Modelling Codicological Sequence With the TEI

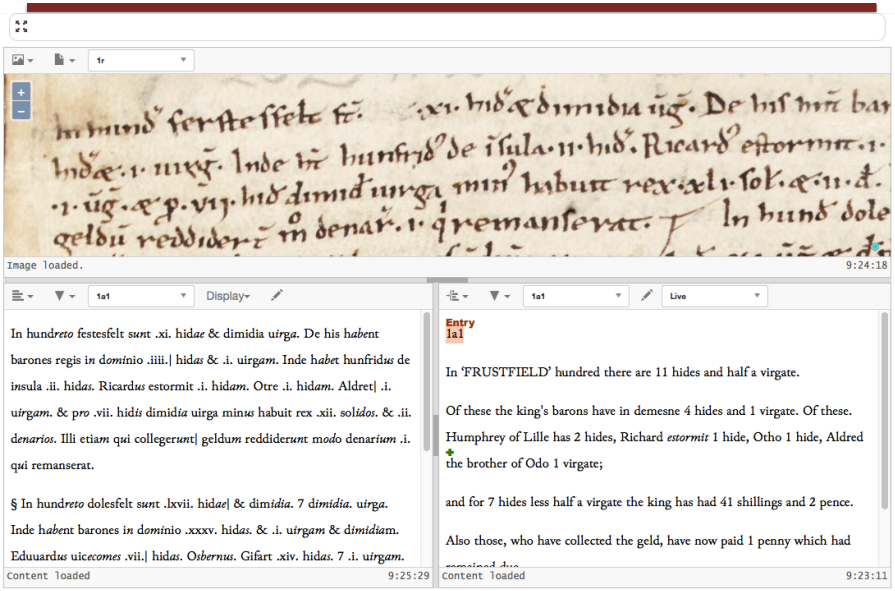
Peter A. Stokes
Department of Digital Humanities
King's College, London
peter.stokes@kcl.ac.uk
@pa_stokes

The Exon Domesday Project

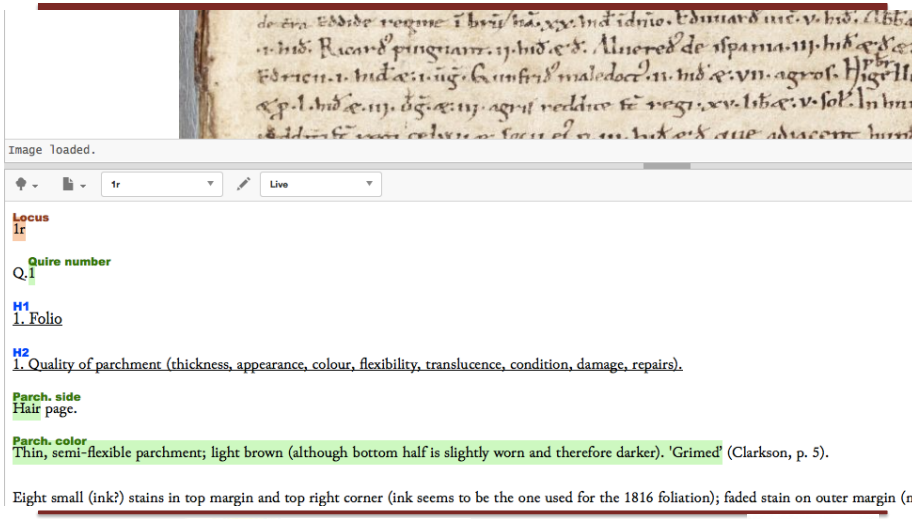
- Three-year, AHRC funded project, began 1 October 2014
- Extending DigiPal framework to provide:
  - Complete high-resolution facsimile
  - Complete text (semi-diplomatic and expanded) and translation
  - Palaeographical analysis of all scribal hands
  - Codicological analysis of complete manuscript
  - Historical analysis of land holdings (drawing on PDE project)
- Project Team: Julia Crick (PI); Stephen Baxter (Oxford, Co-I); Peter Stokes (Co-I); Geoffroy Noël (Lead Analyst Developer); Chris Lewis (Research Fellow); Fran Alvarez-Lopez (Research Associate); Frank Thorne (Visiting Research Fellow); Lois Lane (PhD); Alex Dymond (Oxford, PhD)

KING'S College LONDON | Arts & Humanities Research Council | Exon Domesday exondomesday.ac.uk | DigiPal http://digipal.eu

