



Editorial Work Programme beyond P4

Lou Burnard

(with help from Syd Bauman)

<mailto:editors@tei-c.org>

November 2001



The aim of this talk

➡ Briefly, how we got from P3 to P4



The aim of this talk

- ➡ Briefly, how we got from P3 to P4
- ➡ Less briefly, how we may get from P4 to P5



The aim of this talk

- ➡ Briefly, how we got from P3 to P4
- ➡ Less briefly, how we may get from P4 to P5
- ➡ What we need to do now



The aim of this talk

- ➡ Briefly, how we got from P3 to P4
- ➡ Less briefly, how we may get from P4 to P5
- ➡ What we need to do now

[Yes, these slides are authored in TEI XML]



P3, P4, P5

- ➡ P3, aka *The Big Green Books* first published in April 1994 and last revised May 1999: the final deliverable from the ACH-ALLC-ACL project begun in Poughkeepsie in 1990
- ➡ P4, aka *The XML Version* is due for publication December 2001 (if you approve it!)
- ➡ P5, aka *TEI — the next generation* has no publication date yet. It will not necessarily retain compatibility with P3 but transition aids will be provided.



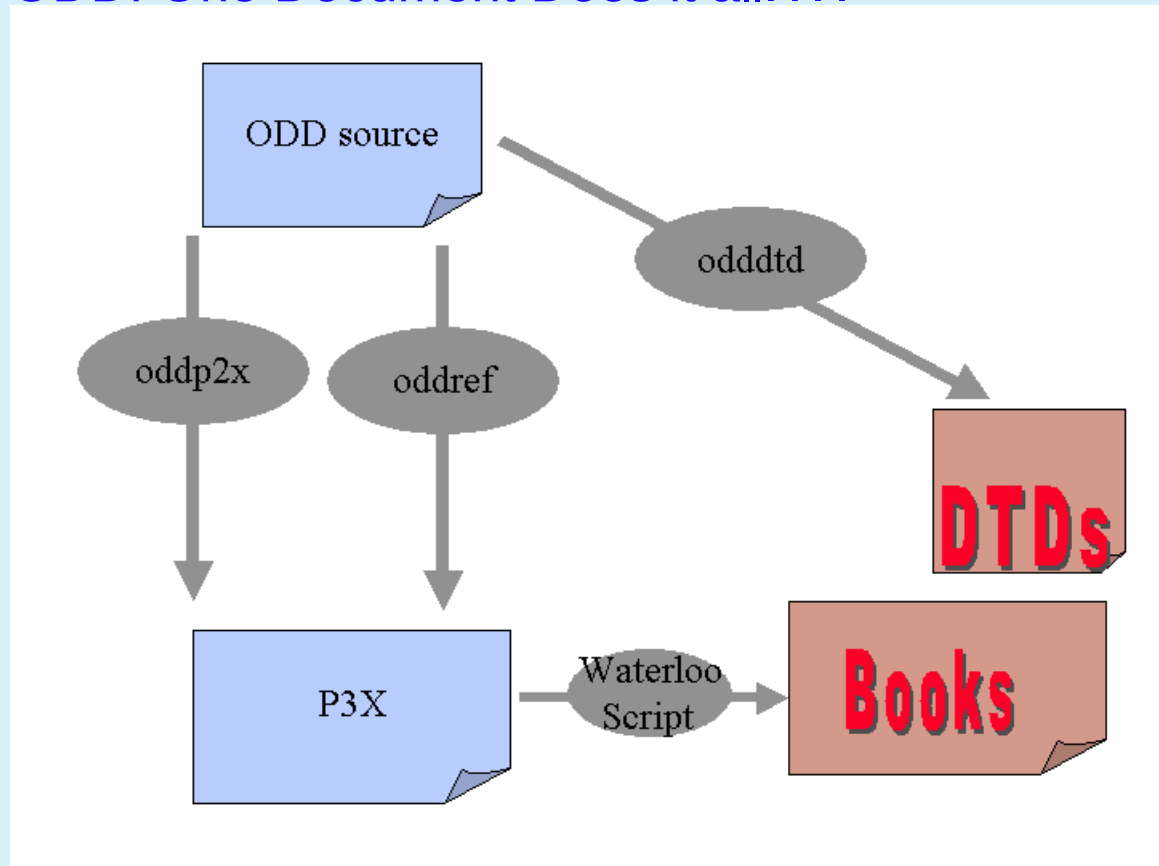
P4: The XML Edition

- ➡ Unlike P3, it is itself expressed (and processed) using XML
- ➡ Its DTD can be configured as SGML or as XML
- ➡ Parts now considered obsolescent are flagged as such, but retained
- ➡ Compatibility with P3 DTDs is also retained: a P3-conformant document is *ipso facto* P4-conformant

