TeiCollator: a TEI to TEI workflow

TEI2022: text as data

Matthias Gille Levenson
École Normale Supérieure de Lyon, France
name [dot] surname1 [-] surname2 [at] ens-lyon [dot] fr

State of the art: automated collation

- Spadini 2016, Nury 2018 and Nury and Spadini 2020: a comprehensive history of automated collation.
- Dekker 2014: CollateX.
- Bleeker et al. 2018: how to collate non textual elements ?
- Camps, Ing, and Spadini 2019 and *CondorCompPhil/Falcon* 2019: first workflow including TEI representation of sources and typology of variants

State of the art: automated collation

- Spadini 2016, Nury 2018 and Nury and Spadini 2020: a comprehensive history of automated collation.
- Dekker 2014: CollateX.
- Bleeker et al. 2018: how to collate non textual elements ?
- Camps, Ing, and Spadini 2019 and *CondorCompPhil/Falcon* 2019: first workflow including TEI representation of sources and typology of variants
- Little has been said about HTR/OCR integration [Kiessling 2019]
- Little has been said about TEI to TEI processing

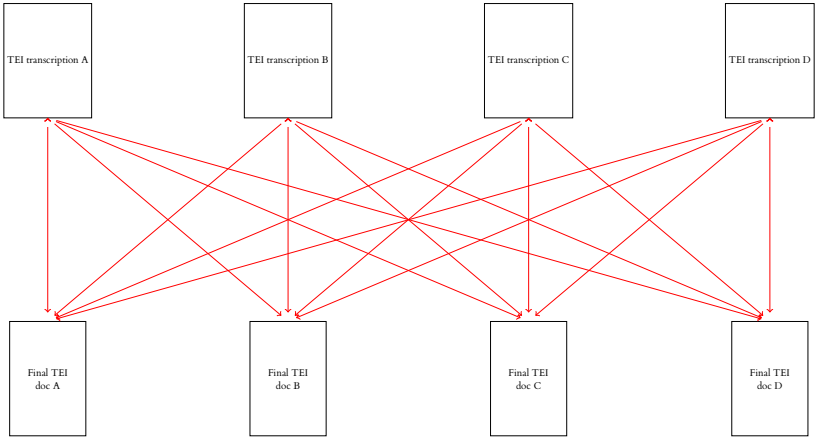
The Gothenburg model

Theoric collation model consisting in several steps[1]:

1. Tokenisation
2. Normalization
3. Alignment
4. Analysis/Feedback
5. Visualisation

[1] Spadini 2016.

Oversimplified view of the workflow

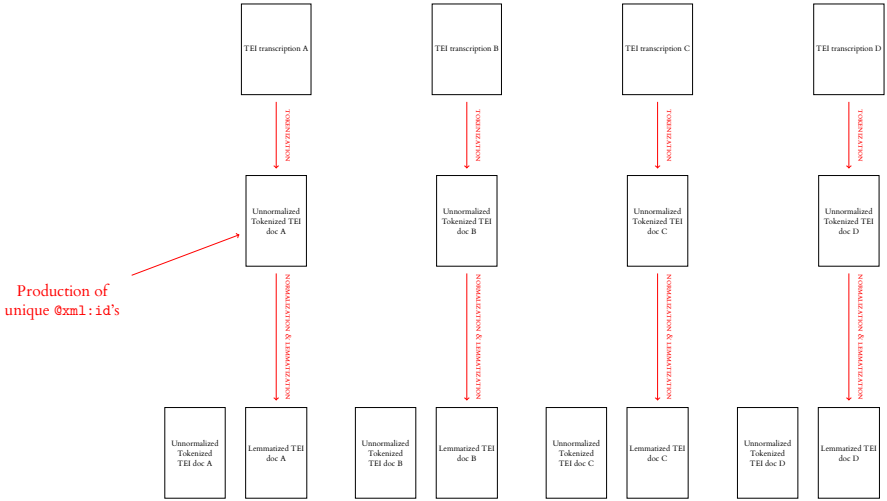


Updating the Gothenburg model model[1,2]

- Individual XML-TEI encoding of sources
- **In-place tokenization and lemmatization [1] [2]**
- **Alignment with CollateX [3]**
- Production of grouped and typed apparatus [4]
- Reinjection of the apparatus in the tokenized TEI documents
- Information transfer
- Detection of simple textual features (omissions, homeoteleuthon transpositions)
- **Transformation into a readable edition [5]**

