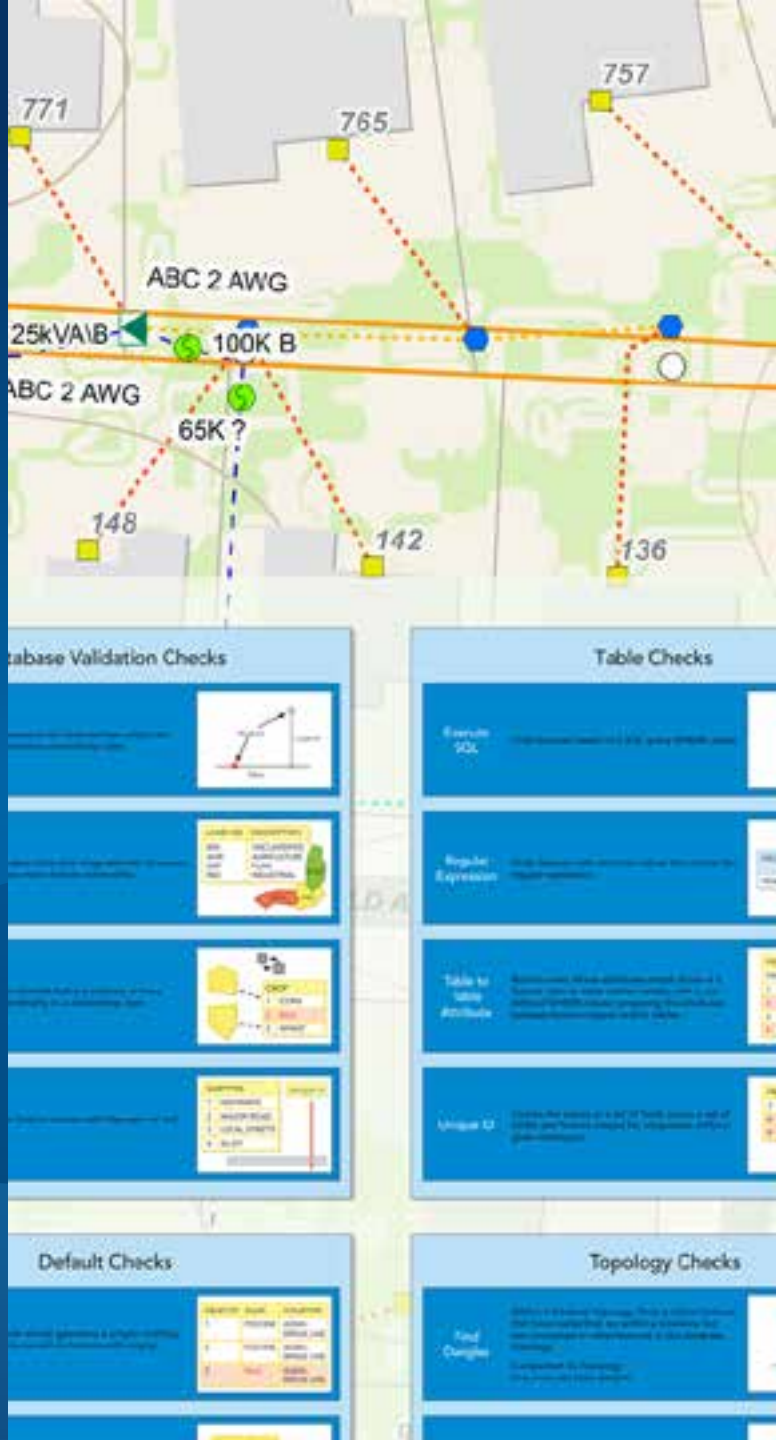


# Managing Quality Control

## Quality Control Processes



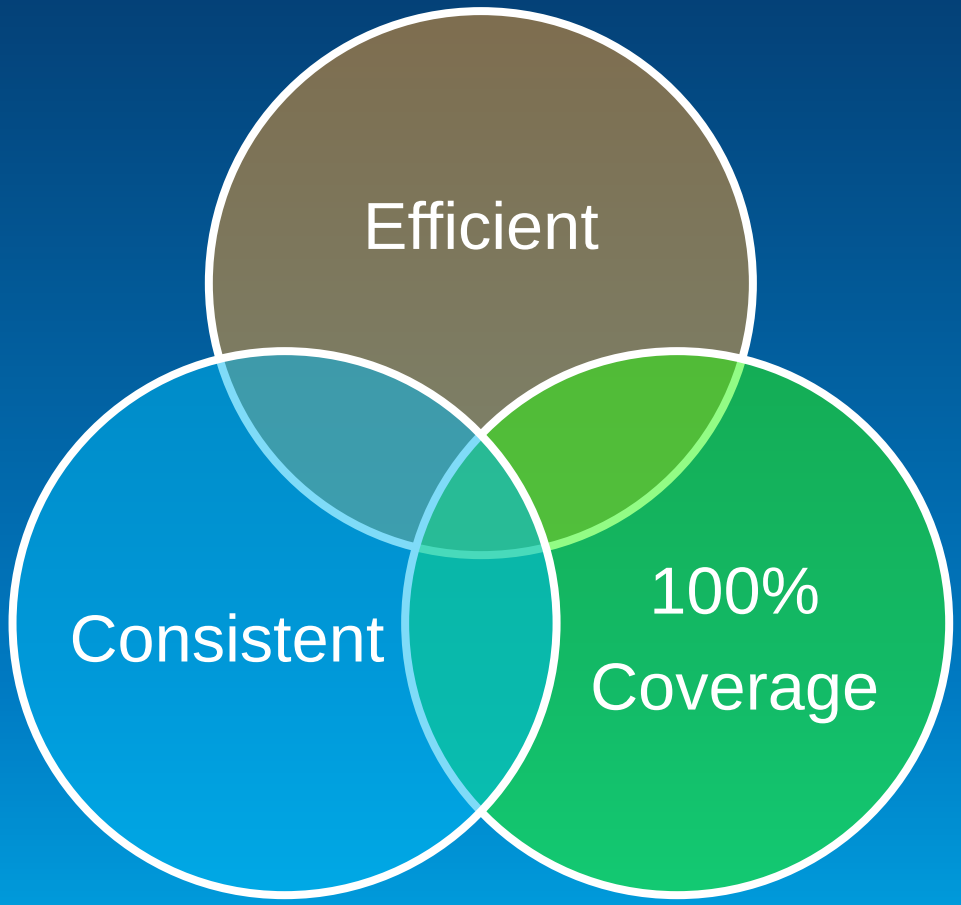
# Automated Data Review



# Defining Quality

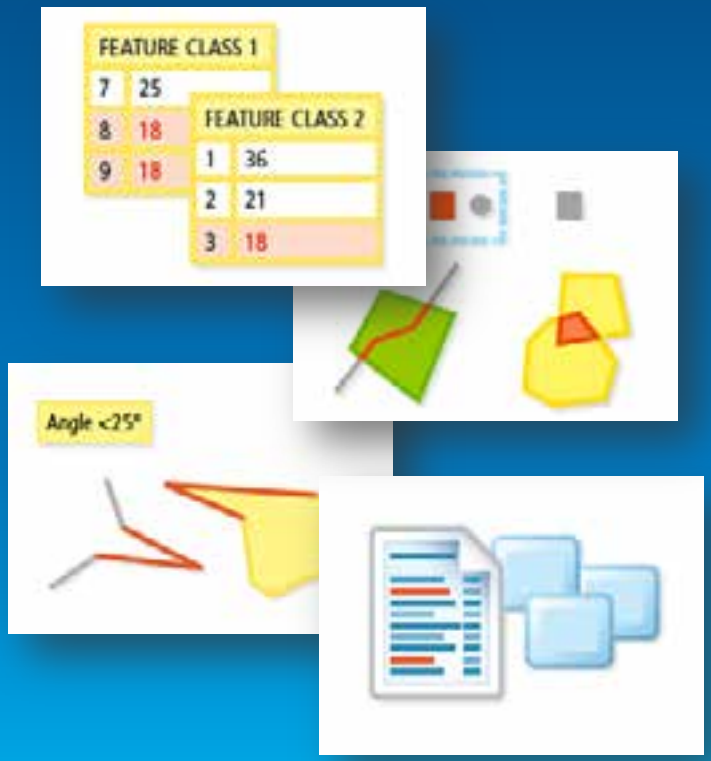
## Sources of Data Quality Requirements

# Automated Review is Valuable!



# Rule-based Automated Review

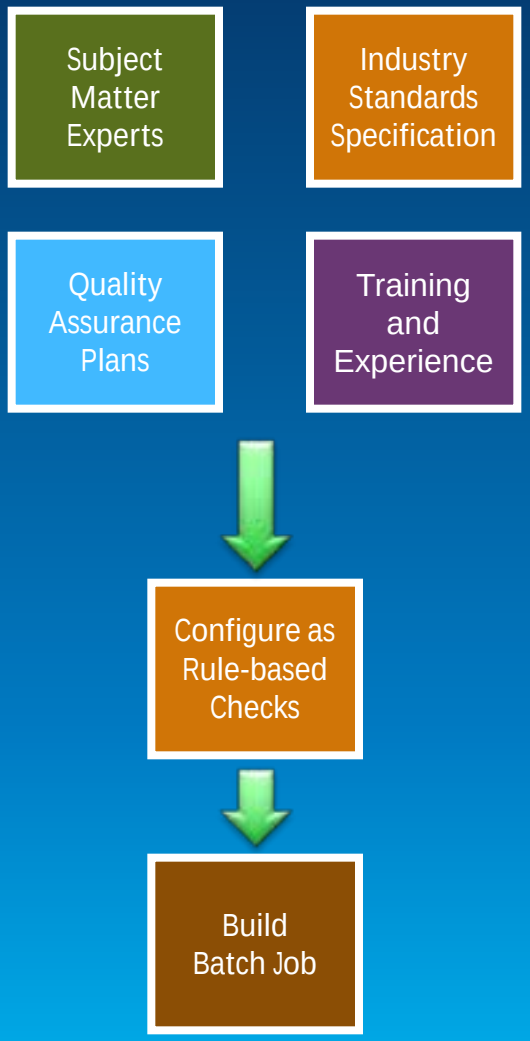
- 40+ configurable checks
- Attribute
