Bodleian TEI Manuscript Description Catalogue Consolidation: Report on Schema Creation

Dr James Cummings, University of Oxford

Version 1.1

2017-08-06

Introduction

This report is a comparison of the various collections of TEI files at the University of Oxford used for recording manuscript descriptions. Only those files provided by the Bodleian have been used. This includes collections for Armenian, Fihrist, Genizah, Georgian, Hebrew, Medieval, and Senmai catalogues. The Tibetan catalogue has not been used for this initial step as this was not supplied. The schema proposed is a permissive schema based on the current usage of the catalogues. A detailed examination of each of the collections have been undertaken. For information on each of the catalogues, the elements they contain, attributes used, and values for those attributes see each of their reports:

Compiled:

- Complete Bodley collections: http://research.it.ox.ac.uk/MSS/oxford.html

Individual collections:

- Armenian catalogue: http://research.it.ox.ac.uk/MSS/armenian.html
- Fihrist catalogue: http://research.it.ox.ac.uk/MSS/fihrist.html
- Genizah catalogue: http://research.it.ox.ac.uk/MSS/genizah.html
- Georgian catalgue: http://research.it.ox.ac.uk/MSS/geogian.html
- Hebrew catalogue: http://research.it.ox.ac.uk/MSS/hebrew.html
- Medieval catalogue: http://research.it.ox.ac.uk/MSS/medieval.html
- Senmai catalogue: http://research.it.ox.ac.uk/MSS/senmai.html
- Tibetan catalogue: http://research.it.ox.ac.uk/MSS/tibetan.html

Cambridge:

- Cambridge catalogue: http://research.it.ox.ac.uk/MSS/cambridge.html

The basic overall statistics for the individual collections above are:

Catalogue	XML Files	Distinct XML Elements	XML Elements	Attributes	Attribute Values
Armenian	124	80	20602	9812	9851
Fihrist	11677	143	1573493	809102	821456

Genizah	225	69	151101	92938	93260
Georgian	94	81	13332	5638	5676
Hebrew	473	101	70372	29027	30589
Medieval	9173	123	810749	462124	463942
Senmai	1051	60	118166	40520	42007
Tibetan	132	94	13748	4860	4975
Total:	22949	165	2771563	1454021	1471756

This does not include the Cambridge collection whose overall statistics at time of writing were:

- **Number of XML Files:** 21408
- **Number of XML Elements:** 2736394
- **Number of Attributes:** 2270576
- **Number of Attribute Values:** 2396549.

Recommendations

I propose the following recommendations for consideration of the project board:

1. The proposed schema should use the following TEI modules: core, figures, header, linking, msdescription, namesdates, tei, textstructure, transcr.
2. All files should have a <schemaRef> pointing to the location of the TEI Customization ODD.
3. All files should have full <authority>, <publisher> and <distributor> details even if these are identical.
4. All files should have a <projectDesc> describing the project which created them.
5. Each <msDesc> should be required to have an xml:id attribute that is unique across the collection based on its shelfmark as stored in msIdentifier/idno which it should also be required to have.
6. Some files, notably from the Genizah, Hebrew, and Senmai collections do not use the <text> element. While this is legal (where a sibling such as <facsimile> is provided), it is recommended that the files have the same structure across the collections and thus provide stub <text> and <body> elements.
7. Some basic textual transcription elements should be provided for those wishing to include some text (for example in quotations in the description) however this should be limited and those wishing to use these files for digital editions should create their own customisations.
8. Files from each collection should have an <idno type=”collection”>collectionName</idno> which may be used as a trigger for catalogue-specific Schematron rules.
9. Elements not used in the Oxford collections but that should be added to the schema include:
    a. am
    b. authority
    c. cit
    d. correspAction
    e. correspContext

    f. correspDesc
    g. dim
    h. distributor
    i. district
    j. ex
    k. expan
    l. geogFeat
    m. l
    n. lg
    o. msFrag
    p. musicNotation
    q. objectType
    r. orig
    s. pb
    t. projectDesc
    u. reg
    v. relatedItem
    w. schemaRef
    x. scriptDesc
    y. scriptNote
    z. sponsor
    aa. surfaceGrp
    bb. zone
10. In general attributes should not be required where the TEI does not require them.
11. Attribute value list should generally remain unconstrained for greater flexibility in the future.
12. At least the following attribute values should be provided with suggested values to encourage convergence but not require it:
    a. availability/@status
    b. bibl/@type
    c. decoNote/@type
    d. dimensions/@type
    e. handNote/@scope
    f. title/@type

