



TEI standards for codicology

Yasmin Faghihi
Head of Near and Middle Eastern Department
Cambridge University Library, UK

Why TEI ?

- The **TEI metadata standard** and **Digital Humanities**
 - History
 - Rational
 - Characteristics

What's new for Manuscript research?



- **TEI/XML**

- Transcription
- Collation and (digital) editing
- Facsimile (digital imaging)
- Manuscript description

- **TEI MSS description**

vs

- **Machine Readable Cataloguing (MARC)**

Example of support descriptions in MARC

MARC

500 : \$aSupport: non-European (Arab) laid paper with 20 laid lines in 32 mm. or 6 laid lines per cm. (horizontal) ; chain lines are difficult to distinguish but a set of 3 with 15-16 mm. between lines can be seen ; replacement folia in another Oriental laid paper, but with lines too indistinct to measure ; some pest damage to some leaves, though not so badly as to obscure text ; edges of a number of leaves are badly damaged.

TEI

```
<objectDesc form="codex"> <support><supportDesc material="chart">  
<material>European (watermark) and Eastern paper, plain and dark buff, mixed  
thickness (thin, regular and thick), one size, glossy / some folios have mending at  
their edges.</material></support><extent>1 + 271 ff. <dimensions type="leaf"  
unit="cm">
```

Example of mark-up for <support>

MARC

500 : \$aSupport: non-European (Arab) laid paper; some pest damage to some leaves, though not so badly as to obscure text ; edges of a number of leaves are badly damaged.

TEI

```
<objectDesc form="codex"> <support><supportDesc material="chart">  
<material> non-European (Arab) laid paper </material><condition>edges of a  
number of leaves are badly damaged.<condition></support>
```

Machine readable:

```
form ="codex">
```

```
material="chart">
```

Schema, standards & practice

- 500 TEI elements 23 chapters
- Schema vs template
- Tailored schema or generic schema?
- Orientation on precedent (i.e. ENRICH schema)
- legacy data
- Evolving state of the art
- Consistency of practice

Physical Description in <MsDesc>

Examples of descriptors : <elements>

- **<PhysicalDesc>**

- <objectDesc>

- <supportDesc><extent><collation><foliation><condition>

- <support><material><watermark><dimension>

- <layout description>

- <layout ruled lines="34">

- <handDesc><handNote><scriptDesc><scribe>

- <decoDesc><decoNotes><seal>

Dates & additions in <MsDesc>

```
<bindingDesc><p>repaired and firmly rebound</p><decoNote  
type="illustration"/></bindingDesc>
```

```
<history><origin><originDate calendar=Hijri-qamari when="1673">dated  
1083</originDate>
```

```
<acquisition><p>Purchased by Ms Enriqueta Rylands, on behalf of the  
John Rylands Library, in <date Calendar="Gregorian"  
when="1901">1901</date><p></acquisition></history>
```


Example: Quire structures

[1] i, 1-9 (8), 10 (6), 11-20 (8), 21 (7), i
[2] I-III⁸, IV¹⁰, V-IX⁸
[3] IV(32), IV-1(40), 9 IV(120), IV-4
[4] 1-48, 52, 64-1, 7-1010

```
<collation>  
<p>  
  <formula>1-3:8, 4:6, 5-13:8</formula>  
</p>  
</collation>
```

```
<collation>  
<p>I (1, 2+9, 3+8, 4+7, 5+6, 10); II (11, 12+17, 13, 14, 15, 16, 18,  
  19).</p>  
</collation>
```

