Digital Scholarly Editions Initial Training Network

Grant Agreement No: 317436

SEVENTH FRAMEWORK PROGRAMME

MARIE CURIE ACTIONS

Text: image and transcription

Georg Vogeler

TEI

- Represent images of the text in a facsimile-structure at the same level as teiHeader and text:
- <tei>
  <teiHeader>…</teiHeader>
  <facsimile> …</facsimile>
  <text>…</text>
  </tei>

<facsimile>

- <surface> = a single viewable component
  - Coordinates as a grid, to which the container elements refer: @uly, @ulx; @lrx, @lry =upper left x/y- and lower right y/x-coordinates, i.e. @ulx and @uly are usually 0
  - <graphic>: image, @url : image file name
- <zone> = a single area on the surface
  - Coordinates as above, in the same measurement

Example

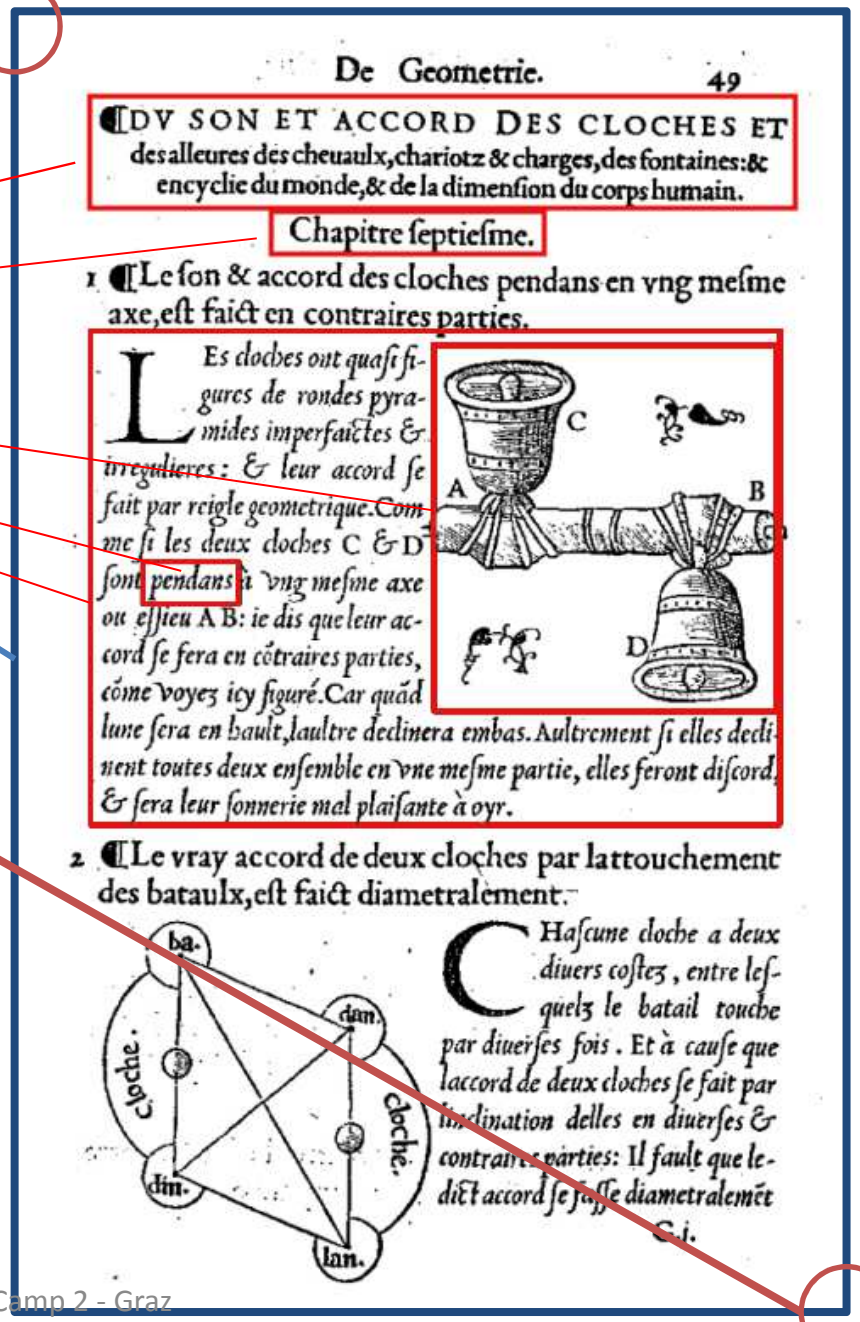
surface

zone

@ulx, @uly

@lrx, @lry

graphic = http://www.tei-c.org/release/doc/tei-p5-doc/en/html/Images/facs-fig1.png



De Geometrie. 49

DV SON ET ACCORD DES CLOCHES ET des alleures des cheuaulx, chariotz & charges, des fontaines: & encyclie du monde, & de la dimension du corps humain.

Chapitre septiesme.

1 Le son & accord des cloches pendans en vng mesme axe, est faict en contraires parties.

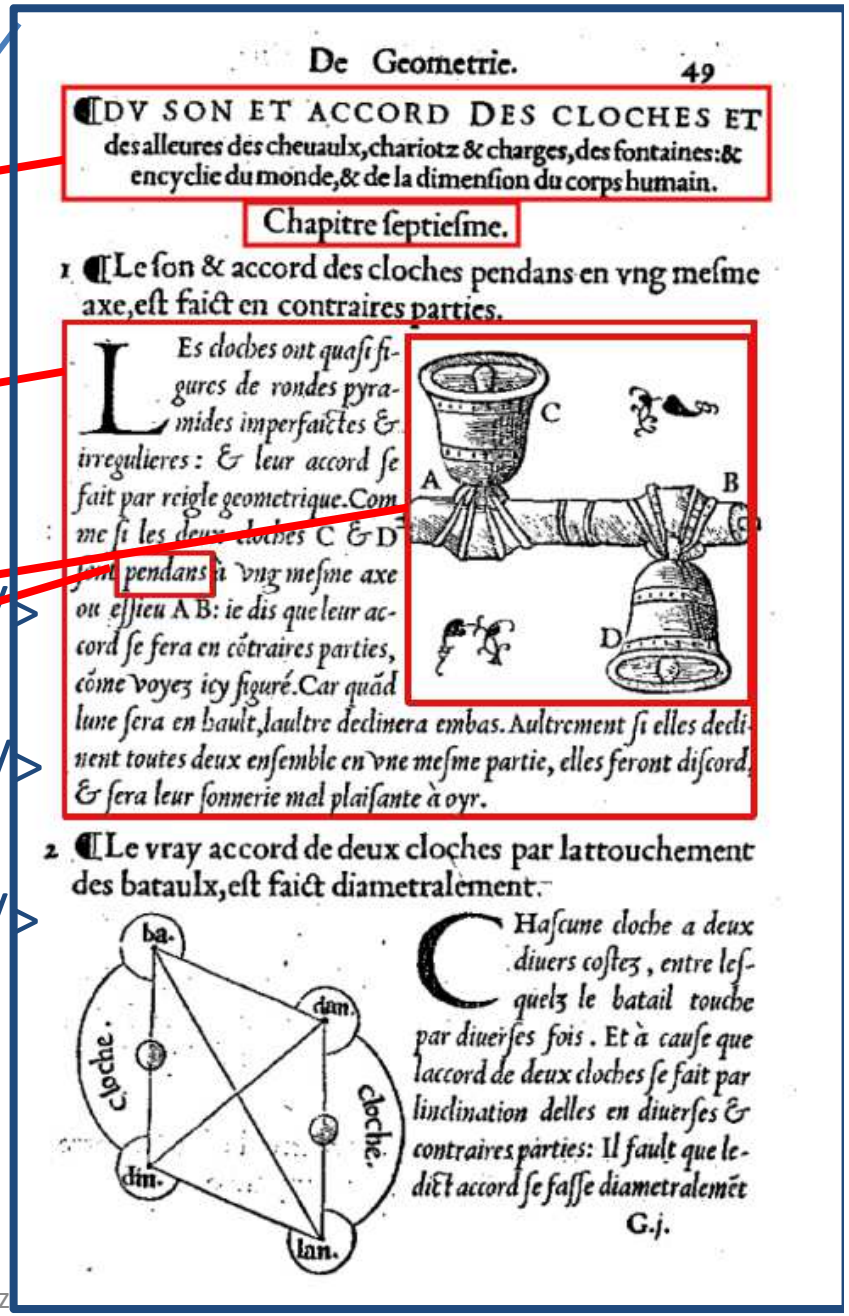
Les cloches ont quasi figures de rondes pyramides imperfaictes & irregulieres: & leur accord se fait par reigle geometrique. Comme si les deux cloches C & D sont pendans à vng mesme axe ou essieu A B: ie dis que leur accord se fera en cõtraires parties, cõme voyez icy figuré. Car quãd lune sera en hault, laultre declinera embas. Aultrement si elles declinent toutes deux ensemble en vne mesme partie, elles feront discord, & sera leur sonnerie mal plaisante à oyr.

