Quality Assurance and Standards for NPSI Datasets

Bruce Riner
NAWC TSD
Orlando, FL
Bruce.Riner@navy.mil

Kerey Howard
NAWC TSD
Orlando, FL
Kerey.Howard@navy.mil

ABSTRACT

The primary objectives of the Naval Aviation Simulation Master Plan Portable Source Initiative (NPSI) are to increase visual database reuse, promote standardization, and lower life cycle acquisition costs for new system acquisitions, legacy platform trainer procurements, and major trainer visual upgrades. The NPSI datasets capture the prepared/corrected/refined visual source data in standard formats for reuse by other platforms. The NPSI datasets include imagery, elevation data, feature data, 3-D models, and metadata. The datasets are stored in the NPSI Archive, which currently contains three NPSI datasets along with additional imagery layers. In addition, there are several procurements underway that will deliver enhanced or new NPSI datasets.

The intent of this paper is to propose quality assurance testing procedures and standards for examining NPSI Datasets for placement into the archive. The quality assurance suite of tests will involve the various layers and the metadata that combine to make a NPSI Dataset. The testing will be utilized to evaluate datasets for compliance, to determine how the data will be archived and to provide information to evaluate the data for future reuse. NPSI datasets, and the results of quality assurance testing, will be made available to contractors at Request For Proposal (RFP) to allow the contractor to better evaluate the NPSI Dataset against program requirements, and make a realistic determination of data quality and potential for reuse, and assess additional effort required for each future program.

ABOUT THE AUTHORS

Bruce Riner is the senior Visual Systems Engineer at the Naval Air Warfare Center Training Systems Division (NAWCTSD) in Orlando, FL. He has a Bachelor of Science degree and Master of Science degree in Computer Science. Bruce has been active in research simulation and training for over 25 years, his specialty areas include visual systems, flight simulators, systems integration, and real-time operating systems. He has also participated in research in the areas of computer graphics, helmet mounted displays, photo-based image simulation, Distributed Interactive Simulation (DIS) networking and forward deployable trainers.

Kerey Howard is an engineer within the Visual and Sensor Simulation Branch at the Naval Air Warfare Center Training Systems Division. He is responsible for providing visual system engineering support to simulation acquisition and R&D programs for the Navy and Marine Corps. He has two years of experience in visual simulation acquisition, and has been working on the NPSI standard for the past 18 months. He received a B.S. in Computer Engineering from the University of Central Florida.

Quality Assurance and Standards for NPSI Datasets

Bruce Riner
NAWC TSD
Orlando, FL
Bruce.Riner@navy.mil

Kerey Howard
NAWC TSD
Orlando, FL
Kerey.Howard@navy.mil

INTRODUCTION

The Naval Aviation Simulation Master Plan (NASMP) Portable Source Initiative (NPSI) primary objectives are to increase visual database reuse, increase standardization, and lower life cycle acquisition costs for new systems acquisition, legacy platform trainer procurements, and major trainer visual upgrades. The NPSI datasets capture the prepared/corrected/refined visual source data in standard COTS formats for reuse (NPSI datasets) by other platforms. An NPSI dataset may include imagery, elevation data, feature data, and three-dimensional (3-D) models. The NPSI data standard also includes support for sensors such as FLIR and Night Vision Devices (NVD), if the program requires material classification. The cost savings and return on investments (ROIs) are achieved through the reuse of costly geospecific imagery, and by capturing the common labor intensive value added effort involved in preparing the various source data layers prior to customization and optimization for a specific image generator's runtime format.

The US Air Force Training System Product Group (TSPG) is using a similar approach for the TSPG Common Dataset Standard (CDS). The USAF TSPG CDS Version 1.0 defines the COTS formats that capture the work in preparing the raw source before the runtime database is produced (CDS, 2006).

The US Army PEO STRI Synthetic Environment Core (SE Core) Database Virtual Environment Development (DVED) is a more ambitious program that also supports the publication of correlated runtime databases. However, the heart of the DVED is the Master Data Base (MDB) which archives the prepared/corrected/ refined visual source data in a form that supports easy import and export in COTS formats (Johnson, Freeman & Perry, 2007).

The approach used by all three services supports exchange of the simulation ready data layers in standard COTS formats for use by the other services.

