

# **AOR: Transcriber's Manual**

Tenth version (August 2016)

By: Jaap Geraerts

## Contents

Introduction.....	3
1. The process of transcribing.....	4
1.1 Information about the page .....	5
1.2. Spelling and interpunction .....	6
1.2 Tagging and transcribing reader's interventions: marginalia.....	7
1.2.1 Marginalia element .....	7
1.2.2 Language element .....	10
1.2.3 Position element.....	10
1.2.4 Marginalia_text element .....	14
1.2.5 Person element .....	14
1.2.6 Book tag.....	15
1.2.7 Location tag .....	16
1.2.8 X-reference tag.....	16
1.2.9 Emphasis tag .....	17
1.2.10 Internal reference tag .....	18
1.2.11 Translation tag .....	21
1.2.12 Links between marginal notes.....	22
1.2.13 Marginalia that run across pages.....	23
1.3 Tagging and transcribing reader's interventions: underline.....	27
1.4 Tagging and transcribing reader's interventions: symbols .....	29
1.5 Tagging and transcribing reader's interventions: marks .....	32
1.5.1 Brackets.....	35
1.5.2 Circumflex .....	38
1.5.3 Est mark .....	39
1.5.4 Hash .....	39
1.5.5 Horizontal bar .....	40
1.5.6 Underlining vs horizontal bars.....	41
1.5.7 Page break.....	43
1.5.8 Pen trial and scribble.....	44
1.5.9 Straight quotation marks vs quotation marks.....	45
1.5.10 Unknown marks .....	45
1.5.11 Rows of marks .....	50
1.6. Numerals .....	51

1.7 Drawings .....	52
1.8 Linking reader's interventions to the printed text I: marginalia.....	52
1.9 Linking reader's interventions to the printed text II: symbols .....	54
1.10 Linking reader's interventions to the printed text III: marks .....	56
1.11 Changing the punctuation .....	61
1.12 Changing the spelling: the errata tag.....	63
1.13 Missing text & Uncertainty.....	67
1.14 Contractions and Abbreviations.....	69
1.15 Strikethrough.....	72
2. Transcribing in XML .....	73
2.1 Using an XML editor .....	73
2.2 Validation.....	73
3. Workflow .....	74
3.1 First phase: transcribing .....	76
3.2 Second phase: checking .....	77
3.3 Third phase: finalised transcriptions .....	78
3.4 How to deal with errors.....	79
4. Spreadsheets.....	80
5. Sources.....	82
Appendix A: DTD.....	83

## **Introduction**

The first phase of ‘The Archaeology of Reading in Early Modern Europe’ focuses on thirteen books annotated by Gabriel Harvey. We transcribe all the interventions made by Harvey in order to provide the end user with a fully searchable dataset of Harvey’s annotations. There are two documents which relate to the transcriptions generated by this project: The ‘Transcription and Encoding Policy’ is primarily meant as an explanation of the XML schema that will be used in this project, whereas this document, the ‘Transcriber’s Manual’, is more tuned towards the actual practice of transcribing and functions as some sort of field guide. (At this point in the project, at the very end of phase I (August 2016), the former document has been subsumed into the Transcriber’s Manual.) The Transcriber’s Manual provides a detailed overview of Harvey’s annotations, discusses a number of ambiguities, and offers some guidelines as to how to deal with complex annotations. Moreover, another important component of this document is the description of the workflow (i.e. the process of generating and checking transcriptions, as well as the internal communication within the project team). In order to be as comprehensive as possible, this document will be continuously updated.

## 1. The process of transcribing

The transcriptions have to be based upon and are validated against an external XML schema (.xsd) and a DTD. The schema contains the various elements, their attributes, and their values, whereas the DTD consists of a set of special characters (e.g. è). The schema and DTD can be found at the project's GitHub repository,<sup>1</sup> yet are also available on more permanent URL's, namely:

Schema: [http://www.livesandletters.ac.uk/schema/aor\\_20141118.xsd](http://www.livesandletters.ac.uk/schema/aor_20141118.xsd)

DTD: [http://www.livesandletters.ac.uk/schema/aor\\_20141023.dtd](http://www.livesandletters.ac.uk/schema/aor_20141023.dtd)

