GIS DATA CONVERSION
QUALITY ASSURANCE AND DATABASE ACCEPTANCE

AN OVERVIEW OF CONCEPTS AND PROCEDURES

Prepared by:

Peter Croswell
PlanGraphics, Inc.
112 East Main Street
Frankfort, Kentucky 40601-1806

2005

SECTION 1
INTRODUCTION

1.1 OVERVIEW

This document contains a general discussion of GIS/AM/FM quality assurance issues and procedures that are important for all data conversion projects. Data conversion project success is dependent upon a sound procedure, independent from the data conversion process itself, to accept and track database deliverables from a designated contractor, perform quality and validity checks, and resolve errors before final data loading. The intent of this document is to describe the overall Quality Assurance (QA) methodology, including the implementation of both manual and automated reviews of database deliverables from designated data conversion contractors.

The quality assurance/acceptance process begins at the time of a delivery by a data conversion contractor. The process assumes that the conversion contractor has put in place sound quality control procedures as part of the data conversion process with an intent of delivering a high-quality database meeting established specifications. The purpose of the independent QA/Acceptance process is to check and validate the data quality to ensure that final deliverables meet contracted terms and quality criteria.

1.2 DATABASE QUALITY CRITERIA

AM/FM/GIS database development is a complex undertaking and a variety of data quality issues must be addressed to achieve the ultimate goal of a high-quality database. Some of the major database quality errors which are the subject to checks and validation are:

- File and Layer Naming: Adherence to standards for naming of files and database layers.

- Map Feature Completeness and Organization: Inclusion of all valid features that appear on source material and inclusion of the features on the proper database layer.

- Graphic Quality: Proper graphic closure and topology, feature segmentation and connectivity, edgematching, and other graphic concerns, including proper adherence to utility rule-base criteria and proper line coincidence of related features (e.g., parcel vs. zoning lines).

- Positional Accuracy: Proper positional placement of map features based on established absolute coordinate accuracy and/or positional accuracy relative to other map features.

- Attribute Completeness: Full population of all database fields for map features.

- Attribute Accuracy: The correctness of entered attributes and proper adherence to database format, established domains, and proper values from source material.

- Symbology, Annotation, and Sheet Format: Adherence to established symbol, line type, color, proper feature labeling and annotation placement, and map sheet layout criteria.

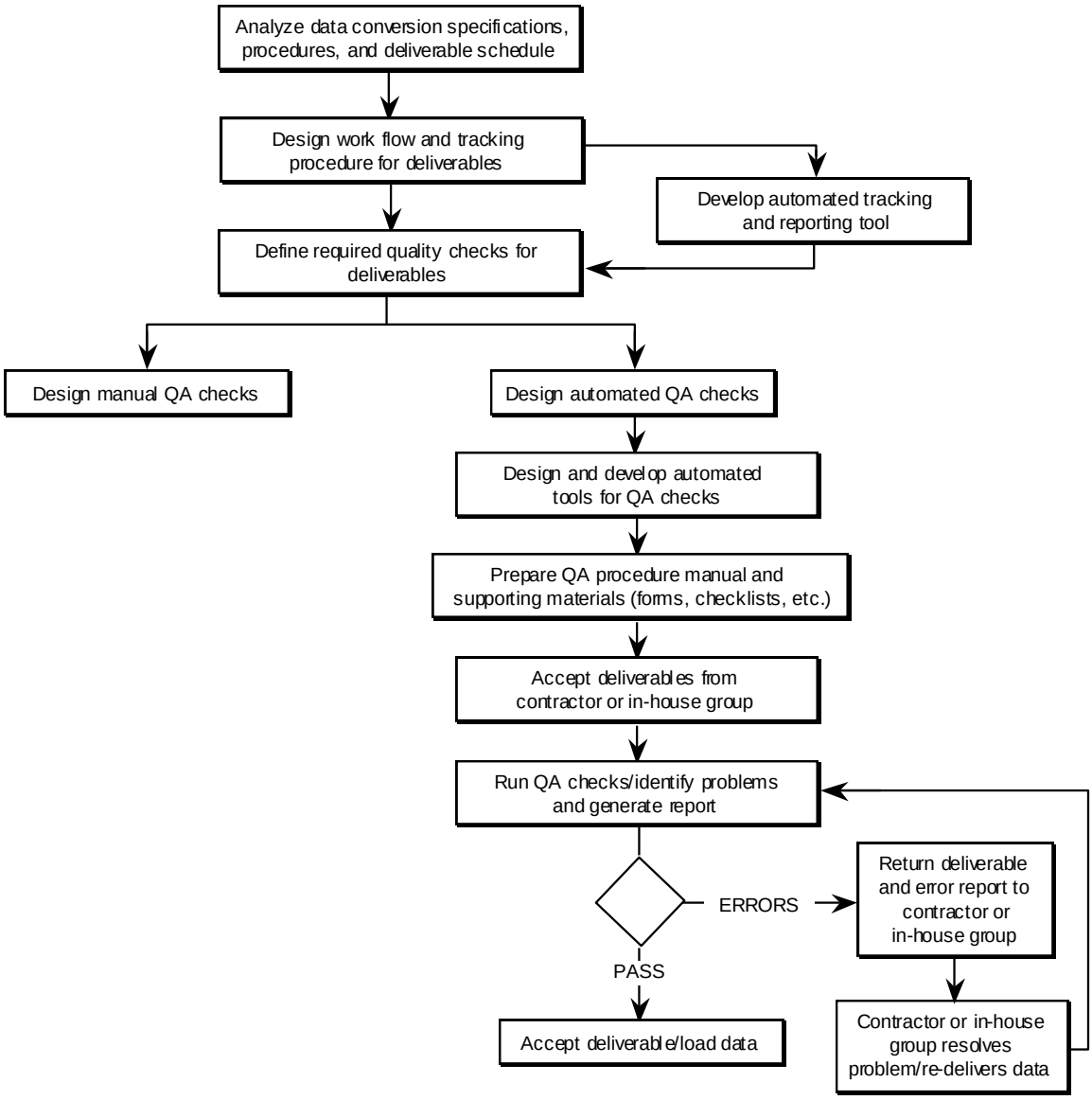
- Image Format and Quality: For data conversion projects involving scanning or drawings, documents, or the production of orthoimagery, this covers such criteria as image rotation and cropping, raster file format, pixel resolution, and overall image quality (contrast, speckles, or artifacts, etc.).