  - Feature and Table values
- Spatial
  - Spatial relationships
- Feature integrity
  - Collection rules
- Metadata check
  - Completeness/Content



# Batch Validation

## Implementing Cumulative Review

- Encapsulate QC model
- Designed once and executed many times
- Complete specification check



# ArcGIS Resource Center

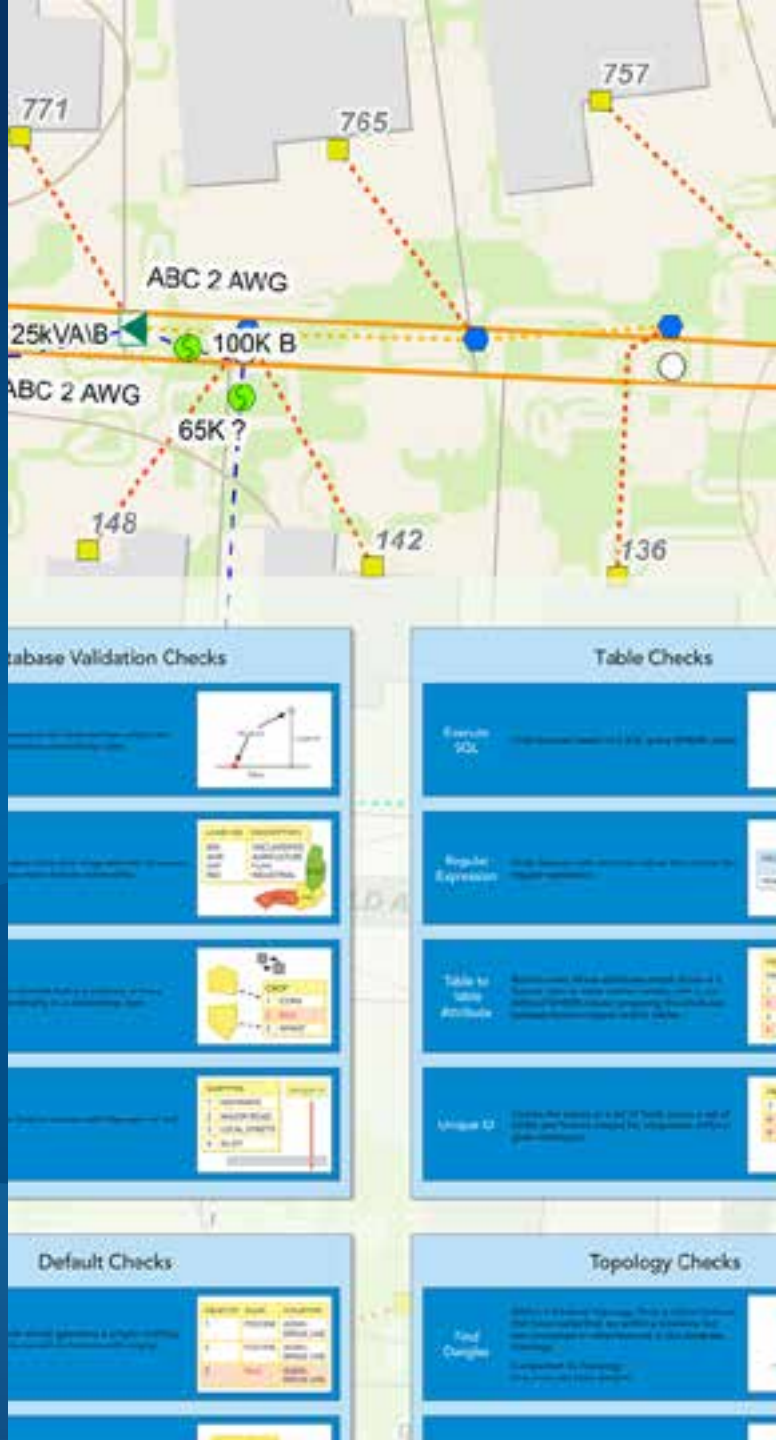
## Leveraging Templates for QC Requirements

