



INTERNATIONAL COMMITTEE E13 on MOLECULAR SPECTROSCOPY AND CHROMATOGRAPHY

Standards Worldwide

Address 100 Barr Harbor Drive
PO Box C700
W. Conshohocken, PA
19428-2959 | USA

Phone 610.832.9500
Fax 610.832.9555
e-mail service@astm.org
Web www.astm.org

Minutes for E13.15 Subcommittee Working Group

10:30 am – 12:30 pm EDT

March 16, 2007

Virtual Meeting

I. Introductions and Welcome: Gary Kramer, E13.15 Chair called the meeting to order at 10:35 am EDT. The ASTM release form for recording the session was displayed on the screen. There being no objections, the meeting was recorded.

II. Attendees:

Tony Davies, GSK
Maren Fiege, Waters
Patrick Gleichmann, NIST
Joe Koury, ASTM
Gary Kramer, NIST
Robert Lancashire, UWI
Peter Linstrom, NIST
David Martinsen, ACS

Jamie McQuay, Scimatic
Dale O'Neill, Agilent
Kordian Placzek, NIST
William Rich, Johnson
& Johnson
Burkhard Schaefer, BSSN
Wei Wang, Pfizer

III. Next Virtual Meeting: April 13, 2007, 10:30 - 12:30 EDT

- The UV/Vis Technique Group will meet on March 22, 2007 at 10:30 am - 12:30 pm EDT.
- The Core Group will meet March 20, 2007 at 10:30 am EDT. They generally meet every other Tuesday.
- Chromatography, March 28, 2007, 10:30 am - 12:30 pm EDT.
- The next business meeting will not be held at EAS, since none of the usual attendees normally attends EAS. The meeting does not have to be face-to-face, but the IUPAC meeting in Torino, Italy in August is a possibility. Tony Davies will check with IUPAC to see what the possibilities are for a joint ASTM/IUPAC meeting in Torino. For those who wish to contribute an abstract, the deadline is March 31. Robert Lancashire has submitted an abstract for the Education section. Other possibilities include the environmental and metrology sections of Analytical Chemistry.

IV. Minutes:

- February 9, 2006 minutes. It was noted that the time should be changed to EST. A motion was made to accept the minutes as amended. Second. The motion carried.
- January meeting minutes: Again, the time should be changed to EST. A motion was made to accept the minutes as amended. Second. The motion carried.

V. Welcome to new attendees:

- Jamie McQuay, developer of scientific software and systems with Scimatic Software.
- William Rich, system administrator and validation expert at Johnson & Johnson. He is an end user of data systems.
- Wei Wang, lead in the Global Analytical Systems at Pfizer.

VI. Pittcon:

- The Pittcon presentations will be posted on the website. Speakers should put them on the wiki, or send to Gary.
- Pittcon 2008 - deadline for submitting a workshop is April 2. There was interest expressed in a session. Potential speakers are asked to submit interest and an abstract to Gary before then, so Gary can submit a proposal to the Pittcon program committee.

VII. Reports from subgroups

- The Uv/Vis group met in March, and discussed units and quantities. They made good progress. The goal is to be able to switch easily to UnitsML when that standard is available. Burkhard came up with keep the units in a single location, so document segments could be signed without signing the units part. This will make it easier to switch to UnitsML without invalidating the signature. For each item in the technique with a quantity, the allowable units will be enumerated. This can easily be extended to add new units. However, it is not possible to replace units. For example, if wavelength and wave numbers are allowable and required, a user could specify quantities in volts, but the wavelength or wave number would still need to be there.
- Core Developer - no meeting since PittCon.

VIII. Miscellaneous Discussion:

- Comment from Robert Lancashire: In the UV/Vis file at Pittcon, from Stuart Chalk, there was an embedded CML structure file and a molfile. CML was there as an XML comment. When he tried to open the file with Jmol, the application did not recognize the CML structure. Because it was present as a comment, the parser ignored it. This demonstrates one of the limitations of

AnIML. Because of the desire to prevent users from introducing arbitrary additions, the ability to add something like CML is also restricted.