- Photogrammetric Data Quality and Integrity: For projects involving aerial photography and photogrammetric compilation, special checks on the results of aerotriangulation, survey control, and vertical accuracy of digital elevation or contour data.

It is typical for a specification document, developed as part of a data conversion contract, to define data quality criteria requirements under application categories above. This establishes the basis for the QA/Acceptance process.

A general work flow in the QA/Acceptance process is outlined in Figure 1.

Figure 1: Database Quality Assurance and Acceptance



SECTION 2
DELIVERABLE RECEIPT AND TRACKING

2.1 DELIVERY TRACKING AND REPORTING

The QA/Acceptance process is driven by a sound procedure, supported by proper tools to track the receipt of deliverables, report the status of QA checks, and return data to the contractor in the event that errors were flagged. This tracking and reporting process is supported by automated tools, normally a custom database developed in a Microsoft Access environment, to allow entry of status information and the generation of standard reports. It is critical that this tracking identify information at the map sheet level and include important data such as the delivery number; dates associated with receipt or return to the contractor; the types of errors, if any; and action taken to resolve them.

2.2 DELIVERY FORMAT/INVENTORY

PlanGraphics will receive converted data in several formats. Each delivery received from the conversion vendor will be inventoried and verified for completeness by the following procedure discussed below.

Each delivery received from the data conversion vendor is expected to include, at a minimum, conversion files and source documents. Additional sources such as drawings to be scanned, assorted index maps, and any other source sent by the client to the conversion vendor will be verified against the original transmittal. It is anticipated that each delivery received from the conversion vendor will contain areas covered by complete delivery boundaries and that the data in these boundaries will be fully converted. If any problems or discrepancies are found in the delivery, the QA Supervisor will contact the conversion vendor and immediately resolve the problem. If a resolution between PlanGraphics and the conversion vendor can not be worked out, the problem can be escalated to the client for direction.

2.3 RECEIPT CONFIRMATION

Once all source documents have been inventoried, the QA Supervisor will complete a Delivery Receipt Letter informing the client that a delivery has been made, received, and certified as ready for QA.

2.4 RECEIPT REPORTING

A standard form will be designed to communicate the receipt of a delivery from the conversion vendor. This form will be generated by the QA Supervisor at PlanGraphics and forwarded to the client. Appropriate information will be entered into the deliverable tracking database.

SECTION 3
QUALITY ASSURANCE PROCEDURES FOR CONVERTED DATA

This section defines the workflow for the verification of the converted database. The combined use of random sampling techniques and automated tools will result in accurately measuring the overall average quality level for each delivery subset. The use of random sampling techniques provides several advantages. The first advantage is a simple means to ensure that the quality of the conversion process is maintained at the required high standard. The second benefit is a simple inspection that provides timely feedback on the quality of the data as measured in smaller subsets.

Random selection of the features for manual inspection gives an indication of the overall quality of the data being converted. This document is intended to guide the quality assurance process and to fulfill the requirements of the AM/FM/GIS project.

3.1 METHODOLOGY

The quality assurance procedures methodology for the converted data follows.

3.2 AUTOMATED QUALITY ASSURANCE

The Quality Assurance Supervisor will run each program on each delivery of converted data. Each automated step in the quality assurance procedures for converted data is described below.