- Data Reviewer Templates for
  - Address Management
  - Electrical Utilities
  - Roads & Highways
  - Tax Parcel Editing
  - Water Utilities
- Based on Esri industry models
- Use as Starting point

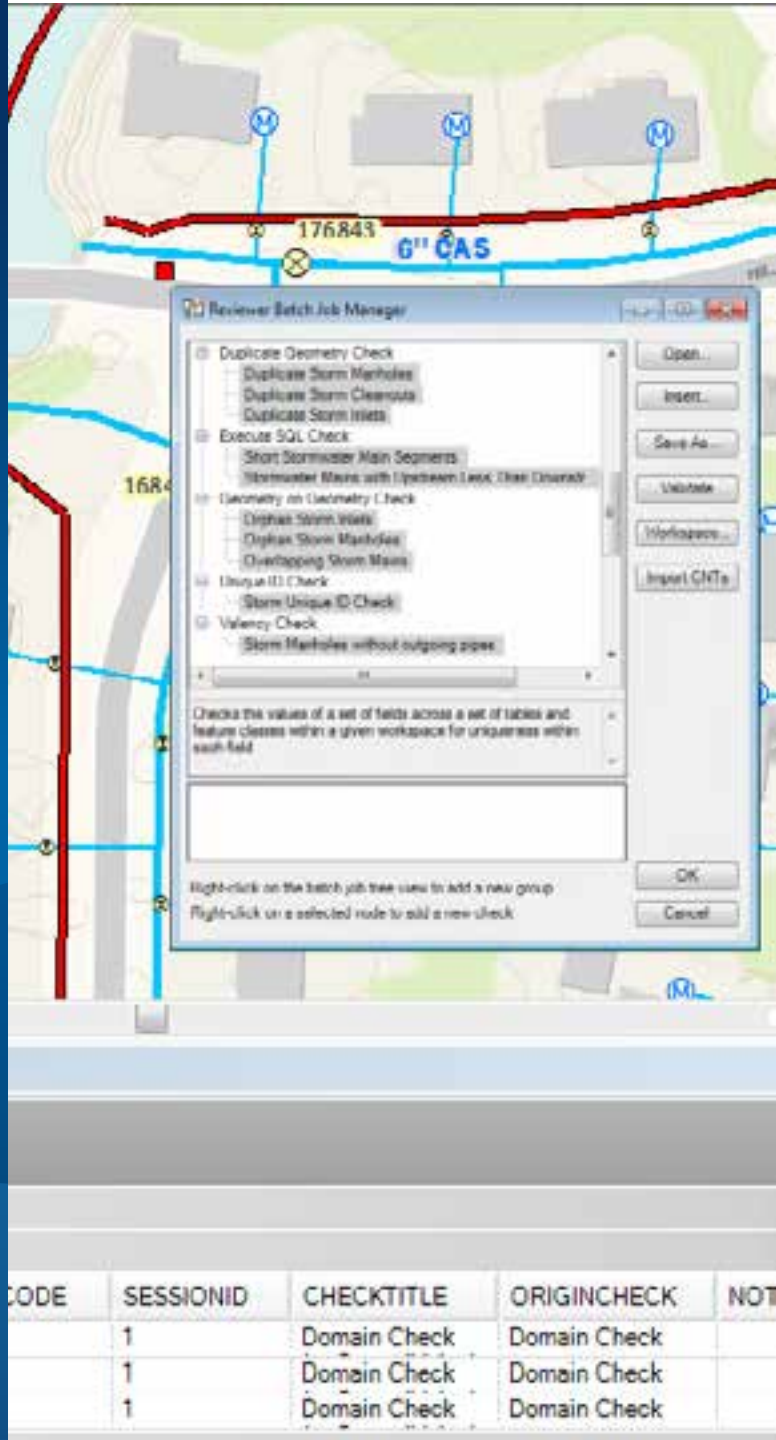
ArcGIS Data Reviewer for Desktop

# Demo:

## Getting Started with Data Reviewer



# Authoring Data Quality Rules



# Authoring Batch Jobs

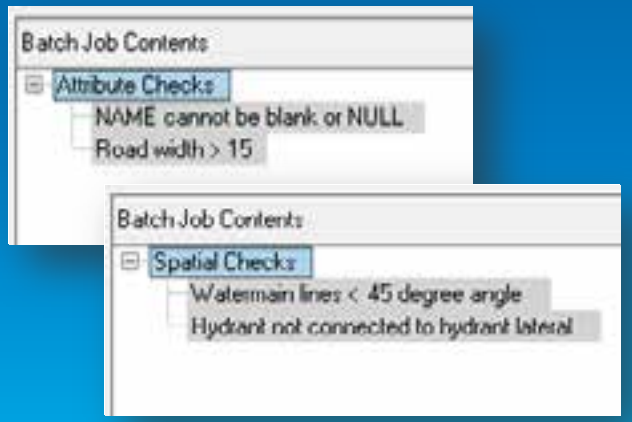
- Batch jobs
  - Container for configured checks
  - Encapsulate QC requirements
  - Reusable and shareable

