DATA COLLECTION GUIDELINES FOR CONSISTENT EVALUATION OF DATA FROM VERIFICATION AND MONITORING SAFEGUARD SYSTEMS

Roberto Lenarduzzi
Oak Ridge National Laboratory
P. O. Box 2008
Oak Ridge, Tennessee 37831-6003
(423)574-5687

Kim Castleberry
Oak Ridge National Laboratory
P. O. Box 2008
Oak Ridge, Tennessee 37831-6003
(423)576-0205

Michael Whitaker
National Security Program Office
Oak Ridge, Tennessee 37831-8206
(423)576-1682

ABSTRACT

One of the several activities the International Atomic Energy Agency (IAEA**)** inspectors perform in the verification process of Safeguard operations is the review and correlation of data from different sources. This process is often complex due to the different forms in which the data is presented. This paper describes some of the elements that are necessary to create a 'standardized' structure for the verification of data. When properly collected and formatted, data can be analyzed with off-the shelf software applications using customized macros to automate the commands for the desired analysis. The standardized-data collection methodology is based on instrumentation guidelines as well as data structure elements, such as verifiable timing of data entry, automated data logging, identification codes, and others. The identification codes are used to associate data items with their sources and to correlate them with items from other data logging activities. The addition of predefined parameter ranges allows automated evaluation with the capability to provide a data summary, a cross-index of all data related to a specific event. Instances of actual databases are used as examples. The data collection guidelines described in this paper facilitate the use of data from a variety of instrumentation platforms and also allow the instrumentation itself to be more easily applied in subsequent monitoring applications

INTRODUCTION

A lot of effort has been and is being applied to the development of instrumentation to measure and log parameters and events that will aid in the verification of international safeguard operations. In order to be applied to safeguard monitoring these measurements are usually designed not only to provide the actual data from the measurement, but also to use the same data to verify the validity of another measurement or of a process declaration entered in a logbook.

Most safeguards systems are developed through a coordinated effort from several teams with each specializing in a particular monitoring discipline. A complication arises from the fact that the teams are almost always from different organizations and have their own established procedures for data analysis and presentation. To further complicate the process, the recipient of the results may also be a team that has its own set of format expectations.

In the monitoring and verification of safeguard operations, the team of inspectors must analyze data gathered by all of the teams. To accomplish their task, the inspectors must sort and correlate results form the various measurements even though the results are often presented in different formats (tables, graphs, photos, etc). Another difficulty that is often encountered is the lack of consistency in the basic parameters, such as time of collection, event names, data source identification, etc.

There are already visible trends in the data acquisition and automation industries to standardize the structure of the data being collected or the commands required to automate processes. The following section describes an example of such a data correlating effort. Subsequently a few, guidelines are presented on how to plan data collection to facilitate automated analysis, correlation, and results presentation.

EXAMPLES OF DATA REVIEW

Early in the verification of Highly Enriched Uranium (HEU**)** down blending experiment at the Portsmouth Gaseous Diffusion Plant, it was realized that the effort required from inspectors to correlate data from the various measurements and logbooks would be complex. One of the tasks the IAEA inspectors performed was the verification of the amount of HEU removed from feed cylinders. To accomplish this, they needed to correlate information from 5 different data sets:

- The operator declarations manually entered into a 'mailbox' computer database, which was not done in real time because of computer accessibility issues.
- Weight measurements from the Load-Cell based Weight Monitoring System (LCWS), which continuously logged the weight of the material in the cylinders at the feed station.
- Still photos from the video surveillance cameras, which stored images triggered by motion in the feed area or by sudden large changes in weight detected by the LCWS and at specified intervals.
- The spot checks on the HEU feed cylinder weights performed by the inspectors.
- Non-Destructive Assay (NDA) measurement of HEU feed cylinder also randomly performed by the inspectors. [1]

The change of feed cylinder provides a good example of how the correlation was accomplished. To verify this event the inspectors:

- printed the operator log entries from the mailbox computer, which provided a description and data related to the entire activity;
- compared the log entries with data from the LCWS along with the list of detected events caused by weight perturbations (old cylinder removed from platform, new cylinder placed on platform, operator stepped on platform, etc);[2] and
- reviewed the images from the video surveillance.[3]

In the above operations, each of the data records had a time stamp, but the correlation was hampered because the various system clocks were not synchronized well. In addition the times listed in the operator log were usually related to the beginning of an operation, and these were frequently off by several minutes with respect to the time listed in computer-logged data.