2 Le vray accord de deux cloches par lattouchement des bataulx, est faict diametralement.

Chascune cloche a deux diuers costez, entre lesquelz le batail touche par diuerses fois. Et à cause que laccord de deux cloches se fait par inclination delles en diuerses & contraires parties: Il fault que ledict accord se fasse diametralemẽt

C.i.

```
<facsimile>
 <surface
  ulx="0" uly="0" lrx="200" lry="300">
 <graphic url="Bovelles-49r.png"/>
 <zone
  ulx="25" uly="25" lrx="180" lry="60">
 </zone>
 <zone
  ulx="28" uly="75" lrx="175" lry="178" />
 <zone
ulx="105" uly="76" lrx="175" lry="160" />
 <zone
  ulx="45" uly="125" lrx="60" lry="130"/>
 </surface>
</facsimile>
```



De Geometrie. 49

DV SON ET ACCORD DES CLOCHES ET des alleures des cheuaulx, chariotz & charges, des fontaines: & encyclie du monde, & de la dimension du corps humain.

Chapitre septiesme.

1 Le son & accord des cloches pendans en vng mesme axe, est faict en contraires parties.

Les cloches ont quasi figures de rondes pyramides imperfaictes & irregulieres: & leur accord se fait par reigle geometrique. Comme si les deux cloches C & D sont pendans à vng mesme axe ou essieu A B: ie dis que leur accord se fera en cõtraires parties, cõme voyez icy figuré. Car quãd lune sera en hault, laultre declinera embas. Aultrement si elles declinent toutes deux ensemble en vne mesme partie, elles feront discord, & sera leur sonnerie mal plaisante à oyr.

2 Le vray accord de deux cloches par lattouchement des bataulx, est faict diametralement.

Chascune cloche a deux diuers costez, entre lesquelz le batail touche par diuerses fois. Et à cause que laccord de deux cloches se fait par linclination delles en diuerses & contraires parties: Il fault que ledict accord se fasse diametralemẽt

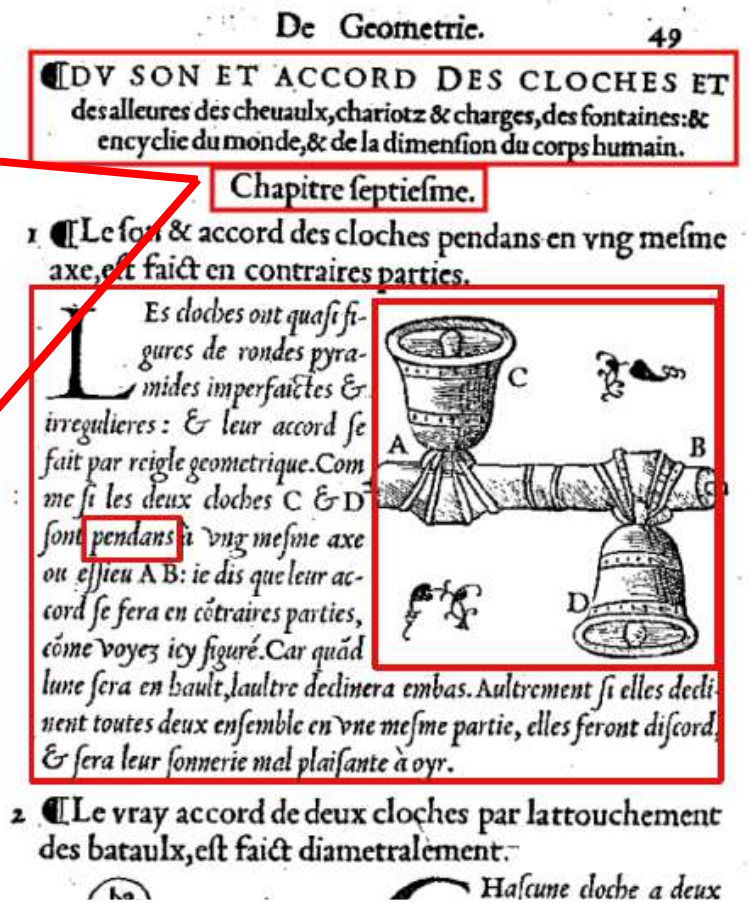
G.j.

Refering from text to image

- facs-attribute corresponding with the xml:id of the facsimile description:

```
<surface xml:id="p49">
  <zone xml:id="p49z2" />
 <graphic url="test.png" />
</surface>

<text><body><div>
 <pb n="49" facs="#p49"/>...
 <head facs="#p49z2">
  Chapitre septiesme </head>
</div></body></text>
```



De Geometrie. 49

DV SON ET ACCORD DES CLOCHES ET des alleures des cheuaulx, chariotz & charges, des fontaines: & encyclie du monde, & de la dimension du corps humain.