**Business Rules**

Attribute rules
1. Name cannot be NULL
2. Road width must be greater than 15

Spatial Rules
1. Water Mains should not have sharp angles
2. Hydrant must be connected to hydrant lateral

**Translated to Batch Job data checks**

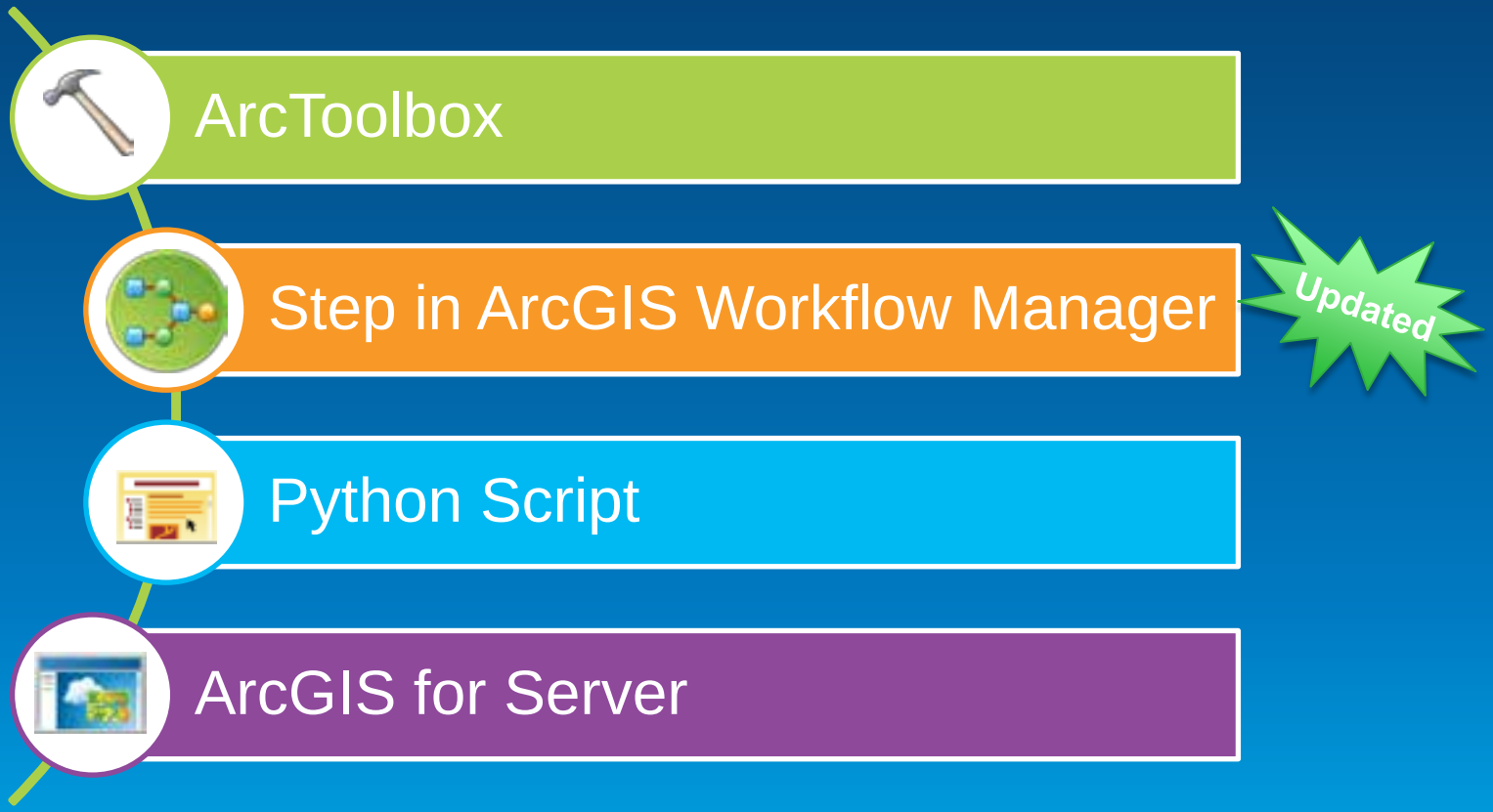


# Authoring Batch Jobs

## A Workflow Perspective

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

# Other Batch Validation Options



# Demo Scenario

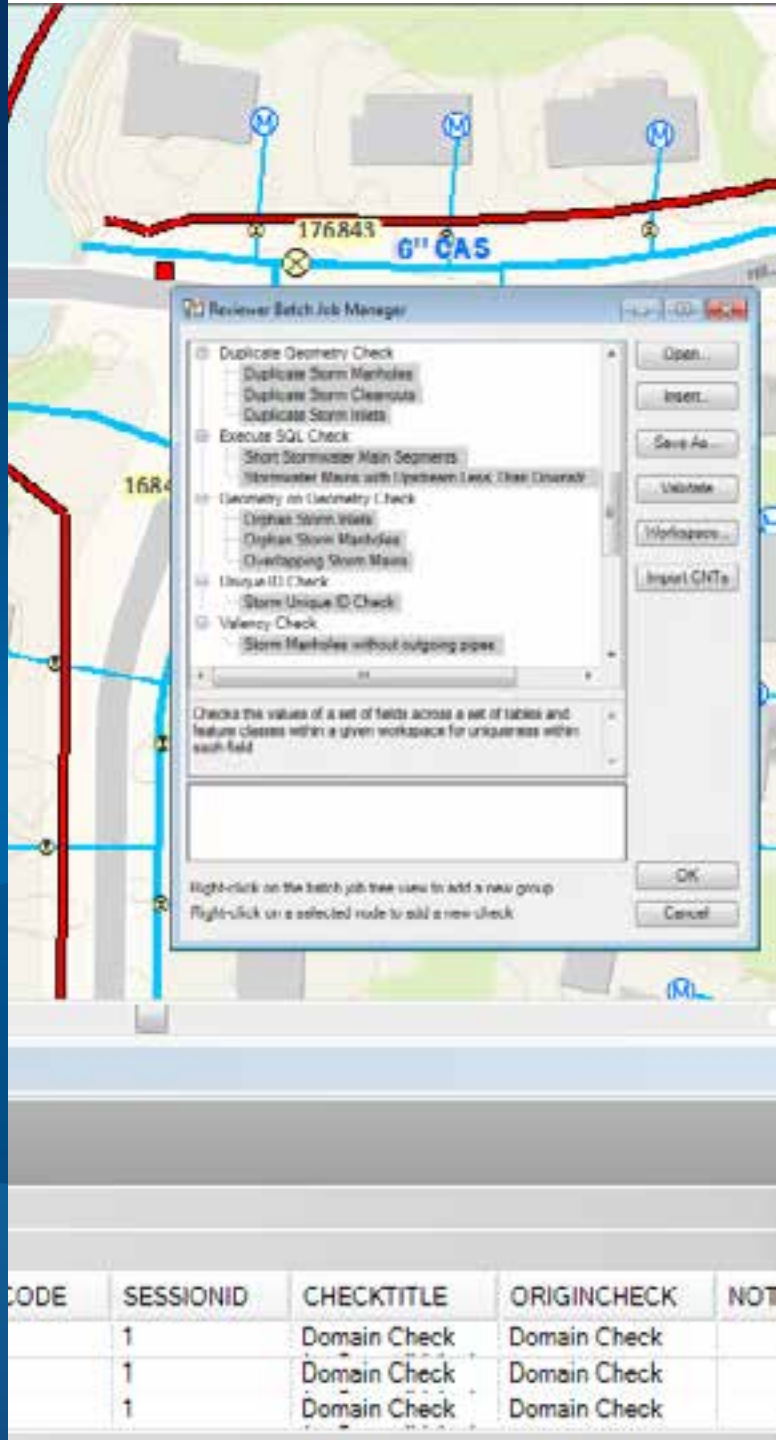