Exon Domesday in (TEI and) New DigiPal

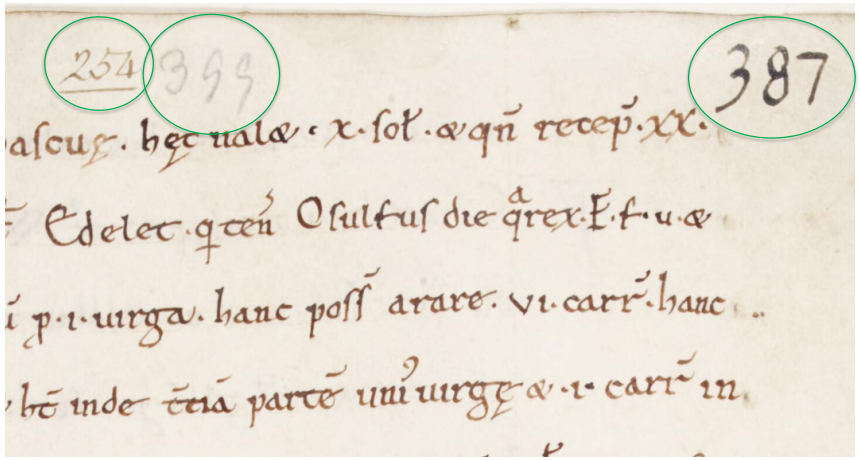


Codicological Profile in (New) DigiPal



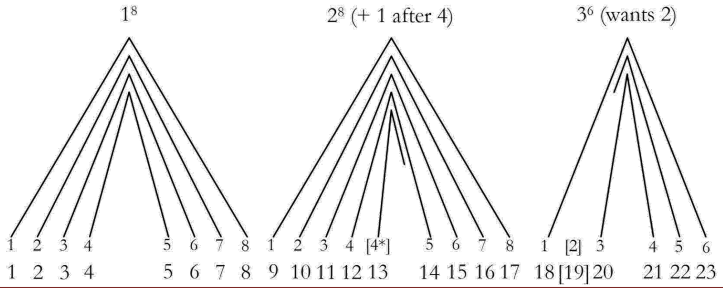
KING'S College LONDON | Arts & Humanities Research Council | Exon Domesday exondomesday.ac.uk | DigiPal http://digipal.eu

Signs of Re-Ordering

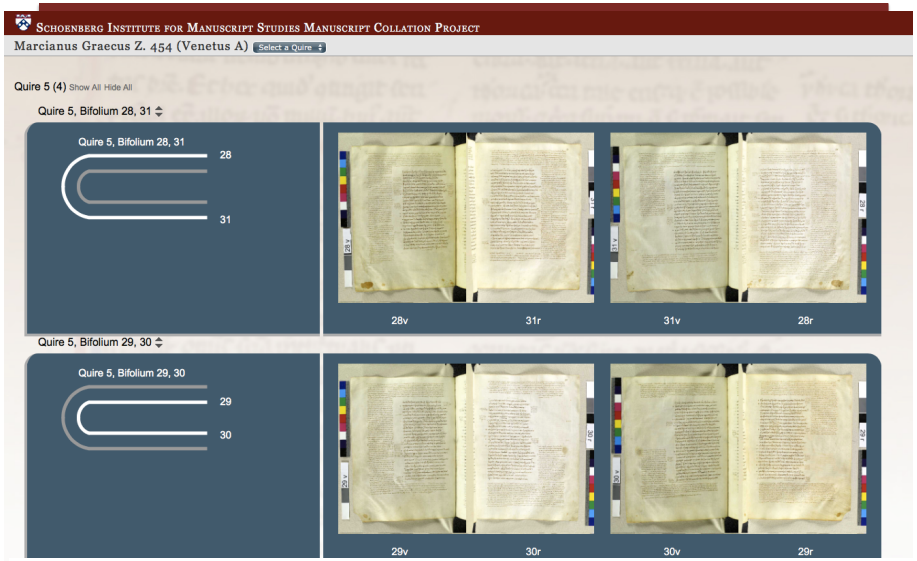


The 'Codicological Model'

- We've built a customisation of the DigiPal framework to capture scribal practices as well as text and translation
- We're now building a codicological model, ideally to allow users to change the order of gatherings, to see what happens when the sequence is changed



