

Partial List of TEI elements for use in ESDS Qualidata XML DTD for transcribed interviews and other qualitative research materials				
Citation information: Draft DTD for qualitative data. ESDS Qualidata. Colchester, Essex: UK Data Archive. March 2005.				
Revised: 28 June 05 lb				
Notes:			List of tagsets used:	
1. These data description elements extend beyond level 4.0 Data Dscr fields of the DDI			Core (available to all)	
2. attributes list is partial, only special cases are elaborated			Spoken (for transcripton)	
3. extensive elements available for conversational analysis, not included			Linking	
4. I didn't pull elements from 18 (Transcription) but we should probably include it.			Analysis	
5. info in Conformance may help with DDI integration			Transcription?	
6. 29 on modifying shows how to suppress elements			names.dates	
Questions:			corpus	
1. there are tags for language profileDesc and langUsage-but isn't this done with DDI level?				
2. do we want DDI to carry all metadata defin, except for what is mandatory within TEI				
3. not clear yet how to use reference systems (6.10)				
4. how will we handle bibl citations - TEI or DDI?				
5. xprt and xref are no longer available; not clear how to do external linking in 14.2				
6. non-hierarchical coding (for thematic coding). Several strategies discussed in 31.Multiple hierarchies				
Ch. Of TEI Guidelines (p5)				
Ch. Of TEI	Element Name	Typical Use	Attributes	Typical use or other comments
5. TEI Header				
	<teiCorpus>			
		contains the whole of a TEI encoded corpus, comprising a single corpus header and one or more TEI.2 elements		
	<TEI>	contains a single TEI-conformant document, comprising a TEI header and a text		
	<teiHeader>	descriptive and declarative information making up an electronic title page		
	<fileDesc>			only required element in teiHeader
	<titleStmt>	groups information about the title of a work and those responsible for its intellectual content.		reqd element in fileDesc
	<editionStmt>	groups information relating to one edition of a text		reqd element in fileDesc
	<extent>	describes the approximate size of the electronic text		reqd element in fileDesc
6. Elements in all TEI docs				
	<p>	prose paragraph	type	any grouping
			id	unique identifier

6.3 Highlight and quotation	<q>	quotation		
	<emph>	marks words or phrases which are stressed or emphasized for linguistic or rhetorical effect.		
6.4 Editorial changes				
	<add>	text added to the document by author, transcriber, etc	place	indicates where the addition has been added by author, scribe, etc. (inline, left[margin], etc.); useful for "end of side n" notations
			resp	person responsible for addition
			hand	hand of the agent which made the deletion
			cert	certainty about attribution of person responsible
	<corr>	substitutes the correction, leaving the original reading as an attribute value	sic	shows original error
	<sic>	leaves the original text untouched, giving the correction as an attribute value	corr	shows corrected form
	<gap>	indicates a point where material has been omitted	desc	description of omitted text
			resp	person responsible for decision not to provide transcription
			reason	e.g. illegible', 'inaudible', 'irrelevant', 'cancelled'
			extent	estimate of amount of material omitted
		any text deleted, or marked at deleted	type	type of deletion
			status	faulty deletions, strikeouts with too much text
			hand	hand of agent who made decision
			resp	person responsible for deletion
			cert	certainty about attribution of person responsible
	<unclear>	any text which cannot be transcribed with certainty	reason	illegible, inaudible source, and so on
			resp	individual responsible for the transcription
	<reg>	contains a reading which has been regularized or normalized		
	<orig>	marked as following the original, rather than being normalized		
6.5 Names, numbers, dates+				
	<rs>	contains a general purpose name or referring string.		
	<name>	contains a proper noun or noun phrase		this can take attribute to handle subcategories: person, place, street, address, etc.
	<num>	any number		
	<date>			elements available for ranges also

	<time>			
	<abbr>	contains an abbreviation of any sort.		
	<expan>	contains the expansion of an abbreviation		
6.6 Links and cross references				
	<ref>	reference to another location in the current document	target	can use URI to point to other docs locally or on other systems
			type	
	<ptr>	pointer to another location in the current document	targType	
			crDate	
			resp	
	<list>	contains any sequence of items organized as a list.		
	<item>	contains one component of a list.		
6.8 Notes, etc				
	<note>	note or annotation	type	
			resp	who is responsible (interviewee or interveiwer)
			certainty	to handle variations of certainty about a text (also uncertainty)
			place	
			target	point of attachment for note
			targetEnd	end of span
			anchored	
6.9 Graphics				
	<graphic>	indicates the location of an inline graphic, illustration, or figure.		
6.10 Reference systems				
			n	need a <refsDecl> declaration statement in TEI header
			xml:id	
6.11 Bibliographic citations and refs				