- Department plans to deploy a new hydraulic modeling application which will leverage the GIS directly.

- New Business Requirement
  - Water utility data should meet a 95% accuracy for network connectivity.

ArcGIS for Desktop

# Demo:

## Authoring Data Quality Rules



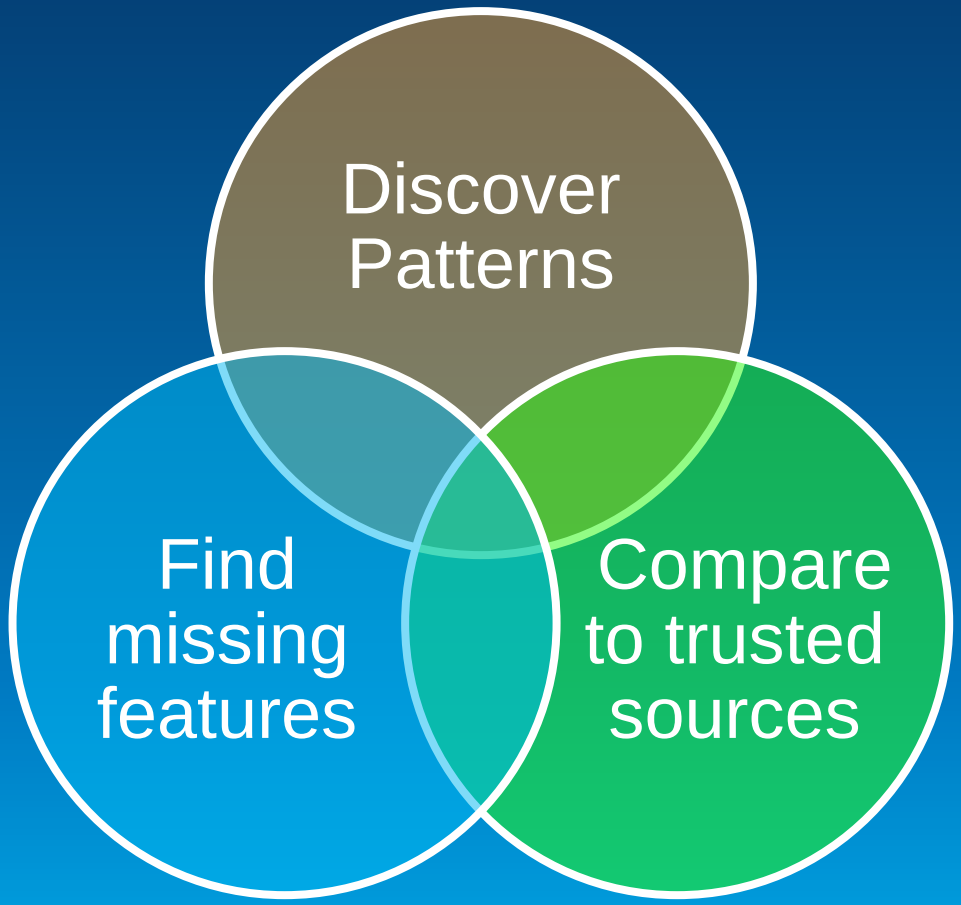
# Authoring Batch Jobs

## Tips and Tricks

- When authoring new checks
  - Configure and execute from Reviewer Toolbar
  - Leverage Mini-Browser to assess results quickly
  - Try validating a pilot area first where known issues are located.

