

---

**Committee E13 on MOLECULAR SPECTROSCOPY AND CHROMATOGRAPHY**

**DRAFT Minutes for E13.15 Subcommittee Working Group DRAFT**

10:35 am – 12:35 pm EDT

October 14, 2005

Virtual Meeting

**I. Introductions and Welcome:** Gary Kramer, NIST, E13.15 Chair called the meeting to order at 10:35 am EDT. Karen Wilson was introduced as the new ASTM/E13.15 liaison.

**II. Review of Agenda and Minutes.** The agenda was approved without change.

Two amendments were requested for the minutes of the August 25, 2005 meeting: 1) the definition of units should be made clearer; 2) just before section VIII, “bests” should be changed to “best”. A motion was made to approve the minutes as amended. Second. The motion passed unanimously.

**III. Attendees:**

Mark Bean, GSK  
Tony Davies, Waters  
Maren Fiege, Waters  
Ronny Jopp, NIST  
Gary Kramer, NIST  
Peter Linstrom, NIST

Dave Martinsen, ACS  
Alex Mutin, Shimadzu  
Jack Newton, Chenomix  
Alexander Roth, NIST  
Burkhard Schäfer, Univ. Kaiserlautern  
Karen Wilson, ASTM

**IV. Publications:** Tony Davies, Maren Fiege, and Peter Lampen have a publication going to SCI. Randy Julian already has one in SCI, Tony has a previous one in SCI. This makes a third publication. Links should be made to the publications from the AnIML web site. Tony will check to see if SCI grants permission to post the articles on the AnIML web site.

**V. Future meetings:**

- a. EAS Task Group. Sunday 3pm-5pm
- b. EAS Business, Monday, 5-7pm
- c. Pittcon: Abstracts should have been received, and notice given to author.
- d. IQPC
- e. Lab Automation: Gary Kramer – accepted. Burkhard Schaefer – accepted
- f. Possible meeting in conjunction with IQPC and Lab Automation in January, if enough people are attending one or the other. In some cases, people need to budget travel 12 months in advance, so travel on short notice is not possible. No decision made on this possible meeting.

**VI. AnIML Requirements Document:** The following points were raised during the discussion:

- a. Too official looking. It would be better if it were labeled as “DRAFT.”
- b. E13.15 – just Analytical Data, not management
- c. 1.4 – references to GAML should be referenced on xml.org.
- d. SpectroML should be referenced as well.
- e. 1.4.6 – is it relevant? Yes.
- f. Section 1 – no comments

- g. Section 2.1
  - i. 2<sup>nd</sup> sentence – “these” needs a modifier. Not clear what it means.
  - ii. Last part – technical competence should be “domain expertise”
- h. Section 2.2
  - i. Analytical data – should it be “analytical chemistry data”? No – the committee name is just analytical data. Leave this as is.
  - ii. The document states that interchange, viewing and archiving are reasons for AnIML. But we are not developing standards for viewing. So maybe viewing should be left out.
  - iii. Next sentence says “all legal requirements”. Perhaps the “all” should be left out. It is difficult to ensure that “all” requirements will be met.
  - iv. In the 4<sup>th</sup> paragraph, the statement is made that “all files necessary to parse” AnIML. Does this include software? Does the parse need to be made freely available? It was pointed out that vendor extensions, for example, would not be available from the official site. Some vendors might make extensions openly available. Others may not. But we can only control those things which are provided within the ASTM environment. If something is balloted by ASTM, it must be freely available in order for documents to be validated. Validation against the non-extended standard should not be a problem. A parser should ignore extensions which are not part of the core/technique standards, and are not publicly available.
  - v. See section 2.6. Take entire section out. Backwards compatible.
  - vi. Move section 2.2 paragraph 4 to section 2.6.
  - vii. Not talking about ASTM application code. Just schema files, samples, documentation. However, some individuals may choose to load application files. In fact, loading application files should be encouraged, if the standard will be used. From pre-xml days, hundreds on meaningless sample files are available. But if we only have syntactic, not semantic, parsing, AnIML will have the same problem. Therefore, we need applications which will read the technique file, and apply it to a given instance file. This is different from the usual semantic validation, and therefore we’ll need a standard validation tool as part of AnIML. One of the requirements is that we have a round-table approach to prove a technique is valid. Therefore, a validation tool is required. Note: see requirement 10, paragraph 3.7. The burden is on the committee to create and maintain. We should be able to define what we require from a validation. Application development may require some upfront funding, but ongoing maintenance is a problem. Open source application development is a possibility, but this is not the only solution.
  - viii. Quotation from Tim Bray – any successful xml implementation has an online validator, which returns problems to the user.
  - ix. If we choose, we can include software. But this will delay the standard.
  - x. Note: this document is not producing the standard. Only the requirements for the standard.
- i. Section 2.6 – take out the 4 digit numbers. This is what the standards document should have.
- j. Take out 2.6.3.1 to 2.6.3.10
- k. Usage of the words “must”, “will”, and “should” should be checked to make sure the use is consistent. These have very specific meanings in ASTM context, and should be put in all-caps to indicate the standard context. See RFC1129. It forces us to be precise. Tony will examine usage and make consistency.
- l. Section 2.2 – last paragraph. Common across a technique and industry? Is it the “irreducible minimum”? What is core? Is it the “core schema”, or the beast (or least)? We are looking for