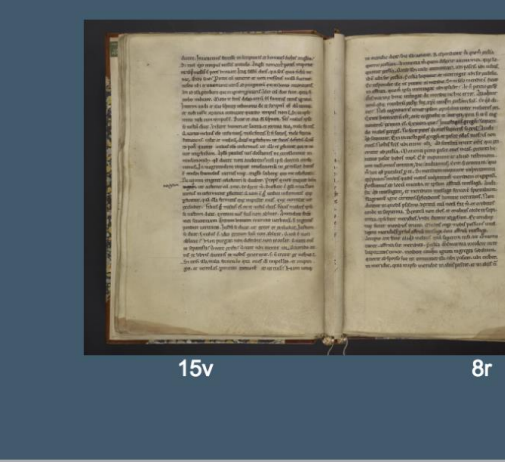
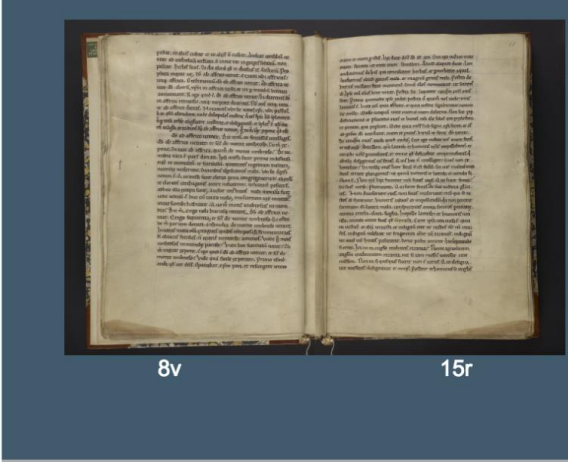
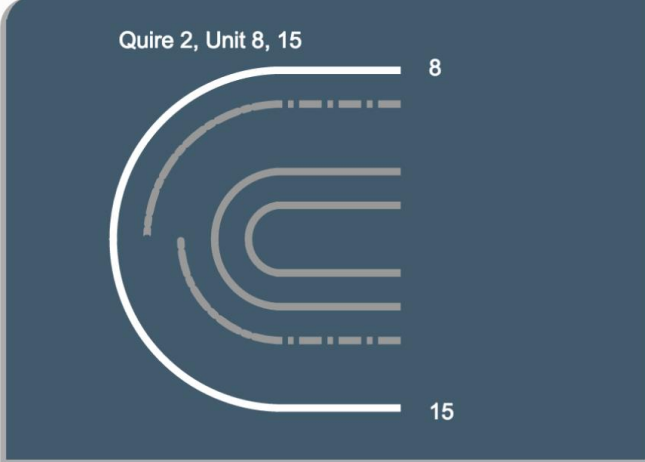
```
<collation>The written leaves preceded by an original flyleaf,  
  conjoint with the pastedown.</collation>
```

Visualisation of quire structure

```
<mapping> <map leaf="1.2" side="r"> <term target="#c2 #b1 #e5"/>
</map> <map leaf="1.2" side="v"> <term target="#c2 #e5 #b1"/> </map>
<map leaf="1.3" side="r"> <term target="#c1 #e5 #b1 #d1"/> </map>
<map leaf="1.3" side="v"> <term target="#c1 #c2 #e5 #b1 #d1"/> </map>
```

 **GENERATED USING THE SCHOENBERG INSTITUTE FOR MANUSCRIPT STUDIES COLLATION MODELER**
University of Pennsylvania Sancti Augustini opera quaedam, Ms.Codex708 [Select a Quire ↕](#)

Quire 2 (10)



VISCOLL

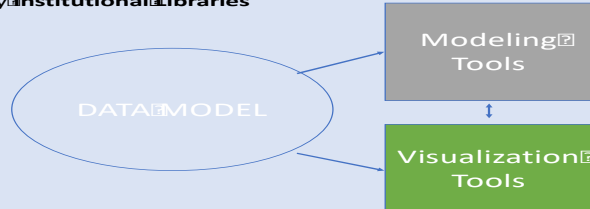
Visualizing the physical structure of medieval manuscripts

Dot Porter, Alexandra Gillespie, Alberto Campagnolo, Laura Mitchell, Rachel Di Cresce

What is VisColl?

VisColl is a system for building models of the physical collation of manuscripts, and then visualizing them in various ways.

VisColl is designed for use by **Individual Scholars** and by **Institutional Libraries**



The **Data Model** is central to VisColl and describes each individual leaf in the textblock, how the leaves are ordered, and most importantly which leaves are physically connected— which leaves form two halves of a sheet, and other words which are **conjoined**



The uncovered spine of a manuscript shows several quires sewn together

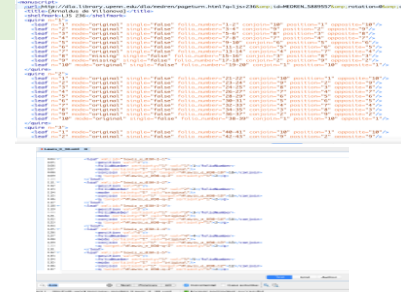
Using the **Data Model** as a guide, projects can develop tools for **building models** and for **visualizing models**, and know that the tools using the model will create models that can be shared between tools.

Data Model 1.0: Leaves are grouped into **Quires** and have a few simple descriptors

Data Model 2.0: Leaves are described individually, and **Quire ID** is an attribute of leaf. This allows for **uncertainty** that is unavailable in DM 1.0. Additionally, DM 2.0 enables the configuration of **taxonomies** that can be mapped onto leaves or groups of leaves.