History of NPSI

The NPSI team was established under NASMP in 2004. The team's first major accomplishment was to develop and publish the NPSI Data Preparation Standard (NPSI DPS) Version 1.0. The NPSI DPS outlined the data layers and formats for the exchange of NPSI datasets. A Draft NPSI DPS Version 2.0 was released in December of 2006, which includes metadata and the schemas for the metadata to assist in using the NPSI datasets. The metadata, or "data to describe the data," includes information such as source, projections, locations, licensing information, and other descriptive data specific to each data element. The NPSI metadata standard is based on eXtensible Markup Language (XML) (Nichols, 2006). The metadata also includes information on material properties.

During 2006 and 2007 the NPSI team has established several Cooperative Research and Development Agreements (CRDAs) with industry to assist in further developing the NPSI DPS to capture additional value added data in formats and structures that are widely used in industry.

The NPSI Archive was established in 2006 and initially populated with the South West USA (SWUSA) and East Coast USA (ECUSA) datasets that were developed for the MH-60 program, and a variety of 3D models from multiple programs. The Straits of Hormuz and Hawaii are more recent additions to the archive. NPSI also has a dataset produced by an AFRL project called Rehearsal Enabling Simulation Technologies (REST) to be used for testing dataset exchanges across the services. The first REST dataset covers the western ranges and includes the NPSI metadata.

The NPSI datasets may contain licensed imagery, NGA data, or other restricted use data. Therefore a Non-Disclosure Agreement (NDA) was developed and is used by NPSI for distribution of the datasets. Through May 2007 we have executed 30 signed NDAs.

The NPSI concept has been incorporated by several programs currently underway at NAWCTSD, including

the MH-60, MV-22, T-45 and C-2. The MV-22 will deliver an enhanced ECUSA and SWUSA dataset in late 2007.

ACQUISITION PROCESS

In order to understand how NPSI affects the acquisition process and the cost savings that can be realized, a brief analysis of the process from the visual system perspective is required. The normal acquisition process is described below:

1. Requirements Analysis – The visual engineer and fleet team establish the visual requirements, including the database extents, areas of interest, minimum resolution of imagery and elevation data, required 3D models and specific training task requirements.
2. The government prepares a Request For Proposal (RFP) which normally includes an appendix for the visual specification and a list of the government furnished information (GFI) that will be provided. Once the RFP is finalized it is published for industry evaluation and proposal.
3. Vendor teams submit proposals, the government team evaluates all proposals and the contract is awarded to the winning team.
4. The contractor (and subcontractors) develop the simulator which normally includes the following visual/sensor database steps:
    - Database design and development monitored through database working groups
    - Visual system and database acceptance tests
    - Final delivery of run-time database and NPSI datasets at Ready For Training (RFT)

THE PROBLEM - EVALUATION OF NPSI DATA CONTENT AND QUALITY

The basic concept of NPSI is to capture the labor intensive effort required to prepare the various source data for publication/compilation to a run-time database, thus allowing data reuse on other programs. In addition, geospecific imagery is archived to be made available for reuse on other programs. The cost to acquire geospecific imagery can be very high, even before considering image preparation costs. Cost savings will be realized when other programs have similar or overlapping area requirements for their visual and sensor databases. However, for cost savings to be realized with the acquisition process described above a few problems must be addressed. Each contractor must be able to determine the effort or cost required for their proposal to meet the program requirements using the supplied NPSI datasets. In order for the contractor to estimate the effort, he must have a reasonable understanding and knowledge of the quality of the NPSI dataset. In addition, the RFP must define the overall requirements for the visual database so that the contractor can reasonably evaluate the amount of reuse of the NPSI dataset against the requirements of the program.

If the contractor assumes little or no data reuse and proposes the full cost of development for the database, little or no cost savings would result if the contract was awarded to that contractor. This scenario may result in contract award to a contractor who projects a higher level of data reuse. The opposite situation involves the contractor assuming too much data reuse. The contractor may then under bid the effort, and be at high risk to meet the program requirements within budget. As a result, the contractor may experience cost overruns, schedule delays, and jeopardize the delivery of a functional training system. To deal effectively with these issues we must develop a process to adequately describe the content and quality of the NPSI data.