Differences between collections

The existing Bodleian collections were compared both manually and programmatically. There are many similarities between the collections, especially when spot comparisons were made with non-Oxford collections, that show a commonality of manuscript description methodologies. This is most likely because some of the catalogues have been developed from earlier projects, notably the ENRICH project whose schema was later used as a starting point by OCIMCO and similar projects such as FIHRIST. That said, the differences are striking and interesting. Although a permissive schema is being designed, it is done so based on existing usage, rather than on some theoretical archetype of manuscript description.

Comparison of TEI modules used

Customizations of the TEI are made using its own literate programming customization language ‘ODD’ (One Document Does-it-all). From these customisation files it is usual to generate both human readable

documentation and schemas which document instances will be validated against. The major divisions in the TEI are that of modules. The TEI modules used by each of the collections are detailed in the The table below:

Proposed	Armenian	Fihrist	Genizah	Georgian	Hebrew	Medieval	Senmai	Tibetan	Cambridge
core	core	core	core	core	core	core	core	core	core
figures	-	figures	-	-	figures	figures	-	-	figures
header	header	header	header	header	header	header	header	header	header
linking	-	linking	-	-	-	linking	-	-	linking
msdescription	msdescription	msdescription	msdescription	msdescription	msdescription	msdescription	msdescription	msdescription	msdescription
namesdates	namesdates	namesdates	namesdates	namesdates	namesdates	namesdates	namesdates	namesdates	namesdates
tei	tei	tei	tei	tei	tei	tei	tei	tei	tei
textstructure	textstructure	textstructure	textstructure	textstructure	textstructure	textstructure	textstructure	textstructure	textstructure
transcr	transcr	transcr	transcr	transcr	transcr	transcr	-	transcr	transcr

These are precisely the modules that one would expect manuscript description to use. The core module contains the basic elements that occur in many TEI documents. The figures model provides elements for dealing with figures and tables. The header module contains elements for the general TEI metadata stored in the teiHeader. The linking module has elements for linking, segmentation, and alignment of texts. The manuscript description, by far the most detailed module of the TEI used for manuscript description catalogues, is for this specialised form of metadata. The names and dates module increases the basic name and date elements with additional elements and attributes.

Unused TEI modules

Not all modules of the TEI are used in the proposed schema. In specific the following ones are not included at all:

- analysis (Simple analytic mechanisms)
- certainty (Certainty and uncertainty)
- corpus (Corpus texts)
- dictionaries (Dictionaries)
- drama (Performance texts)
- figures (Tables, formulae, notated music, and figures)
- gaiji (Character and glyph documentation)
- iso-fs (Feature structures)
- nets (Graphs, networks, and trees)
- spoken (Transcribed speech)
- tagdocs (Documentation of TEI modules)
- textcrit (Critical apparatus)
- verse (Verse structures)

Comparison of elements used (by module)

These comparisons list the use of elements by collection, including the Cambridge files, and the proposed list of elements to include. Elements suggested for inclusion in the Bodleian schema which are not currently used in any of its collections are highlighted in bold.

TEI ‘Core’ Module

The TEI ‘Core’ module contains general elements found in lots of textual transcriptions. While this is not the focus of manuscript descriptions, many of these elements are also used in manuscript descriptions.