# Visual Data Review

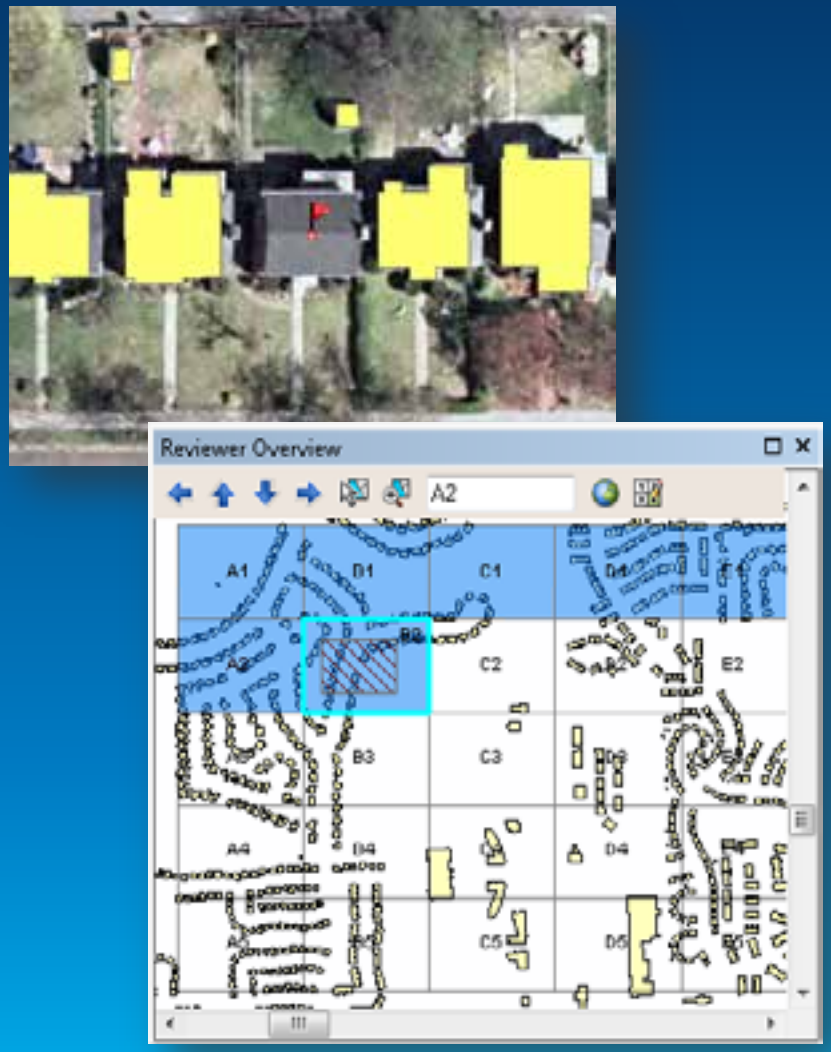
# Value of Performing Visual Review



# Visual Review

## Leveraging ArcGIS for Desktop

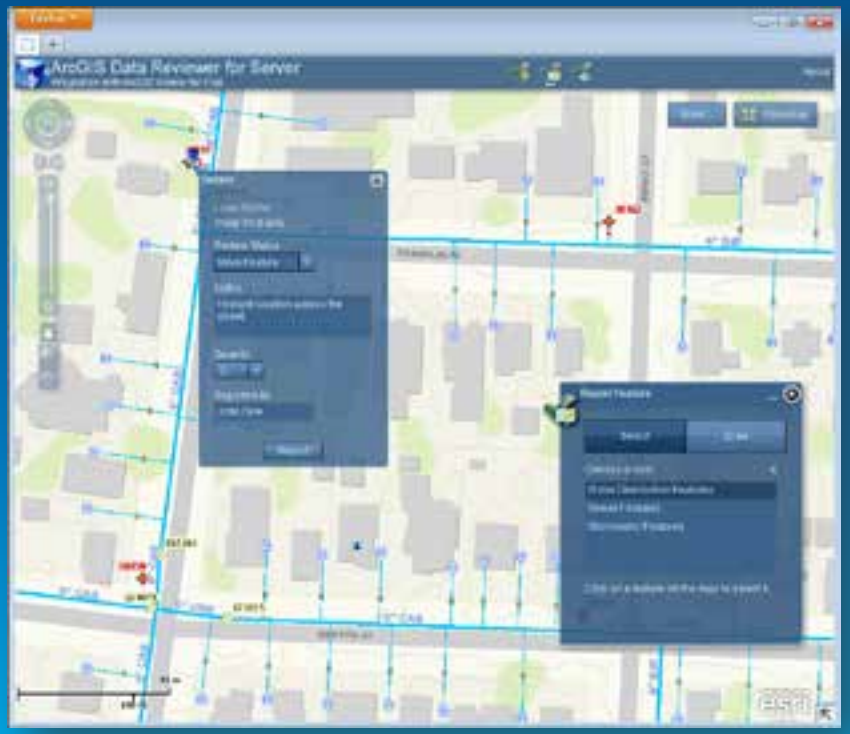
- Tools supporting
  - Feature Selection
  - Browsing
  - Flagging features
- Reviewer Overview window
- Random Sampling
- Positional Accuracy Assessment Tool (PAAT)



# Visual 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 Scenario

- Water utilities database will be leveraged by the Fire Department to support incident response and to minimize data duplication.

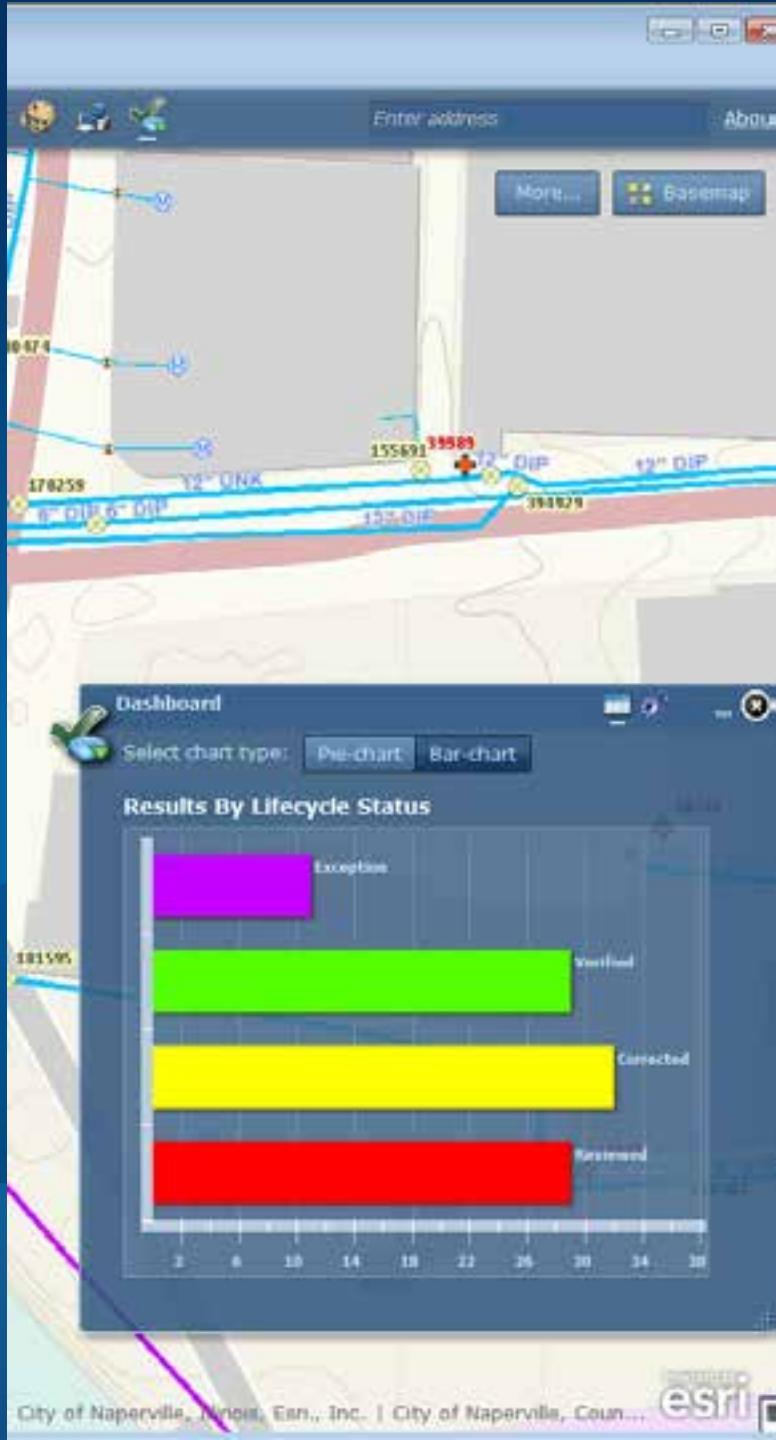
- New Business Requirements
  - Hydrant data should meet 98% accuracy for placement (+/- 1 meter).