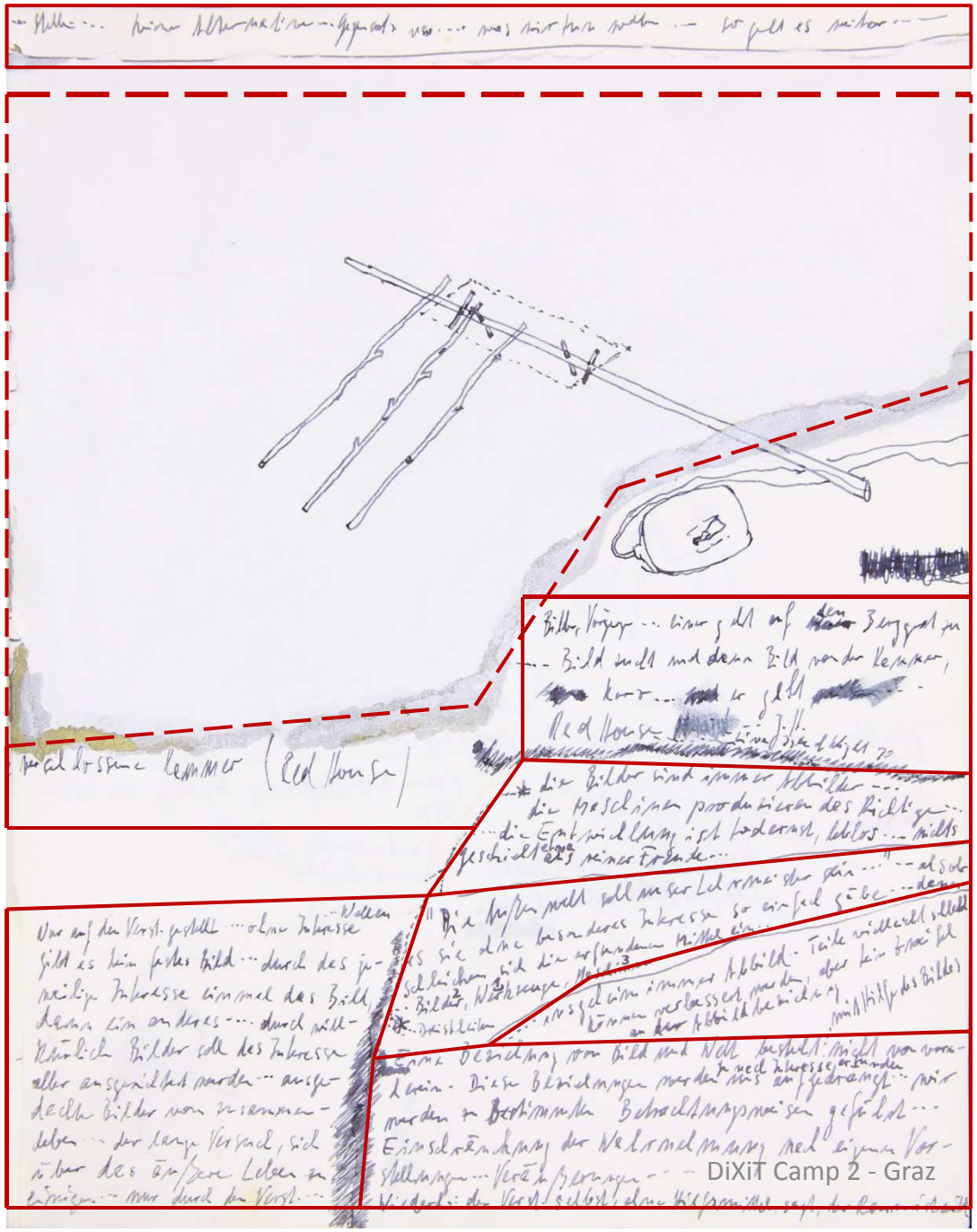
Chapitre septiesme.

1 Le son & accord des cloches pendans en vng mesme axe, est faict en contraires parties.

Les cloches ont quasi figures de rondes pyramides imperfaictes & irregulieres: & leur accord se fait par reigle geometrique. Comme si les deux cloches C & D sont pendans à vng mesme axe ou essieu A B: ie dis que leur accord se fera en cõtraires parties, cõme voyez icy figuré. Car quãd lune sera en hault, laultre declinera embas. Aultrement si elles declinent toutes deux ensemble en vne mesme partie, elles feront discord, & sera leur sonnerie mal plaisante à oyr.

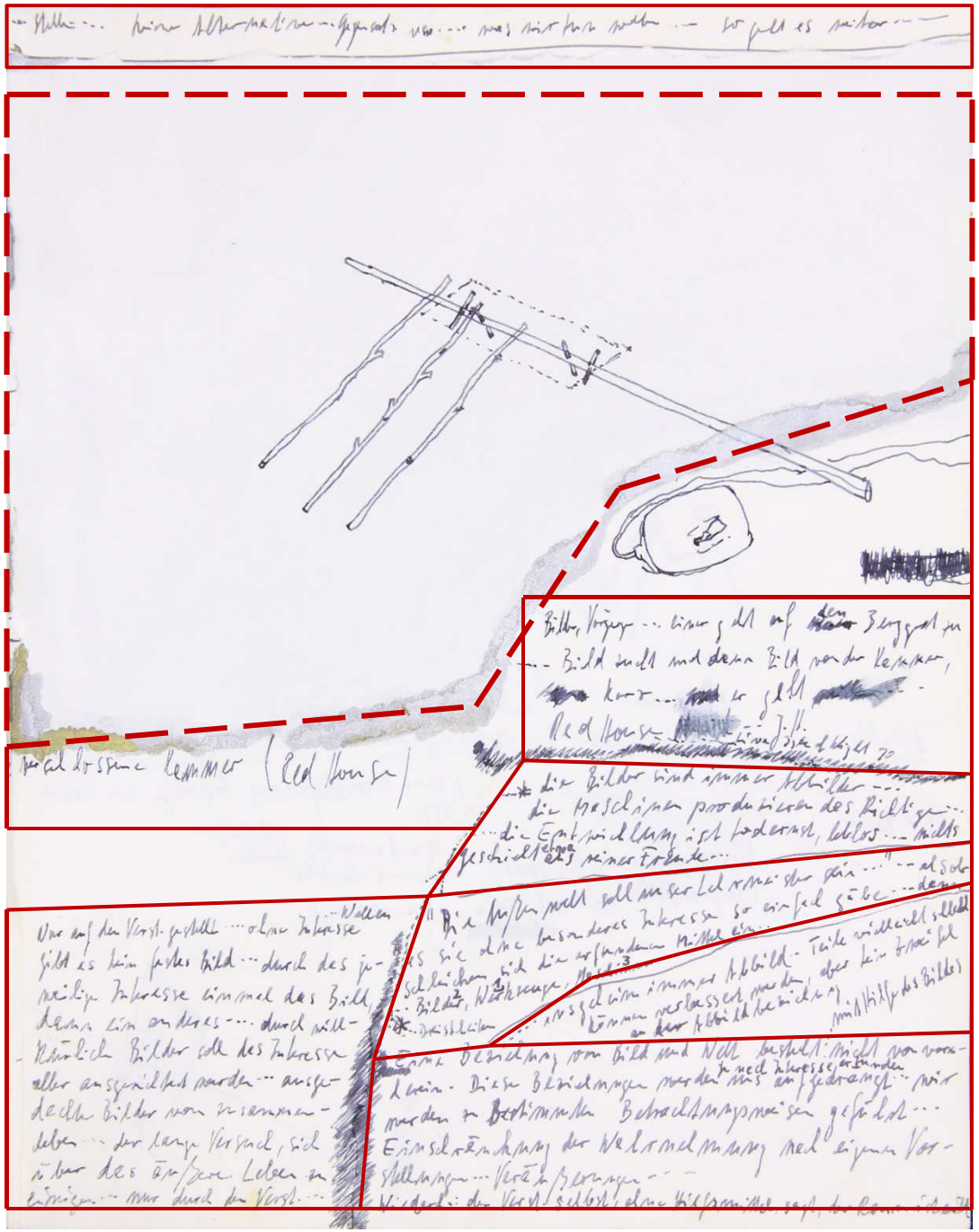
2 Le vray accord de deux cloches par lattouchement des bataulx, est faict diametralement.