Proposed	**Armenian**	**Fihrist**	**Genizah**	**Georgian**	**Hebrew**	**Senmai**	**Tibetan**	**Medieval**	**Cambridge**
abbr	-	-	-	-	-	-	-	abbr	abbr
add	-	add	-	-	-	-	add	-	add
address	address	address	address	address	address	address	address	-	address
addrLine	addrLine	addrLine	addrLine	addrLine	addrLine	addrLine	addrLine	-	addrLine
-	-	-	-	-	-	-	-	-	analytic
author	author	author	author	author	author	author	author	author	author
bibl	bibl	bibl	bibl	bibl	bibl	-	bibl	bibl	bibl
biblScope	-	biblScope	-	-	-	-	-	biblScope	biblScope
-	-	-	-	-	-	-	-	-	biblStruct
-	-	-	-	-	-	-	-	-	cb
choice	-	-	-	-	-	-	-	choice	choice
citedRange	-	-	-	-	-	-	-	citedRange	-
cit	-	-	-	-	-	-	-	-	-
corr	-	-	-	-	-	-	-	corr	corr
date	date	date	date	date	date	date	date	date	date
del	-	-	-	-	desc	desc	-	del	del
desc	-	desc	desc	-	-	-	-	-	
editor	editor	editor	-	editor	editor	-	editor	editor	editor
email	email	email	email	email	email	email	email	-	email
expan	-	-	-	-	-	-	-	-	expan

foreign	-	foreign	-	-	foreign	foreign		foreign	foreign
gap	-	-	-	-	-	-	-	gap	gap
graphic	graphic	graphic	graphic	graphic	graphic	-	graphic	graphic	graphic
head	-	head	-	-	-	-	head	head	head
hi	-	-	-	-	hi	-	-	hi	hi
-	-	-	-	-	-	-	-	-	imprint
item	item	item	-	item	item	-	-	item	item
l	-	-	-	-	-	-	-	-	l
label	-	label	-	-	-	-	-	label	label
lb	-	-	-	-	-	-	-	lb	lb
lg	-	-	-	-	-	-	-	-	lg
list	list	list	-	list	list	-	-	list	list
listBibl	listBibl	listBibl	-	listBibl	listBibl	listBibl	listBibl	listBibl	listBibl
measure	-	-	-	-	-	-	-	measure	-
-	-	-	-	-	-	-	-	-	monogr
name	-	name	name	name	name	-	name	name	name
note	note	note	note	note	note	note	note	note	note
orig	-	-	-	-	-	-	-	-	orig
num	-	-	-	-	-	-	-	num	-
p	p	p	p	p	p	p	p	p	p
pb	-	-	-	-	-	-	-	-	pb
postCode	postCode	postCode	postCode	postCode	postCode	postCode	postCode	-	postCode
publisher	publisher	publisher	publisher	publisher	publisher	publisher	publisher	publisher	-
pubPlace	pubPlace	pubPlace	pubPlace	pubPlace	pubPlace	pubPlace	pubPlace	pubPlace	pubPlace
q	-	-	-	-	-	-	-	q	q
quote	-	-	quote	-	quote	-	-	quote	quote
ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
reg	-	-	-	-	-	-	-	-	reg

relatedItem	-	-	-	-	-	-	-	-	-
resp	-	resp	resp	-	resp	-	-	resp	resp
respStmt	-	respStmt	respStmt	-	respStmt	-	-	respStmt	respStmt
series	-	-	-	-	-	-	-	series	series
sic	-	-	sic	-	-	-	-	sic	sic
street	street	street	street	street	street	street	street	-	street
term	-	term	term	-	term	-	-	-	term
textLang	textLang	textLang	textLang	textLang	textLang	textLang	textLang	textLang	textLang
title	title	title	title	title	title	title	title	title	title
unclear	-	unclear	-	-	-	-	-	unclear	unclear

TEI ‘figures’ module

The TEI figures module provides markup for figures, formulae (sometimes used in collation formulas), as well as tables (not suggested).

Proposed	Armenian	Fihrist	Genizah	Georgian	Hebrew	Senmai	Tibetan	Medieval	Cambridge
-	-	-	-	-	-	-	-	-	cell
figDesc	-	figDesc	-	-	-	-	-	figDesc	-
figure	-	figure	-	-	-	-	-	figure	-
formula	-	formula	-	-	formula	-	-	formula	-
-	-	-	-	-	-	-	-	-	row
-	-	-	-	-	-	-	-	-	table

TEI ‘header’ module

The TEI header module provides metadata for the file as a whole. Some elements, for example those for the description of correspondence, did not exist in the TEI at the time that most of the Bodleian manuscript description schemas were constructed.