ArcGIS API for JavaScript

# Demo:

## Visual Data Review

# Managing and Reporting Quality



# Managing Quality Control

## QC lifecycle management

Find &
Record Errors

Perform Edits
*or*
Note Exceptions

Acceptable
*or*
Unacceptable

# Reporting

## ArcGIS for Desktop

- Automated reporting of quality control results
- Available Reports
  - Automated Check (Origin Table, Subtype, Check Group)
  - Total Record Count
  - Sampling

| Date/Time | Origin Table | Subtype | Total Records | Confidence Level | Margin of Error | Sampling Number | Acceptable Error Percentage | Acceptable Error | Number of Errors | Pass/Fail | QC Complete |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 7/4/2010 11:05:59 AM | | | | | | | | | | | |
| | AtlasMap | | 153 | 98% | 3% | 0 | 2% | 0 | 0 | Pass | Yes |
| | | | | | | 0 | | | 0 | | |
| | wControlValve | | 302 | 98% | 3% | 9 | 2% | 1 | 1 | Pass | Yes |
| | | AirRelease | | | | 2 | | | 0 | | |
| | | BackflowControl | | | | 4 | | | 0 | | |
| | | BlowOff | | | | 3 | | | 1 | | |
| | wFitting | | 297 | 98% | 3% | 9 | 2% | 1 | 2 | Fail | Yes |
| | | Reducer | | | | 7 | | | 0 | | |
| | | Tap | | | | 2 | | | 2 | | |
| | wLateralLine | | 18690 | 98% | 3% | 544 | 2% | 17 | 9 | Pass | Yes |
| | | Hydrant Lateral | | | | 50 | | | 0 | | |
| | | Service Lateral | | | | 493 | | | 8 | | |
| | | Service Lateral To Meter | | | | 1 | | | 1 | | |

# Reporting

## ArcGIS for Server

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

# Scenario

As a GIS manager, I need to manage the quality of the water utilities database to ensure that it can support existing and new staff operations.