Hascune cloche a deux



<zone>

@points:
a list of coordinates (pairs of numbers) which, if connected by lines, circumvent the textarea

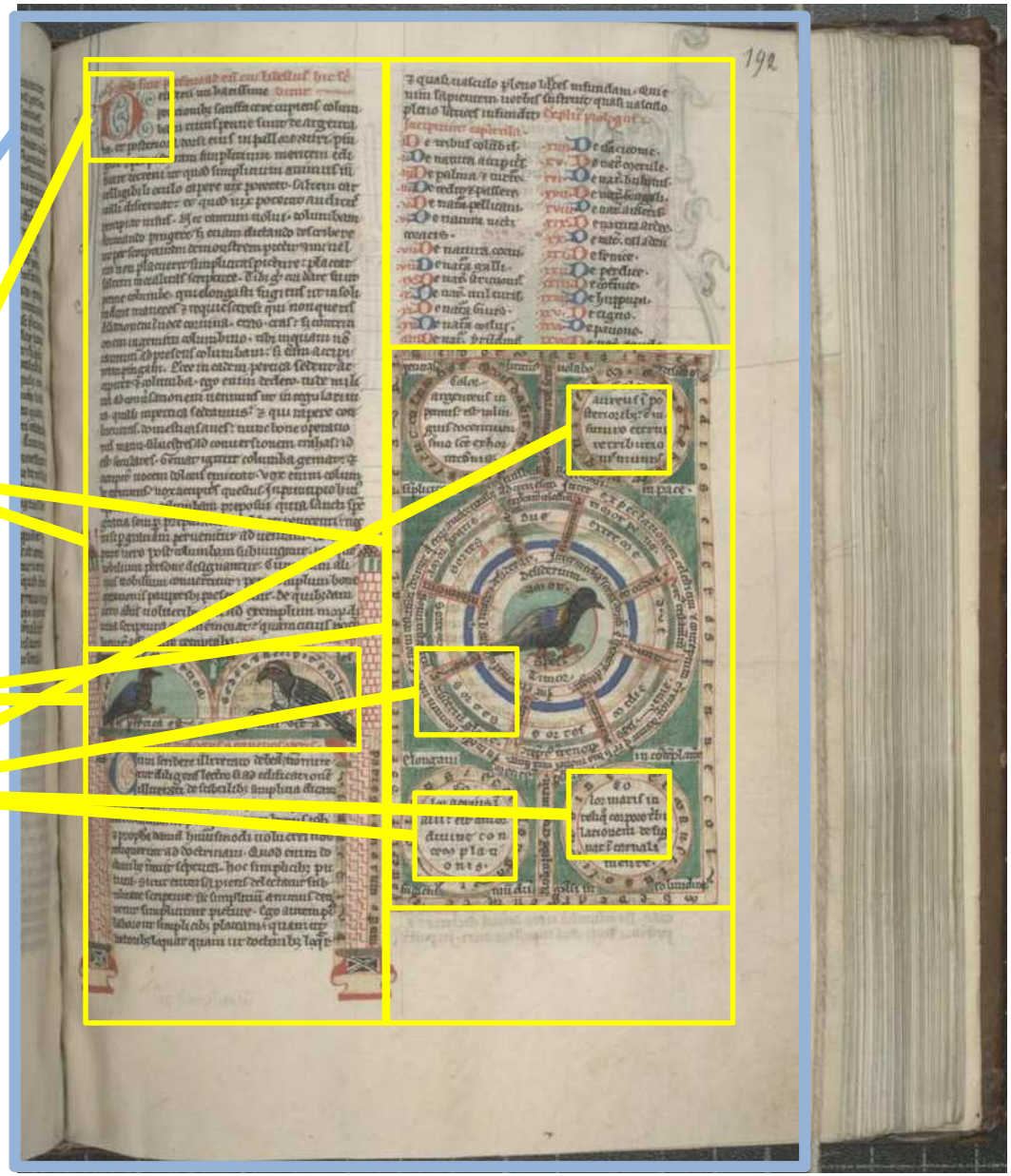


<zone points="0,29 534,20 536,215 334,282 259,376 0,409"/>

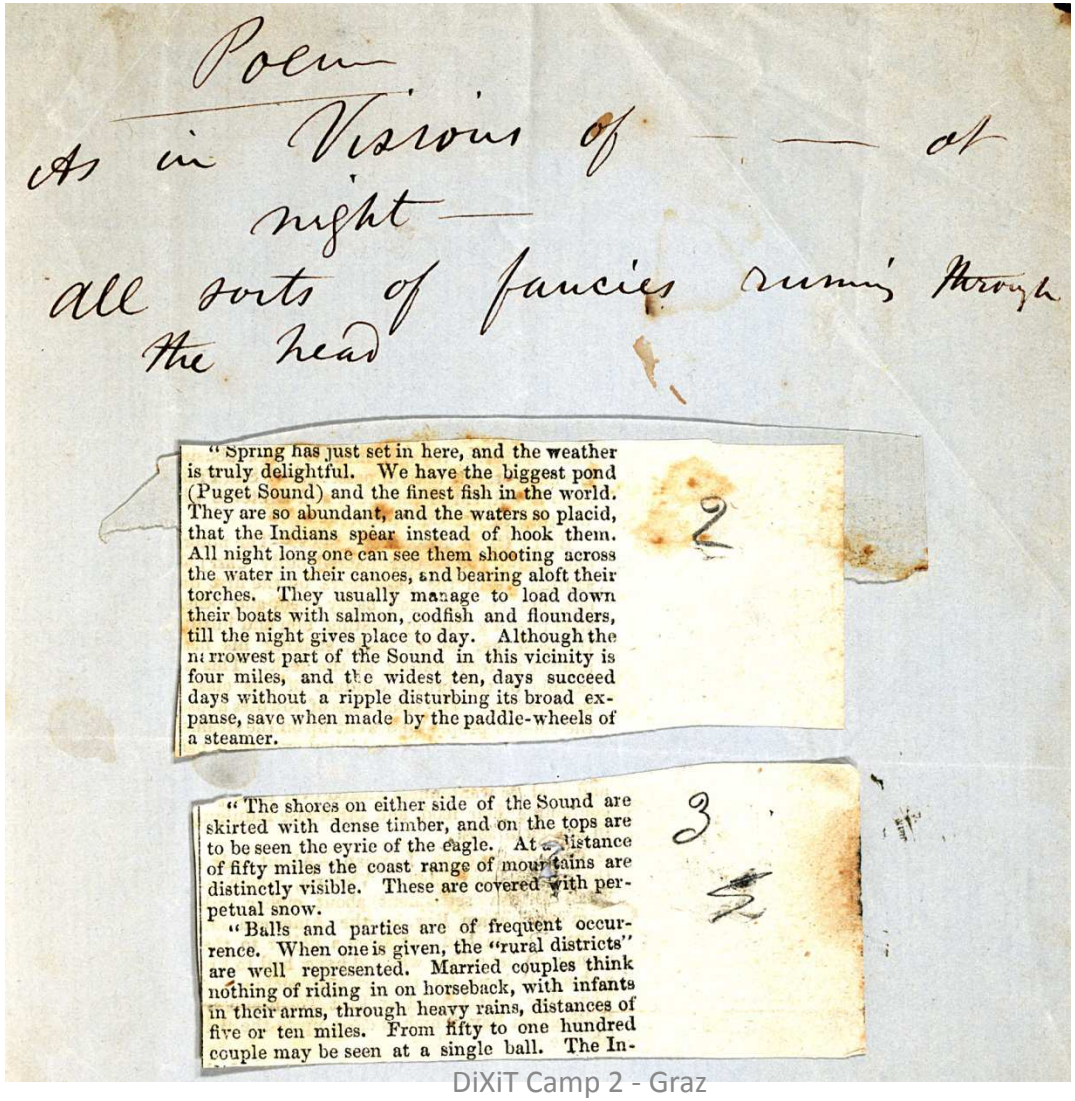
Cambrai, BM Ms. A 259
fol. 192r: Hugo von Folieto, De avibus