Evaluation of NPSI Datasets Against Program Requirements

Each simulation program must evaluate available NPSI datasets to determine the level of correlation between the data and their requirements. Airfield requirements are an example that demonstrates possible enhancements that may be needed to allow an available NPSI dataset to meet a new program's requirements. The current archived NPSI datasets were developed to meet a rotary wing platform's requirements. That particular platform does not require the runway thresholds to be exact, so the airport models were built flat for simplicity in model development and the elevation data was flattened under the airports. A fixed wing platform will likely require the runway thresholds to have a lower tolerance for error. In this situation, the runway models with imagery and buildings should provide some cost savings; however the airport model will require significant work to meet the fixed wing platform's requirements. Another example is low level routes. Different platforms have different requirements for low level routes even within the same database extents. The current ECUSA dataset would require the addition of new low level routes to meet a new platform's requirements. Different target areas and

different sensors for the various platforms will also require the NPSI dataset to be enhanced to meet new requirements.

The examples described above clearly show that a NPSI dataset can meet one or several program's requirements without meeting all requirements for every program. In order for a new simulation program to be successful, the government team and prospective contractors must understand the requirements for all major enhancements to the NPSI dataset prior to proposal development.

The first problem described above deals with the quality of the NPSI dataset the government supplies to the contractor and the requirements of the new program. The next major problem is how the government checks the quality of the new or enhanced NPSI dataset to be delivered back to the government at the end of a contract. The government must ensure the value added work the contractor has done is captured in the NPSI dataset and meets the NPSI DPS. The large size of NPSI datasets further complicates the acceptance or quality assurance testing of the NPSI dataset before it is placed in the archive.

CRDA Results and Quality Assurance

The NPSI team is working with several industry partners using CRDAs to improve the NPSI DPS document and increase the opportunity for reuse of the NPSI datasets. Two of our industry partners have evaluated an NPSI dataset and reported some areas of concern that are listed below:

- Some imagery in a dataset was 'pre-feathered', while other imagery was not
- The imagery in a dataset contains multiple null value colors - most COTS tools are limited to one value to indicate "do not process"
- One scene was not color-corrected to an acceptable level
- Snow imagery was included for one area without the matching summer imagery
- The vector shapefiles for power lines were not consistent:
    - Linears were not connected
    - Points were not equally spaced
- 3-D Models issues:
    - Missing LOD beads
    - Missing materials
    - Missing textures
    - Missing footprints

Several of the concerns listed above are associated with providing quality assurance for the NPSI dataset and the ability of the contractor to reasonably estimate the amount of work that must be done to publish and enhance the NPSI dataset to meet a new program's requirements.

PROPOSED SOLUTION

The proposed solution to realizing maximum cost savings, while meeting program requirements is broken into the several areas:

- RFP
    - RFP Preparation
    - Providing NPSI Dataset for Evaluation
- QA
    - QA of NPSI datasets
    - QA Process

RFP Preparation

The first part of the solution will be for the government team to establish the major requirements for the visual and sensor databases, and then evaluate these requirements against the existing NPSI archive. This effort should include a fleet team evaluation of the key areas of the NPSI dataset. The government team can identify any known required enhancements or new work that must be done. Some examples of additional required effort are listed below:

- Primary airports
- Low level routes
- Areas of interest
- Different imagery resolution requirements for specific areas
- Different geographic extents
- New or enhanced models
- Sensor support

The government will prepare the RFP and include their specific visual and sensor database requirements in the visual appendix. An additional section will be added to the specification describing known issues and/or necessary enhancements to the NPSI datasets. It will be difficult to develop an exhaustive list of every problem or required enhancement, but this section should assist the contractor in evaluating the NPSI dataset.

Provide NPSI dataset

The NPSI dataset should be provided for evaluation to vendors as soon as possible. Several representatives from industry have expressed their opinion that providing the NPSI dataset at RFP, or possibly RFI, would greatly assist them in preparing their proposals, and allow them to develop a more accurate cost estimate of the required visual and sensor database development effort.

The second and key part of the solution is to deal with the quality of the data in the NPSI archive and to accurately describe the data in the archive. The NPSI team developed an XML metadata schema to describe the data to assist with consumers using NPSI data. NPSI-STD-021-2006 NASMP Portable Source Initiative Standard for Reusable Source Dataset Metadata defines the metadata schema.