Another correlation example involves the spot-check weight measurement performed by the inspectors. These measurements had to be compared to both the printed operator log entries and the weight data provided by the LCWS.

It is evident from these brief examples that the inspectors needed to understand the structure of each measurement database in order to correlate the proper data and thus verify the validity of the claimed operations. Since verifications requirements can vary greatly from site to site, starting the process can involve a significant learning curve for the inspectors.

SUGGESTED DATA COLLECTION AND PRESENTATION GUIDELINES

In designing measurement systems for safeguard operations, significant computing power is often applied to collect large amounts of data, but very little thought is put into using the same power to transform this data into useful information. The following paragraphs suggest a few guidelines that can facilitate the use of computers to automate the correlation of data sets. Although the implementation of some of the guidelines may be impractical or not cost effective, their intent can help in planning a more efficient data structure.

1. Off-the-Shelf Software Readable Data

All electronic data should be stored in a format that is readable by commercial off-the-shelf software such as spreadsheet applications (e.g. Microsoft Excel). Tab delimited ASCII files can be imported into spreadsheet applications a well as text editors, and word processors. This file type affords the widest possible capability with existing office automation software. With the addition of custom macros, most off-the-shelf applications can perform repetitive and even complex operations with ease.

2. Database Record Structure

Planning the structure of data records is usually necessary if automated analysis is desired. Even though monitoring or experimental data needs will vary significantly from task to task, there are usually some common entries, such as date, time, location, experiment ID, etc., which can be positioned consistently in the record sequence. In addition, the depth of decimal resolution for numeric data should be specified, as well as the format for basic data and for the time stamp associated with events and measurement readings.

3. 'Standardized' Variables

Data variables representing the same parameter should be consistently identified and represented. While it is typical for a parameter such as weight to be identified with the same name in different experiments, the use of Metric units in one case and English units in another should be avoided.

Industry has encountered compatibility problems in large data acquisition and system automation projects where products from different manufacturers must be integrated into a single monitoring or control system. This is especially true in cases where it would not be practical or cost effective for a single manufacturer to produce the entire system.

An example in variable standardization that can be referenced as a model is the definition of standard variables developed by Echelon Corporation, creator of LonWorks. These standard variables were defined in conjunction with a series of companies that have adopted the use of Echelon products. LonMark is the association of LonWorks based product manufacturers and users, and their goal is to establish guidelines for product interoperability at the application layer. Part of these guidelines is a list of Standard Network Variable Types (SNVT) that are associated with measurements, status information, or commands to be exchanged among products. These variable types are identified by a uniquely assigned number. Table 1 shows the entries for a few of the presently defined variables.

The total list is managed and published by the LonMark association. There is also a published petitioning process for the official approval of new data types. The related interoperability guidelines along with the variable type definitions have been in effect for several years and have probably had a significant positive influence in helping the automation industry to remain competitive.[4]

Table 1 Example of entries in LonMark's Master SNVT List

Measurement	Name	Range (resolution)	SNVT #
Mass	SNVT_mass	0..6,553.5 grams (0.1 g)	23
	SNVT_mass_f	0 .. 1E38 g	56
	SNVT_mass_kilo	0..6,553.5 Kg (0.1 Kg)	24
	SNVT_mass_mega	0..6,553.5 metric tons (0.1 ton)	25
	SNVT_mass_mil	0..6,553.5 milligrams (0.1 mg)	26
....			
Pressure-gauge	SNVT_press	-3,276.8 .. 3,276.7 kilopascals (0.1 kPa)	30
Pressure-absolute	SNVT_press_f	0 .. 1E38 pascals	59
Pressure-gauge	SNVT_press_p	-3,2768 .. 3,2767 pascals (1 Pa)	113
....			

4. Data Source / Measurement Identification

In general, each sensor/transducer and data source (manual or automated) should be assigned a unique identification (ID). This ID is helpful for two important reasons:

- A history about that particular sensor can be kept regarding its performance, calibration status, especially useful if these same sensors are subsequently used in other installations.
- The data collected from a sensor can be more easily related to a particular location.