Physical Collation

Medieval manuscripts are made of **Sheets** of parchment or paper that are **Stacked** and **Folded** into booklets called **Quires**. Quires are stitched together to form the **Textblock**. VisColl models the **Textblock**



VisColl currently has two implementations

VisColl Web Application @ University of Toronto

The University of Toronto, through the Mellon-funded project entitled **Digital Tools for Manuscript Study**, is developing a robust VisColl web application which implements the **Data Model 2.0**, and allows users to visually manipulate and present diagrams and metadata in real time.

With the VisColl application scholars will be able to:

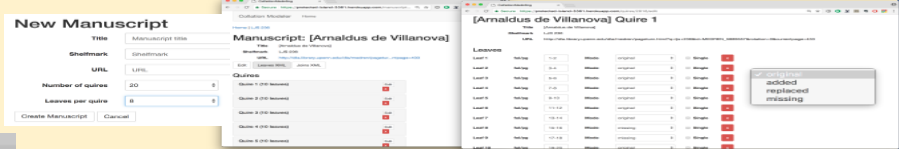
- create basic collation diagrams
- add sub-quires and booklets
- include metadata at the leaf level
- batch edit diagrams
- add manuscript images
- share visualizations
- export diagrams as image files for use in publications

In addition, for better web compatibility, we hope to integrate the application with popular data standards such as IIIF

VisColl Collation Modeler & Collation Visualizer @ University of Pennsylvania

The Schoenberg Institute for Manuscript Studies at the University of Pennsylvania has developed the **Collation Modeler** and **Collation Visualizer**, currently using the **Data Model 1.0**, that provide a very simple interface for creating models and visualizing them. **SIMS** is currently using these tools to generate models for the CLIR-funded **Bibliotheca Philadelpiensis** project.

The **Collation Modeler** provides a tabular interface for creating quires and describing leaves.



The **Collation Visualizer** generates three different views:

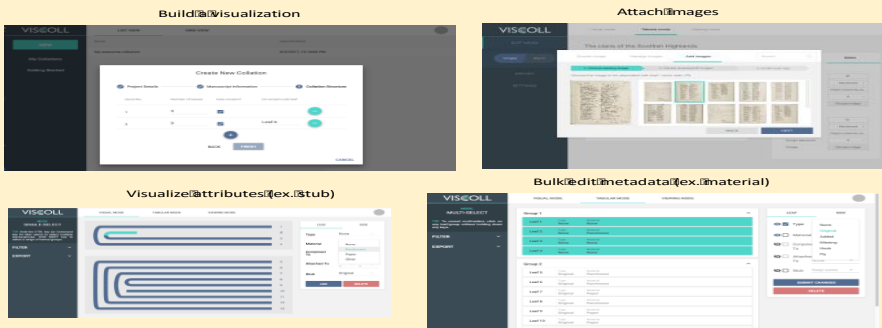
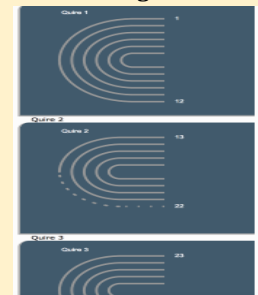
1. Collation formula: Multiple variants are possible

Collation Formula for Book of Hours, Use of Calixtus (Lection, Latin 9 699)
 Previous: 1 (open) unpaired
 1 10, 2 10, 3 10, 4 10, 5 10, 6 10, 7 10, 8 10, 9 10, 10 10, 11 10, 12 10, 13 10, 14 10, 15 10, 16 10, 17 10, 18 10, 19 10, 20 10, 21 10
 Previous: 1 (open) unpaired
 1 10, 11 10, 12 10, 13 10, 14 10, 15 10, 16 10, 17 10, 18 10, 19 10, 20 10, 21 10
 Previous: 1 (open) unpaired
 1 10, 11 10, 12 10, 13 10, 14 10, 15 10, 16 10, 17 10, 18 10, 19 10, 20 10, 21 10
 Previous: 1 (open) unpaired
 1 10, 11 10, 12 10, 13 10, 14 10, 15 10, 16 10, 17 10, 18 10, 19 10, 20 10, 21 10
 Previous: 1 (open) unpaired
 1 10, 11 10, 12 10, 13 10, 14 10, 15 10, 16 10, 17 10, 18 10, 19 10, 20 10, 21 10

3. Bifolia view: showing images arranged as conjoined leaves rather than facing pages



2. Diagrams



Opportunity and challenge



Research is still at the



of the process

Any data set is skewed by your own interest

Digital data openly reveals the intention behind the data

You can take the machine out of the human but you can't take the human out of the machine.