Every transcription has to refer to the schema and the DTD by using these URL's, and the starting lines in the XML files always have to look like this (although the filenames of the DTD and schema can be different, as these reflect the latest version):

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>

<!DOCTYPE transcription SYSTEM
"http://www.livesandletters.ac.uk/schema/aor_20141023.dtd">

<transcription xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:noNamespaceSchemaLocation="http://www.livesandletters.ac.uk/schema/aor_20141118.
xsd">
```

In order to make sure that Oxygen, the XML editor we use in this project, does not validate the transcription against the DTD but against the schema, go to the menu (in Oxygen) Document > Validate > Validation options, and tick the box 'Ignore the DTD for validation if schema is specified'.

Once the link with the external DTD and schema has been established, Oxygen will recognize the elements and the attributes, the values of which can be easily selected in the right section of the screen, all of which greatly increases the speed with which transcriptions can be generated. Oxygen also forces the transcriber to follow the structure of the XML schema, and if the hierarchy established in the schema is not respected (i.e. if the elements and their child elements in the XML file do not follow the sequence laid out in the schema), Oxygen will give an error. Therefore, transcribers should start attaching the necessary information to the page by making use of the page element and its attributes.

---

<sup>1</sup> <https://github.com/livesandletters/aor/tree/master/XMLSchema>

Thereafter the reader's interventions should be tagged, starting with the marginalia, followed by all underlining, symbols, marks, numerals, errata and, finally, drawings.

The external DTD ensures that special characters are declared and can be used in the transcription. The list of special characters that are included in the DTD is certainly not exhaustive, and if you need to use a special character that is not on the list, contact me (Jaap Geraerts) and I will add it to the DTD.

## 1.1 Information about the page

Once the link with the schema and DTD has been established, the transcriber can start to add information about the page by making use of the `page` element. This element has four attributes, filename, pagination, signature and reader. The filename should be an **exact copy** of the filename of the digital surrogate of the particular page, including the extension (or suffix), for example: Ha2.001r.tif. The pagination and signature attributes refer to the place of a page within a book. Pagination refers to any printed numbering on the page: whether they be Roman numerals, Arabic numerals, whether they are consistent throughout the volume or whether paratextual material is given a separate pagination apart from the main body of the text, as well as including errors introduced by the printer. Signature refers to the standard bibliographical practice of giving an identifier for each leaf of a book using a method inherited from the book-printing and binding methods of the early modern era. For example, signature A1r would be followed by signature A1v. The convention for capturing the signatures is as follows:

- Inferred information is put between brackets. For instance, if A1 is printed, the 'r' (indicating the recto side of the paper) is put between brackets: A1[r].
- Often the first instance of a signature does not contain a number (e.g. A, aa), but we silently attribute a number in case there isn't one. This number is also put between brackets: A[1r].