The approach addressed by the committees working on the family of IEEE standards regarding the Smart Transducer Interface for Sensors and Actuators can be used as a model for this identification definition. The committee's original goal was to define how sensor activation and polling could be accomplished over a network. Part of this effort is the development of TEDS (transducer electronic data sheet) which is a read only table included or associated with a sensor or transducer. It completely defines the properties of that device (referred to as channel TEDS) and the related association in the system (referred to as MetaTEDS). The fields of TEDS and how to access them are defined in the standard.[5]

5. Synchronization of Data Collection/Logging Times

Correlating data form different sets relies heavily on the correct time of data entry. This element is probably the most difficult to implement, and to be effective it may require more effort in its definition than in its implementation. The basic requirement is that the clocks of all data collecting systems, whether automated or manual, must be synchronized with 'one' reference clock. The internal clock of all data-collecting equipment, computers, or other instruments must somehow be synchronized with a 'project' reference clock.

The second requirement is that a necessary synchronization resolution be determined at the beginning of the project. It may be impractical to expect very fine time resolutions, especially if some of the data is manually entered. A time slot resolution, which in some cases could be in minutes, can be defined; e.g. a 5-minute slot resolution for a set of measurements. In this case any data from the associated databases falling within this 5-minute period could be easily retrieved, correlated and verified. If a fine resolution can be achieved, the time slot may include fewer events to be sorted and correlated. For activities lasting longer than the defined resolution, an event start time and an event end time should be reported. Alternatively, the activity should be broken down into sub activities each with their own reported times.

6. Association Table

To facilitate the automated correlation of data sets, it is necessary to define an association table for the measurements. For example, in the data collected at Portsmouth, the load cell data supported by the images of the surveillance cameras was used to verify the operator declarations in the mailbox computer. Thus, an association had to exist among the three data systems. To verify the change of a cylinder for feed station number 5, the inspector would select the time slot declared by the operator, and the master computer application would search the associated databases and find the corresponding load cell data and the surveillance images for the event in question.

7. Minimization and/or facilitation of manual data entry

Whenever there is manual entry there is a potential for human error. Therefore, whenever possible, manual data entry should be minimized and replaced by automated or semi-automated means. This is especially important with regard to the operator log entries. Some examples of automation are:

- Unchanging information that is entered with every record should be bar-coded wherever possible. Items such as data collection station, instrumentation ID number, etc. could fall into this category. The bar code could be located either at the instrument or in a folder by the log computer. The operator can scan the bar code just prior to logging an entry.

- If a hand-held scanning and logging device is available, the operator or inspector could scan the ID and enter the related data at each measurement location. Such a device also records date and time automatically and provides good time synchronization with automated readings. The data from the hand-held device would periodically be transferred to the master computer via a wired or infrared link.

CONCLUSIONS

Current safeguard systems typically collect data in format that is measurement or logging system dependent. Since data in a variety of formats does not lend itself to general automated analysis, either special software applications have to be developed to correlate and analyze such data, or the task is performed manually. The data collection and formatting guidelines presented here are intended to assist project teams to structure their recorded data in a widely compatible format. Even if direct implementation of each guideline is not possible, the information presented can help promoting the adoption of consistency in recorded data sets. The main intent here is to promote data compatibility between measurements and systems. It is important to realize that selection of instruments that produce data formatted according to some standard will make that instrument more reusable in other installations. The overall objective is minimization of project costs and the reduction of the learning curve of the verification authority.

ACKNOWLEDGMENTS

The authors thank Melissa Byrd for the effort and support she provided in getting this paper ready for publication.

REFERENCES

1. D. M. Gordon, M. Subudhi, O. L. Calvert, T. N. Bonner, J. G. Adams, R. C. Cherry and N. E. Whiting, *IAEA Verification Experiment at the Portsmouth Gaseous Diffusion Plant*, INMM 39th Annual Meeting, Naples, FL, July 1998.

2. R.L. Martinez, K. Tolk, N. E. Whiting, K. Castleberry, R. Lenarduzzi, *Verification Experiment on the Downblending of High Enriched Uranium (HEU) at the Portsmouth Gaseous Diffusion Plant*, INMM 39th Annual Meeting, Naples, FL, July 1998.

3. R. Lenarduzzi, K. Castleberry, M. Whitaker, R. Martinez, *An Integrated Video- and Weight-Monitoring System for the Surveillance of Highly Enriched Uranium Blend Down Operations*, INMM 39th Annual Meeting, Naples, FL, July 1998.

4. Echelon Corporation, *The SNVT Master List and Programmer's Guide*, Document # 005-0027-01 Rev J, March 1996

5. IEEE Std 1451.2-1997, *IEEE Standard for Smart Transducer Interface for Sensor and Actuators-Transducer to Microprocessor Communication Protocols and Transducer Electronic Data Sheet (TEDS) Formats*, IEEE, NY, NY, Sept 25, 1998.