Y-12

OAK RIDGE
Y-12
PLANT

SGML Metadata for Nuclear Weapons Information

James David Mason

April 15, 1998,

Preprint for submission to
InForum 1998
Office of Scientific and Technical Information
Oak Ridge, TN
May 6–7, 1998

Prepared by the
Oak Ridge Y-12 Plant
Oak Ridge, Tennessee 37831
managed by
Lockheed Martin Energy Systems, Inc.
for the
U.S. DEPARTMENT OF ENERGY
under contract DE-AC05-84OR21400

MANAGED BY
LOCKHEED MARTIN ENERGY SYSTEMS
FOR THEUNITED STATES
DEPARTMENT OF ENERGY

SGML Metadata for Nuclear Weapons Information

James David Mason
Lockheed Martin Energy Systems
Information Technology Services
1060 Commerce Park, M.S. 6480
Oak Ridge, Tennessee 37831-6480
masonjd@ornl.gov
423 574 6973
http://www.ornl.gov/sgml/WG8/wg8home.htm

ABSTRACT

With underground nuclear weapons testing ended, maintenance of the nation's stockpile depends in large part on tracing the as-built parameters of individual units. Besides the construction histories, there are other records associated with weapons design, ranging form technical reports on design issues to the actual construction procedures and design drawings. All of these together constitute an immense base of records that must be maintained and catalogued. Within the Department of Energy, a general application, defined in a SGML Document Type Definition (DTD) for records metadata has been developed to help with metadata collection across many installations. To meet the more particular requirements of the Oak Ridge Y-12 Plant, additional DTDs are being developed, and a catalog of records will be developed using these DTDs. Primary searching of the catalogue within the Plant will be of this metadata. The records will be converted to the more generic DTD for export to other sites. Within the Plant the metadata will eventually also be linked to actual records in an electronic library.

James D. Mason, originally trained as a mediaevalist, is a systems architect in the Information Technology Services organization of Lockheed Martin Energy Systems. Dr. Mason has been active in standards development since 1981 and has convened ISO/IEC JTC1/SC18/WG8, the group responsible for SGML, since 1985. For his work on SGML, Dr. Mason has received the Gutenberg Award from Printing Industries of America and the Tekkie Award from the Graphic Communications Association.

Dr. Mason is leader of the SGML Publishing project in the Information Technology Services organization of Lockheed Martin Energy Systems in Oak Ridge, Tennessee. His other interests involve weapons records management and manufacturing process improvement.

Weapons Building and Records

With the termination of underground nuclear weapons testing, the Department of Defense and the Department of Energy must maintain the nation's enduring stockpile on the basis of data about the weapons as they have actually been built, as opposed to how they were designed. For the foreseeable future, computer simulations will replace underground testing, and this calls for accumulating and analyzing even more as-built data as simulation input. The primary repositories of as-built data are the DOE manufacturing facilities such as the Oak Ridge Y-12 Plant. The simulations, however, will be run by the research and design laboratories, such as those at Los Alamos, Lawrence Livermore, and Sandia.

In many ways, nuclear weapons are typical high-technology manufactured products. Their design and production involves metallurgy and materials science, and the process includes typical industrial processes like casting, forming, and precision machining. Fabrication processes are followed by quality-assurance processes like inspection and radiography. Except for the particular materials used and the requirement for high levels of security, a weapons factory differs little from other precision manufacturing facilities. Weapons manufacturing does differ from other manufacturing, of course, in that it has never been possible to pull a sample off the production line and crash test it like an automobile. Even in the days when undergournd testing was permitted, test devices were uniquely designed and specially instrumented.

In such a manufacturing environment, quality assurance and collection of data from the manufacturing process becomes especially important. Every batch of raw material is analyzed for its composition and material properties. Machining and assembly are guided by exacting procedures. Each component coming from machining must be measured and certified on computerized coordinate-measuring machines before it can proceed to the next assembly step. As subassemblies come together, they are tested further, and radiographs (X-rays) are made after almost every major step. All of the procedures and design drawings for each step are kept under configuration control, and certification data is coordinated with the versions of the process documents. Assembly histories for regular production units, for this reason, run to several hundred pages, and those for special test devices can be even longer.

The manufacturing and maintenance process has always been records intensive. Records exist for all the numerous warheads in the U.S. stockpile. Some of these units are actually quite old, and their records exist only on paper. More-recent units have more electronic records, but even for the latest production some records (such as machinists' checklists) exist only in paper.

Although the products of the manufacturing process are unique, the problems faced by records managers at the design and manufacturing sites are not. In some ways, weapons records resemble other collections of diverse records, such as medical records. In both cases, records may be made by many people. Records may include laboratory reports, handmade notes, X-rays, and interpretations of the data that has been collected. Some older records may be fading or damaged and thus in need of restoration.