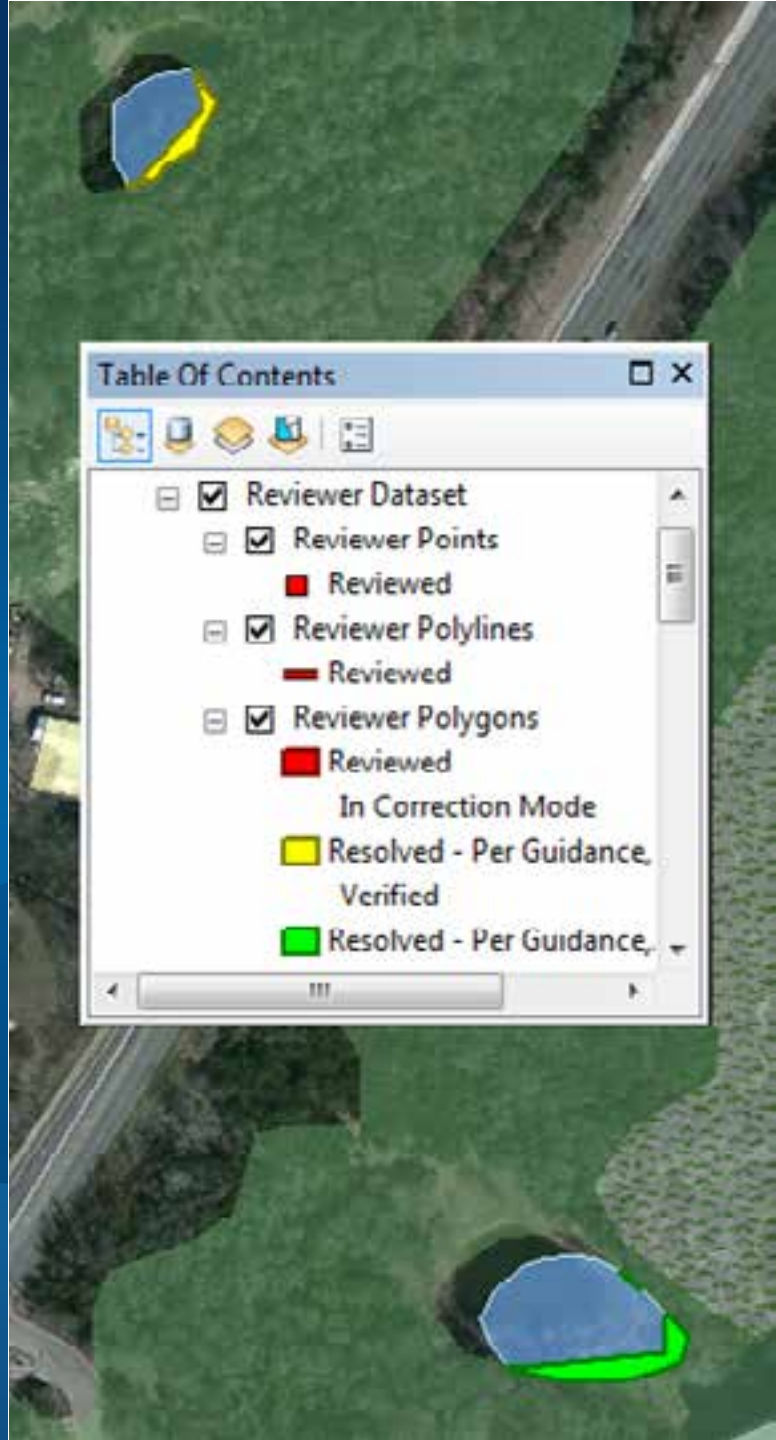
Areas to be addressed

- GIS shall support hydraulic modeling tasks (95% accuracy).
- Fire hydrant locations shall be positionally accurate and correctly attributed to support use by the Fire Department for emergency response (98% accuracy).

ArcGIS for Desktop

# Demo:

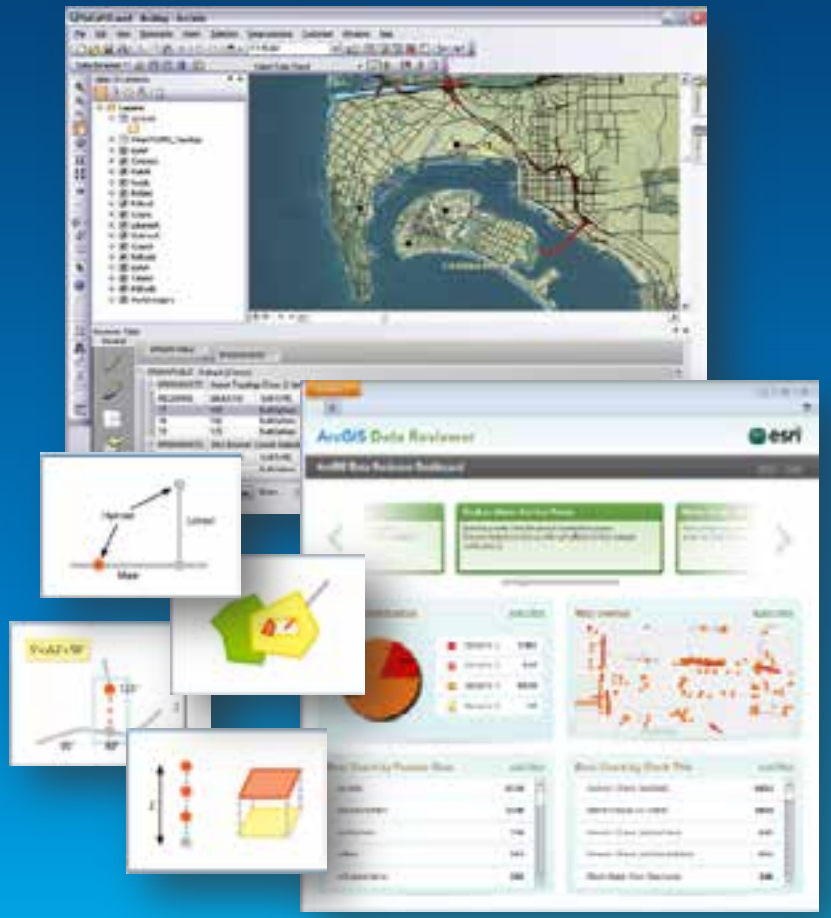
## Managing and Reporting Quality



# ArcGIS Data Reviewer

## Automate, Simplify, and Improve your Quality Control Process

- Extension to ArcGIS for Desktop and Server
- Supports complete QC process
- Provides
  - Rule-based validation
  - Interactive tools
  - Track errors



# Resources

- Product page
  - www.esri.com/datareviewer
- Resource Center
  - http://resources.arcgis.com/
- Training
  - www.esri.com/training
- Questions & comments
  - datareviewer@esri.com

# *Thank you…*

Please fill out the session evaluation

*Offering ID:  1276*

**Online** – www.esri.com/ucsessionsurveys

**Paper** – pick up and put in drop box

# Other Data Reviewer Sessions

| Day and Description | Type | Time | Location |
|---|---|---|---|
| Tuesday July 9 | | | |
| ArcGIS Data Reviewer: Validating Item Metadata | Demo Theater | 3:30 – 4:00 PM | Geodatabase Mgmt Theater |
| ArcGIS Data Reviewer for Server - Deploying Distributed Quality Control Services | Demo Theater | 3:30 – 4:00 PM | Web & Server Theater |
| Using ArcGIS Data Reviewer to Validate the Quality of your Community Maps Content | Demo Theater | 4:00 – 4:30 PM | Online GIS Theater |
| ArcGIS Data Reviewer for Server - Leveraging Distributed Quality Control Services in Web Clients | Demo Theater | 4:30 – 5:00 PM | Web & Server Theater |
| Wednesday July 10 | | | |
| ArcGIS Data Reviewer Special Interest Group Meeting | SIG | 12:00 – 1:00 PM | Room 28B |
| ArcGIS Data Reviewer: An Introduction | Technical Workshop | 1:30 – 2:45 PM | Room 31C |
| Integrating ArcGIS Data Reviewer & ArcGIS Workflow Manager to Automate Data Quality Control Workflows | Demo Theater | 1:30 – 2:00 PM | Geodatabase Mgmt Theater |
| ArcGIS Data Reviewer for Server - Deploying Distributed Quality Control Services | Demo Theater | 3:00 – 3:30 PM | Web & Server Theater |
| ArcGIS Data Reviewer for Server - Leveraging Distributed Quality Control Services in Web Clients | Demo Theater | 3:30 – 4:00 PM | Web & Server Theater |
| Thursday July 11 | | | |
| Using ArcGIS Data Reviewer to Validate the Quality of your Community Maps Content | Demo Theater | 10:00 – 10:30 AM | Online GIS Theater |

Understanding our world.