surface?
zone?

```
<surface xml:id="p192r" ulx="0"
uly="0" lrx="798" lry="922">
<zone xml:id="p192r-Initial1"
ulx="202" uly="76" lrx="260"
lrx="119"/>
<zone xml:id="p192ra" ulx="202"
uly="61" lry="442" lry="809"/>
<zone xml:id="p192rb" ulx="442"
uly="55" lrx="713" lry="765"/>
<zone xml:id="p192r-fig1" ulx="204"
uly="525" lrx="421" lry="608"/>
<zone />....
</surface>
```



Surfaces? Zones?



Poem
As in Visions of — — at night —
all sorts of fancies running through the head

"Spring has just set in here, and the weather is truly delightful. We have the biggest pond (Puget Sound) and the finest fish in the world. They are so abundant, and the waters so placid, that the Indians spear instead of hook them. All night long one can see them shooting across the water in their canoes, and bearing aloft their torches. They usually manage to load down their boats with salmon, codfish and flounders, till the night gives place to day. Although the narrowest part of the Sound in this vicinity is four miles, and the widest ten, days succeed days without a ripple disturbing its broad expanse, save when made by the paddle-wheels of a steamer.

2

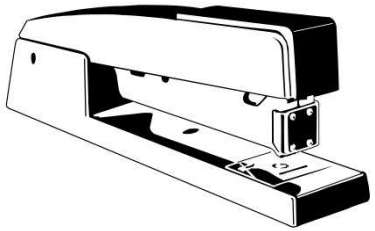
"The shores on either side of the Sound are skirted with dense timber, and on the tops are to be seen the eyrie of the eagle. At a distance of fifty miles the coast range of mountains are distinctly visible. These are covered with perpetual snow.

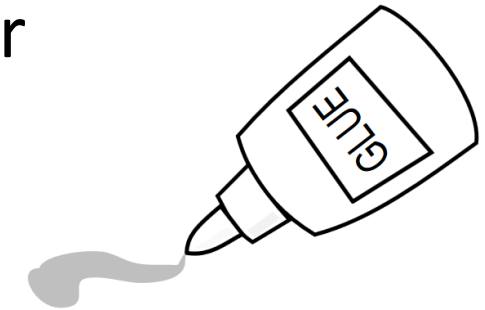
"Balls and parties are of frequent occurrence. When one is given, the "rural districts" are well represented. Married couples think nothing of riding in on horseback, with infants in their arms, through heavy rains, distances of five or ten miles. From fifty to one hundred couple may be seen at a single ball. The In-

3

"Patch" as <surface>

The attribute *@attachment* describes the method by which a surface (for example a newspaper clipping) is or was connected to the main surface. For example glued, pinned, stapled or

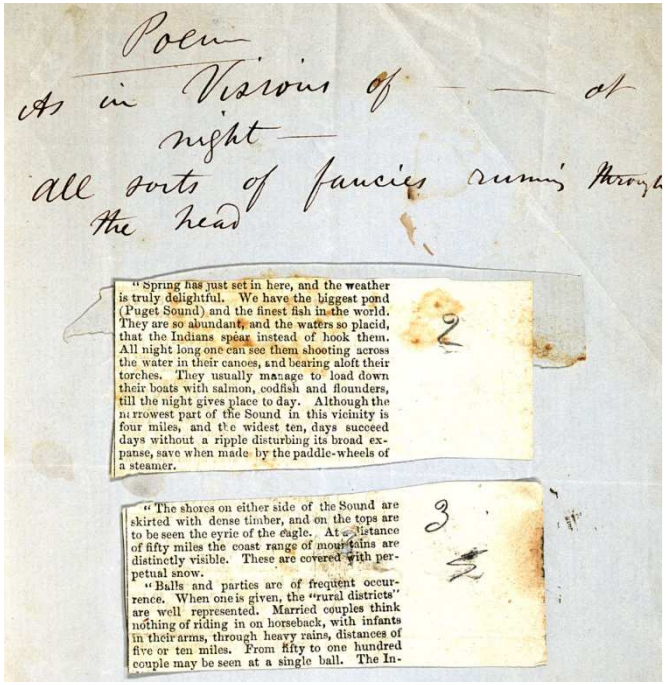




The attribute *@flipping* indicates whether the surface is attached and folded in such a way as to provide two writing surfaces.