[1] The Gothenburg model steps are between square brackets

[2] The analysis and feedback steps – check for inconsistency and error – are performed at each step of the workflow

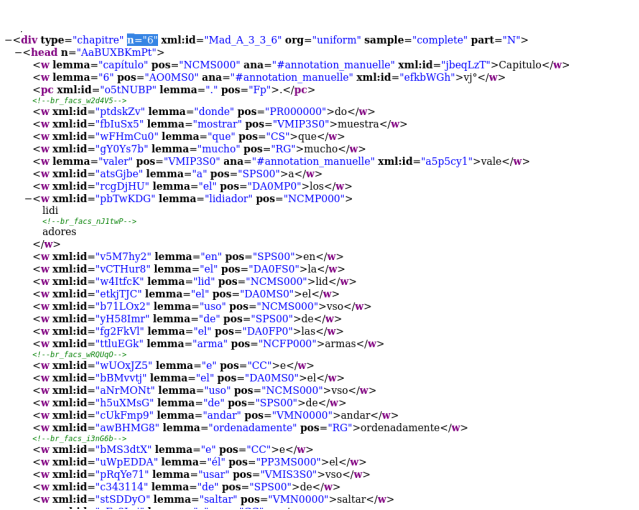

TEI transcription A
TEI transcription B
TEI transcription C
TEI transcription D
TOKENIZATION
Unnormalized Tokenized TEI doc A
Unnormalized Tokenized TEI doc B
Unnormalized Tokenized TEI doc C
Unnormalized Tokenized TEI doc D
Production of unique @xml:id's
NORMALIZATION & LEMMATIZATION
Unnormalized Tokenized TEI doc A
Lemmatized TEI doc A
Unnormalized Tokenized TEI doc B
Lemmatized TEI doc B
Unnormalized Tokenized TEI doc C
Lemmatized TEI doc C
Unnormalized Tokenized TEI doc D
Lemmatized TEI doc D

```
<div type="chapitre" n="6" xml:id="Mad_A_3_3_6">
  <head n="AaBUXBKmPt"><w lemma="capítulo" pos="NCMS000">Capitulo</w>
    <w lemma="6" pos="AO0MS0">vj°</w>.<lb break="yes" facs="#facs_w2d4V5" xml:id="elem_w2d4V5"
      copyOf="file:/home/mgl/Bureau/These/Edition/hyperregimiento-de-los-principes/Dedans/XML/analyse_linguistique/Mad_A.xml#facs_w2d4V5"
    />do muestra que mucho <w lemma="valer" pos="VMIP3S0">vale</w> a los lidi<lb break="no"
      facs="#facs_nJ1twP" xml:id="elem_nJ1twP"
      copyOf="file:/home/mgl/Bureau/These/Edition/hyperregimiento-de-los-principes/Dedans/XML/analyse_linguistique/Mad_A.xml#facs_nJ1twP"
    />adores en la lid el vso de las armas<lb break="yes" facs="#facs_wRQUqO"
      xml:id="elem_wRQUqO"
      copyOf="file:/home/mgl/Bureau/These/Edition/hyperregimiento-de-los-principes/Dedans/XML/analyse_linguistique/Mad_A.xml#facs_wRQUqO"
    />e el vso de andar ordenadamente<lb break="yes" facs="#facs_i3nG6b" xml:id="elem_i3nG6b"
      copyOf="file:/home/mgl/Bureau/These/Edition/hyperregimiento-de-los-principes/Dedans/XML/analyse_linguistique/Mad_A.xml#facs_i3nG6b"
    />e el vso de saltar e de correr.</head>
  <div type="traduction">
    <p n="gnPRxGrfip"><lb break="yes" facs="#facs_fPoJly" xml:id="elem_fPoJly"
        copyOf="file:/home/mgl/Bureau/These/Edition/hyperregimiento-de-los-principes/Dedans/XML/analyse_linguistique/Mad_A.xml#facs_fPoJly"
        /><w><hi type="initiale" facs="#facs_rjS6KV" xml:id="elem_rjS6KV">T</hi>Odos</w> los
      lidiadores <lb break="yes" facs="#facs_iWqcyE" xml:id="elem_iWqcyE"
        copyOf="file:/home/mgl/Bureau/These/Edition/hyperregimiento-de-los-principes/Dedans/XML/analyse_linguistique/Mad_A.xml#facs_iWqcyE"
      />deuen ser vsados e a<lb break="no" facs="#facs_qyKM5d" xml:id="elem_qyKM5d"
        copyOf="file:/home/mgl/Bureau/These/Edition/hyperregimiento-de-los-principes/Dedans/XML/analyse_linguistique/Mad_A.xml#facs_qyKM5d"
      />costunbrados a estas<lb break="yes" facs="#facs_aGugFB" xml:id="elem_aGugFB"
        copyOf="file:/home/mgl/Bureau/These/Edition/hyperregimiento-de-los-principes/Dedans/XML/analyse_linguistique/Mad_A.xml#facs_aGugFB"
      />tres cosas que aqui pone<lb break="yes" facs="#facs_b3jRku" xml:id="elem_b3jRku"
        copyOf="file:/home/mgl/Bureau/These/Edition/hyperregimiento-de-los-principes/Dedans/XML/analyse_linguistique/Mad_A.xml#facs_b3jRku"
      />.La primera es que anden muy orde<lb break="no" facs="#facs_dheQM3"
```