Proposed	Armenian	Fihrist	Genizah	Georgian	Hebrew	Senmai	Tibetan	Medieval	Cambridge
authority	-	-	-	-	-	-	-	-	authority

availability	availability	availability	availability	availability	availability	-	availability	-	availability
catDesc	catDesc	catDesc	-	catDesc	-	-	-	catDesc	catDesc
category	category	category	-	category	-	-	-	category	category
change	change	change	change	change	change	-	change	-	change
classDecl	classDecl	classDecl	classDecl	classDecl	classDecl	-	classDecl	classDecl	classDecl
correspAction	-	-	-	-	-	-	-	-	-
correspContext	-	-	-	-	-	-	-	-	-
correspDesc	-	-	-	-	-	-	-	-	-
distributor	-	-	-	-	-	-	-	-	-
edition	-	-	-	-	-	-	-	edition	edition
editionStmt	-	-	-	-	-	-	-	editionStmt	-
encodingDesc	encodingDesc	encodingDesc	encodingDesc	encodingDesc	encodingDesc	-	encodingDesc	encodingDesc	encodingDesc
extent	extent	extent	-	extent	extent	extent	extent	extent	extent
fileDesc	fileDesc	fileDesc	fileDesc	fileDesc	fileDesc	fileDesc	fileDesc	fileDesc	fileDesc
funder	funder	funder	-	funder	-	funder	-	-	funder
handNote	handNote	handNote	-	handNote	handNote	handNote	handNote	handNote	handNote
idno	idno	idno	idno	idno	idno	idno	idno	idno	idno
keywords	keywords	keywords	keywords	keywords	keywords	-	-	-	keywords
principal	principal	principal	-	principal	-	principal	principal	-	principal
profileDesc	profileDesc	profileDesc	profileDesc	profileDesc	profileDesc	-	profileDesc	-	profileDesc
projectDesc	-	-	-	-	-	-	-	-	-
publicationStmt	publicationStmt	publicationStmt	publicationStmt	publicationStmt	publicationStmt	publicationStmt	publicationStmt	publicationStmt	publicationStmt
revisionDesc	revisionDesc	revisionDesc	revisionDesc	revisionDesc	revisionDesc	-	revisionDesc	-	revisionDesc
schemaRef	-	-	-	-	-	-	-	-	-
scriptNote	-	-	-	-	-	-	-	-	scriptNote

sourceDesc	sourceDesc	sourceDesc	sourceDesc	sourceDesc	sourceDesc	sourceDesc	sourceDesc	sourceDesc	sourceDesc
sponsor	-	-	-	-	-	-	-	-	
taxonomy	taxonomy	taxonomy	taxonomy	taxonomy	taxonomy	-	taxonomy	taxonomy	taxonomy
teiHeader	teiHeader	teiHeader	teiHeader	teiHeader	teiHeader	teiHeader	teiHeader	teiHeader	teiHeader
textClass	textClass	textClass	textClass	textClass	textClass	-	-	-	textClass
titleStmt	titleStmt	titleStmt	titleStmt	titleStmt	titleStmt	titleStmt	titleStmt	titleStmt	titleStmt
typeNote	-	-	-	-	-	-	typeNote	-	

TEI ‘linking’ module

The TEI linking module provides various elements for linking of portions of text. The <seg> element used here is for marking segments of text with no specific semantics other than those given by its @type attribute.

Proposed	Armenian	Fihrist	Genizah	Georgian	Hebrew	Senmai	Tibetan	Medieval	Cambridge
seg	-	seg	-	-	-	-	-	seg	seg

TEI ‘msdescription’ module

As would be expected with collections of manuscript descriptions, the majority of the msdescription module is evident in the existing descriptions. The recommendation with respect to this is to include all TEI elements for manuscript description, including newer ones like <msFrag>. Indeed, it is recommended that the Bodleian TEI MSS customisation include the msdescription module as a whole, not removing any elements that may be added in the future. Some suggested customisations can be made for attribute values, and potentially removal of some attributes unlikely to be used.