The Y-12 Plant is not the only DOE manufacturing facility; there are other major sites like Kansas City and Pantex, all of which generate records similar to those created at Y-12. Besides these sites, the research and design agencies produce other documents, frequently of different sorts from the manufacturing facilities. The laboratories, for example, have more than half a century's collections of research reports. Many of the earliest reports obviously were available only in paper form, but the laboratories have engaged in a massive scanning effort that has resulted in the creation of collections of PDF files. Many of these papers, like almost all of the manufacturing information, are classified and under considerable restrictions on access.

Maintaining and tracking the records calls for a massive records-management project. Because of the volume of data resulting from this process and the sensitivity of its content, it will be necessary to develop extensive catalogues of weapons records. Records are often used by more than one site. At Y-12, we need the records as part of our assurance to our customers (ultimately the Department of Defense) about the products we make for them. We use the records ourselves when warheads return to be refurbished or dismantled as part of our treaty obligations. The design laboratories have a great interest in our records to see how products vary from unit to unit and what differences there are between theoretical design and practical manufactured products.

Making Records Available

For the past few years, records in the Nuclear Weapons Complex (NWC) have been among the topics of study in the Nuclear Weapons Information Group (NWIG). Although many of the records NWIG has studied will probably never be generally available, there is considerable interest in sharing them among the NWC organizations.

The weapons-records projects actually show considerable similarity to the initiatives in HL7 concerning medical records: In both cases it is necessary to gather and organize large numbers of heterogeneous documents. Many of the documents are legacy items, perhaps decades old, with handwritten entries and other data for which there is no electronic source. Some of the actual record types are common to both fields (e.g., radiographs/X-rays). Even the issues of access control have similarities. Only authorized individuals should be allowed to see either patient records or weapons data. Within a particular set of records, not all records may have the same access conditions. (Within a patient's records, certain details of diagnosis should be restricted to medical professionals, while cost data may be of interest only to a hospital's business office. Similarly, designers and manufacturing engineers may not see the same subsets of data on a given warhead or test device.) The problems of cataloging, linking records to catalogues, and developing controlled vocabularies for searching catalogues can perhaps benefit from sharing of ideas between the disciplines.

Because of the volume of records and the condition of many of them, as well as access restrictions, it is not generally feasible for one site to search another's records directly. Furthermore, many types of records are site specific, and finding a given piece of data may require the assistance of a subject-matter expert. However, as a first step towards allowing inter-site searching, the NWIG has proposed cataloging records, whatever their media and locations. To support interchange of catalogues among sites in the NWC, NWIG has developed an SGML DTD for metadata about weapons records. Interchange of catalogues may eventually lead to searching of records collections and perhaps to interchange of individual records.

NWIG Catalog Format

The NWIG catalog format collects metadata in the form of an SGML document, defined by a document type definition (DTD, which is a collection of “declarations” such as those for SGML elements). This document can contain an entry about a single record or a whole collection of records. The first portion of the metadata in a catalog concerns the classification and origins of the metadata itself. Then follows the metadata about records being cataloged, including their classification, references to controlled terminology, sources (bibliographic data), abstracts, and other reference data

NWIG catalogs are relatively simple series of records. The critical portion of the NWIG DTD establishes the basic record structure:

```
<!ELEMENT NWIG - - (REC+)>
<!ELEMENT REC - - (M-ID, M-LOAD, M-LEV, M-CAT?, M-WDC*, M-CAV*, M-SRC*, M-REV*, TI+,
CR-DT+, LEV, CAT?, WDC*, CAV*, NTK+, ORG+, SUBJ+, TYPE+, SUBTYPE*, MED-FMT*, SIZE*,
ATTR*, LOC+, CON*, EX-ID*, AU*, KW*, AB*, SEE*, VER*, COM*, C-TI*, BROWSE*, EX-EL*)>
```

That is, the catalog consists of repeatable records, each of which contains a prescribed list of components, some of which are required, others of which are optional, and some of which may be repeated.

Each record carries metadata about its origin and classification. (The closing portion of each element declaration, enclosed in “--” is an explanatory comment, which should be sufficiently self explanatory for the purposes of this paper.) The NWIG application does not itself attempt to define need-to-know (NTK) methodology; it merely provides a means for collecting information. The first part of the data applies to the metadata itself

```
<!ELEMENT M-ID - - (#PCDATA)>
<!ELEMENT M-LOAD - - (#PCDATA) -- Load Date in YYYYMMDD format - numbers only-->
<!ELEMENT M-LEV - - (#PCDATA) -- Metadata Classification Level -->
<!ELEMENT M-CAT - - (#PCDATA) -- Metadata Classification Categories -->
<!ELEMENT M-WDC - - (#PCDATA) -- Metadata Weapon Data Categories -->
<!ELEMENT M-CAV - - (#PCDATA) -- Metadata Classification Caveats -->
<!ELEMENT M-SRC - - CDATA -- Author and or process that created the metadata-->
<!ELEMENT M-REV - - CDATA -- Metadata Revision Notes -->
```

