

ASKING FOR AN ELECTRONIC DATA DELIVERABLE IS EASIER THAN GETTING IT

Timothy J. McCracken, LMITCO
Lockheed Martin Idaho Technologies Company
Sample Management Office

An Electronic Data Deliverable (EDD) is the magic bullet most sample management offices have sought to reduce costs for data entry and enable electronic validation of data. An effort to implement a format of a particularly flexible specification was difficult but proven possible in 1995. Most of the problems associated with the format have since been reflected in the reluctance by commercial laboratories to produce EDDs written to this specification and format. It has become painfully obvious that the format originally piloted by the Idaho National Engineering and Environmental Laboratory (INEEL) was causing an unnecessary economic and resource burden for the analytical laboratories. Many laboratories made contingent promises that delivery of the new EDDs would become routine. However, as time passed very few labs could prove competency in production of the data format.

My father had a knack for car trading (not always to his best interests I'm afraid) that illustrates why I think our current methods of requesting EDDs may be faulty. When my dad decided that the 1936 Chevrolet Pickup (a classic in today's market) was causing him more grief than he cared to endure, he would go see his used car sales buddy, Ron. Now the pickup had a bent drive shaft that was annoying above 25mph (unless you had a tooth ache you wanted to ignore) and had no heater (Idaho winters rarely get below -20 F) but I thought it was heavenly and I couldn't fathom why my dad wanted to part with it. Ron would let my dad look around at the field of perspective fishing vehicles (my dad was a devote member of the Seventh Day Lumbricidae Baptists) and when a particular mode of transportation caught his eye (in this case I think it was the '58 Rambler Ambassador stationwagon) the appropriate dialogue would probably be, "Ron, I don't know if you can swing it but if you'll give me \$200 for that hunk of junk (probably worth close to the national debt today) I'll take that fishing car over there." Ron knew he only had to add \$300 onto the "rock bottom" price (Ron wasn't stupid either) he was going to offer for the Rambler and he would have a deal. What I came to realize is that if you want to give someone money they will probably find a way to accept it, I know that may seem earth shaking to many people but I find it very true.

I have also (unfortunately for the analytical laboratory community) slowly begun to realize that the quest for the magic bullet EDD was another instance of our willingness to give some people some money while hoping that they could tell us what we wanted to hear. If analyses of an organic specimen (dead mouse) is deemed clearly necessary, and Ron's Environmental Lab and Used Cars (hereafter referred to as RELUC) is contacted to test for oogy-goo 246 and yucky-uk 158, then of course they will take the money and give you accurate results. What has become a real problem with getting analytical data is not that RELUC doesn't want to take our money or

can't do the analysis accurately or even fails report the results in a timely manner, its that we have decided we need to receive the data (the EDD) in a specifically difficult format.

Ron and my dad had a good workable relationship. My dad expected the title to the Rambler and a set of keys. Ron expected the title to the pickup (which immediately went on sale for a reduced price of \$600), keys and a contractually specified amount of money. They both were smiling when the transaction ended and my dad continued to buy cars from Ron for many more years. My dad would never have told Ron that the deal was only possible if the seat covers were professionally cleaned, a fresh coat of wax was applied to the rusted body, and year's supply of pine tree air fresheners were thrown in. Ron sold his cars as is and to my knowledge never lied or tried to deceive my dad. To my father's credit, I now realize that he accepted the cars from Ron, as is, and then my dad would wax it and buy air fresheners and install a set of green seat covers (the Rambler was red and white, I thought we appeared to driving a watermelon turned inside out). My dad accepted what was possible and then he made the necessary changes to make it fit his perspective of a fishing car.

What we need from the analytical laboratory is accurate, honest and timely reporting of analytical data. If a reasonable payment for the services can be settled on then we may need to accept responsibility for waxing and purchasing our own air fresheners. Because the absolute requirements for EDDs reside in contractual agreements a more familiar data format to the lab may need to be agreed upon.