Proposed	Armenian	Fihrist	Genizah	Georgian	Hebrew	Senmai	Tibetan	Medieval	Cambridge
accMat	-	-	-	-	-	-	accMat	accMat	accMat
acquisition	acquisition	acquisition	acquisition	acquisition	acquisition	acquisition	acquisition	acquisition	acquisition
additional	additional	additional	additional	additional	additional	-	additional	additional	additional
additions	additions	additions	-	additions	additions	-	-	additions	additions
adminInfo	adminInfo	adminInfo	adminInfo	adminInfo	adminInfo	-	adminInfo	adminInfo	adminInfo
altIdentifier	altIdentifier	altIdentifier	altIdentifier	altIdentifier	altIdentifier	altIdentifier	altIdentifier	altIdentifier	altIdentifier
binding	-	-	-	-	-	-	-	binding	binding

bindingDesc	-	bindingDesc	-	-	bindingDesc	bindingDesc	bindingDesc	bindingDesc	bindingDesc
catchwords	-	catchwords	-	-	catchwords	-	-	catchwords	catchwords
collation	-	collation	-	-	collation	-	collation	collation	collation
collection	collection	collection	-	collection	collection	collection	collection	-	collection
colophon	-	colophon	-	-	colophon	colophon	colophon	colophon	colophon
condition	-	condition	-	-	condition	condition	condition	condition	condition
custEvent	-	-	-	-	-	-	custEvent	-	-
custodialHist	-	-	-	-	-	-	custodialHist	-	-
decoDesc	-	decoDesc	-	-	decoDesc	decoDesc	decoDesc	decoDesc	decoDesc
decoNote	-	decoNote	-	-	decoNote	decoNote	-	decoNote	decoNote
depth	depth	-	-	-	-	depth	-	depth	depth
dim	-	-	-	-	-	-	-	-	-
dimensions	dimensions	dimensions	-	dimensions	dimensions	dimensions	dimensions	dimensions	dimensions
explicit	-	explicit	-	-	explicit	explicit	explicit	explicit	explicit
filiation	-	filiation	-	-	-	-	filiation	filiation	filiation
finalRubric	-	-	-	-	-	-	-	finalRubric	finalRubric
foliation	-	foliation	-	-	foliation	foliation	-	foliation	foliation
handDesc	handDesc	handDesc	-	handDesc	handDesc	handDesc	handDesc	handDesc	handDesc
height	height	height	-	height	height	height	height	height	height
heraldry	-	-	-	-	-	-	-	heraldry	-
history	history	history	history	history	history	history	history	history	history
incipit	incipit	incipit	-	incipit	incipit	incipit	incipit	incipit	incipit
institution	institution	institution	institution	institution	institution	institution	institution	institution	institution
layout	layout	layout	-	layout	layout	-	layout	layout	layout
layoutDesc	layoutDesc	layoutDesc	-	layoutDesc	layoutDesc	-	layouDesct	layoutDesc	layoutDesc
locus	locus	locus	locus	locus	locus	-	locus	locus	locus
locusGrp	-	locusGrp	locusGrp	-	-	-	-	-	-

material	-	material	material	-	material	material	-	material	material
msContents	msContents	msContents	msContents	msContents	msContents	msContents	msContents	msContents	msContents
msDesc	msDesc	msDesc	msDesc	msDesc	msDesc	msDesc	msDesc	msDesc	msDesc
msFrag	-	-	-	-	-	-	-	-	-
msIdentifier	msIdentifier	msIdentifier	msIdentifier	msIdentifier	msIdentifier	msIdentifier	msIdentifier	msIdentifier	msIdentifier
msItem	msItem	msItem	msItem	msItem	msItem	msItem	msItem	msItem	msItem
msName	-	msName	msPart	-	msName	-	-	msName	-
msPart	-	msPart	-	-	msPart	-	msPart	msPart	msPart
musicNotation	-	-	-	-	-	-	-	-	musicNotation
objectDesc	objectDesc	objectDesc	-	objectDesc	objectDesc	objectDesc	objectDesc	objectDesc	objectDesc
objectType	-	-	-	-	-	-	-	-	-
origDate	-	-	origDate	-	origDate	-	-	-	-
origin	origin	origin	origin	origin	origin	origin	origin	origin	origin
origPlace	-	origPlace	origPlace	-	origPlace	-	-	origPlace	origPlace
physDesc	physDesc	physDesc	physDesc	physDesc	physDesc	physDesc	physDesc	physDesc	physDesc
provenance	provenance	provenance	provenance	provenance	provenance	provenance	provenance	provenance	provenance
recordHist	recordHist	recordHist	recordHist	recordHist	recordHist	-	recordHist	recordHist	recordHist
repository	repository	repository	repository	repository	repository	repository	repository	repository	repository
rubric	-	-	-	-	-	-	-	rubric	rubric
scriptDesc	-	-	-	-	-	-	-	-	scriptDesc
seal	-	-	-	-	-	-	seal	seal	-
sealDesc	-	-	-	-	-	-	sealDesc	sealDesc	sealDesc
secFol	-	-	-	-	-	-	-	secFol	secFol
signatures	-	-	-	-	signatures	-	-	signatures	signatures
source	source	source	source	source	source	-	source	source	source
stamp	-	stamp	-	-	-	-	-	-	stamp
summary	summary	summary	summary	summary	summary	summary	summary	summary	summary