Following the metadata about the metadata, the record includes the classification data for the actual record. This information might potentially be passed to further systems to determine whether the records themselves might be made available.

```
<!ELEMENT LEV - - (#PCDATA) -- Data Classification Level -->
<!ELEMENT CAT - - (#PCDATA) -- Data Classification Category -->
<!ELEMENT WDC - - (#PCDATA) -- Data Weapon Data Categories -->
<!ELEMENT CAV - - (#PCDATA) -- Data Classification Caveats -->
<!ELEMENT NTK - - (#PCDATA) -- Data Object Need-to-know -->
```

Further elements locate data according to agreed semantics, such as those in thesauri maintained by OSTI.

```
<!ELEMENT ORG - - CDATA -- Originating organization -->
<!ELEMENT SUBJ - - (#PCDATA) -- Subject Codes from a controlled Taxonomy -->
<!ELEMENT TYPE - - (#PCDATA) -- Type of data object -->
<!ELEMENT SUBTYPE - - (#PCDATA) -- Subtype of data object -->
<!ELEMENT MED-FMT - - CDATA -- Media/Format -->
<!ELEMENT SIZE - - CDATA -- Data type specific size -->
<!ELEMENT ATTR - - CDATA -- Data type specific attributes -->
```

After the classification elements, the remaining elements pertain to the specific records. Most of these are conventionally bibliographic.

```
<!ELEMENT TI - - CDATA -- Title -->
<!ELEMENT CR-DT - - (#PCDATA) -- Date in YYYYMMDD format or text-->
<!ELEMENT LOC  - - (#PCDATA) -- Archive  Location -->
<!ELEMENT CON - - (#PCDATA) -- Contact-->
<!ELEMENT EX-ID - - CDATA -- External Identifiers-->
<!ELEMENT AU - - CDATA -- Author/Process -->
<!ELEMENT KW - - CDATA -- Keywords -->
<!ELEMENT AB - - CDATA -- Abstract/Description -->
<!ELEMENT SEE - - CDATA -- See Also Refs -->
<!ELEMENT VER - - CDATA -- Version-->
<!ELEMENT COM - - CDATA -- Comments-->
<!ELEMENT C-TI - - CDATA -- Collection Title -->
<!ELEMENT BROWSE - - CDATA -- URL to Browse Representation-->
<!ELEMENT EX-EL - - CDATA -- Bucket for externally defined elements-->
```

Elements can be designated as having default values, being unclassified or withheld or can be “labelled”. (The SGML “Attlist” declarations define attributes of the elements declared above.) This latter feature, the “lbl” attribute, becomes critical for adapting diverse data to the NWIG model, since there is no other place to capture metadata not otherwise accounted for.

```
<!ATTLIST (TI | CR-DT | LEV | CAT | WDC | CAV | NTK | ORG | SUBJ | TYPE | SUBTYPE |
MED-FMT | SIZE | ATTR | LOC | CON | EX-ID | AU | KW | AB | SEE | VER | COM |  C-TI |
BROWSE | EX-EL)
        wh (WH) #IMPLIED -- nothing withheld
            unless specifically set--
        def (DEF) #IMPLIED -- not defaulted
            unless specifically set--
```

```
cls  (U) #IMPLIED -- can be designated
    as Unclassified--
lbl  CDATA #IMPLIED -- label for the entry-->
```

The NWIG catalog works best for report-like records because its elements are typically bibliographic, such as author, title, and version. It also provides support for abstracts and other easily searchable content. However, other types of data must be indicated by adding label attributes to “keyword” or bibliographic elements

Here is a sample from the OSTI “OpenNet” collection (http://www.doe.gov/html/osti/opennet/opennet1.html) which shows how the application works for report-like records:

```
<NWIG DTD-VER=1.2.1>
<REC><M-ID>OSTI-ON-AL92066403 </M-ID>
<M-LOAD>19960503</M-LOAD>
<M-LEV>U</M-LEV>
<M-CAV>DIST-A</M-CAV>
<M-SRC>OpenNet Entry</M-SRC>
<M-REV>19940906 </M-REV>
<TI CLS=”U”>Plutonium ignition in perchlorethylene vapors </TI>
<CR-DT>19691211 </CR-DT>
<LEV>U</LEV>
<CAV>DIST-A</CAV>
<NTK>NA</NTK>
ORG>Dow Chemical Co. </ORG>
<SUBJ def>15</SUBJ >
<SUBJ def>7002</SUBJ>
<TYPE>Document</TYPE>
<SUBTYPE>Report</SUBTYPE>
<MED-FMT>Paper</MED-FMT>
<SIZE>3 p</SIZE>
<LOC>&ntis;</LOC>
<EX-ID>AL92066403 ; RFP4560 </EX-ID>
<AU>Musgrave, L.E. </AU>
<KW>PLUTONIUM ignition; CHLORINATED ALIPHATIC
HYDROCARBONS  ignition; PLUTONIUM; IGNITION;
SOLVENTS; CARBON TETRACHLORIDE; VAPORS;
CLEANING </KW>
<AB>The possible reaction of liquid carbon tetrachloride with hot  plutonium chips
makes the use of an alternative cleaning solvent desirable.  Presently,
perchloroethylene is being considered. However, no study has been conducted on the
reaction of perchloroethylene vapors and plutonium. Since carbon tetrachloride
vapors significantly affect the ignition  properties of plutonium, a study has been
carried out in order to find what  affect perchloroethylene vapors may have on
plutonium ignition. </AB>
<EX-EL>
<Addressee>
<Declassification Status>Declassified document available
</EX-EL></REC>
```