- There was also a suggestion at Pittcon to create AnIML 1.0 techniques using the JCAMP and ANDI dictionaries, in the absence of domain experts. In addition to the ability to handle new data, this approach would have the advantage of being able to handle the legacy information without losing anything.
- In order to provide more information for domain experts to get involved, the web site needs some redesign. Patrick is willing to update the website. He will contact Mark Bean to get access.
- There was some discussion of the problem of indexing, specifically how to handle multi-dimensional techniques, and techniques where you record a series of spectra or chromatograms which create another axis, e.g., chromatograms as a function of temperature, or non-continuous axes like pH, or a material might be added as a function of time. In UV/Vis, one section is used for 2D or 1D spectral information, then a section on temperature or pH. But each new section replicated what was in the spectral section, and caused problems. Going back to earlier discussions, would a timer technique be useful? This approach could be useful for many types of variables - time, temperature, pH, pressure.
- Patrick and Kordian presented an approach adding ExperimentStep references, in which each Experiment would have an ExperimentStep, and each ExperimentStep could have an ExperimentStepReference. This could lead to three vectors. The page would be valid without referencing, meaning that we would have optional independent axes. This should be avoided, since it would lead to confusion in the processing application. Comments: if you have a temperature run, why can't you just have multiple pages, or multiple experiment steps, with a different temperature on each? This may not be contradictory to what Kordian and Patrick are proposing. If you refer to another experiment step, the data is now overlaid; you can superimpose axes that come from somewhere else. The spectrum page can gain a 3rd dimension, such as spectra observable over temperature gradient - 3D plot of spectrum vs. temperature. The indexing techniques need to be referenced in the technique definition for UV, in order to combine the two. We should be able to have a list of indexing techniques outside of the technique definition itself. We also need to be able to combine techniques. Consider two approaches: a) each technique definition lists its acceptable extensions. This makes things simpler for applications, but reduces extensibility; b) the technique doesn't pre-define extensions. This is more flexible, but applications could then have things show up which they don't understand and can't handle. However, it may be all right if a generic viewer can't view the 3rd axis, as long as it can view the "standard" axes. Neither solution seems to be ideal. Both have problems. Perhaps 80/20 rule work. 80% of the experiments can be handled by the generic indexing techniques, 20% will need specialized techniques. The specialized experiments will be outside the scope of the subcommittee. If

there is a large enough community interested in the technique, that community will need to develop the technique definition.

- It was suggested that there is a requirement in AnIML that once the basic techniques have been defined, we need to be able to combine them to create multidimensional techniques. This can be done either through combining multiple techniques, or through the use of indexing techniques. But we need to be able to combine experiment steps from different techniques easily, because that's what people do. The problem arises if a variable is measured in place of the measurement commonly used for the technique. Then you are fundamentally doing a different type of experiment. Consider fluorescence - excitation emission. One step on excitation axis, the emission spectrum is measured. Then the excitation energy is changed, and another emission spectrum is scanned. Peter Linstrom suggested that there are two ways to approach this in AnIML. First, an extension could be written with a parameter to identify excitation level of each scan. Alternately, a data set could be created which recorded the excitation, and each child would then refer to that data set as its sample source. In neither case does the UV/Vis technique definition need to be modified.
- Tony Davies suggested that as long as it's documented, and the application software knows what to do with it, it doesn't matter. But the question arises as to how to organize the indexes, so they are handled the same way for temperature and spectroscopy, chromatography and spectroscopy, or two spectroscopies, e.g., UV/UV. One approach would be that as long as you understand where the spectrum is coming from, you can use the sample source mechanism to point to a parent. The problem arises with parallel experiments, for example, two spectrometers on the same sample at the same time with temperature program applied. This causes problems. Is it possible to write these indexing techniques so we don't need to redefine for each analytical technique? How do you define the indexing techniques, with rules for time, pH, metadata associated with them, and create the infrastructure to build that in? The UV/Vis group first set out the data layout and metadata for simple x-y UV spectra, then needed to repeat it all temperature to do kinetics, and then needed to do it again for pH. Where does all of the structure for that metadata go? It's not in the UV/Vis technique? It could be put in an extension, for example, a pH extension. You don't know x-axis in this case. You could define a second technique, such as UV/Vis with arbitrary scanning. Then define the independent axis as a sequence number, with an arbitrary range. Then you could add on pH, or anything else you want to measure that occurs over that sequence, in addition to wavelength and intensity, as an extension. The problem is if you want to do something not coupled to wavelength, or wave numbers. In that case, you will need a new UV/Vis technique definition that doesn't define wavelength or wave numbers. This might work for an arbitrary index, but a concrete index, like pH, temperature, or pressure, you have metadata that goes with it. An arbitrary structure doesn't support standard metadata. AnIML cannot allow you to create a variable which is indexing.

- It was suggested that we concentrate on the type of experiments which most vendors' instruments can run, not on experiments of academic interest. However, in some in some cases, these complex experiments are possible, such as the kinetics experiment. Many instruments have ability to index temperature. If there is a technique that says these different measurements happen in sequence, then everything could become children of the sample inlet technique, or there could be an alternate UV/Vis technique with sequence numbers as independent variable. In that case, any additional dependent variable can be added as extensions.
- Time ran out. The discussion on this topic will be continued next time.

IX. Adjourn: 12:40 pm EDT.

Minutes prepared by David Martinsen, ACS
ASTM E13.15 Secretary