all the information you need to store, info for an experiment, across techniques. Need some explanation of technique/core relationship. Tony will expand.

- i. Last sentence – did we reach consensus? Is this correct? E.g. NMR, sweep width? Vendors should be required to include it if we need it to interpret the experiment. A new AnIML file should still work with existing applications, and obey rules set down by the technique. Note: vendors leave out certain parameters in JCAMP which are required. You shouldn't be able to modify the way you store required parameters. Most attractive selling point: if you follow technique definitions and extensions, and base defs are still there. This make files interchangeable. If we delete, we throw this advantage away.
- m. Section 2.3
  - i. Paragraph 2 – difficult to fulfill. People use xml every day without knowing about it. “Must” implies something we can't deliver. Doc for standard should be understandable by non-computer phrasing, but the standard itself? Work on phrasing, get rid of “must”.
  - ii. Go to “System features” or “design rules”.
- n. A version number should be added to the header.

## VII. Core Group Update

- a. Took Visio diagram from SSI/Mark Mullins – implemented AnIML in corporate environment in chromatography.
- b. Looked at simplest way to express
- c. Next meeting – Burkhard joined and put into sourceforge.net using tracking database. Most are addresses. Work in xmlspy, but not visual studio
- d. Almost ready to deliver.
- e. Question: are we using visual studio as a standard tool? A lot of people are using it. Therefore, we should be verifying with both.
- f. Long term solution – Naming and design rules. Take encoded part of design rules, check correspondence to encoded design rules. Is it being done right? Note that W3C schema validation can also be used. We should use as many as we can find. If we turn up errors, we'll fix them.

## VIII. Naming and Design Recommendations. Alexander Roth and Ronny Jopp discussed some proposals for naming and design rules. Important for interoperability. See PowerPoint.

- a. ISO 11179 explains metadata and naming metadata. XML.org – see 2 from earlier.
- b. 5 parts
- c. GK will send out url to ISO doc.
- d. Example: naming convention for data types
  - i. Upper camel case
    1. do not use hyphens, dashes, spaces
    2. put first letter in caps:
      - a. upper camel: SlitWidth
      - b. lower camel: slitWidth
    3. data type and data elements: upper camel
    4. attributes: lower camel
  - ii. Don't make data types plural
- e. Other considerations
  - i. URIs vs. URLs

- ii. Do URIs or URLs need to resolve or not?
  - f. Should AnIML use local or global variables?
  - g. The XML community does not agree on which conventions to use, but within AnIML, we should choose a convention, and follow it consistently within AnIML. NDR (Naming and Design Recommendations) is a buzzword to attach to these concepts.
- IX.** Next meeting: October 28, 2005, 10:30 am – noon EDT
- X. Adjournment.** Motion to adjourn. Second. Carried unanimously at 12:35 pm EDT.

Submitted by David Martinsen, ACS, ASTM E13.15 Secretary