Y-12’s Extensions to the NWIG Application

The NWIG efforts were driven at first largely from the research laboratories, and so the DTD has a strong bias towards cataloging their data. The Y-12 Plant has decided that since the NWIG DTD concentrates on research reports, additional DTDs will be required for manufacturing documents. As part of the Weapons Records Archiving and Preservation Project (WRAP) at Y-12 and in coordination with DOE-wide projects such as the Accelerated Strategic Computing Initiative (ASCI), Y-12 is developing a Weapons Information Locator System around SGML metadata records.

Much of the metadata will be extracted from existing catalogues that reside in relational databases. However, the SGML entries will become the primary internal catalogue for searching records at Y-12. The choice of the SGML records rather than the original database for searching is influenced by several factors. First of all, the databases were designed not so much to be records catalogues as to be part of the configuration-control and access-monitoring system for production documents like drawings and procedures. Furthermore, many records, such as assembly histories, are not indexed in these databases. Having the catalogue in SGML, of course, simplifies the task of generating interchange records according to the NWIG DTD: all that is required is a few OmniMark programs. The WRAP project hopes to link the records to as many electronic documents as possible (and to make many of these SGML documents themselves). For external sharing of catalogue records, the Y-12 metadata will be converted to the more-generic NWIG metadata.

The WRAP and WILS projects are still largely in the planning phases. We have done some sample catalogues with dummy data for demonstration purposes. We have already recognized the need to deal with multiple document types, rather than just research reports. We need to catalogue, for example, drawings, procedures, radiographs, various kinds of laboratory analyses of materials, dimensional certifications, and assembly histories. Although Y-12 is primarily a manufacturing plant, it also does some research, particularly into materials properties and advanced manufacturing techniques. (Besides manufacturing nuclear-weapons components, we have also made parts for submarines and spacecraft and such specialized tools as sample-collecting boxes for lunar exploration.) We do have our own collections of research reports, though not such extensive ones as the western laboratories. Our DTD differs from the NWIG one because it defines a collection of many different types of documents, rather than a collection of uniform records.

A typical collection of records at Y-12 might include assembly histories (composite documents), analytical laboratory reports, dimensional certifications, radiographs, procedures (assembly, QA), as well as fairly conventional reports

Accordingly, Y-12 has supplemented the NWIG SGML catalog so that it becomes a catalog of record-type specific entries. It allows for a detailed hierarchical structure of assembly history as composite document that is based on individual records types and follows structure of assemblies. It remains, however, a record-oriented catalog like the NWIG application:

```
<!ELEMENT %doctype; - - (title? ,  (work_request | z-history | roa | afs |
dimen_cert | radio_rpt | chem_rpt | drawing)*) >
```

The individual records in the catalog are specific to actual document types. The entry for any particular type of record includes some common types of information shared by all records and then information that is specific to that record type:

```
!ELEMENT roa - - (roa_id , %common_info;) -- Record of Assembly -->
```

The common information is bibliographic and largely based on the NWIG pattern (e.g., subject_code, title, abstract). A few elements are added from Y-12's existing metadata structures that talk about things like part numbers and test events:

```
<!ENTITY % common_info "(subject_code , title+ ,  (%locat;)* , author*, abstract?,
keyword* , comment* , related_object* , ref_object* ,  ship_part_no? , event_name?)
">
<!ELEMENT work_request_id - - (work_request_full , process_no+ , prod_part_no+ ,
device_name? , purchase_order_no , customer_part_id?) -- Work Request -->
<!ELEMENT roa_id - - (assy_part_no? , work_request_full+ , subassy_part_no+ ,
piece_part_no* , responsible_area) -- Record of Assembly -->
<!ELEMENT chem_rpt_id - - (chem_req_no , work_request_full , process_no+ ,
prod_part_no+ , chem_lab_req_no , material_type+) -- Chemical Report -->
<!ELEMENT drawing_id - - (da_dwg_no , dwg_no?, designer*, draftsman* ,
group_tech_code , process_no* , equipment_no* , building_no*) --Y-12 Drawing -->
```

This latter, record-type-specific information is typical of the data that can be expressed in the NWIG application only by repeated invocations of the “lbl” attribute.