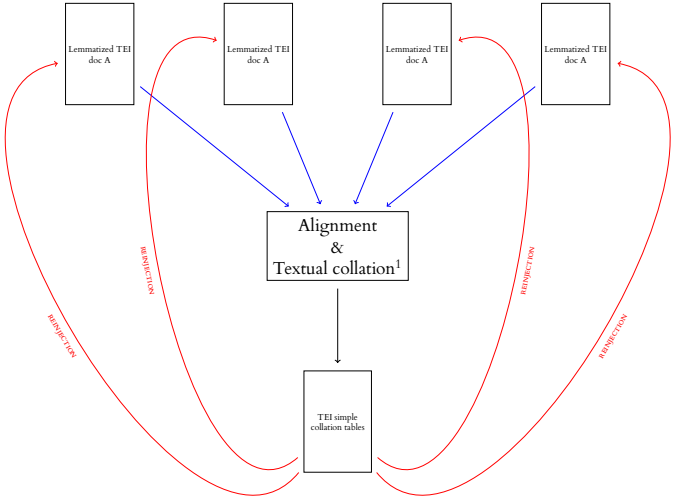
TEI transcription (HTR output)

```
<div type="chapitre" n="6" xml:id="Mad_A_3_3_6" org="uniform" sample="complete" part="N">
  <head n="AaBUXBKnPt">
    <w lemma="capitulo" pos="NCMS000" ana="#annotation_manuelle" xml:id="jbeqLzT">Capitulo</w>
    <w lemma="6" pos="A0OMS0" ana="#annotation_manuelle" xml:id="efkbWGh">vj*</w>
    <pc xml:id="o5tNUBP">.</pc>
    <lb break="yes" facs="#facs_w2d4V5" xml:id="elem_w2d4V5" copyOf="file:/home/mgl/Bureau/These/Edition/hyperregimiento-de-los-principes/Dedans/XML/analyse_linguistique/Mad_A.xml#facs_w2d4V5"/>
    <w xml:id="ptdskZv">do</w>
    <w xml:id="fbIuSx5">muestra</w>
    <w xml:id="wFHmCu0">que</w>
    <w xml:id="qY0Ys7b">mucho</w>
    <w lemma="valer" pos="VMIP3S0" ana="#annotation_manuelle" xml:id="a5p5cy1">vale</w>
    <w xml:id="atsGjbe">a</w>
    <w xml:id="rcgDjHU">los</w>
    <w xml:id="pbTwKDG">
      lidi
      <lb break="no" facs="#facs_nJ1twP" xml:id="elem_nJ1twP" copyOf="file:/home/mgl/Bureau/These/Edition/hyperregimiento-de-los-principes/Dedans/XML/analyse_linguistique/Mad_A.xml#facs_nJ1twP"/>
      adores
    </w>
    <w xml:id="v5M7hy2">en</w>
    <w xml:id="vCTHur8">la</w>
    <w xml:id="w4ItfcK">lid</w>
    <w xml:id="etkjTJC">el</w>
    <w xml:id="b7ILOx2">vso</w>
    <w xml:id="yH58Imr">de</w>
    <w xml:id="fg2FkVl">las</w>
    <w xml:id="tUuEGk">armas</w>
    <lb break="yes" facs="#facs_wRQUqO" xml:id="elem_wRQUqO" copyOf="file:/home/mgl/Bureau/These/Edition/hyperregimiento-de-los-principes/Dedans/XML/analyse_linguistique/Mad_A.xml#facs_wRQUqO"/>
    <w xml:id="wUOxJZ5">e</w>
    <w xml:id="bBMvvtj">el</w>
    <w xml:id="aNrMONt">vso</w>
    <w xml:id="h5uXMsG">de</w>
    <w xml:id="cUkFmp9">andar</w>
    <w xml:id="awBHMG8">ordenadamente</w>
    <lb break="yes" facs="#facs_i3nG6b" xml:id="elem_i3nG6b" copyOf="file:/home/mgl/Bureau/These/Edition/hyperregimiento-de-los-principes/Dedans/XML/analyse_linguistique/Mad_A.xml#facs_i3nG6b"/>
    <w xml:id="bMS3dtX">e</w>
    <w xml:id="uWpEDDA">el</w>
    <w xml:id="pRqYe71">vso</w>
    <w xml:id="c343114">de</w>
    <w xml:id="stSDDyO">saltar</w>
    <w xml:id="xFe8Lej">e</w>
    <w xml:id="vDiMKwt">de</w>
```

