

QA/QC for GIS Data: Recording and Tracking Errors

Transcript

Copyright © 2006 ESRI

All rights reserved.

The information contained in this document is the exclusive property of ESRI. This work is protected under United States copyright law and other international copyright treaties and conventions. No part of this work may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying and recording, or by any information storage or retrieval system, except as expressly permitted in writing by ESRI. All requests should be sent to Attention: Contracts and Legal Services Manager, ESRI, 380 New York Street, Redlands, CA 92373-8100, USA.

The information contained in this document is subject to change without notice.

@esri.com, 3D Analyst, ADF, AML, ARC/INFO, ArcAtlas, ArcCAD, ArcCatalog, ArcCOGO, ArcData, ArcDoc, ArcEdit, ArcEditor, ArcEurope, ArcExplorer, ArcExpress, ArcFM, ArcGIS, ArcGlobe, ArcGrid, ArcIMS, ArcInfo Librarian, ArcInfo, ArcInfo-Professional GIS, ArcInfo-The World's GIS, ArcLocation, ArcLogistics, ArcMap, ArcNetwork, ArcNews, ArcObjects, ArcOpen, ArcPad, ArcPlot, ArcPress, ArcQuest, ArcReader, ArcScan, ArcScene, ArcSchool, ArcSDE, ArcSdl, ArcStorm, ArcSurvey, ArcTIN, ArcToolbox, ArcTools, ArcUSA, ArcUser, ArcView, ArcVoyager, ArcWatch, ArcWeb, ArcWorld, Atlas GIS, AtlasWare, Avenue, BusinessMAP, Database Integrator, DBI Kit, ESRI, ESRI-Team GIS, ESRI-The GIS Company, ESRI-The GIS People, FormEdit, Geographic Design System, Geography Matters, Geography Network, GIS by ESRI, GIS Day, GIS for Everyone, GISData Server, InsiteMAP, JTX, MapBeans, MapCafé, MapObjects, ModelBuilder, MOLE, NetEngine, PC ARC/INFO, PC ARCPLOT, PC ARCSHELL, PC DATA CONVERSION, PC STARTER KIT, PC TABLES, PC ARCEDIT, PC NETWORK, PC OVERLAY, PLTS, Rent-a-Tech, RouteMAP, SDE, SML, Spatial Database Engine, StreetEditor, StreetMap, TABLES, the ARC/INFO logo, the ArcCAD logo, the ArcCAD WorkBench logo, the ArcCOGO logo, the ArcData logo, the ArcData Online logo, the ArcEdit logo, the ArcExplorer logo, the ArcExpress logo, the ArcFM logo, the ArcFM Viewer logo, the ArcGIS logo, the ArcGrid logo, the ArcIMS logo, the ArcInfo logo, the ArcLogistics Route logo, the ArcNetwork logo, the ArcPad logo, the ArcPlot logo, the ArcPress for ArcView logo, the ArcPress logo, the ArcScan logo, the ArcScene logo, the ArcSDE CAD Client logo, the ArcSDE logo, the ArcStorm logo, the ArcTIN logo, the ArcTools logo, the ArcView 3D Analyst logo, the ArcView Business Analyst logo, the ArcView Data Publisher logo, the ArcView GIS logo, the ArcView Image Analysis logo, the ArcView Internet Map Server logo, the ArcView logo, the ArcView Network Analyst logo, the ArcView Spatial Analyst logo, the ArcView StreetMap 2000 logo, the ArcView StreetMap logo, the ArcView Tracking Analyst logo, the Atlas GIS logo, the Avenue logo, the BusinessMAP logo, the Data Automation Kit logo, the ESRI ArcAtlas Data logo, the ESRI ArcEurope Data logo, the ESRI ArcScene Data logo, the ESRI ArcUSA Data logo, the ESRI ArcWorld Data logo, the ESRI Digital Chart of the World Data logo, the ESRI globe logo, the ESRI Press logo, the Geography Network logo, the MapCafé logo, the MapObjects Internet Map Server logo, the MapObjects logo, the MOLE logo, the NetEngine logo, the PC ARC/INFO logo, the Production Line Tool Set logo, the RouteMAP IMS logo, the RouteMAP logo, the SDE logo, The World's Leading Desktop GIS, Water Writes, www.esri.com, www.geographynetwork.com, www.gisday.com, and Your Personal Geographic Information System are trademarks, registered trademarks, or service marks of ESRI in the United States, the European Community, or certain other jurisdictions.

Other companies and products mentioned herein are trademarks or registered trademarks of their respective trademark owners.

Welcome to our ESRI Instructional Series podcast: Recording and Tracking Errors During QA/QC. I'm Keith Mann from Educational Services at ESRI in Redlands, California. Today, I will be providing you with some ideas for recording and tracking errors during QA/QC. This discussion is tailored to GIS managers and coordinators, and GIS data producers.

