

2014 Washington GIS Conference Proceedings – Individual Presentation

Title

How Good are Your Data and Analyses? Communicating Quality. Part 3 of 3: Analysis

Presenter

David Howes

Abstract

If you're in the business of providing clients or members of the public with the results of GIS-based analysis, it's in your interests to maintain records of exactly how your deliverables were developed. Public scrutiny and the threat of litigation are strong motivators for keeping such records, yet, fortunately, many of us are rarely, if ever, faced with them. Of course, that doesn't lessen their importance and, if a request did come up, how effective would you be at explaining precisely what you did months or years ago? The reality is that, compared to generating a map or creating an application, documenting processes is not very stimulating. A framework for thinking about the requirements may help.

In "Executing Data Quality Projects: Ten Steps to Quality Data and Trusted Information" (Morgan Kaufmann, 2008), Danette McGilvray provides a set of "data quality dimensions," each of which may be used to "define, measure, and manage the quality of data and information." Examples include "ease of use and maintainability", "accuracy" and "perception, relevance and trust." Would it not make sense to apply this way of thinking to our analyses and processes, as well as to our data? Could such an approach help us be more accountable to ourselves and those we serve? Could it make us more efficient? The purpose of this presentation is to explore these questions and, potentially, help prepare us for the moment when the call for details comes. That call may not be a legal one, which could be rather frightening, but could be just an attempt to remember what we did a couple of weeks ago. Regardless of the situation, any attempt to make the response less onerous and more efficient seems worthwhile.

Proceedings Entry

The main point of this presentation is to emphasize the importance and value of good documentation, especially with respect to GIS analyses and processes. At the 2014 Washington GIS Conference, the presentation was the third in a set of three emphasizing the communication of quality, with the other presentations, by Duncan Munro and Stephen Beimborn of Seattle Public Utilities, focusing on GIS data. The content of the presentations was arranged to ensure a logical sequence, while also allowing for each presentation to be received independently.

Depending on the type of work being undertaken, it is often tempting to start clicking on the keyboard and performing operations without keeping track of what's being done. Of course, this may not matter if the job is small and relatively straightforward, but it can become a problem as scope and complexity increase. Often, clients or recipients of GIS operations don't require detailed documentation and a focus on deliverables, especially in a somewhat pressured environment, can result in there being no record of the steps taken to produce such deliverables. But, as producers of products, it is our professional duty to be sure that at any point we can readily explain our approaches in any necessary level of detail. That said, until we're asked to provide explanation, it can be difficult to convince ourselves that the effort involved in documentation is worthwhile. Computer backups can serve as a suitable analogy. They can be time-consuming, but it only usually takes one hard drive failure to reinforce their value. When it comes to GIS documentation, it's much better to be prepared than for it to be revealed, potentially in a public forum, that you really can't say exactly what you did to develop something of importance to you

or someone else. The preservation of institutional knowledge is another important reason for maintaining good documentation, as was emphasized in the session entitled “They’ll Stone You When You’re Trying to Build Your GIS: The Multi-Dimensional Role of the GIS Coordinator.” Among other things, the contributors to that session were keen to address a common scenario whereby a GIS professional takes a job in a new organization only to find that they have to start largely from scratch because there are no documents explaining the existing GIS operations. Of course, it’s a little unreasonable to complain about this scenario if you haven’t left any documentation for your successor either.