Tokenized TEI transcription



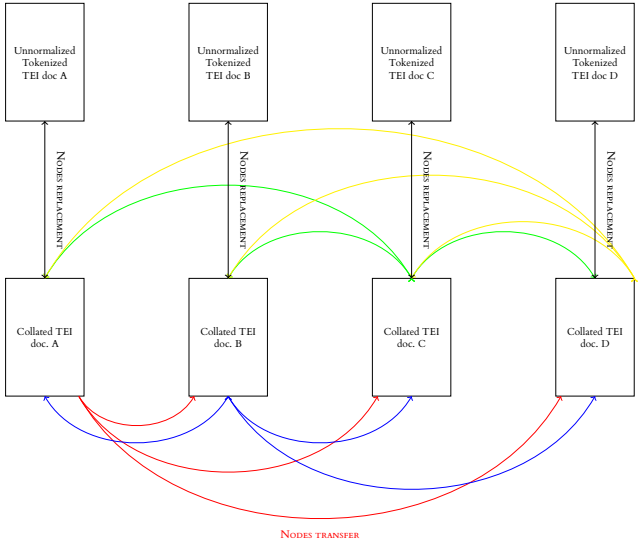
Normalized and lemmatized TEI transcription



[1]production of typed [Camps, Ing, and Spadini 2019] and grouped apparatus

```
<div type="chapitre" n="6" xml:id="Mad_B_3_3_6">
 <head n="AaBUXBKmPt">
  <app ana="#normalisation">
   <rdgGrp>
    <rdg id="ahC8hRq" lemma="capítulo_capítulo_capítulo_capítulo_capítulo" pos="NCMS000 NCMS000 NCMS000 NCMS000 NCMS000" wit="#Esc_Q #Mad_G #Sev_R #Mad_B #Sal_J"
     n="jWzm2cW_wCal5qE_jTzRf8r_yDg5fs1_zly5iLw">
     <w xml:id="jWzm2cW_wCal5qE_jTzRf8r_yDg5fs1_zly5iLw">Capítulo</w>
    </rdg>
    <rdg id="kVeyRAo" lemma="capítulo" pos="NCMS000" wit="#Mad_A" n="yuiMoC1">
     <w xml:id="yuiMoC1">Capítulo</w>
    </rdg>
    <rdg id="uQgju1k" lemma="capítulo" pos="NCMS000" wit="#Sev_Z" n="jhpk4UZ">
     <w xml:id="jhpk4UZ">capítulo</w>
    </rdg>
   </rdgGrp>
  </app>
  <app ana="#lexicale">
   <rdgGrp>
    <rdg id="kG55vaQ" lemma="6" pos="AO0MS0" wit="#Mad_A" n="yCMSqux">
     <w xml:id="yCMSqux">vj°</w>
    </rdg>
    <rdg id="zexCTPT" lemma="6_6" pos="AO0MS0 AO0MS0" wit="#Mad_G #Sal_J" n="v5pSxd2_hDGnhNg">
     <w xml:id="v5pSxd2_hDGnhNg">vi°</w>
    </rdg>
   </rdgGrp>
   <rdgGrp>
    <rdg id="bdqpuqo" lemma="ver_6_6" pos="VMIS1S0 Z Z" wit="#Esc_Q #Sev_R #Mad_B" n="aYIRC45_cMVjbre_o5E8N86">
     <w xml:id="aYIRC45_cMVjbre_o5E8N86">vi</w>
    </rdg>
   </rdgGrp>
   <rdgGrp>
    <rdg id="hNax50n" lemma="sexto" pos="AO0MS0" wit="#Sev_Z" n="qeQOLoi">
     <w xml:id="qeQOLoi">sexto</w>
    </rdg>
   </rdgGrp>
  </app>
  <pc xml:id="uMsCgvq">.</pc>
```