Separating the specification of an EDD from the format of the EDD was an important step for the Idaho Engineering and Environmental Laboratory (INEEL), the Los Alamos National Laboratory (LANL) and the Lawrence Livermore National Laboratory (LLNL) SMO folks. After discussions with current analytical laboratory data management personnel were carried out a perspective of what was reasonable to expect in an EDD began to take shape. In the current economic situation that commercial laboratories have found themselves in, it will be necessary for the customer to carefully consider buying their own pine scented air fresheners. Customers need to work with their labs to agree on what data elements can be reasonably supplied with in an EDD, in a format the labs consider easy and timely to produce.

When I obtained the ultimate badge of manhood, my drivers license, I tried to convince my dad to trade the Rambler in on a new 1968 Chevelle Super Sport, bloody nose red with white hood strips, 396ci engine, 4.11 differential, 3 speed turbo hydromatic. A real bad boy fishing car I perceived. His enthusiasm was predictable and the outcome predestined. His insightful wisdom (now that I have teenagers of my own) was to logic that driving up to a stream in it would scare the fish away. It could have, too. He saw the Rambler as reliable, affordable transportation without any excessive features that would distract the occupants from life's true purpose, fishing.

Many of the labs have expressed a preference for producing an EDD in a row and column "flat file". Not stylish, but it is reliable and affordable in the general aspect of tabular data formats. INEEL, LANL, and LLNL have as a result proposed a relational data format by does not look much different from other EDD formats currently in use. However, this format has also been designed specifically to take advantage of the broad based analytical data dictionary of the

Department of Energy Environmental Management Electronic Data Deliverable Master Specification (DEEMS). It is believed that the impact to labs can be reduced by taking advantage of a familiar tabular data format and minimizing the number of data elements used. Keeping the DEEMS as the eventual goal for analytical data will allow the extension of the format. This can be a win, win, win solution for DOE (and other federal agencies), commercial laboratories, and sample management offices. This should be of interest to analytical laboratory managers, data management personnel, sample management personnel and others who create, store, transmit, and report analytical data.

The effort to support the DEEMS as a future standard for creating analytical based electronic data formats is not precluded by any effort to create an EDD format other than the original hierarchical format proposed in DEEMS version 1.0. As new technologies in hardware and software are incorporated in the analytical laboratories, the specification defined by DEEMS can provide a guidepost for developers of Laboratory Information Management Systems and analytical instruments. It is realistic to create an extensible EDD format based on a broad specification (such as DEEMS) that can be routinely produced by labs.

GENERAL ELECTRONIC DATA DELIVERABLE (GEDD) FORMAT

This is a draft definition of a format for the data tables file to be used in transmitting analytical data from a laboratory to its client.

This format intends for the analytical data to be transmitted in a single file. The data contained within this file should be variable length ASCII records terminating with an end of line character.

The contents of the analytical data file is of multiple relational data tables. Currently defined tables are Data_Package, Samples, Analyses, Batch, Results and QC. The order of the tables within the file is not specified, however, it is recommended that the Data_Package table be first for ease of use.