Sometimes a framework to support our thinking can be helpful, be it conceptual or, perhaps, prescriptive in nature and regardless of whether we’re concerned with data, processes or our thought processes in general. The Capability Maturity Model, for example, emerged as a result of the procurement activities of the U.S. Department of Defense in the late 1980s/early 1990s and has been adapted for many purposes since its introduction. The descriptions could be seen as slightly provocative in a positive sense with level 1, the “Initial” level, referring to “ad hoc and chaotic process” and a dependence on “heroes and luck.” We usually like to think that, as professionals, our work is well thought out and properly structured in terms of our methods and approaches, and this may very well be so, but there’s a good chance that without any form of documentation to record what we did, our intentions are questionable. Likewise, the description for level 2 includes the word “reactive.” There may be nothing wrong with reacting to some needs, but, taken to the extreme, a generally reactive strategy might not be a good one. The value of the model is that it makes us think about what we do and how we do it and helps us adopt beneficial tactics and strategies that not only improve the quality of our deliverables, but also provide protective explanatory cover. Level 2 of the model is the “Repeatable” step, so if we think about what it takes to ensure that the production of our deliverables is repeatable, documentation is often a key component. If one were feeling contrarian, one could argue that developing a simple tool satisfies the repeatability requirement, but it helps to explain how you built the tool in the first place.

Another potentially useful framework is provided by Danette McGilvray (2008). This one is a set of “Data Dimensions,” intended to provide a basis for assessing and measuring data quality. It is not a stretch, however, to replace “data” with “process” and use the steps as a basis for thinking about analyses and processes for the purpose of improving their quality as well as describing them in documentation. These data or process dimensions can be related to another framework, the “Information Lifecycle,” as a way of tying them closely to a, perhaps, more familiar structure for thinking about organizational activities.

It’s all very well having frameworks for thinking about things, but sooner or later we need to do something concrete. As the presentation slides record, we need to come up with methods for keeping track of what we’re doing that (1) aren’t onerous, (2) have strong positive benefits and (3) can be measured to assess the extent of 1 and 2. These three items constitute the key message of the presentation. Think of the converse: we tend to avoid doing things that are tedious and not particularly helpful. But, if the value of what you’re doing quite clearly outweighs the effort it takes to do it, then there’s a good chance you’ll continue to do whatever it is you’re doing. In this vein, the concept of “incorporated documentation” is a suggestion that’s meant to fit the three requirements stated above.

“Incorporated documentation” simply refers to the idea of embedding documentation within (the development of) analyses and processes and a number of options are presented in the slides. Regardless of which approaches are taken, the goal remains the same: to be able to explain your work to yourself and others such that you can, to the greatest extent possible, account for what you did and, if required, largely repeat it. Meeting this goal as often as possible helps you become more efficient because, for

example, you no longer have to remember procedures and you're building a repository of techniques that you can apply or re-use with relative ease.

One example of incorporated documentation from the author's own experience pertains to the development of species protection areas for a client. In this case, the deliverables are polygons, produced using custom tools and processes within Esri's ArcMap program. The rules for creating these polygons vary from state to state within the U.S. and maintaining a clear picture of how each polygon was constructed is critical for both legal and professional reasons. Documentation of the steps is accomplished in various ways, such as through the creation of neat file and folder structures and names, the use of simple setup files for running the tools, thus maintaining a record of all parameter values, and the development of tidy code supported by clear, consistent and descriptive naming conventions. The clarity that approaches of this sort can provide cannot be understated.

In conclusion, there are various reasons why it's important to maintain good documentation, some of which have been mentioned as the rationale for this presentation. Getting into the habit of documenting your work carefully yields strong positive benefits, makes you more efficient and increases your overall level of professionalism. It's worth the effort, therefore, to make documentation an integral part of your GIS practice.

Presenter Details

Email: david@dhowes.com

Phone: 253-312-3441

Job Title: Geospatial Information Scientist

Agency and/or Department: David Howes, LLC

Biography (100 word max.):

David specializes in the development of GIS tools, processes and supporting infrastructure for a variety of clients from small operations to multinational corporations (www.dhowes.com). With 23 years of academic and private sector experience in the UK and US, including an M.Sc. in GIS from the University of Edinburgh and a Ph.D. in Geomorphology from SUNY Buffalo, David's background is well-suited to developing innovative solutions to spatial problems. He is the founder of the Lone GIS Professional Initiative, helping GIS professionals working on their own or in small groups help each other, and is a Washington URISA Board member.