Proposed Outline for GeoMAPP Intrastate Data Transfer Design

The following outline can be followed for documenting intrastate data transfer for the different types of data included in the demonstration portion of the project: (Local Govt, Ortho, Centralized Vector, Projects, and Digitized maps)

I. Introduction/ Background
   a. Short background on timelines (when systems procured, etc)
   b. One or two sentences on types of data to be moved, sources etc
   c. One or two sentences on hardware/ software associated with the system (i.e., what kind of servers, what kind of storage, size of disk for environment, tools used in process, etc.)
   d. One or two sentences on demonstration dataset sizing
II. Data Preparation Workflow
   a. Describe process of data ingest into central GIS organization from outside entities (state agencies, local govt, feds, etc)
      i. One or two sentences on the data sources (who's providing data, how is it displayed/made available by central GIS)
      ii. What happens when data comes in, for example:
          1. Validating that file opens?
          2. Checking projections?
          3. Checking metadata?
   b. How data is prepared for being archived
      i. Converted to new format? (SDE→ shp or FGDb)
      ii. Any updates to metadata?
      iii. Is it compressed before transferring?
      iv. Are checksums or manifests created?
III. Data Transfer
   a. How is data physically transferred from State GIS to State Archives?
      i. FTP/ Across the network?
      ii. Hard Drive?
      iii. DVD/ CD?
   b. Are Orthos handled differently than vector?
   c. Though not part of the design, transfer/loading rates for different media will be important to capture (see section V).
IV. Ingest/ tracking processes
   a. What happens once the data is received by state Archives?
      i. Are checksums run?
      ii. Are manifests confirmed?
      iii. Is preservation metadata created or is geospatial metadata updated?
   b. How are data managed in the archives environment

i. How are archived GIS data stored?
        1. Are archived files put in file system or in a GIS database?
            a. Describe how folder hierarchy is set up
                i. ISO Categories→ Themes → Years?
                ii. By data creator agency?
            b. Is standalone metadata stored with the record
    ii. How are data tracked or referenced
        1. Is data included in a catalog system?
        2. Finding Aides?
    iii. Are processes set up to validate data over time checksums, etc?
  c. Data Access
    i. How can data be located/ discovered?
    ii. Who has access to the data?
V. Demonstration Validation
  a. Now that are system is designed, how are we going to document the movement of each of these types of data as they go through the process. How are we going to measure success? What info is important to capture?
    i. Checksums?
    ii. Transfer rates?
    iii. File validation?
    iv. Ease of access of getting out of the archives?
        1. Have someone unfamiliar with the system try and find particular datasets?