Result of the textual collation

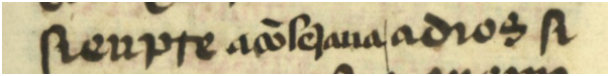

Unnormalized Tokenized TEI doc A
Unnormalized Tokenized TEI doc B
Unnormalized Tokenized TEI doc C
Unnormalized Tokenized TEI doc D
Nodes replacement
Nodes replacement
Nodes replacement
Nodes replacement
Collated TEI doc. A
Collated TEI doc. B
Collated TEI doc. C
Collated TEI doc. D
Nodes transfer

```
<anchor type="citation" xml:id="ELlsvMUhrH"/>La <w lemma="2" pos="A0OFS0">segunda</w> es
ser sacramentado, ca <w lemma="deber" pos="VMIP3S0">deve</w> fazer sacramento e jura <w
lemma="el" pos="DA0MS0">el</w> día que <w lemma="lo">lo</w> fazen <w lemma="caballero"
pos="NCMS000">cavallero</w><note xml:id="OzfmfIvIpC" type="sources"><quote
type="primaire" xml:lang="lat">Miles enim dicitur quare unus ex mille electus, uel
miles mille valens, ait Papias. Romulus <unclear>gratia</unclear> prius e populo
sumpsit e appellauit. Et sic miles est nomen laboris <unclear>ita</unclear> honoris,
ait Policraticus l. vi, c. viii. Secundo, videndus est de hiis qui militem faciut, de
quibus Policraticus, l. vi, c. v et vii, bene ubi ait quod duo sunt principua que
militem fatiunt: electio e sacramentum. <ref type="biblio"
target="#juandegales_SummaCollectionumSeu_1401">I, 9, 2, <title type="section">De
institutione militie, e militis nuncupatione sine effectione</title>, fol.
50v</ref></quote>Le glosateur vient ici gloser et expliciter les deux concepts de
sélection et de sacrement: il ne s'agit pas simplement d'une traduction du texte de
Jean de Galles.</note><!--<note xml:id="pPvhQMlmUP"
```

```
−<app ana="#graphique">
 −<rdgGrp>
  −<rdg wit="#Esc_Q #Mad_G #Sev_R #Mad_B #Sal_J">
     <w>cavallero</w>
   </rdg>
  −<rdg wit="#Mad_A #Sev_Z">
     <w>cauallero</w>
   </rdg>
 </rdgGrp>
</app>
−<note xml:id="OzfmfIvIpC" type="sources" ana="#injected" corresp="#Sal_J">
 −<quote type="primaire" xml:lang="lat">
    Miles enim dicitur quare unus ex mille electus, uel miles mille valens, ait Papias. Romulus
    <unclear>gratia</unclear>
    prius e populo sumpsit e appellauit. Et sic miles est nomen laboris
    <unclear>ita</unclear>
    honoris, ait Policraticus l. vi, c. viii. Secundo, videndus est de hiis qui militem faciut, de quibus Policraticus, l. vi, c. v et vii, bene ubi ait quod duo sunt principua que militem fatiunt: electio e
    sacramentum.
 −<ref type="biblio" target="#juandegales_SummaCollectionumSeu_1401">
    I, 9, 2,
   −<title type="section">
      De institutione militie, e militis nuncupatione sine effectione
    </title>
    , fol. 50v
   </ref>
 </quote>
 Le glosateur vient ici gloser et expliciter les deux concepts de sélection et de sacrement: il ne s'agit pas simplement d'une traduction du texte de Jean de Galles.
</note>
<pc xml:id="qexUnSY">.</pc>
```

Nodes transfer between witnesses (J > collated B)

Another example: non material collation



Ms. II/215, Real Biblioteca, Madrid, fol. 418r

Ms 15304, Fundación Lázaro Galdiano, Madrid, fol. 245v

```
    >fazienda</w> con temor e con tremor, a enxenplo de <w lemma="david"
    pos="NP00OP0">David</w>, que sienpre <w><subst>
        <space/>
        <add hand="identique" place="inline" type="correction">aconsejava</add>
    </subst></w><note type="particulier" xml:id="HgAJWEnofe"
    facs="/home/mgl/Bureau/These/Edition/hyperregimiento-de-los-principes/Dedans/Facsimiles/Mad_G/aconsejava_add.png"
    >Ici, le terme <quote>aconsejava</quote> semble avoir été ajouté après, par
    une main similaire: il est écrit dans la ligne, mais avec une taille de
    caractères plus petite, comblant un emplacement qui semble avoir été laissé
    vide pour une raison inconnue.</note> a Dios <w lemma="si" pos="CS">si</w>
<w lemma="acometer" pos="VMIC3S0">acometería</w> a sus enemigos, segund que de
```

TEI encoding

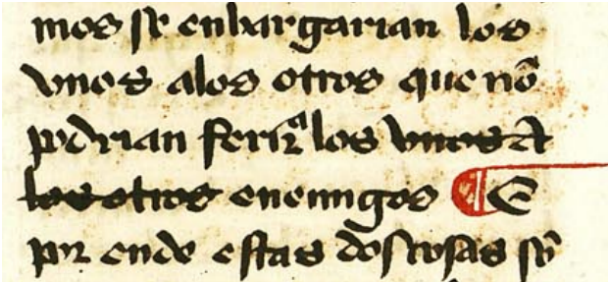
Result

```
−<app ana="#graphique">
 −<rdgGrp>
  −<rdg wit="#Esc_Q #Mad_A #Mad_G #Sev_R #Sal_J">
    <w>sienpre</w>
    <space xml:id="OarJbBUnCZ" type="in_witness" ana="#injected" corresp="#Mad_A"/>
   +<note xml:id="NZrEOmzJZs" type="variante" anchored="true" ana="#injected" corresp="#Mad_A"></note>
  </rdg>
 +<rdg wit="#Mad_B"></rdg>
 +<rdg wit="#Sev_Z"></rdg>
 </rdgGrp>
</app>
−<app ana="#omission #lexicale">
 −<rdgGrp>
  −<rdg wit="#Mad_G">
   −<w ana="#annotation_manuelle">
    −<subst>
      <space xml:id="gW4OVGi"/>
      <add hand="identique" place="inline" type="correction">aconsejava</add>
     </subst>
    </w>
   +<note type="codico" xml:id="HgAJWEnofe" facs="/home/mgl/Bureau/These/Edition/hyperregimiento-de-los-principes/Dedans/Facsimiles/Mad_G/aconsejava_add.png" ana="#injected" corresp="#Mad_G">
   </rdg>
 </rdgGrp>
 −<rdgGrp>
  −<rdg wit="#Esc_Q #Sev_R #Mad_B">
    <w>demandava</w>
   </rdg>
 +<rdg wit="#Sal_J #Sev_Z"></rdg>
 </rdgGrp>
 −<rdgGrp>
    <rdg wit="#Mad_A"/>
 </rdgGrp>
</app>
```

Adaptability and modularity

```
"reinjection": {
    "tei:del": {"position": "after", "level": "witness"},
    "tei:add[@type='commentaire']": {"position": "after", "level": "witness"},
    "tei:note[@type='general']": {"position": "after", "level": "work"},
    "tei:note[@type='codico']": {"position": "after", "level": "witness"},
    "tei:note[@type='variante']": {"position": "after", "level": "witness"},
    "tei:note[@type='sources']": {"position": "after", "level": "work"},
    "tei:pb[@break='yes']": {"position": "after", "level": "witness"},
    "tei:cb[@break='yes']": {"position": "after", "level": "witness"},
    "tei:milestone[@unit][ancestor::tei:div[contains(@xml:id, 'Sev_Z')]]": {"position": "before", "level": "work"},
    "tei:anchor[@type='citation']": {"position": "before", "level": "work"},
    "tei:anchor[@type='ligne']":{"position":"before", "level":"work"},
    "tei:pc":{"position":"after", "level":"witness"},
    "tei:figure":{"position":"after", "level":"work"},
    "tei:space[@type='in_witness']":{"position":"after", "level":"witness"}
},
"exclude_descendant_of": [],
```