support	-	support	-	-	support	-	-	support	support
supportDesc	supportDesc	supportDesc	-	supportDesc	supportDesc	supportDesc	supportDesc	supportDesc	supportDesc
surrogates	-	-	-	-	surrogates	-	surrogates	surrogates	-
typeDesc	-	-	-	-	-	-	typeDesc	-	-
watermark	-	-	-	-	watermark	-	-	watermark	watermark
width	width	width	-	width	width	width	width	width	width

TEI ‘namesdates’ module

The TEI namesdates module provides additional name elements (and metadata for people, places, and organisations). The most common name-related elements should be added to the schema, but since the manuscript descriptions are not compiling metadata about individuals or places these named entity elements should not be added.

Proposed	**Armenian**	**Fihrist**	**Genizah**	**Georgian**	**Hebrew**	**Senmai**	**Tibetan**	**Medieval**	**Cambridge**
addName	-	addName	-	-	-	-	addName	-	addName
country	country	country	country	country	country	-	country	country	country
district	-	-	-	-	-	-	-	-	-
forename	-	forename	-	-	-	-	-	-	forename
geogFeat	-	-	-	-	-	-	-	-	-
geogName	-	-	-	-	geogName	-	-	-	-
orgName	-	orgName	orgName	-	-	-	-	-	orgName
persName	persName	persName	persName	persName	persName	persName	persName	persName	persName
placeName	-	placeName	placeName	placeName	placeName		placeName	placeName	placeName
region	region	region	region	region	region		-	region	region
settlement	settlement	settlement	settlement	settlement	settlement	settlement	settlement	settlement	settlement
surname	-	surname	-	-	-	-	-	-	surname

TEI ‘tei’ module

No elements are provided by the TEI ‘tei’ module, it is instead used to create much of the class infrastructure used by the TEI. It should be considered as required.

TEI ‘textstructure’ module

The text structure module of the TEI provides general textual structure elements. It is recommended that a text and body element be used, even if a stub, to provide a consistent format across all manuscript descriptions. docAuthor, usually used in title pages, is used by one collection for specialised use inside manuscript items.

Proposed	Armenian	Fihrist	Genizah	Georgian	Hebrew	Senmai	Tibetan	Medieval	Cambridge
body	body	body	-	body	-	-	body	body	body
docAuthor	-	docAuthor	-	-	-	-		-	
div	-	-	-	-	-	-	div	-	div
TEI	TEI	TEI	TEI	TEI	TEI	TEI	TEI	TEI	TEI
text	text	text	-	text	-	-	text	text	text

TEI ‘transcr’ module

The transcr module from the TEI provides a significant number of elements for document transcription, genetic editing, and image facsimiles. It is the last of these categories which is used in some collections. A full complement of related elements for surfaces, groups of surfaces, zones on surfaces should be provided. The basc elements for transcription that may appear inside descriptions when expanding abbreviations or editorially supplying text are also proposed.