Dot Porter, Manuscript Collator



http://dorpdev.library.upenn.edu/collation/

Dot Porter, Collation Modelling

```
<quire n="2">
  <leaf n="1" mode="original" single="false" folio_number="8" conjoin="8" position="1"/>
  <leaf n="2" mode="original" single="false" folio_number="9" conjoin="7" position="2"/>
  <leaf n="3" mode="original" single="false" folio_number="10" conjoin="6" position="3"/>
  <leaf n="4" mode="original" single="false" folio_number="11" conjoin="5" position="4"/>
  <leaf n="5" mode="original" single="false" folio_number="12" conjoin="4" position="5"/>
  <leaf n="6" mode="original" single="false"
  <leaf n="7" mode="original" single="false"
  <leaf n="8" mode="original" single="false"
</quire>
```

```
<quire n="2" positions="6"/>
<quire n="3" positions="8">
   <less>6</less>
</quire>
<quire n="4" positions="8">
   <less>4</less>
</quire>
```

github.com/leoba/VisColl/ See also github.com/demery/collation-modeling/

Customising the TEI (after Dot Porter)

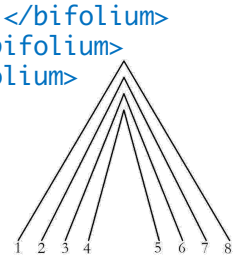
```
<exon:quire n="1" xml:id="q1">
   <exon:folio xml:id="1a1" n="1" conjoins="2b1" single="false" ruled="before" hair="hf"/>
   <exon:folio xml:id="1b1" n="2" single="true" ruled="unruled" hair="fh"/>
   <exon:folio xml:id="2a1" n="3" single="true" ruled="after" hair="hf"/>
   <exon:folio xml:id="2b1" n="4" conjoins="1a1" single="false" ruled="before" hair="fh"/>
</exon:quire>
<exon:quire n="2">
   <exon:folio xml:id="3a1" n="1" conjoins="4b1" single="false" ruled="before" hair="hf"/>
   <exon:folio xml:id="3b1" n="2" conjoins="4a1" single="false" ruled="unruled" hair="fh"/>
   <exon:folio xml:id="4a1" n="3" conjoins="3b1" single="false" ruled="after" hair="hf"/>
   <exon:folio xml:id="4a1*" n="3*" single="true" ruled="" hair="hf"/>
   <exon:folio xml:id="4b1" n="4" conjoins="3a1" single="false" ruled="before" hair="fh"/>
</exon:quire>
```

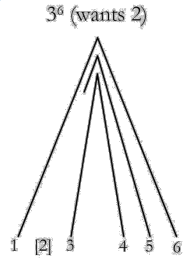

Codicological Model

Gatherings can be modelled as OHCO…

```
<bifolium>
   <bifolium>
      <bifolium>
         <bifolium>
         </bifolium>
      </bifolium>
   </bifolium>
</bifolium>
```

```
<bifolium>
   <bifolium>
   </bifolium>
   <singleton/>
</bifolium>
```





Gerrit Brüning, 'faustedition collation visualisation'

encoding of the collation (non-TEI)

```
<archivalDocument>
   <disjunctLeaf>
      <page/>
      <page/>
   </disjunctLeaf>
   <sheet>
      <leaf>
         <page/>
         <page/>
      </leaf>
      <sheet> [9 lines]
      <leaf> [3 lines]
   </sheet>
```


Mapping to the TEI: Text

```
<exon:quire n="1" xml:id="q1">
   <exon:folio xml:id="1a1" n="1" conjoins="2b1" single="false" ruled="before" hair="hf"/>
   <exon:folio xml:id="1b1" n="2" single="true" ruled="unruled" hair="fh"/>
   <exon:folio xml:id="2a1" n="3" single="true" ruled="after" hair="hf"/>
   <exon:folio xml:id="2b1" n="4" conjoins="1a1" single="false" ruled="before" hair="fh"/>
</exon:quire>
<exon:quire n="2">
   <exon:folio xml:id="3a1" n="1" conjoins="4b1" single="false" ruled="before" hair="hf"/>
   <exon:folio xml:id="3b1" n="2" conjoins="4a1" single="false" ruled="unruled" hair="fh"/>
   <exon:folio xml:id="4a1" n="3" conjoins="3b1" single="false" ruled="after" hair="hf"/>
   <exon:folio xml:id="4a1*" n="3*" single="true" ruled="" hair="hf"/>
   <exon:folio xml:id="4b1" n="4" conjoins="3a1" single="false" ruled="before" hair="fh"/>
</exon:quire>
```

```
<surfaceGrp type="quire" corresp="#q1">
   <surfaceGrp type="bifolium">
      <surfaceGrp type="folio" corresp="#1a1"> [20 lines]
      <surfaceGrp type="singleton">
         <surfaceGrp type="folio" corresp="#1b1"> [19 lines]
      </surfaceGrp>
      <surfaceGrp type="singleton">
         <surfaceGrp type="folio" corresp="#2a1"> [14 lines]
      </surfaceGrp>
      <surfaceGrp type="folio" corresp="#2b1"> [15 lines]
   </surfaceGrp>
</surfaceGrp>
```