"Patch" as <surface>

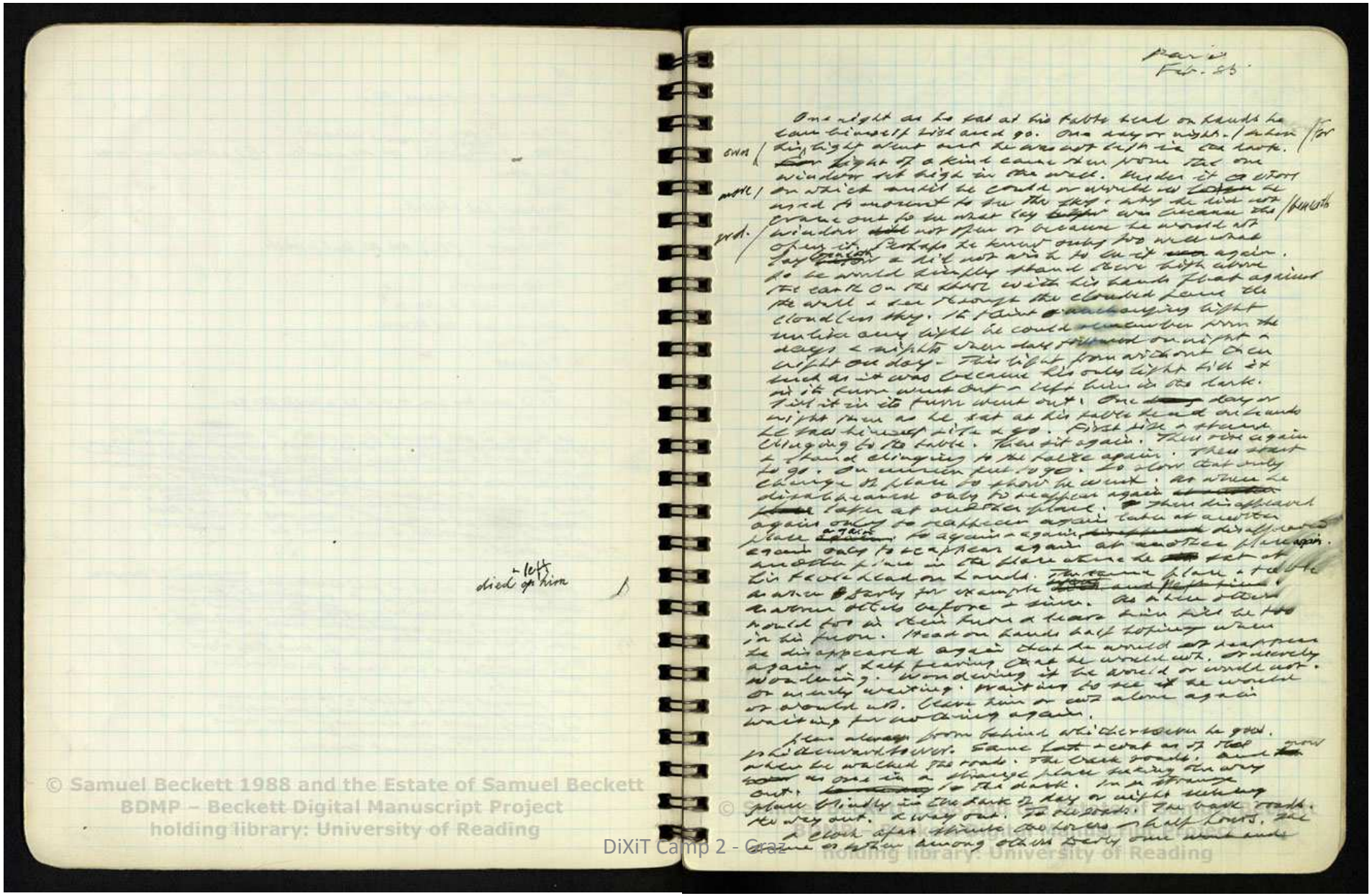
```
<surface>
  <zone>
    <line>Poem</line>
    <line>As in Visions of — at</line>
    <line>night —</line>
    <line>All sorts of fancies running through</line>
    <line>the head</line>
  </zone>
  <zone>
    <surface type="newsprint" attachment="glue" flipping="false">
      <zone>Spring has just set in here, and the weather.... a steamer
</zone>
      <metamark function="sequence">2</metamark>
    </surface>
  </zone>
  <zone>
    <surface type="newsprint" attachment="glue" flipping="false">
      <zone>"The shores on either side of the Sound are... The In-
</zone>
      <metamark function="sequence">3</metamark>
    </surface>
  </zone>
</surface>
```



Exercise I

- What is a surface/zone in these two examples?

Exercise Ia



© Samuel Beckett 1988 and the Estate of Samuel Beckett
BDMP – Beckett Digital Manuscript Project
holding library: University of Reading

Excercise Ib

What is a surface/zone in this example?

Graham Bond killed in fall

GRAHAM BOND, a leading pioneer of the British r-and-b movement, died last week when he fell in front of a tube train at Finsbury Park Station in London. The tragedy occurred last Wednesday, and he was identified from his fingerprints on Friday. He was 37.

First gaining attention as a jazz saxist with the Don Rendell Quintet, Bond fronted various bands during the r-and-b boom of the 60's. Among his sidemen were Ginger Baker, Jack Bruce, John McLaughlin and Jon Hiseman. He formed his first band in 1963 out of Alexis Korner Inc., and later progressed to the renowned Graham Bond Organisation. After its break-up, he concentrated for a time on session work and writing.

He subsequently joined Ginger Baker's Airforce, then formed a new band with his wife Diane Stewart and Pete Brown. Last year, following the end of his marriage, Bond formed a new group with folk singer Carolanne Pegg called Magus. It played a few gigs, but folded a few months ago due to financial problems. He was recently in hospital, but was planning a major comeback at the time of his death.

● A full appreciation of Graham Bond will appear in next week's issue.

Linking the transcription to the text

- From text elements to facsimile or zone: @facs

  <facsimile><surface xml:id="***p192r***"/></facsimile>
  <text> <body><pb facs="***#p192r***"/><head>Epistola sine prefatio ad eum cui libellus hi scribitur</head>
  </body> </text>

- From facsimile or zone to text: @start

  <facsimile><surface start="***#p192r***" start?"/></facsimile>
  <text> <body><pb xml:id="***p192r***"/><head>Epistola sine prefatio ad eum cui libellus hi scribitur</head>
  </body> </text>

Some tools for Image Linking

- Image markup tool (Martin Holmes)
  http://www.tapor.uvic.ca/~mholmes/image_markup/index.php

Non-TEI

- TextGridLab
  http://www.textgridlab.de
- Faust-edition:
  https://github.com/faustedition/ext-imageannotation
- TILE (Text Image Linking Environment)
  http://mith.umd.edu/tile/
- T-PEN
  http://www.t-pen.org
- Image coordinates
  http://imagecoordinates.com

EXCERSISE II

Exercise II

- Add the image called 2935-1-5-1r.jpg to a 2935-1-5.xml
- and create a surface/zone encoding; use zone for at least each paragraph
- http://imagecoordinates.com will help you to get the coordinates
- Create links between the text in 2935-1-5.xml and the surface/zone.

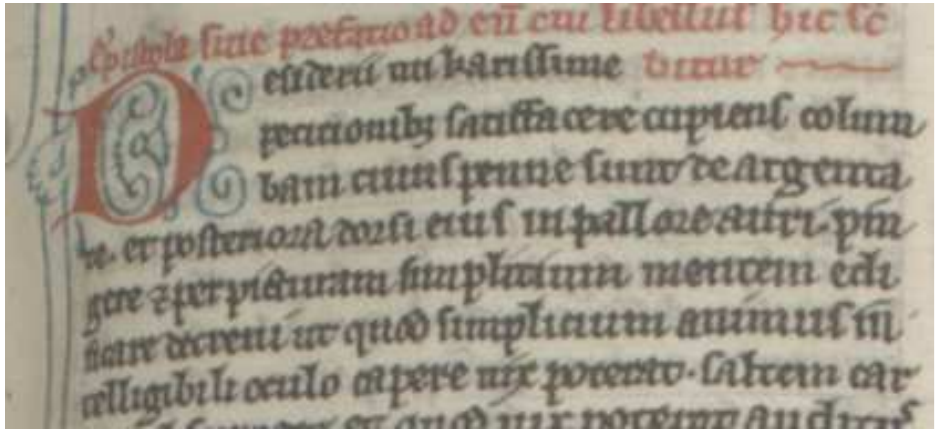
facultative

EXERCISE III

Create a transcription and link it to the image

Epistola sine prefatio ad eu(m) cui libellus hi sc(ri)bitur
Desiderii uti karissime petitionib(us) satisfacere cupiens columbam cuius penne sunt de argentate et posteriora dorsi eius in pallore altri pingere 7 per picturam simplitium mentem edificare decrevi ut quod simplicium animus intelligbili oculo capere vix poterat saltem car[nali …]

using: p, head, ex, choice, abbr, expan



() stand for abbreviations, the 7 is an abbreviation for et

Full image at http://guillelmus.uni-koeln.de/images/Ca/max/Ca_192r.jpg

Digital Scholarly Editions
Initial Training Network

Grant Agreement No:
317436

Embedded Transcription

What do we do, when we transcribe?

- **Look** at an image, identify text areas, find the first line
- grab the keyboard and start to **type**, seeing special characters, highlighting, strike throughs, text above line, etc.,
- and **identify** textual structure, writing activities, named entities, propositions, etc.

What do we do, when we transcribe?

- **Look** at an image, identify text areas, find the first line
- grab the keyboard and start to **type**, seeing special characters, highlighting, strike throughs, text above line, etc.,
- and **identify**
  - textual structure => div, p, w, sp, l, head, list, …
  - writing activities => add, del, abbr, …
  - named entities => rs, name, measure, …
  - propositions => index, salute, …
  - etc.

What do we do, when we transcribe?

- **Look** at an image, => facsimile/sourceDoc, graphic, surface
- identify text areas, => zone
- find the first line, => line
- grab the keyboard and start to **type**, seeing special characters (g), highlighting (hi), strike throughs (seg@rend), text above line (@place), etc.,
- and **identify** writing activities, textual structure, named entities, propositions, etc.

Transcription of primary sources

Parallel transcription

- Transcription in text
- Use primarily structural markup
- Potentially overlapping markup of layout can be encoded with empty elements
- Link to facsimiles and text areas with @facs

Embedded trancription

- Transcription in sourceDoc
- Use layout oriented markup
- Insert accidentally interesting text structure as stand-off markup (milestone/anchor or span)
- Link to text structure and interpretations with @start

Embedded Transcription

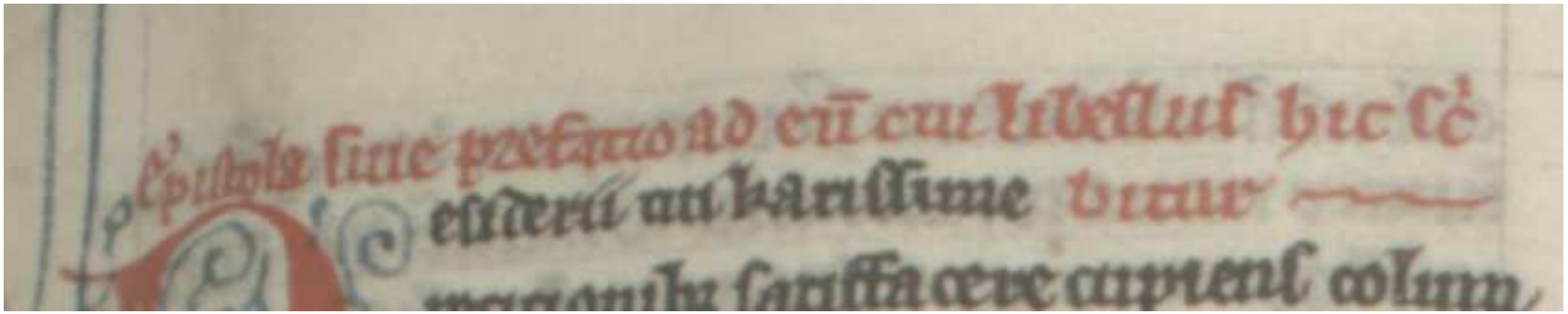
```
<sourceDoc>
  <surface xml:id="p192r" ulx="0" uly="0" lrx="798" lry="922">
    <graphic url="http://guillelmus.uni-koeln.de/images/Ca/small/Ca_192r_small.jpg"/>
    <zone xml:id="p192r-Initial1" ulx="202" uly="76" lrx="260" lry="119"/>
    <zone xml:id="p192ra" ulx="202" uly="61" lrx="442" lry="809">
      <line>Epistola sive prefatio ad eum cui libellus hic scri</line>
      <line><hi facs="#p192r-Initial1">D</hi>esiderii uti karissime bitur</line>
      <line>petitionibus satifsfacere cupiens colum</line>
      <line>bam cuius penne sunt de argenta</line>
      …
    </zone>
  </surface>
</sourceDoc>
```

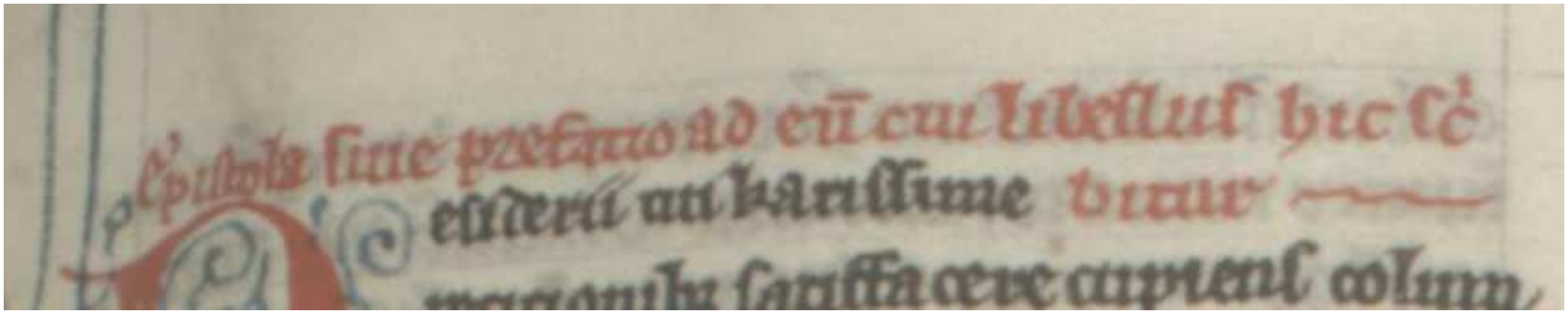
Visual Properties

- @rend: e.g.

  <line rend="*color:red*">Epistola sive prefatio ad eum cui libellus hic scri</line>

  <line><hi facs="*#p192r-Initial1*">D</hi>esiderii uti karissime <seg rend="*color:red*">bitur</seg></line



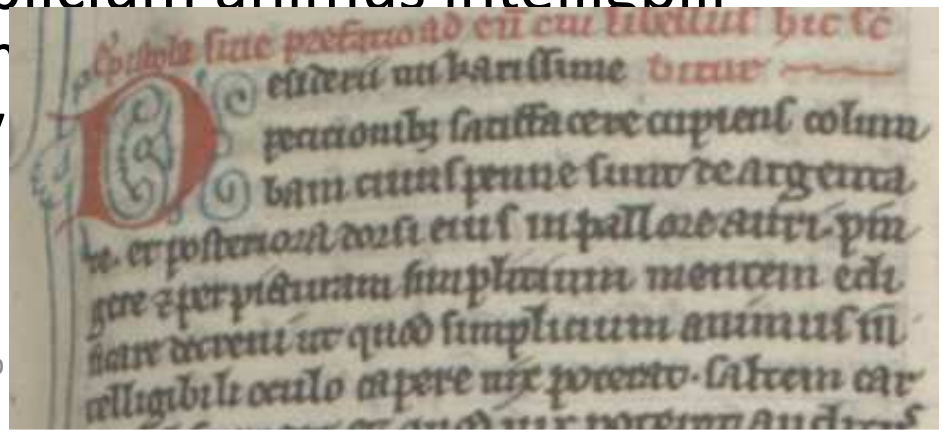


```
<sourceDoc>
  <surface xml:id="p192r" ulx="0" uly="0" lrx="798" lry="922">
      <graphic url="http://guillelmus.uni-
      koeln.de/images/Ca/small/Ca_192r_small.jpg"/>
      <zone xml:id="p192ra" ulx="202" uly="61" lry="442" lry="809">
          <line>Epistola sive prefatio ad eum cui libellus hic scri</line>
          <line>Desiderii uti karissime bitur</line>
          …
      </zone>
  </surface>
</sourceDoc>
```

EXERCISE IV

Convert your transcription of Cambrai MS A 259 to an embedded transcription

```
<text>
    <body>
      <head>Epistola sine prefatio ad eu<ex>m</ex> cui
libellus hi sc<ex>ri</ex>bitur</head>
      <p>Desiderii uti karissime petitionib<ex>us</ex>
satisfacere cupiens columbam cuius penne sunt de
argentate et posteriora dorsi eius in pallore altri pingere
<choice><am><g type="tironian et"/></am>
<ex>et</ex></choice> per picturam simplitium mentem
edificare decrevi ut quod simplicium animus intelligbili
oculo capere vix poterat sltem
vix poterat auditus percipiat v
    </body>
 </text>
```



What about headings, paragraphs, names etc.?

- Boundary Marking with Empty Elements

  <milestone unit ="*tei:head*" spanTo="***#d1f24***"/>De inscriptione<anchor xml:id="***d1f24***"/>

- Transcription-structure linking

  <sourceDoc>...<seg corresp="***#head1***">De inscriptione</seg> ...</sourceDoc>

  <text> <head xml:id="***head1***"/><p xml:id="*h1p1*"/><p xml:id="*h1p2*">

- Text-image-linking

Boundary Marking with Empty Elements

```
<sourceDoc>
  <surface xml:id="p192r" ulx="0" uly="0" lrx="798" lry="922">
      <zone xml:id="p192ra" ulx="202" uly="61" lry="442" lry="809">
          <zone>
              <line><milestone unit="head" spanTo="#d2ad34"/>Epistola sive
              prefatio ad eum cui libellus hic <milestone unit="w"
              spanTo="#d2ad34"/>scri</line>
              <line>bitur<anchor xml:id="d2ad34"/></line>
          </zone>
          <zone>
              <line>Desiderii uti karissime</line>
              <line>peticiontibus satisfacere cupiens colum</line>
          </zone>
  </surface>
</sourceDoc>
```

Transcription structure linking

```
<sourceDoc>
  <surface xml:id="p192r" ulx="0" uly="0" lrx="798" lry="922">
      <zone xml:id="p192ra" ulx="202" uly="61" lry="442" lry="809">
          <zone corresp="#head1">
              <line>Epistola sive prefatio ad eum cui libellus hic scri</line>
              <line>bitur</line>
          </zone>
          <zone start="#h1p1">
              <line>Desiderii uti karissime</line>
              <line>peticiontibus satisfacere cupiens colum</line>
          </zone>
  </surface>
</sourceDoc>
<text>
      <body><div><head xml:id="head1"/><p xml:id="h1p1"/></div></body>
</text>
```

Text-Image-Linking

```
<facsimile>
  <surface xml:id="p192r" ulx="0" uly="0" lrx="798" lry="922">
    <zone xml:id="p192ra" ulx="202" uly="61" lry="442" lry="809">
      <zone xml:id="p192ra-l1" .../>
      <zone xml:id="p192ra-l2" .../>
      <zone xml:id="p192ra-l3" .../>
    </zone>
  </surface>
</facsimile>
<text>
    <body>
      <div>
        <cb facs="#p192ra"/>
        <head><lb facs="#p192ra-l1"/>Epistola sive prefatio ad eum cui libellus hic
        scri<lb break="no" facs="#p192ra-l2"/>bitur</head>
        <p><lb facs="#p192ra-l2"/>Desiderii uti karissime</p>
      </div>
    </body>
</text>
```

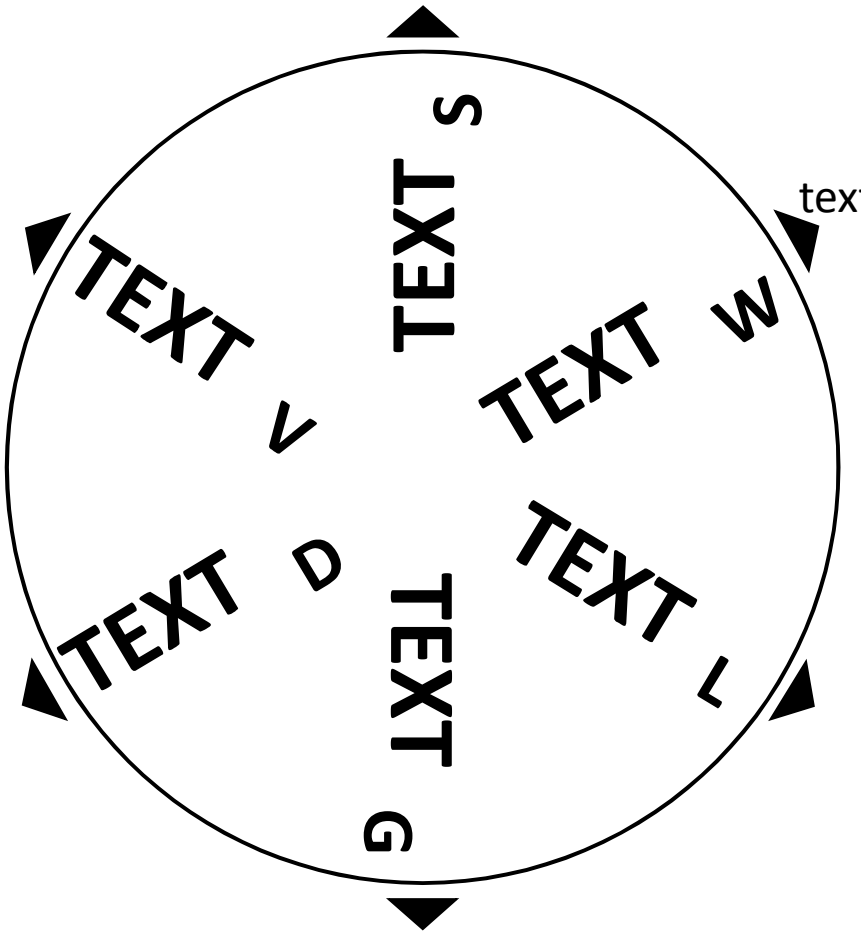
Prefering intellectual or visual structure?

- pb@facs or milestone@unit ?
- pb@facs="#..." or milestone@corresp="#.."/@start="#..."?
- What would you like better?

Pluralistic view on text

text as idea, intention, meaning, **s**emantics, sense, content

text as a **v**isual object, as a complex sign



text as a **w**ork, as rhetoric structure

text as **d**ocument: physical, material, individual

text as **l**inguistic code, as series of words, as speech

text as a version of ..., as a set of **g**raphs, graphemes, glyphs, characters, etc. (... having modes ...)

Digital Scholarly Editions
Initial Training Network

Grant Agreement No:
317436

SEVENTH FRAMEWORK PROGRAMME

MARIE CURIE ACTIONS

Repeat

Excercise V

- Take 2935-3-10-1r.jpg from the Beckett folder
- Create facsimile/surface encoding (http://imagecoordinates.com)
- Add this encoding to 2935-3-10-1r.xml
- Link it to the text
- Convert it to an embedded transcription
- Insert structural markup
  - With milestones
  - With links to empty structural markup