Brief Description of NPSI Metadata Schema

The NPSI metadata is based on XML. The purpose of the NPSI metadata is to assist the consumer in understanding the NPSI data and its value in producing a run time database to meet a program's requirements. The NPSI metadata schema is used to validate and load the metadata files.

The NPSI metadata schema starts with the Base, which is the root element shown in Figure 1.

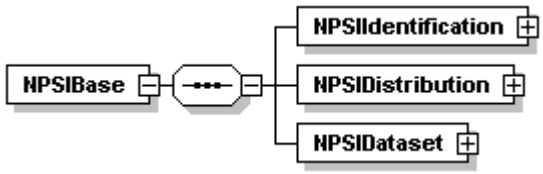

Figure 1. The NPSIBase element

The NPSIIdentification element contains the dataset identification element including more detailed information such as title, version, and license summary. This element is shown in Figure 2:

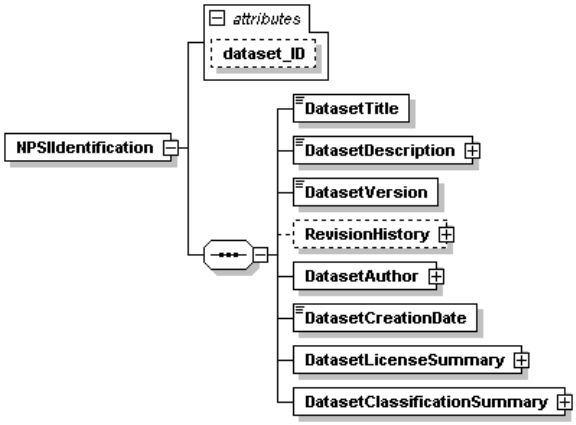

Figure 2. The NPSIIdentification Element

The NPSIDataset element continues to describe the data down to the various layers, such as Raster Data, Vector Data, Cultural Features and the Dictionary. The Generic Raster File Type is shown in Figure 3 as an example.

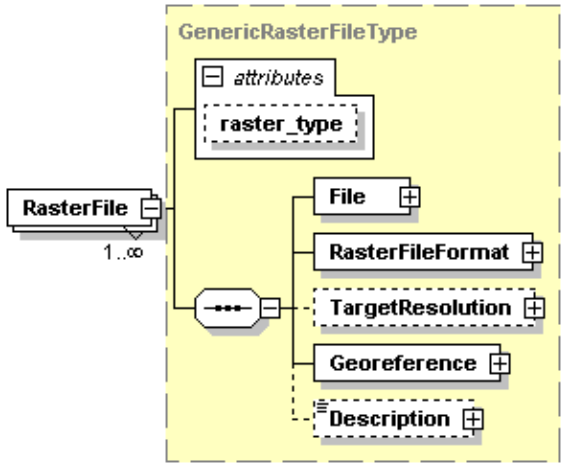

Figure 3. The RasterFile Element

The NPSI metadata describes the NPSI dataset in both a human readable and machine readable format.

Brief description of MPRD Schema

The Material Properties Reference Dictionary was established to provide a common set of material properties and a common mechanism for referencing them. There is currently no standard COTS format for material properties. The MPRD-STD-021-2006 NASMP Portable Source Initiative Standard for Material Properties Reference Database defines the initial schema. The NPSI MPRD is a component of the NPSI metadata architecture. MPRD is based on the

Material Markup Language (MatML) that was initiated by NIST in 1999. As an example, the Material element from the MPRD is shown below:

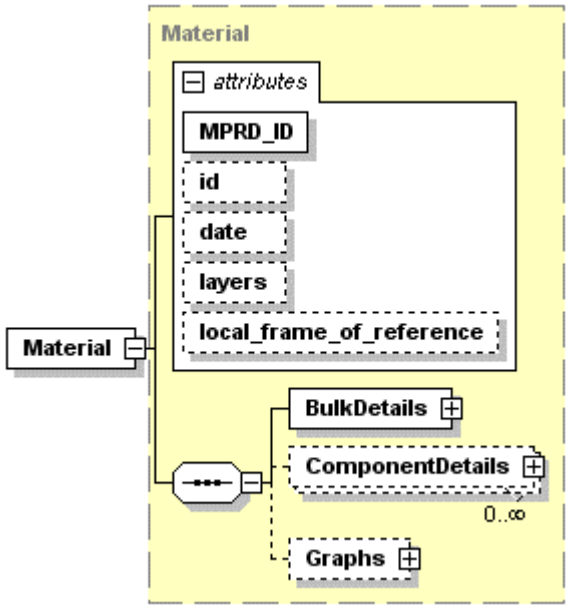

Figure 3. The MPRD Material Element

The MPRD has been included in the REST dataset and will be evaluated by industry and the Government.