Let's start by explaining why you should record errors during QA/QC. One reason is that having a record of errors provides you with a means for measuring data quality. You can quantify errors by comparing the number of errors found against the number of features inspected. This is known as an error rate. So, for example, if you inspected one hundred features in a dataset, and found five errors, the error rate would be five percent.

Another reason is that recording errors allows you to evaluate and improve your QA/QC processes and your data production workflows. A quality assurance program is designed to be proactive, and error recordation helps you to identify and change data production, or data collection processes that are causing you problems.

Another reason is that, in some cases, you may not be the person who produces or edits the database, where the QA/QC project involves many people. Recording and tracking errors helps you establish accountability for data quality. It allows you to verify that errors are corrected, and to preserve a history of the work. What type of information should you collect when you record and track errors during QA/QC?

To help answer this question, think about an error as having a life cycle. In the first phase of an error life cycle, errors are detected or exposed through a QA/QC review process. This is the point where you want to record information about the error, so you can respond to it.

In the next phase of an error life cycle, the error gets corrected. Typically, this is done by a GIS technician, or by the original data producer, if possible. Once the error has been corrected, you, or another QA/QC person verifies that the error has been fixed. Verification is the last phase of an error life cycle.

At each of these phases in the error life cycle, you should consider gathering information about the location of the error, geographically, as well as its system location. There should be a detailed description of the error, and a solution for fixing it. You should also record who reviewed,

corrected, and verified the error, and when they did it. The error should also be given some sort of unique identifier, so that you can track it through its life cycle.

How should you record and track errors? There are several methods for recording errors. For example, you could keep track of errors in a spreadsheet, or you could design a table in your word processing application. For this broadcast, let me suggest a more effective solution. Think about designing a geodatabase point feature class for recording and tracking errors during QA/QC. A point feature class will allow you to automatically store the geographic location of the error. You can add fields to the feature attribute table, so that you can store the necessary error information. You'll be able to symbolize the error points in a map, based on any of the error information stored in the attribute table, and because the feature class is stored in a geodatabase, you'll be able to query the database, and perform analysis on the point features and attributes. You'll also be able to use attribute domains to standardize field values, and you can set default field values, or use the field calculator, to populate records.

Let me provide an example of what I'm talking about. First, within a geodatabase, create a set of attribute domains to help standardize the values for certain fields in the error table. For example, for the field used to describe the inspection method used to detect the error, I could create a domain with values, like *feature completeness*, *feature accuracy*, *valid attribute value*, and so on. Next, I'd create a new, empty point feature class.

Finally, I'd add the necessary fields to the new point feature class attribute table, and design the appropriate domains for each field. (Not all fields will be associated with the domain.) Each of these steps: creating a geodatabase, creating attribute domains and domain values, creating a point feature class, and adding fields to an attribute table, can all be accomplished using geoprocessing tools in ArcToolbox.

You can use geoprocessing tools by themselves, or you can incorporate them into a model or script to automate these processes. Once you've created a point feature class for capturing errors, you can begin using the quality control methods defined in your quality assurance plan. To add point features to the point feature class, you'll need to be in an edit session in ArcMap.

When you discover an error in the feature class you're reviewing, you simply click on the location of the error in the map to create a new point feature, and then by using the attributes

window or the attributes table, populate the error tracking fields. When you populate fields with domains assigned, you get a drop-down list of acceptable domain values to pick from. In the attribute table, you can use the Field Calculator to calculate date fields, or x and y coordinate fields. Click the Help button in the Field Calculator to learn more about creating expressions for x and y coordinates.

When you're finished capturing errors, save your edits. You can provide the error location point feature class to the person who will correct the errors, and then use it again to verify that the corrections have been made.

Let's review the concepts in this broadcast. Recording errors during QA/QC provides you with a means for measuring data quality. It also allows you to evaluate and improve your QA/QC processes and your data production workflows. The information you collect about an error should reflect the error life cycle, which is review the error, correct the error, and verify that the error has been corrected.

You could use the point feature class as a way to record errors for QA/QC. You edit the point feature class during quality control to mark the location of errors in a map. If creating and managing your own error recordation system is not for you, or you work in a production environment, you should consider using ESRI's GIS Data Reviewer. GIS Data Reviewer is the data quality control management application that simplifies many aspects of automated and visual spatial data quality control tasks.

To learn more about GIS Data Reviewer, go to www.esri.com, and click on the Products tab. For more resources, please check out our instructor-led training courses at www.esri.com/training.

This discussion touched on topics that are covered in our two-day instructor-led course: *QA/QC for GIS Data*.

Thank you for tuning in to this session of our ESRI Instructional Series podcast. Stay tuned for future broadcasts.