The following example (of invented data) is typical of what Y-12 expects to put in its online catalog. The "work request number" is one of the keys to tracking many things at Y-12, including costs of our products as well as documents related to those products. These numbers, which are associated with test devices, are recycled, so a fully qualified number, including a year as well as a work request, is needed to identify a particular device. For our regular production of devices for the weapons stockpile, identification of the particular weapons program is a key identifier. The "process number" is another Y-12 specific identifier around which information is collected. A process might involve a particular manufacturing step, including design drawings, manufacturing or testing procedures, and checklists associated with the procedures. The combination of a work request and a series of process numbers might be used to track all the manufacturing records for components associated with a product built at Y-12.

```
<!DOCTYPE CATALOG PUBLIC "-//Lockheed Martin Energy Systems//DTD ASCI Catalog//EN"
"ASCIBASE.DTD">
<CATALOG CLASSLEV="U" SIGMAS="NONE">
<TITLE>Test Catalogue</TITLE>
<WORK_REQUEST CLASSLEV="U" SIGMAS="sigma1" ENTRY_DATE="19970425"><WORK_REQUEST_ID>
<WORK_REQUEST_FULL>
<WORK_REQUEST_NO>111111</WORK_REQUEST_NO>
<YEAR>1987</YEAR></WORK_REQUEST_FULL>
<PROCESS_NO>1111-11</PROCESS_NO>
<PROCESS_NO>1112-11 </PROCESS_NO>
<PROD_PART_NO>1111-11-1111</PROD_PART_NO>
<PROD_PART_NO>1112-11-1234 </PROD_PART_NO>
<DEVICE_NAME>dummy device</DEVICE_NAME
<PURCHASE_ORDER_NO>dummy purchase order</PURCHASE_ORDER_NO>
<CUSTOMER_PART_ID>dummy customer ID</CUSTOMER_PART_ID></WORK_REQUEST_ID>
<SUBJECT_CODE>0101 Arms Control and Nonproliferation</SUBJECT_CODE>
<TITLE>Dummy WorkRequest </TITLE>
<LOCATION VERIFIED="Y"><NAMED_LOC>EFM </NAMED_LOC></LOCATION>
<MEDIA>TIFF </MEDIA>
<SIZE>3 pages </SIZE>
<OWNER>cli</OWNER>
<OWNER>jdv</OWNER>
<OWNER>chm </OWNER>
<ORIGORG>DSRD </ORIGORG>
<CONTACT>cli </CONTACT>
<AUTHOR>Cobb, C.K. </AUTHOR>
<ABSTRACT>dummy abstract </ABSTRACT>
<COMMENT>dummy comment </COMMENT>
```

```
<RELATED_OBJECT>dummy related object </RELATED_OBJECT>
<REF_OBJECT>dummy reference object </REF_OBJECT>
</WORK_REQUEST>
<Z-HISTORY CLASSLEV="U" SIGMAS="sigma1" ENTRY_DATE="19970425"><Z-HISTORY_ID>
<YPG_NO>dummy number</YPG_NO>
<ASSY_PART_NO>2222-22-2222</ASSY_PART_NO>
<DEVICE_NAME>dummy zh device</DEVICE_NAME>
<PURCHASE_ORDER_NO>dummy zh po</PURCHASE_ORDER_NO>
<ASSY_WORK_REQUEST_FULL>
<WORK_REQUEST_NO>222222</WORK_REQUEST_NO>
<YEAR>1992</YEAR></ASSY_WORK_REQUEST_FULL>
<WORK_REQUEST_FULL><WORK_REQUEST_NO></WORK_REQUEST_NO><YEAR></YEAR></WORK_REQUEST_FU
LL></Z-HISTORY_ID>
<SUBJECT_CODE>010101 Policy, Negotiations, and Legislation</SUBJECT_CODE>
<TITLE>Dummy Z-History </TITLE>
<LOCATION VERIFIED="Y"><NAMED_LOC>9111 </NAMED_LOC></LOCATION>
<MEDIA>paper </MEDIA>
<SIZE>200 pages </SIZE>
<OWNER>lzd </OWNER>
<ORIGORG>DP </ORIGORG>
<CONTACT>chm </CONTACT>
<RESPONSIBLE_AREA>Inspection </RESPONSIBLE_AREA>
<AUTHOR>DeMarotta, L. J.</AUTHOR>
<ABSTRACT>Dummy Z-History Abstract </ABSTRACT>
<COMMENT>Dummy Z-History Comment </COMMENT>
<RELATED_OBJECT>Another dummy Z-History </RELATED_OBJECT>
</Z-HISTORY>
```

