Quality Assurance II break-out session
January 13, 2010

Attendees:

USGS	*Other Government*	*Commercial*
David Terrill	Tim Trainor	Mike Thompson
Larry Moore (notes)	Becky Gleim	Allen Carroll
Jenny Runyon	Tom Terry (moderator)	Paul Harwig
Katrina Burke		
Morgan Bearden		

Moderator Tom Terry passed out a pre-prepared handout (last page), consisting of a matrix of quality attributes by National Map layer. The handout was a starting point for discussion, and didn't get completely filled out. Most discussion revolved around the questions What is quality? How do you assure it? And which aspects of quality are most important?

Tom Terry observed that on the old topo maps, rigorous content and quality standards were followed, and we were assured that all quads met the same quality standards.

Larry Moore said he is convinced this old model is obsolete in the context of partner-supplied data, and certainly in the VGI context. We need a new model of quality, but it isn't clear what that should be.

Allen Carroll said National Geographic sees the issues shifting in the same way as for USGS. Most people haven't thought about this to the same level we are thinking about it now. In the right context, having thousands of people providing content raises quality.

Tom said that certain layers probably aren't appropriate for VGI, such as elevation and cadastral. Can we say that some datasets meet standard and some are best available data? The consensus of the group was that this was true, and the following categories were created, with the caveat that these declarations are valid only for the types of data and maps of interest to the USGS National Geospatial Program:

- Structures, transportation, VGI can definitely contribute.
- Hydro, maybe.
- Geographic names and land cover, VGI can provide supplemental support.
- Boundaries, elevation, cadastral, ortho are for the most part not appropriate for VGI.

Paul Harwig and Becky Gleim said that all quality considerations are use-case sensitive. The definition of quality depends on what the map or data will be used for.

Tim Trainor said that feature-level metadata is possible today, so we don't have to make blanket statements. In the context of VGI, feature-level metadata might very well be a requirement.

Alan said, consider the OSM model. If a feature persists over time, does that make it more credible?

Tom said, unless there is a safety-of-life issue involved, does it really matter whether or not a particular feature is accurate, or if the accuracy is measurable? In the OSM model, you just wait to see if someone detects an error, and let them fix it.

Tom encouraged the group to assign quality priorities to the basic TNM themes. On a scale of 1-5, with 5 being the most critical need for quality assurance, the group assigned the following values to quality importance:

- Hydro -- 4.
- Elevation (contours only; no spot or monumented points) – 5
- Orthoimager -- 5. However, this isn't a practical issue, as modern orthoimagery is almost uniformly very high quality.
- Geograph names – 3. Not life-critical, but USGS is responsible agency. There is an existing body of law, regulation, and procedure, which uses volunteered information and is consistent with involvement of volunteers on a larger scale. Among most of our users, names don't really register as a critical layer. "Geographic names" refers to natural features of the land and populated place names, and does not include street names.
- Land cover – 1, with some dissent. Tom related an anecdote about a USFS map that didn't contain any forest green. Paul noted this is another example of intended use. David said that for general-pupose maps, land cover isn't critical, and data sources are pretty course anyway. Within broad limits, what we have is useful enough. Big categories (timber, urban) can be taken from LANDSAT. VGI can probably help a lot with orchards, plantations, swamp, etc.
- Boundaries – VGI not very especially relevant.
- Structures – 4. Larry said this should be high, but the reality now is that quality and quality assurance is very low for the country as a whole.
- Transportation – 5.

Is data of unknown quality better than no data? This might be a false dichotomy…or maybe not.

The following was put to the group by Larry.

> Hypothesis: VGI will improve quality in the long run, but at any given time, quality is variable and partially indeterminate.
>
> Question: If this is true, are we willing to accept this trade-off in the National Mapping Program?

The group agreed that the hypothesis is probably true, and the answer to the question is Yes.

Alan said that in the long run, quality of VGI might be quantifiable. Over enough time, with enough volunteer input, we can make definite statements about quality.

Worksheet used for session, partially filled in.

Quality Assurance Session II

From a User Perspective:

Goal: Develop a list of user requirements related to data quality. How important is it to users to have knowledge of data quality. What is the priority of this information vs currentness and completeness of the data?

		Rank relative importance of:						
	How important do you consider an assurance of quality from the data provider	Currentness	Completeness	Positional Accuracy	Attribute Accuracy	Are there some features that have more stringent accuracy requirements than others?	Is data of unknown quality better than no data? In what situations?	Other topics?
Hydro	*4*			*x*				
Elevation	*5*							
Orthoimagery	*5*							
Geo Names	*3*			*x*				
Land Cover	*1*			*x*				
Boundaries	*NA*							
Structures	*4*			*x*				
Transportation	*5*			*x*				

Other topics / issues: