

QA/QC for GIS Data: Starting a Quality Assurance Program

Transcript

Copyright © 2005 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: Introduction to QA/QC for GIS Data.

I'm Keith Mann from Educational Services at ESRI in Redlands, California. Today, I will be providing you with a brief introduction to QA/QC, as well as a few concepts for starting a quality assurance program. This discussion is tailored to GIS managers and coordinators and GIS data producers.

Let's start by defining what QA/QC means. QA, or quality assurance, is an all-encompassing management approach combining technical aspects of quality, qualitative methods, and human resources, in a system designed to meet the customer's expectations. In other words, you can think of quality assurance as a management program that encompasses all aspects of your GIS workflows, procedures, and standards. QC, or quality control, is an individual task or set of tasks performed at a given level of the production process that is measured and aimed at ensuring integrity of a product, output, or action. Quality control tasks often involve using a combination of GIS tools and methods, along with visual inspection of the data, to find features and attributes that don't conform to a specified standard or other criteria.

So why do you need QA/QC for your GIS data? First of all, consider the vast array of data sources available to you. There's the data that you produced from scratch. There's data collected in the field, say with GPS. You can also purchase data or download it online. In fact, data can be migrated, converted, inherited, and derived. However you acquire GIS data, a QA program allows you to match the expectations you or your customers have for it with the workflows you require. These workflows include getting data into your GIS, spatial analysis, and making maps. Essentially, QA/QC is a proactive effort. By setting up a QA program, you are taking action to protect and sustain your GIS. Think of a QA program as a firewall against data that doesn't meet your needs.

How do you get started with a QA program? Designing quality assurance and quality control for your GIS is very similar to database design. In fact, a quality assurance program can be developed at the same time as you design your database. However, this isn't always possible. You can get started by first acknowledging all of your workflows where GIS data is used and processed. This helps you define the expectations for each dataset. Second, establish the criteria you want or need for each dataset. You may be able to rely on pre-designed data models or data dictionaries for most of this information. Finally, once you define expectations and have established a set of data

standards, you can design quality control methods to measure or compare the data against the criteria. ESRI software is composed of functionality and tools to help you inspect and validate features and attributes.

How do you assess quality in GIS data? One way to assess quality is to find and record errors. However, measuring error is difficult to quantify because it is very subjective. Here are some suggestions for getting started with quality control.

- Check for feature completeness. This means inspecting the dataset for missing or incomplete features.
- Check for feature accuracy. This means checking that the data feature properly describes the on-the-ground feature.
- Check for attribute value consistency. This means making sure attribute values conform to defined naming conventions, or established attribute domain values, or field data types.
- Check for attribute value precision. This means determining the real level of precision possible for a measurement value.

If a feature is incomplete or inaccurate, or an attribute value is inconsistent or implies false precision, you can count that feature as an error. The number of errors you find gives you an indication of the quality of your data.

If you're interested in starting a QA program, here are a few recommendations for implementing QA/QC in your next GIS project.

- First, create a quality assurance plan. Initially, this may be a simple document where you record your data criteria and the procedures used to check the data against that criteria. Later, this document can evolve into a complete guideline for your quality assurance program.
- Second, develop a method for recording and tracking errors. Recording errors allows you to assess data quality. You can use a spreadsheet, or even better, a point feature class, that allows you to locate the errors in a map.
- Third, create a list of quality control tasks you plan to use for each dataset. Start with an initial set of quality control checks that measure minimum requirements. These are typically automated methods that help you assess data quality quickly and cheaply. If the dataset passes the initial checks, describe a set of more intensive inspection methods. These are

typically visual inspection methods, which are manual processes. Visual inspection is more labor intensive, which means it's more expensive.

Let's review the concepts in this broadcast. QA equals quality assurance, which is a management program for ensuring data quality and meeting customer expectations. QC stands for quality control, which is a set of tasks for measuring data quality. These tasks can include automated and manual methods including visual inspection. Together QA/QC helps you protect and sustain your GIS. Finding and recording error in your data is a way to measure quality. Measuring error, however, can be difficult, so it's important to define the levels of quality you expect. Suggested basic elements for implementing QA/QC are:

1. Create a quality assurance plan
2. Devise a method for recording and tracking error
3. Create a list of quality control tasks you plan to use for each dataset.

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 into this session of our ESRI Instructional Podcast series. Stay tuned for future broadcasts.