Mapping to the TEI: Text

```
<surfaceGrp type="singleton">
   <surfaceGrp type="folio" corresp="#1b1">
      <surface type="page">
         <zone xml:id="p3" next="">§ In hundreto cicimethorne sunt .c. lxix. hidae &amp;
            &amp; .vi. hidas &amp; dimidiam. &amp; dimidiam. uirgam. Abbas malmesberiens
            hidas. Rotbertus nepos gloecestre .iii. hidas &amp; dimidiam. uirgam. de qu
            dimidia. uirga minus sunt reddite regi in constitutis terminis .xxvii. libr
            super eos qui collegerunt geldum recuperauerunt Walterus &amp; focii eius .
            solidis &amp; .ix. denariis quos inuenerunt. episcopus &amp; socii eius.</z
      </surface>
      <surface type="page">
         <zone xml:id="p4" prev="#g1p1z2">§ In hundreto Aluuartberie sunt .lxv. hidae &
            in dominio .xxix. hidas. &amp; dimidiam. &amp; .i. uirgam .iii. agras minus
            &amp; dimidiam. Eduuardus uicecomes .i. hidam. Vluricus .i. hidam. Godescal
            hidam &amp; dimidiam uirgam. Saricus .i. hidam &amp; dimidiam uirgam. Gaufr
            .ii. hidas quas rex accepit de eo. Suauinc .i. uirgam. Vluietus dimidiam hi
            presbiter .iii. hidas. Ediua .ii. hidas. Walerannus .ii. hidas &amp; .i. ui
            .i. uirgam. Gode- man .i. uirgam. Nigellus dimidiam uirgam. Eded .i. hidam.
            pro .xxxv. hidis &amp; dimidia. &amp; dimidia uirga sunt reddite regi .x. l
      </surface>
   </surfaceGrp>
</surfaceGrp>
```


Modelling Singletons and Stubs

```
<surfaceGrp type="singleton">
   <surfaceGrp type="folio" corresp="#1b1">
      <surface type="page">
         <zone xml:id="p3" next="">§ In hundreto cicimethorne sunt .c. lxix. hidae &amp;
            &amp; .vi. hidas &amp; dimidiam. &amp; dimidiam. uirgam. Abbas malmesberiens
            hidas. Rotbertus nepos gloecestre .iii. hidas &amp; dimidiam. uirgam. de qu
            dimidia. uirga minus sunt reddite regi in constitutis terminis .xxvii. libr
            super eos qui collegerunt geldum recuperauerunt Walterus &amp; focii eius .
            solidis &
      </surface>
      <surface type="
         <zone xml:id
            in domini
            &amp; dim
            hidam &am
            .ii. hida
            presbiter
            .i. uirga
            pro .xxxv
      </surface>
   </surfaceGrp>
</surfaceGrp>
```

```
<surfaceGrp type="bifolium">
   <surfaceGrp type="folio" corresp="#3a1"> [13 lines]
   <surfaceGrp type="singleton">
      <surfaceGrp type="stub">
         <surface type="stub">
            <zone/>
         </surface>
         <surface type="stub">
            <zone>Traces of writing</zone>
         </surface>
      </surfaceGrp>
```


Codicological Constraints

1. All Folios comprise exactly two Pages.
2. For parchment, Pages must be either Hair side (H) or Flesh side (F). A Folio must comprise one H Page and one F Page.
3. For parchment, Pages must be one of Ruling side, Non-Ruling side, or Unruled. A Folio must comprise either one Ruling and one Non-Ruling Page, or two Unruled Pages.
    1. Pages normally have further properties, for example a given color in the case of parchment.
    2. Folios normally have further properties, for example thickness and stiffness; potentially color in the case of paper.
4. A Folio might stand on its own or might be conjoint with another Folio. A standalone folio is called a *singleton*; the pair of conjoint Folios together is called a *bifolium* (plural *bifolia*).
5. …

exondomesday.ac.uk > 'Modelling Codicology I'

Example Constraints

- Hard Constraints:
    1. A page *must* have exactly two sides
    2. A page *must not* have two hair or two flesh sides
    3. A gathering *must* contain only bifolia and/or singleton(s) (etc.)
- Soft Constraints:
    1. Two folios w/ sequential text are *likely* to be physically sequential
    2. A bifolium ruled before folding is *likely* to have the same ruling on both folia (etc.)
- 'Transformational' Constraints:
    1. Two pages in the same folio *must* remain in that folio (hard)
    2. Two folios in the same bifolium *must* remain in that bifolium (hard)
    3. Reversing order of folios in a bifolium requires reversing order of pages in each folio (hard) (etc.)