Witness level transfer



```
<w lemma="se" pos="PP3CN000">se</w>
<w lemma="embargar">enbargarían</w> los unos a los otros que non podrían <w lemma="ferir"
  >ferir</w>
<add place="above"><w>a</w></add> los <del>unos a los otros</del> enemigos. E por
ende, estas dos cosas son menester en la <w lemma="haz" pos="NCFS000">haz</w>: <w
```

Witness level transfer

```
<app ana="#not_apparat">
  <rdgGrp>
    <rdg wit="#Esc_Q #Mad_A #Mad_G #Sev_R #Mad_B #Sal_J #Sev_Z">
      <w>ferir</w>
    </rdg>
  </rdgGrp>
</app>
<app ana="#not_apparat">
  <rdgGrp>
    <rdg wit="#Esc_Q #Mad_A #Mad_G #Sev_R #Mad_B #Sal_J #Sev_Z">
      <w ana="#same_word">a</w>
    </rdg>
  </rdgGrp>
</app>
<app ana="#not_apparat">
  <rdgGrp>
    <rdg wit="#Esc_Q #Mad_A #Mad_G #Sev_R #Mad_B #Sal_J #Sev_Z">
      <w ana="#same_word">los</w>
      <del xml:id="k0sRK75" ana="#injected" corresp="#Sal_J">unos a los otros</del>
    </rdg>
  </rdgGrp>
</app>
<app ana="#not_apparat">
  <rdgGrp>
    <rdg wit="#Esc_Q #Mad_A #Mad_G #Sev_R #Mad_B #Sal_J #Sev_Z">
      <w>enemigos</w>
    </rdg>
  </rdgGrp>
</app>
<pc xml:id="cYRlu9o">.</pc>
```

Example: the omissions

```
% We iterate over all witnesses represented by {sigle}
% Match all apps with empty reading who follow and app with empty reading
//tei:app[contains(@ana,'#omission')]
[preceding::tei:app[1][contains(@ana,'#omission')]
[descendant::tei:rdg[not(node())][contains(@wit, '{sigle}')]]
[descendant::tei:rdg[not(node())][contains(@wit, '{sigle}')]]]
% Retrieve their position
% get all adjacent apps
% filter with given threshold
omissions = [(a - 1, b) for a, b in ranges
if b - a >= (lacuna_sensibility - 1)]
% Insert witStart and witEnd at given indices
```

```
<witEnd xml:id="eojSEAY_uhshqI9_jSTC1Ly_pKhR3I8_tx7Iq1G" corresp="#Mad_A #Mad_G #Sal_J #Sev_R #Sev_Z" next="#nVgfSVr" ana="#homeoteleuton" cert="high"/>
<!--
  <witEnd xmlns="http://www.tei-c.org/ns/1.0" xmlns:tei="http://www.tei-c.org/ns/1.0" xml:id="uhshqI9" corresp="#Mad_G" next="#hZ6QVu1" ana="#homeoteleuton" cert="high"/>
-->
<!--
  <witEnd xmlns="http://www.tei-c.org/ns/1.0" xmlns:tei="http://www.tei-c.org/ns/1.0" xml:id="jSTC1Ly" corresp="#Sev_R" next="#yNNnpny" ana="#homeoteleuton" cert="high"/>
-->
<!--
  <witEnd xmlns="http://www.tei-c.org/ns/1.0" xmlns:tei="http://www.tei-c.org/ns/1.0" xml:id="pKhR3I8" corresp="#Sal_J" next="#fJ6XOQI" ana="#homeoteleuton" cert="high"/>
-->
<!--
  <witEnd xmlns="http://www.tei-c.org/ns/1.0" xmlns:tei="http://www.tei-c.org/ns/1.0" xml:id="tx7Iq1G" corresp="#Sev_Z" next="#gbUVWBf" ana="#homeoteleuton" cert="high"/>
-->
<app ana="#not_apparat">
  <rdgGrp>
    <rdg wit="#Esc_Q #Mad_B">
      <w ana="#same_word">en</w>
    </rdg>
  </rdgGrp>
</app>
<app ana="#not_apparat">
  <rdgGrp>
    <rdg wit="#Esc_Q #Mad_B">
      <w ana="#same_word">su</w>
    </rdg>
  </rdgGrp>
</app>
<app ana="#not_apparat">
  <rdgGrp>
    <rdg wit="#Esc_Q #Mad_B">
      <w>grado</w>
    </rdg>
  </rdgGrp>
</app>
<app ana="#not_apparat">
  <rdgGrp>
    <rdg wit="#Esc_Q #Mad_B">
      <w ana="#same_word">e</w>
    </rdg>
  </rdgGrp>
</app>
<witStart xml:id="gbUVWBf_fJ6XOQI_yNNnpny_hZ6QVu1_nVgfSVr" corresp="#Mad_A #Mad_G #Sal_J #Sev_R #Sev_Z" previous="#tx7Iq1G"/>
```