	<bibl>	contains a loosely-structured bibliographic citation		
7 Default text structure				
	<front>	any prefatory matter, interviewer's comments and observations		
	<body>	whole body of text, including front and back matter		
	<back>	appendices, analytic thoughts following interview		
	<text>	the entire text		
	<group>	contains the body of a composite text, grouping together distinct texts		
	<div>	subdivision of body of text	type	to handle diverse subcategories within a division; can also be numbered, divX, to achieve hierarchies, but can't use both numbered and unnumbered
10.2 Elements unique to spoken texts				
	<u>	Complete utterance (turn) of one speaker	who	
	<pause>	a pause either between or within utterances		
	<vocal>	vocalized, but not necessarily lexical phenomena	who	
	<kinesic>	any communicative phenomenon, not necessarily vocal	who	
	<event>	any phenomenon, not necessarily vocalized or communicative	desc	description of the phenomena
	<tei.timed>	defines a set of attributes common to those elements which have a duration in time		start, end, and dur(ation)
14 Linking, segmentation and alignment				
	<link>	defines an association or hypertextual link among elements	target	
	<anchor>	specifies a location so that it may be pointed to		
	<seg>	identifies a segment of text so it may be pointed to	type	distinguishes amongst different purposes of segments

14.2 Advanced pointing				need to determine what system we will use for overlapping coding - themes, etc.
14.5 Synchronization				
	<when>	indicates a point in time either relative to other elements or absolutely.		these and related features enable synching of text with audio (example in 14.5.2 of TEI guidelines)
	<timeline>	provides a set of ordered points in time which can be linked to elements of a spoken text		
	<multiMedia>	indicates where audio, video or other media are available	type	audio; video and audio; video only (this is a new element)
15 Simple analytical mechanisms				
	<s>	sentence (or sentence-like) unit; does not allow nesting	id	unique identifier
	<w>	word	id	unique identifier
17.1.1 Certainty and responsibility				we are using <note> in 6.8 with appropriate type attributes to handle these occurrences. Cert and resp are also attributes available in most of the basic editing elements such as <add>, , etc.
18 Transcription				There are elements here for complex transcription (e.g., to differentiate levels of certainty. Most cases can be handled with tags in 6.4
20 Names and dates				needed to codes parts of names
	<persName>	a proper noun or proper-noun phrase referring to a person		these are some of the elements available
	<surname>	contains a family (inherited) name		
	<foreName>	contains a forename, given or baptismal name		

	<roleName>	a name component which indicates that the referent has a particular role or position in society		
	<placeName>	contains an absolute or relative place name.		these are some of the elements available
	<country>			
	<region>			
	<bloc>	a geo-political unit containing one or more nation states		
	<geoName>	place reference with a geographic feature, Thames River		
	<orgName>			
	<orgTitle>			
	<orgType>			
	<orgDiv>			
23.2 Contextual information				Great find! This is the start of a set of elements for describing context
23.2.1 Text description				
	<textDesc>	full description of the situation within which a text was produced		
	<channel>	describes the medium or channel by which a text is delivered or experienced		
	<domain>	describes the most important social context in which the text was realized		
	<interaction>	describes the extent, cardinality and nature of any interaction among those producing and experiencing the text		
	<purpose>	characterizes a single purpose or communicative function of the text		
23.2.1 Participants description				
	<particDesc>	additional information about the participants in a spoken text		
	<person>	describes a single participant	role	
			sex	

			age	
	<personGrp>	group on individuals treated as a single person for analysis		
	<birth>	date and place of birth	date	
	<education>	prose		
	<occupation>	informal description	scheme	classification sytem or taxonomy used
			code	occupation code defined within above system
	<socecStatus>		scheme	classification sytem or taxonomy used
			code	occupation code defined within above system
	<residence>	describes a person's present or past places of residence		
	<affiliation>	contains an informal description of a person's present or past affiliation with some organization		
	<particLinks>	way to describe any linkages among participants		
	<relation>		type	
			desc	
			active	
			passive	
			mutual	
23.2.3 Setting Description				
	<setting>	describes one particular setting in which a language interaction takes place		
	<locale>	contains a brief informal description of the nature of a place for example a room		
	<activity>	description of what a participant in a language interaction is doing other than speaking		

Interface with Atlas		This section still very much in draft stage		
Thematic codes				
	<codes size>	number of codes used in HU (could be single transcript?)		
	<code name>	name of code or theme	id	unique number identifier for the code
			au	code author
			cDate	creation date
			mDate	modified date
			software	caqdas tool used to create code
			cCount	
			qCount	
	<memos size>	number of memos in HU		
	<memo name>		id	unique number identifier for the code
			au	code author
			cDate	creation date
			mDate	modified date
			type	Method, Commentary, etc
			mime	eg text/plain
			fn	eg me_4