This sort of data can be translated to NWIG format. EFM, the Electronic File Manager, is a major collecting point for both metadata and electronic documents at Y-12. As this example shows, Y-12's internal metadata also includes revision tracking and the ability to accommodate multipage drawings.

```
<!DOCTYPE NWIG PUBLIC "-//USDOE Nuclear Weapons Information Group//NWIG Metadata DTD
1.3//EN">
<NWIG DTD-VER="1.3">
<REC>
<M-ID>Y-12 M2A921200A003     ,   0   </M-ID>
<M-LOAD>19970131</M-LOAD>
<M-LEV>U   </M-LEV>
<M-CAT>     </M-CAT>
<M-SRC>EFM DB</M-SRC>
<TI LBL="REV-U">MATERIAL HANDLING,DETAIL</TI>
<TI LBL="SHT  1-U">MATERIAL HANDLING,DETAIL</TI>
<TI LBL="SHT  1-U">MATERIAL HANDLING,DETAIL</TI>
<CR-DT LBL="REV-REL">19890120</CR-DT>
<CR-DT LBL="SHT  1-REL">19890120</CR-DT>
<CR-DT LBL="SHT  1-REL">19890120</CR-DT>
<LEV LBL="REV">U   </LEV>
<CAT LBL="REV">     </CAT>
<NTK>NA</NTK>
<ORG>Y12</ORG>
<SUBJ>2002</SUBJ>
<TYPE>DRAWING</TYPE>
```

```
<SUBTYPE>TOOLING</SUBTYPE>
<MED-FMT LBL="SHT  1">PAR</MED-FMT>
<MED-FMT LBL="SHT  1">PLT</MED-FMT>
<LOC>Y-12 EFM DB</LOC>
<CON>J. D. MCCLANAHAN</CON>
<EX-ID>M2A921200A003     ,   0   </EX-ID>
<AU>STAPEL           FE</AU>
<AB>Y-12 TOOL DRAWINGS</AB>
<C-TI>PD   DT    </C-TI>
</REC>
```

The Y-12 Z-History Extensions

We have recently started a further extension of the cataloging process by dissecting an assembly history. A major history document is a composite document. It is, in effect, a whole document collection in its own right. It may contain almost all the other types of Y-12 documents except research reports. A "Z-History" is Y-12's central manufacturing record. It is the assembly (or "Z") history for shipped item. A Z-History typically follows the manufacturing flow, where a fabrication, machining, or assembly procedure, with a checklist, is followed by one or more certification procedures with their checklists and collections of test results. The typical Z-History is a gathering place for many kinds of data such as forms, printouts, strip charts, notes, and sign-off stamps. Above all, Z-Histories tend to be very large documents.

We have, therefore, developed a unitary DTD for a history collection that treats it as a hierarchical collection of documents, organized along the major component subassemblies. The DTD collects metadata about only a single type of record, the Z-History, but then dissects that composite document into its components, which are in many cases the same sorts of documents that are described in our more generic catalog. The Z-history is typically broken down into various assembly processes, each with its own record of assembly (ROA).

```
<!ELEMENT asci-zh - - (title? ,  ( z-history )*) >
<!ATTLIST asci-zh %classification; entry_date CDATA #IMPLIED>
<!ELEMENT z-history     - - (z-history_id , %common_info; , roa* ) -- Z-History
Collection -->
```

The inner layers of the application, within the ROA extract their structure from patterns in the catalog DTD. Record types thus look like those in discrete catalog entries. However, much metadata for individual entries is inherited from upper layers rather than entered for the items themselves. Since the overall structure of the application is hierarchical, an individual item may inherit metadata from more than one level in the overall catalog. More subsidiary record types are cataloged in this application than may appear as independent documents in Y-12's high-level catalog application. The design of this application is still evolving.

```
<!ELEMENT roa  - - (roa_record)* -- Record of Assembly -->
<!ELEMENT roa_record           - - ( (work_request_no , (part_id , customer_part_id?
, ship_part_no? , part_name))+  , product_doc*) >
<!ELEMENT product_doc - - (( afs | dimen_cert | NDT_rpt | physical_test_rpt |
chem_rpt | drawing | roa | gage_report) , page_no+ , attachment* )  -- production
document types -->
```

However, as the layers of nesting increase, so does the amount of detail in the metadata.

```
<!ELEMENT dimen_cert    - - (dimen_cert_id ) -- Dimensional Certification Report -->
```

```
<!ELEMENT dimen_cert_id - - (responsible_area ,  ( (part_id | (process_no ,
serial_no?) | product_id) , part_name?)* , dwg_no* ) -- Dimensional Certification
Report -->
<ATTLIST dimen_cert_id
        stages (1|2|3|4) #IMPLIED
        stage NUMBER #IMPLIED>
```

In this case, for a dimensional certification we record both the total number of stages and which stage is in question for a particular record.

All real Y-12 Z-History documents are classified. However, I have constructed a sample catalog for a fictitious device, revealing no information about our actual processes. The example shows how a typical history might consist of ROA sections, in some cases with ROAs nested within others. The nesting notwithstanding, the records can still be expanded into NWIG records.