Since the data being checked by these programs is programmatically generated, this check is run on 100 percent of the data and, in general, must be 99 percent correct for acceptance, although specific error limits may be defined in contractor specifications. Acceptability will be based on the number of errors reported divided by the number of attributes checked.

The types of automated checks used will vary from project to project, but the database quality criteria that lend themselves to automated checks include:

- Graphic Quality
- Attribute Completeness/Rule Base Adherence
- Attribute Accuracy
- Symbology, Annotation, and Sheet Format
- Photogrammetric Data Quality and Integrity.

3.3 MANUAL QUALITY ASSURANCE CHECKS

Some QA checks cannot easily be automated and require manual checks, usually based on a sound random sampling approach, although for some criteria, 100 percent checks may be possible. Quality criteria that are most suited to manual checks include those below. While these checks are manual in nature requiring operator inspection, they are aided by automated tools to facilitate access and viewing of automated files and the recording of error notes by QA technicians. In some cases, manual checking will require the comparison of data to source material using generated checkplots from the digital data.

- File and Layer Naming
- Map Feature Completeness and Organization
- Positional Accuracy
- Symbology, Annotation, and Sheet Format
- Image Format and Quality
- Photogrammetric Data Quality and Integrity.

Some important details about the manual QA checking process are described below.

3.3.1 Overall Cursory Scan

This general visual scan will be an attempt to identify any missing or incorrectly placed data. This check will be conducted at the beginning of the random sampling inspection. The QA/QC Technicians will look over the converted data related to the facility source map currently being inspected, comparing the converted data to the source maps to verify that all features are captured and spatially correct. Any features found to be missing will be recorded on the report. Those objects found to be missing will be added to the entity count for the drawing and then divided by the modified entity total. If the resulting value is .015 or lower (99 percent acceptance), the converted data related to the facility source map being inspected will be considered passed for this portion of the inspection process.

3.3.2 Random Selection

The QA Supervisor will run the random sampling selection program to generate a report consisting of a random sample of converted objects based on the total number of populated record attributes per grid area. The acceptance criteria (allowable errors) are also listed on the report based on the sample size for each design file. The allowable error count is determined by applying the ANSI Single Sampling Plan for Normal Inspection formulas to achieve an Average Outgoing Quality Limit of 1 percent. Using this procedure, the client is assured the conversion data check by this method will be, on average, 99 percent correct. The 99 percent level of quality was chosen as a generally accepted industry standard.

The QA Technician locates the object and reviews the graphical placement. Correct placement is determined by adherence to the conversion specification and subsequent guidelines, the capture of all relevant components, and location relative to other objects. The QA Technician will then verify the accuracy and completeness of all tabular attribution through the comparison of the data preparation attribute sheets and the object attributes reviewed through the object editor menus. Discrepancies are recorded on the QA Review Report sheets.

Attributes for each object are included on the QA review report. The Technicians will check the sources for the delivery (including maps, detail drawings, etc.) to verify the completeness of the data capture.

If the number of allowable errors, as defined by the ANSI Single Sampling Plan for Normal Inspection formulas, is exceeded, the subset is rejected. If the data set does not exceed the allowable errors, it will be considered accepted.

Data converted from a facility source map that fails any of the three inspections (Cursory, Design, and Random) may be returned to the conversion vendor for correction. Inspections may end any time at or after the point of failure.

3.3.3 Missing or Omitted Data

At any time during the inspection process, the QA Technicians will be looking for missing or omitted data. Missing graphic data identified during the Cursory Inspection will be cause for immediate failure. During the Random Sampling Inspection, the QA Technicians will note and account for any facilities prepared on the Facility Source Maps but not captured in the converted database identified during the normal course of the inspection process. These facilities will be added to the overall object and attribute count and will be considered when judging the overall acceptance of a particular tile.

3.3.4 Disposition of Inspected Data

Completion of QA for a data set will be based on standard delineation. QA will be considered complete based on the full inspection of a standard subset area. Once all tiles in a subset are accepted, the area will be recorded as "QA Accepted," and no updates or corrections will be performed in that area prior to delivery to the client. Delivery of partial subsets will be organized based on limitations of the converted data. Delivery of accepted data to the client will be scheduled based on mutually agreeable intervals.

Redelivery of rejected data will begin by reporting the results of the inspection to the conversion vendor. Based on the nature of the errors and what is required to correct the errors, redelivery of rejected data may consist of simply issuing the *QA Summary Report* and supporting documents but may also require the return of all source documents to facilitate the rework process. Decisions regarding redelivery will be made on a delivery subset basis.

3.4 REPORTING

Reporting of errors identified by the QA Technicians will be documented by several means. A brief description of all errors will be recorded on the QA review report. Additionally a screen print and detailed description will be made of any area to clarify descriptions of graphic placement and graphic capture errors.

This documentation will be assembled by delivery subset and delivered to both the conversion vendor and the client for review and corrective action if necessary.