

BIG DATA Summary
Earl F. Burkholder – May 2014

BIG DATA is a generic term applied to aggregation of electronic information that is typically stored for future use and re-use. Of the many kinds of BIG DATA out there, these comments are addressed specifically to spatial data – that is, information that is location based and referenced to some framework.

- I. Kinds of BIG DATA
 - A. Spatial data
 - B. Economic data
 - C. Population/genealogical data
 - D. Medical data
 - E. Legal/legislative/administrative data
 - F. Sociological/cultural data
 - G. Scientific/mathematical data
 - H. Other
- II. Data are generated by:
 - A. Measurement sensors –
 - 1. Fundamental physical quantities
 - 2. Converted to electronic bits and bytes
 - 3. Referenced to standard models
 - B. Purchase transactions
 - C. Conversations/telephone/internet traffic
 - D. Prescriptions/usage
 - E. Court/legislative/administrative proceedings
 - F. Conferences/meetings/directives
- III. **Spatial data are used for**
 - A. Land records/ownership**
 - B. Engineering design/construction**
 - C. Photogrammetric mapping/remote sensing (LiDAR etc)**
 - D. Transportation/aviation/navigation - drones**
 - E. Mining – oil/gas/coal/minerals - both underground and open pit**
 - F. Economic/commerce modeling**
 - G. Population migration studies**
 - H. Resources inventory**
 - I. Other . . .**
- IV. Some of the biggest issues/challenges include:
 - A. What reference frame/system is used – absolute/relative?
 - B. Are the data referenced to flat-earth or real-world?
 - C. What is the integrity/lineage of the data?
 - D. What is the best way to describe/handle/use spatial data accuracy?
- V. The global spatial data model (GSDM):
 - A. Defines location geometry with rectangular coordinates world-wide.
 - B. Accommodates a 3-D digital spatial data standard for all disciplines.
 - C. Includes provision for handling spatial data accuracy at all levels.
 - D. See – www.globalcogo.com/challenge.pdf