The outer layer of a Z-History looks rather like an entry for a document in the original Y-12 catalog:

```
<!DOCTYPE ASCI-ZH PUBLIC ""-//Lockheed Martin Energy Systems//DTD ASCI Z-History
Catalog//EN">
<ASCI-ZH CLASSLEV="S" CLASSCAT="RD"><TITLE>Dummy History</TITLE>
<Z-HISTORY CLASSLEV="S" CLASSCAT="RD">
<Z-HISTORY_ID>
  <YPG_NO>no such number</YPG_NO>
  <DEVICE_NAME>No Such Device</DEVICE_NAME>
  <PURCHASE_ORDER_NO>no such number</PURCHASE_ORDER_NO>
  <WORK_REQUEST_FULL>
    <WORK_REQUEST_NO>no number</WORK_REQUEST_NO><YEAR>never</YEAR>
  </WORK_REQUEST_FULL>
  <ASSY_PART_NO>xxxxxxxxxxx</ASSY_PART_NO>
  <PART_NAME>Unnamable</PART_NAME>
</Z-HISTORY_ID>
```

The metadata about a Z-History also includes information a about the location of the records being cataloged:

```
<SUBJECT_CODE>undefined</SUBJECT_CODE>
<TITLE>The History that Never Was</TITLE>
<LOCATION VERIFIED="N">
  <STRUCTURED_LOC>
    <BUILDING>No such building</BUILDING>
    <ROOM>Vacuum</ROOM>
  </STRUCTURED_LOC>
</LOCATION>
<RESPONSIBLE_AREA>Invisible Place</RESPONSIBLE_AREA>
<CONTACT>Nobody</CONTACT>
```

The entry for a ROA step can be relatively simple. In this example, I have used indentation to show the nesting of structure.

```
<ROA>
  <ROA_RECORD>
    <WORK_REQUEST_NO>zzzzz</WORK_REQUEST_NO>
    <PART_ID>yyyyyyyyyy</PART_ID>
    <PART_NAME>The Assembly that Wasn't</PART_NAME>
```

```
      <PRODUCT_DOC ACCEPTED="1">
        <DIMEN_CERT>
          <DIMEN_CERT_ID STAGES="1" STAGE="1">
          <RESPONSIBLE_AREA>
            Someplace Else</RESPONSIBLE_AREA>
          <PART_ID>Nonesuch </PART_ID>
          <DWG_NO>dddddddddd</DWG_NO>
          </DIMEN_CERT_ID>
        </DIMEN_CERT>
        <PAGE_NO>x</PAGE_NO>
      </PRODUCT_DOC>
    </ROA_RECORD>
```

The ROA entry can, however, become quite complex, as more processing steps are added. This example is rather long, but it illustrates how records can be nested. In this case, a ROA record includes a "product document" which includes another entire ROA, made up of several ROA records. This structure accurately reflects the structure of the imaginary device at a level where subassemblies are brought together. Each subassembly has its own ROA, and these are incorporated into the ROA for the completed device.

```
  <ROA_RECORD>
    <WORK_REQUEST_NO>zazaza</WORK_REQUEST_NO>
    <PART_ID>ybybybybyb</PART_ID>
    <PART_NAME>An assembled composite layer inside</PART_NAME>
        <PRODUCT_DOC ACCEPTED="1">
          <ROA>
            <ROA_RECORD>
              <WORK_REQUEST_NO>zazaza</WORK_REQUEST_NO>
              <PART_ID>ybybybybyc</PART_ID>
              <PART_NAME>A nonlayer inside that</PART_NAME>
                <PRODUCT_DOC ACCEPTED="1">
                  <DIMEN_CERT>
                    <DIMEN_CERT_ID STAGES="1" STAGE="1">
                      <RESPONSIBLE_AREA>Yet Another Place</RESPONSIBLE_AREA>
                      <DWG_NO>dxdxdxdxdxd</DWG_NO>
                    </DIMEN_CERT_ID>
                  </DIMEN_CERT>
                  <PAGE_NO>xyz</PAGE_NO>
                  <ATTACHMENT>
                    <GAGE_REPORT>
                      <NC_TAPE>ncncnc</NC_TAPE>
                      <PROGRAM_TAPE>ptptpt</PROGRAM TAPE>
                      <MACHINE_NO>mnmnmnm</MACHINE_NO>
                      <PAGE_NO>xya</PAGE_NO>
                    </GAGE_REPORT>
                  </ATTACHMENT>
                </PRODUCT_DOC>
                <PRODUCT_DOC ACCEPTED="1">
                  <NDT_REPORT>
                    <NDT_RPT_ID>
                      <NDT_REPORT_NO></NDT_REPORT_NO>
                        <NDT_TEST_NAME>A Kind of Test</NDT_TEST_NAME>
                        <OPERATION_NO>abc</OPERATION_NO>
                    </NDT_RPT_ID>
```

```
            </NDT_REPORT>
            <PAGE_NO>xya</PAGE_NO>
          </PRODUCT_DOC>
          <PRODUCT_DOC ACCEPTED="1">
            <CHEM_RPT>
              <CHEM_RPT_ID>
                <CHEM_LAB_REQ_NO>clrno</CHEM_LAB_REQ_NO>
              </CHEM_RPT_ID>
            </CHEM_RPT>
            <PAGE_NO>xyb</PAGE_NO>
          </PRODUCT_DOC>
      </ROA_RECORD>
      <ROA_RECORD>
        <WORK_REQUEST_NO>zazaza</WORK_REQUEST_NO>
        <PART_ID>ybybybybyd</PART_ID>
        <PART_NAME>A second thing inside that</PART_NAME>
        <PRODUCT_DOC ACCEPTED="1">
          <DIMEN_CERT>
            <DIMEN_CERT_ID STAGES="1" STAGE="1">
              <RESPONSIBLE_AREA>Yet Another Place</RESPONSIBLE_AREA>
              <DWG_NO>dxdxdxdxdyf</DWG_NO>
          </DIMEN_CERT_ID>
          </DIMEN_CERT>
          <PAGE_NO>xyc</PAGE_NO>
        </PRODUCT_DOC>
        <PRODUCT_DOC ACCEPTED="1">
          <NDT_REPORT>
            <NDT_RPT_ID>
              <NDT_REPORT_NO>uql</NDT_REPORT_NO>
              <NDT_TEST_NAME>A Kind of Test</NDT_TEST_NAME>
              <OPERATION_NO>abc</OPERATION_NO>
            </NDT_RPT_ID>
          </NDT_REPORT>
            <PAGE_NO>xyd</PAGE_NO>
        </PRODUCT_DOC>
        <PRODUCT_DOC ACCEPTED="1">
          <CHEM_RPT>
            <CHEM_RPT_ID>
              <CHEM_LAB_REQ_NO>clrno</CHEM_LAB_REQ_NO>
            </CHEM_RPT_ID>
          </CHEM_RPT>
          <PAGE_NO>xye</PAGE_NO>
        </PRODUCT_DOC>
      </ROA_RECORD>
    </ROA>
  </PRODUCT_DOC>
 <ROA_RECORD>
</ROA>
<Z-HISTORY>
<ASCI-ZH >
```