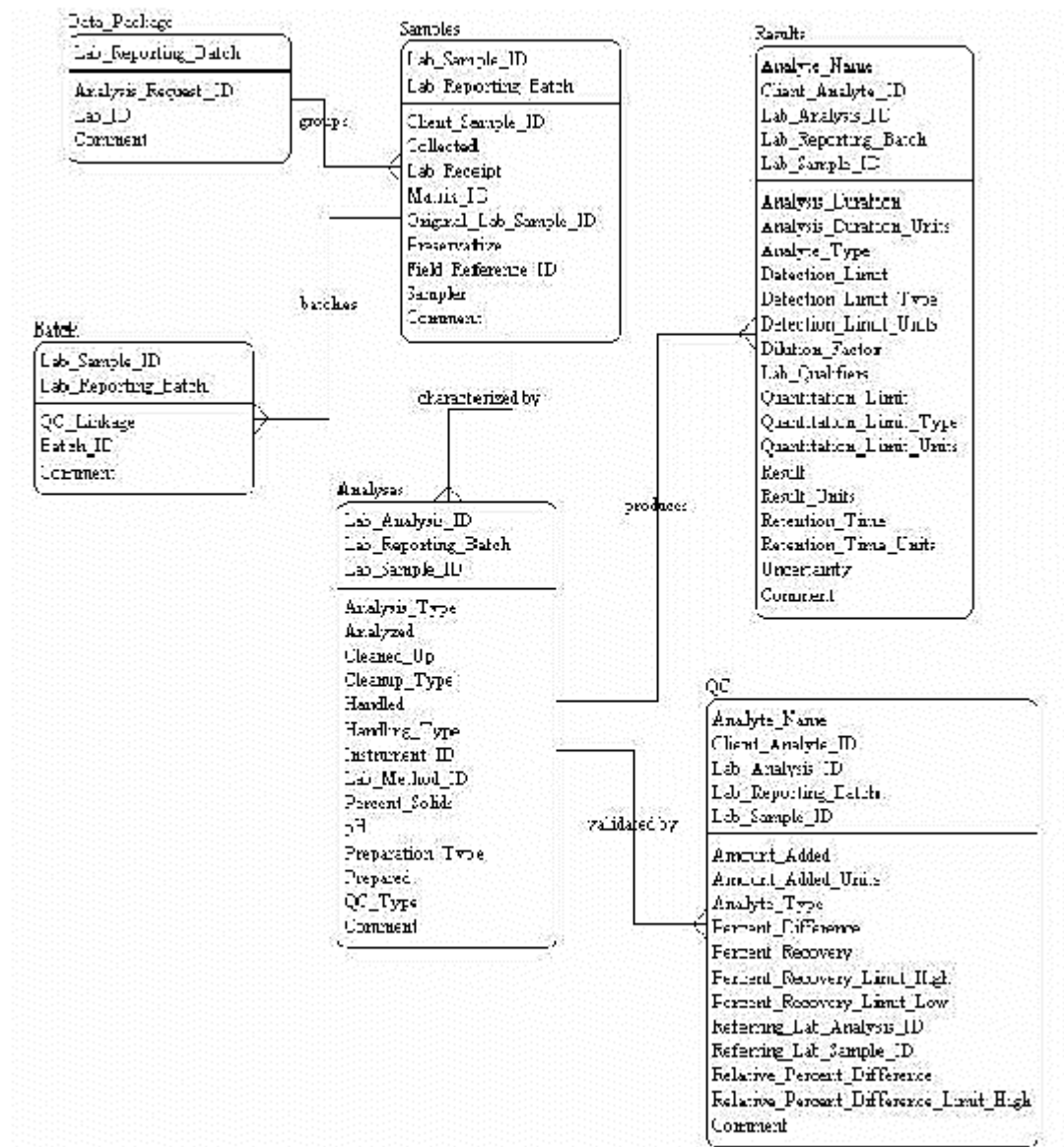
Each data table consists of at least three data records; one table name record, one data element list record, and one or more detail records.

The table name record identifies which table it is by beginning with the word Table, a colon and the table name (i.e. Table: Data_Package). White space (spaces, tabs, etc.) within this record is ignored. The table name record must be the first record in the table. The table name is not case sensitive.

The data element list record contains the data element names in the same order that the data element values occur in the detail records. The data element list record must be the second record, following the table name record. The data elements of the current table are specified by this record. The order of the data elements in the current table's detail records is specified by this record. The data element names are not case sensitive. The data element names are separated

with an ASCII character 124 (|) as a delimiter. White space characters will be ignored in this record.

The Data Element Definitions section defines the current data elements. Within the data tables certain data elements are required for relational linking of the tables. The Entity Relationship Diagram illustrates the data tables and the currently associated data elements in their relationships. This format should allow the flexibility of using only those data elements needed (or leaving out those un-needed or un-obtainable ones).



The data element detail record(s) contain the ordered data values that correspond to the data element names listed in the data element name record for the current table. Data values are separated with an ASCII character 124 as a delimiter. Null data values should be zero characters but the delimiter will still be used as a place holder. Data values do not have a specific length or data type defined by this format but will be left to the lab and client to contractually define. Blank records in the file will be ignored.