QUALITY ASSURANCE OF NPSI DATA

There are three levels of quality assurance that will be used for reviewing NPSI data. The first level of QA is validation by the producer. A producer may be a contractor or a government agency. Quality assurance checklists are under development for the data layers in the NPSI dataset and NPSI metadata. These producer checklists will set a standard for initial data evaluation and allow the generation of a qualitative overview report for future assessment. Checklists may be focused at producers to assist in best practice database generation, and if used could also assist the government during database reviews and IPT meetings.

After the contractor has completed the database the second level of quality assurance will be program acceptance testing. This generally includes database review allowing the content to be verified by the Fleet subject matter experts and government visual engineers. After any database discrepancies generated during acceptance tests have been closed, and the final database is found acceptable, the required NPSI datasets will be produced. Compliance tests will be performed on the data to insure that it meets NPSI format requirements and the content will be validated for completeness. A small test sample may be required to verify reproducibility. The test sample may be used to generate a database, and may be compared visually to check for any discrepancies when compared to the final accepted database. While rapid database development may produce minor artifacts, the focus will remain on reproducibility and content. The final test for acceptance will be verification that the entire gaming area has been provided.

When the database has been accepted and the NPSI data delivered, the final level of NPSI quality assurance tests will be performed. Delivered NPSI data will be of two types: new and modified. In either case, data that has been provided back to the government will have decisions made about data quality, and a determination made to add data to the archive as new data or to replace existing archive data. NPSI metadata will assist in change detection at this level of review. The focal point of the review will remain around the areas of interest defined for the specific program; however, all areas can be reviewed by a high-level fly through.

Producer QA Sample Checklist

The Multispectral Environment Engineering Team (MEET) at Patuxent River has played a key role in developing and implementing NPSI, and has developed some quality assurance guidelines for data producers. A brief sample of the checklist developed for Open Flight airports is shown below:

- Master file
  - Texture Paths
  - Hierarchy
  - No Concurrent Vertices
- Runway File
  - No unused textures
  - Hierarchy
  - LOD
- Buildings
  - Light groups
  - No Concave Faces
  - Footprints
- Lights File
  - LOD
  - Lightgroup codes

Future NPSI Producer Checklist Development

The sample checklist above, and the foundation QA effort by MEET, are good starting points but the checklist and process must be enhanced to include all

layers of the dataset. The following areas need to be addressed:

- Validating Metadata and required fields
  - NPSI
  - MPRD if required
- Imagery
  - Projection - Lat Long
  - Resolution
  - Color balanced
  - Different seasons if required
- Sensor data
- Shape files
  - Basic correlation
  - Properly attributed
- Elevation data
  - Correlation
  - Projection
  - Format
- Models
  - Open Flight Format
  - Textures
  - Material Encoded if required
  - Lighting
  - LOD
    - Airports
    - Fixed models
    - Moving Models

Parts of the producer QA process listed above can be automated. As an example, a check for required fields in shape files or in metadata can be automated. Obviously wherever an automated process can be developed, it will be capable of a much more thorough evaluation of large datasets. However some quality checks will require manual review and validation of the data.

NPSI Dataset Acceptance Testing

For database developments on current contracts, contractors are required to deliver the NPSI datasets at RFT. The government will evaluate and accept the final runtime database before the NPSI data is produced. The acceptance level QA of the dataset will verify that all layers of the dataset meet the requirements and formats defined in the NPSI DPS. The government program's visual engineer will verify the content of the NPSI data:

- Airfields
- Moving Models
- Imagery
- Elevation
- Shape files

The data will be analyzed for compliance with the NPSI specification. Due to the large size of the dataset an exhaustive check will not be possible. However, all detailed areas of interest such as primary airfields, moving models, and target areas will be evaluated thoroughly. The data will be checked for completeness and a test data set will be evaluated.

