

Power to the ODDs. Or: Why tei_all isn't everything

Peter Stadler

University of Paderborn, Germany; stadler@weber-gesamtausgabe.de

There is a general tension in the TEI between the freedom of encoding and the ideal of interoperability [Baumann 2011]. It's at the core of the TEI Guidelines that they are just a 'proposal', rather than normative technical guidelines. Yet, TEI conformance is quite strictly defined to what boils down to "Is your document valid against tei_all?" [TEI Guidelines, chapter 23.4 Conformance]. There had been some discussion about the issue last year on TEI-L, triggered by a collection of TEI examples which mostly failed to validate against tei_all [TEI-L 2013]. Some attempts had been made to validate those documents against their respective Customizations and there had been a strong opposition to disregarding P4 documents as TEI---but still, there had been no real attempts to overcome the constraints of an XML schema, which in fact only validates the surface structure rather than the abstract model.

To illustrate the point, consider the renaming of elements: It's in total agreement with the abstract model but results in an officially unclean modification. The same is true for the addition of e.g. the element <metropolis> and making it a member of model.placeNamePart. (This stresses the fact, that all members of model.placeNamePart are considered 'syntactic sugar' for the generic name element. [TEI Guidelines, chapter 10.4 The Manuscript Identifier])

While I try to argue in favor of the TEI abstract model as the basis (and sole?) criterion for TEI conformance, we need at least to push two things:

1. Tool support
2. Development of the abstract model

While the latter is continuously developed (e.g. by trying to move away from locally defined attributes and to group those in classes), charging the abstract model with new responsibilities will without doubt lead to new requirements. But my focus lies on the former issue, since I believe that with very little effort a lot can be achieved here. Inspired by last year's TEI workshop Perspectives on querying TEI-annotated data at the Members' Meeting in Rome, I developed a simple proof of concept of a generic TEI query, relying on the TEI class model as documented in an according ODD file [<https://github.com/peterstadler/Query-TEI-data-by-class>]. This XQuery module makes it possible to search for e.g. model.placeNamePart in a given document (given the according ODD file is readable), thus abstracting from the surface element names and allowing blind searches over heterogenous TEI documents.

The next thing I intent to tackle is the class based validation (there is no public repo yet, but will be for the conference). The idea is to first analyze the ODD file whether it denotes a subset of the TEI abstract model and then validate the given document against the resulting schema. This validation could be a first step in freeing the TEI from the constraints of tei_all, giving back some freedom of encoding to the scholars while preserving the facilities of data interchange.

References

[Baumann 2011] Bauman, Syd. "Interchange vs. Interoperability." Presented at Balisage: The Markup Conference 2011, Montréal, Canada, August 2 - 5, 2011. In Proceedings of Balisage: The Markup Conference 2011. Balisage Series on Markup Technologies, vol. 7 (2011). doi:10.4242/BalisageVol7.Bauman01.

[TEI-L 2013] "TEI Examples." Mailing List TEI-L. 6 Nov. 2013. <http://tei-l.970651.n3.nabble.com/TEI-Examples-tp4024809.html>