Modelling - Limitations

- The XML-TEI input cannot be too complex.
- Cannot deal with large transpositions yet (whole paragraph)
- Can we really transfer any information from document to document ?

Technology

- Python/LXML is really powerfull for XML processing and analysis
- It should be taught more...
- ...Even if it's not the best tool for some tasks (tokenization, transformation into LaTeX or web-based interfaces)

Further works

- Improve pre-processing validation (please come and test the scripts with your documents!)
- Improve the alignment (global alignment?)
- Add semantic computation to textual collation

Link to the gitlab repo

Thank you !

`https://gitlab.huma-num.fr/mgillelevenson/tei_collator`

las prinçipales que fazen al cavallero. **1** La una es elecçión, que deve ser escogido de todos los otros
asý commo es mejor et para más **2** La ii° es[R] sacramento, ca deve fazer[r] sacramento e jura el dia 75R
que le fazen cavallero[iii] .
Onde cuenta Vegeçio en el libro de la cavallería, fablando de la elección de los cavalleros, e dize
que en esto está la salud de la Rrepública, o de toda comunidat : que los cavalleros non solamente
sean escogidos por bondat de sus cuerpos, mas por nobleza de sus coraçones ; ca la virtud de todo
el rregno de Rroma e el çimiento del su nonbre en la primera examinaçión de la elección déstos 80R
está. E por que este ofiçio non sea tenido por liviano nin por pequeño, nin sea dado[12]a qualquier

[12] Saut du même au même ici pour QABGJZ.

[iii] (J)

Miles enim dicitur quare unus ex mille electus, uel miles mille valens, ait Papias. Romulus gratia(?) prius e populo sumpsit e appellauit. Et sic miles est nomen laboris ita(?) honoris, ait Policraticus l. vi, c. viii. Secundo, videndus est de his qui militem faciut, de quibus Policraticus, l. vi, c. v et vii, bene ubi ait quod duo sunt principua que militem faiiunt: electio e sacramentum. [A 332/111, I, 9, 2, « *De institutione militie, e militis nuncupatione sine effectione* », fol. 50v]

Le glosateur vient ici gloser et expliciter les deux concepts de sélection et de sacrement : il ne s'agit pas simplement d'une traduction du texte de Jean de Galles.

81 dado *QAGRBJZ* | Saut du même au même pour le témoin base : dado a aventura *R*

70 fueron QABZ|fueran *GRJ* 71 cavallero *AZ'GRBJ* |cavalleros *Q* 71 dixo QB|dize *AGRJZ* 71

Bibliography

Bleeker, Elli et al. “Including XML Markup in the Automated Collation of Literary Text.”. In: XML Prague 2018. Prague, République Tchèque, 2018.

Camps, Jean-Baptiste, Lucence Ing, and Elena Spadini. “Collating Medieval Vernacular Texts: Aligning Witnesses, Classifying Variants”. In: *DH2019 Digital Humanities Conference*. DH2019: Complexities. Utrecht, 2019.

CondorCompPhil/Falcon. CondorCompPhil, July 12, 2019.

Dekker, Ronald Haentjens. “Computer Automated Collation with CollateX and Python”. In: *Presentation given at DH Benelux in The Hague* (2014), pp. 12–13.

Kiessling, Benjamin. “Kraken - an Universal Text Recognizer for the Humanities”. In: DH2019 : Complexity. Utrecht, 2019.

Nury, Elisa. “Automated Collation and Digital Editions From Theory to Practice”. 2018.

Nury, Elisa and Elena Spadini. “From Giant Despair to a New Heaven: The Early Years of Automatic Collation”. In: *it-Information Technology* 62.2 (2020), pp. 61–73.

Spadini, Elena. “Studi Sul Lancelot En Prose”. PhD thesis. PhD thesis, Sapienza Università di Roma, 2016.