Proposed	Armenian	Fihrist	Genizah	Georgian	Hebrew	Senmai	Tibetan	Medieval	Cambridge
-	-	-	-	-	-	-	-	-	addSpan
am	-	-	-	-	-	-	-	-	
ex	-	-	-	-	-	-	-	-	
-	-	-	-	-	-	-	-	-	damage
-	-	-	-	-	-	-	-	-	damageSpan
facsimile	facsimile	facsimile	facsimile	facsimile	facsimile	facsimile	facsimile	facsimile	facsimile
-	-	-	-	-	-	-	-	-	handShift
-	-	-	-	-	-	-	-	-	subst

supplied	-	-	-	-	-	-	-	supplied	-
surface	-	-	-	-	-	-	-	surface	-
surfaceGrp	-	-	-	-	-	-	-	-	-
zone	-	-	-	-	-	-	-	-	-

Attribute value lists

In TEI customisations one is able to constrain the value lists provided to attributes, and indeed this is encouraged on a project basis to reduce human error. Across the collections surveyed there are a number of attribute value lists which while open in the TEI could be either more tightly controlled or provided as suggested values. Below is a table of teidata.enumerated class of attributes and how they are used in the Oxford collections. The table contains the name of the element, the attribute in question, the existing list of distinct-values, a proposal and notes. In the majority of the cases there is no contraint proposed, in no cases is it suggested that the attribute be made compulsory. 'Suggested values' means that a subset of the list should be provided as suggested values but the actual list not constrained.

Element Name	Attribute	Existing Value List	Proposed	Notes
addName	@type	khitab, kunyah, laqab, nisbah	No constraint	
altIdentifier	@type	SCN, external, former, internal, other, palimpsest, partial	No constraint	
availability	@status	restricted	free, restricted, unknown	
bibl	@type	LYELL, MS, QUARTO, SC, abridgement, bible, commentedOn, digitised-version, edition, extract, extracts, intervening, ms, realted-volumes, referred, related, related-items, related-volumes, repertory, text-relations, textual-relations,	Suggested values	• MS, • QUARTO, • SC, • OC, • bible, • commentary, • edition, • extract, • related, • text-relations, • translation

		translated, translation		
collection	@type	CUL_collection, main, sub	No constraint	
custEvent	@type	check, other	No constraint	
date	@type	approxDate	No constraint	
decoNote	@type	border, bordersInitials, decoration, diagram, diagrams, frieze, illustration, initial, initial_border, initials, marginal, marginalSketches, micrography, miniature, other paratext, printmark, rubrication, secondary, unspecified, unwan	Suggested values	• border, • diagram, • illustration, • initial, • marginal, • micrography, • miniature, • other, • rubrication,
dimensions	@type	binding, book, folia, leaf, leaves, line-height, membrane, 'number of folia', photograph printed, ruled, ruledColumn, ruling, unknown, written	Suggested values	• binding, • folia, • leaf, • line-height, • ruled, • written, • other
explicit	@type	commentedOn, preface, prologue, text	No constraint	
handNote	@scope	major, minor, sole	Constrained by TEI	Keep constraint
idno	@type	LDAB, PR, SCN, TM, call_number, cat_no, microfilm, part, shelfmark, sum_cat_no	No contraint	
incipt	@type	argument, basmalla, commentedOn, dedication, lemma, preface, prologue, text, verses	No constraint	

listBibl	@type	related-items, related-volumes, text-relations	No constraint	
list	@type	gloss	No constraint	
measure	@type	leavesCount, ruledSpace	No constraint	
name	@type	artist, church, corporate, dedicatee, org, person, place, unknown	Suggested values	Person, place, org, unknown, other
note	@type	content, dimensions, extent, gloss, music, text-relations, versions	No constraint	
orgName	@type	ment, supplied	No constraint	
persName	@type	acr, alt, ara, aut, author, desc, heb, ment, org, own, owner, par, parallel, patron, scribe, standard, standatrd, supplied, wit	No constraint	
placeName	@type	aut, heb, ment	No constraint	
quote	@type	acr, beg, col, ebd, end, head, heb, own, post	No constraint	
ref	@type	#a5, page	No constraint	
region	@type	country, supranational	No constraint	[could use 'bloc' rather than 'region'?]
rubric	@type	commentedOn, preface	No constraint	
seg	@type	gathering, genreform, subject	No constraint	

sourceDesc	@default	false	No constraint	
title	@type	alt, alternative, collection, desc, general, main, parallel, sub, supplied, uniform	Suggested values	Keep TEI sample values of