<http://www.tei-c.org/P4X/> ... and on your CD ... and in print

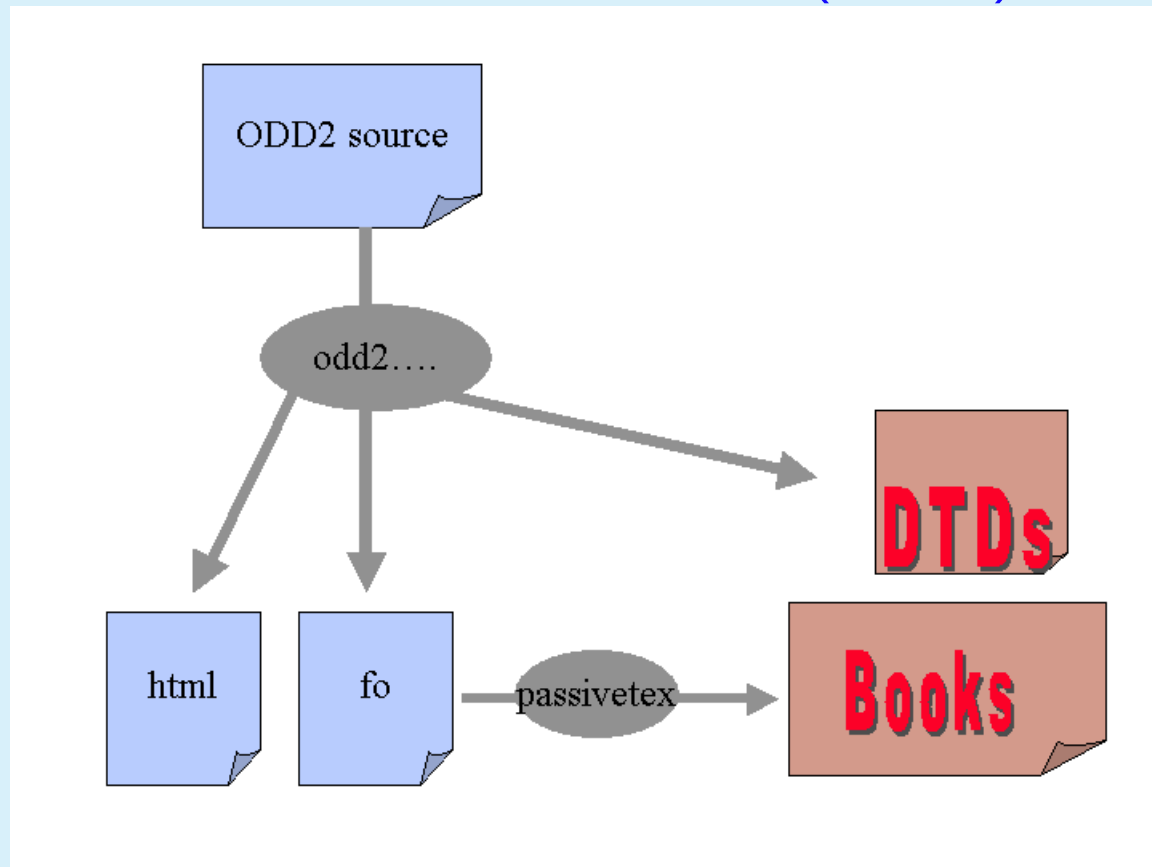
From P3 to P4

ODD: One Document Does it all....



From P3 to P4

ODD2: One Document Does it all (in XML)





Some infrastructural issues

- ➡ How can we make the Guidelines more accessible?
- ➡ Should we continue to support {SG,X}ML?
- ➡ How do we inter-operate with other standards?



Accessibility issues: suggestions

- ➡ Micro-manuals
- ➡ Project showcase
- ➡ Teaching packages
- ➡ TEI SWAT Team
- ➡ User feedback/PFC mechanism
- ➡ Translations



Work Is Needed In These Areas

1. Character set issues
 2. Core tags
 3. Metadata, mini-headers
 4. Manuscript description and transcription
 5. Standoff markup; XLink and TEI Xpointers
 6. Termbanks, dictionaries, ontologies
 7. Schemas, enhanced validation services, Son of ODD
 8. New areas I haven't thought of
- Evangelization and training as above



Character set issues

- ➡ Unicode only?
- ➡ W(h)ither the WSD?
- ➡ Glyph definition mechanism?
- ➡ Redefine LANG? What price xml:lang?