NPSI Archive QA Process

After an NPSI dataset is accepted by the government it will be evaluated for inclusion in the NPSI archive. To begin data analysis, a high level fly through would be created to quickly look at the mosaicked data files. Data noted as changed or updated would be reviewed using COTS tools and evaluated against existing archive data. The data will be flagged as improved, or no change. Any data discrepancies will be flagged for correction. While color balancing and image normalization are key criteria for image analysis, image registration is of the highest priority. Likewise, elevation and vector data should be registered to the database imagery data. Vectors should be properly attributed using a common standard such as DIGEST FACC, Census CFCC, or SEDRIS EDCS. Configuration management will be essential for archive success.

For the initial datasets the government will be responsible for producing or upgrading the metadata to the NPSI-STD-021-2006 NASMP Portable Source Initiative Standard for Reusable Source Dataset Metadata schema, since the NPSI DPS version 1.0 did not include the previous metadata definitions. Future contracts will reference the latest NPSI DPS version. The contractor will be responsible for producing the NPSI metadata for any required NPSI deliveries.

Industry Participation in NPSI DPS and QA

The NPSI DPS and the quality assurance testing will evolve and improve as we gain more experience and insight over time. Industry involvement in this process is necessary to allow us to improve the quality of data available for reuse. Together we can deliver the best and most cost effective solution for the war fighter.

SUMMARY

The NPSI standard has been adopted by multiple Navy Aviation platforms. The adoption of the standard has

provided the following results.

- Cost savings / Cost avoidance (Haberman, 2007)
- Larger gaming areas
- NPSI archive of datasets
- Archive of airfields
- Archive of moving models
- Capability to produce multiple runtime databases from the same source data for correlation.
- Rapid production of a runtime database from an existing NPSI dataset.

The NPSI archive will continue to grow as more NPSI datasets are delivered back to the government from the various programs with different requirements. As NPSI datasets are reused on other programs and continue to be enhanced, the quality assurance process will be used to improve the quality of the data. As various discrepancies are detected they will be documented and corrected. In order to achieve these results, an initial process has been developed to QA the data and to provide NPSI datasets during RFP/Proposal development. The QA process will continue to mature as we get feedback on the initial QA results, and gain expertise in QA process.

The final result will be a growing NPSI archive with datasets of known quality.

ACKNOWLEDGEMENTS

We would like to thank NAVAIR PMA205 for sponsoring the NPSI effort under the NASMP umbrella. We would also like to thank the CRDA industry partners for their valuable input and effort as well as the NAWCTSD Command and the legal department support in creating the CRDA partnerships. Finally, I we would like to thank the NPSI team for their efforts in making the NPSI program a reality.

REFERENCES

Haberman, F. (2007). Naval Aviation Simulation Master Plan Portable Source Initiative (NPSI) A Study in Software Reuse Success, and Positive Return on Investment (ROI), *Human Systems Integration Symposium (HSIS).*

Jacobs, M. (2004). Portable Source Initiative– An Industry Perspective. *Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC).*

Nichols, W. K. (2005). The NASMP Portable Source Initiative - Shifting the Contracting and Acquisition Paradigm. *Interservice/Industry Training, Simulation and Education Conference (I/ITSEC).*

Nichols, W. K. (2004). The Portable Source Initiative - Building Reusable Databases. *Interservice/Industry Training, Simulation, and Education Conference(I/ITSEC).*

Nichols ,W. K. (2006). Building Cross-Platform Sensor Databases The Portable Source Initiative. *Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC), December 2006.*

Johnson, M., Freeman, J. & Perry, C. (2007). SE Core DVED – An Introduction to the Standard Rapid Database Generation Capability (STDGC), *IMAGE Conference.*

Sieverding, M. (2006). Almost Accidental Database Standards. *Image Conference.*

Stokes, T. & Stephens, B. (2005). Using Open Industry Standards to Share Visual Simulation Databases, *Interservice/Industry Training, Simulation, and Education Conference(I/ITSEC).*

NPSI-STD-021-2006 NASMP Portable Source Initiative Standard for Reusable Source Dataset Metadata. *NAVAIR, June 2006.*

MPRD-STD-021-2006 NASMP Portable Source Initiative Standard for Material Properties Reference Database. *NAVAIR, June 2006.*

USAF Training Systems Product Group Common Dataset Standard Version 1.0 -1, TSPG, December 2006.