Although this pattern looks complex, it actually reduces to many repetitions of a small number of units. The example consists mostly of standard test results: dimensional certifications, chemical reports, and NDT reports. The structure gathers the assorted tests together for each step of assembly. With a processing application that supports the inheritance of metadata from higher layers of the catalog, this type of structure is actually much more efficient than the NWIG pattern, which requires entry of metadata for each object in the catalog.

What actually convinced us to develop this application was an attempt to enter NWIG metadata for a 500-page Z-History. We kept having to repeat the same information over and over, and the repeated information quite overwhelmed the actual new data for each section in the document.

We have actually cataloged the entire Z-History for the Hermosa test device. The resulting catalog occupies about 84 KB on disk. It has about 2500 words; given the high density of markup even in this application, that is still a lot of original data—about 1900 distinct items. Had this data been captured in the NWIG application, the markup-to-data ratio would have been much less efficient. I suspect the file would have been over 2000 KB for the same actual amount of information.

What this type of catalog does is allow tracking of individual kinds of information: we can search it and find, for example, all the radiographs that were made of the Hermosa device or all the dimensional contour tracings. Another group has scanned in the Hermosa Z-History, but since their results are in the form of images (and most of the scanned pages would not be suitable for OCR processing), their document is not directly searchable. They have, however, generated VRML images from the contour plots and linked these to a very attractive user interface that allows visual browsing of the device. If we were able to get funding for linking the two applications, our searchable data would provide yet another interface to the live models of the device.

A Future for Y-12 Metadata?

The Y-12 catalog applications are still under development. We have only a few sample collections of data in the per-document catalog format and one major test device in the Z-History catalog. In the short term, our greatest need is for a full-text search engine that is aware of SGML/XML structures. Then we could approach linking our catalogs with other live data, such as the device-visualization application.

In the long run, we can see large benefits from linking generation of metadata to manufacturing process: If we could put the many procedures on line and enable generation of checklists from procedures, we would have a start on creating electronic Z-histories. The first part of that proposition is not too difficult to do. We have DTDs for several types of procedures already developed. Our primary SGML/XML processor, OmniMark, could easily be programmed to generate the checklists and data-entry forms. As various weapons programs come up for life extension, the procedures are having to be revised, presenting us with an opportunity to capture them. At the same time, DOE is urging us do more electronic data capture on units returned for rework as they pass through our shops for life extension or revalidation. Somewhere in between the two processes is an opportunity to bring all the processes together. What is holding us back is largely the lack of computer workstations on the actual shop floors and the lack of funding for infrastructure improvement.

If we were to bring all these elements together, we would be considerably further along towards supporting our mandated goal of science-based stockpile management.