Core issues

(See forthcoming EDW76 *Editors' summary of proposed revisions*)

- ➡ An expandable list
- ➡ Things which have annoyed at least one TEI user enough for them to have proposed a fix
- ➡ For example:
 - ➡ consistent treatment for dates and times
 - ➡ review divtop and divbot elements
 - ➡ simple extensions for authoring
 - ➡ linking page images
 - ➡ etc.
- ➡ cf. the Technical Core Committee which produced the May 1999 revision.



Metadata matters!

The TEI Header is rooted in the library world, and as such is one of our most successful contributions to the digital library. But:-

- ➡ non-librarians need metadata too
- ➡ we need to provide better tools for interoperability with current (post-MARC) standards:
 - ➡ RDF, Topic Maps
 - ➡ Dublin Core
 - ➡ OLAC
 - ➡ METS
- ➡ we need to make good our claims with respect to long term digital preservation issues

We cannot afford to miss this boat!



Manuscripts, their description and transcription

- ➡ a key TEI community
- ➡ much work has been done; it simply needs to be completed
- ➡ emphasis should be on usability, promotion, demonstration
- ➡ expansion beyond medieval western tradition: include incunables, early print works, non-European traditions. . .
- ➡ digital scholarly editions
- ➡ modern manuscript transcription

The ESTATE project, and others



Standoff Markup, Xpointers etc.

- ➡ Existing TEI mechanisms need to be made more visible, and better tested
- ➡ Urgent need to review TEI extended pointer syntax and decide how to interoperate with W3C standards
- ➡ Critical to usage of TEI in the linguistics community
 - ➡ multimodal annotation
 - ➡ annotation graphs

Can we afford to miss this boat?



Termbanks, dictionaries, ontologies

- ➡ Original TEI proposals for terminology are obsolete; new standards are in town
- ➡ TEI proposals for dictionaries require extensive revision for technical reasons: we need to decide whether we are aiming to encode existing dictionaries, or new ones
- ➡ A new set of TEI proposals for computational lexica, ontologies, and termbanks would solve both problems
- ➡ Interoperability with other standards is also crucial here: cf TC37/SC4

Potentially, a key TEI constituency.



Schemata and enhanced validation

- ➡ In P3, we had DTDs only, with some prose specification of additional constraints (for TEI conformance. . .)
- ➡ In P4, XMLification mostly meant removing such few constraints as were available in DTD content models
- ➡ In P5, we could introduce full schematron-like content validation
- ➡ This would require further modification to the ODD system, but nothing much beyond what is already implicit
- ➡ We could also generate multiple schemas, with specific associated documentation

We could also consider other formalisms, e.g. UML



New areas we haven't thought of

By definition, we haven't anything much to say about these, but. . .

- ➡ tags for authoring
- ➡ tags for éditions génétiques
- ➡ tags for multimodal communicative acts
- ➡ develop feature structure annotation
- ➡ develop metrical annotation scheme
- ➡ develop image annotation scheme



How do we get there?

- ➡ How viable or well-defined are our procedures for
 - ➡ collecting and collating feedback



How do we get there?

- ➡ How viable or well-defined are our procedures for
 - ➡ collecting and collating feedback
 - ➡ setting up workgroups



How do we get there?

- ➡ How viable or well-defined are our procedures for
 - ➡ collecting and collating feedback
 - ➡ setting up workgroups
 - ➡ getting work out of workgroups



How do we get there?

- ➡ How viable or well-defined are our procedures for
 - ➡ collecting and collating feedback
 - ➡ setting up workgroups
 - ➡ getting work out of workgroups
 - ➡ incorporating work into the Guidelines



How do we get there?

- ➡ How viable or well-defined are our procedures for
 - ➡ collecting and collating feedback
 - ➡ setting up workgroups
 - ➡ getting work out of workgroups
 - ➡ incorporating work into the Guidelines



How do we get there?

- ➡ How viable or well-defined are our procedures for
 - ➡ collecting and collating feedback
 - ➡ setting up workgroups
 - ➡ getting work out of workgroups
 - ➡ incorporating work into the Guidelines

- ➡ The Council is the key



How do we get there?

- ➡ How viable or well-defined are our procedures for
 - ➡ collecting and collating feedback
 - ➡ setting up workgroups
 - ➡ getting work out of workgroups
 - ➡ incorporating work into the Guidelines

 - ➡ The Council is the key
- Politics will always confound us, but that's no reason to give up.