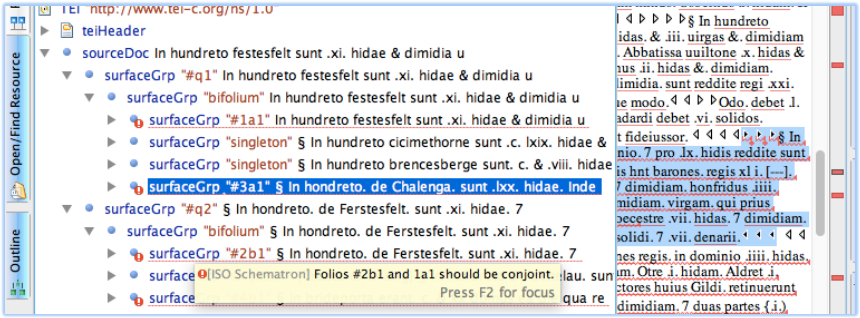
Codicological Constraints

- Enforce the codicological constraints using Schematron

```
<rule context="tei:surfaceGrp[@type='folio']">
    <assert test="count(tei:surface[@type='page']) = 2">Folio <value-of select="@xml:id"/>
        has <value-of select="count(tei:surface[@type='page'])"/> pages.</assert>

    <report test="descendant::tei:surfaceGrp">Folio <value-of select="@xml:id"/> contains
        further surface groups of type <value-of select="descendant::tei:surfaceGrp/@type"/>.</report>

    <!-- Check that the folio does correspond to something valid -->
    <let name="corresp" value="@corresp"/>
    <assert test="not(@corresp) or (//exon:folio[@xml:id= substring-after($corresp, '#')])">Folio
        <value-of select="@xml:id"/> has invalid@corresp pointer (<value-of select="@corresp"/>).</asse

    <let name="corresp" value="@corresp"/>

    <!-- Check that no two surfaceGrps correspond to the same exon:folio -->
    <assert test="count(//tei:surfaceGrp[@type='folio' and @corresp=$corresp])=1">Folios <value-of
            select="//tei:surfaceGrp[@type='folio' and @corresp=$corresp]/@xml:id"/> correspond to
        the same exon:folio <value-of select="$corresp"/>.</assert>

    <!-- Check that conjoint folis are indeed conjoint -->
    <let name="conjoint" value="//exon:folio[@xml:id= substring-after($corresp, '#')]/@conjoins"/>
```


Model in Practice: Verification and Testing

- Use to verify encoding of manuscript state.
  - Already revealed unusual practices in violation of soft constraints.
- Project team can use Oxygen to re-order codicological units, see the new text, and see if constraints are violated.



Next Steps

- Develop UI for more intuitive way of reordering quires
- Use transpose to record different orderings
- Develop algorithms to find orderings for given criteria (minimise changes of scribe; etc.)

```
<profileDesc>
   <listTranspose resp="PAS">
      <transpose>
         <ptr target="#1b1"/>
         <ptr target="#2a1"/>
         <ptr target="#3b1"/>
         <ptr target="#2b1"/>
         <ptr target="#1b1"/>
         <ptr target="#4a1*"/>
         <ptr target="#3a1"/>
         <ptr target="#4b1"/>
         <ptr target="#4a1"/>
      </transpose>
   </listTranspose>
</profileDesc>
```


Conclusions / What's Missing from the TEI?

- surfaceGrp is a bit verbose but does the job
  - Might want to consider recommendations for how to use, e.g. must folios be inside singletons; is it a quire or a gathering, etc.
- A structured collation is essential
  - Proposing an overall format seems valid and useful
  - Some basic attributes could be helpful
  - Can't specify all possible attributes people might want (consider paper, parchment, papyrus, palm leaf, amate, …)
- It is possible to model codicology with TEI
- But, conceptually and pragmatically it's a bit messy...
  - Trying to reconcile different models of text and document
  - Probably useful more as format for interoperability and exchange

Thanks to

The Exon Domesday Team (see project website).

The Arts and Humanities Research Council

(and the European Research Council for previous funding).

Various other people for discussion

(esp. E.Pierazzo, G.Brüning, D. Porter)

www.exondomesday.ac.uk

(still v. basic for now; see also initial thoughts at gist.github.com/pastokes/)

This work is supported by the Arts and Humanities Research Council (UK), grant no. AH/L013975/1.

The research leading to these results has also received funding from the European Union Seventh Framework Programme (FP7) under grant agreement n° 263751 (DigiPal), and the AHRC under grant no. AH/L008041/1 (Models of Authority).