- Usually the signature is only given on the recto side of the page, hence for the verso side all information about the signature is inferred, hence: [A1v]. In some cases, for the sake of clarity, the letter indicating the side of the paper should be bracketed separately. For instance, [ai[v]], since otherwise the indication of the verso side could be read as forming the Roman number 4 (iv).
- The white spaces between the different components of the signature are omitted: not A 1 [r] or A 1[r], but A1[r].
- We only infer the signature of the verso side of the page. If the recto side does not have a signature, we do not infer the signature based on the extrapolation of the signature of a preceding page. The only exception to this rule is when the page numbers are flawed or non-

existent; in that case, using an inferred signature actually helps the end users to navigate through the book.

- The pagination should be treated similarly. For instance, in the Princeton Machiavelli, the folio number is only given on the recto side. Hence, on the transcription of the verso side we put the folio number in brackets, with the verso indication bracketed separately: [livi[v]].

We can also use terms such as boards, endpaper, pastedown, flyleaf, as descriptions where necessary. The reader attribute contains the name of the reader.

## 1.2. Spelling and interpunction

Although the project will make transcriptions and translations of the original text available, we do not aim to provide a (scholarly) edition of the text. As a result, we can be a bit more pragmatic about which “rules” to follow, primarily serving the easiness with which the user can work with the transcriptions, rather than becoming enmeshed in academic discussions about what constitutes the best edition.

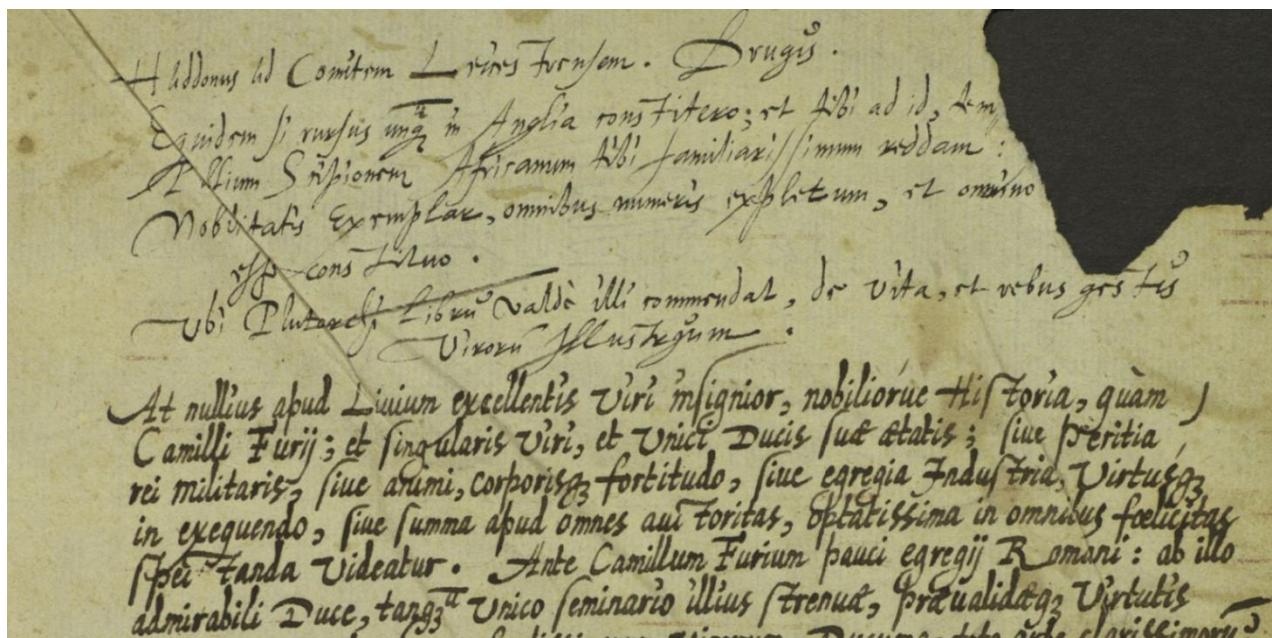
As a result, we do have some freedom in setting rules for our transcription, yet the more interventions we make, the more time it will cost. In general, therefore, the transcriptions will closely follow the manuscript, thus retaining its interpunction and its spelling. Unlike other projects, such as ABO, we will not adept or modernize the spelling. For example, in Latin marginalia we will not transcribe the “u” as a “v” when standing between vowels, nor will the “j” be changed to an “i” (e.g. egregii instead of egregij). Special characters, such as the “&” and the various sorts of accents will be retained as well, using the appropriate XML-code (& et cetera). It also occurs that at the end of a line a word has been divided in two by using a dash, something which will not be retained in the transcription.

Whereas the transcriptions closely follow the original text, the translations should be optimized for clarity and “usability”. Therefore, translations follow modern standards of interpunction and the use of capitals. Names should be standardized - as the original spelling is given in the transcription, users can search for names in both the original and modern, standardized variant. English annotations will not be translated, nor will English phrases be standardized.

## 1.2 Tagging and transcribing reader's interventions: marginalia

### 1.2.1 Marginalia element

The marginalia element consists of five main attributes: hand, date, other reader, topic, and text. The hand attribute specifies the hand in which the reader's comment was written (in case of Harvey, either English secretary or Italian). Although within these hands there was some variation, we only make a distinction between his English Secretary and Italian hand.



(Source: ABO, Livy, *Romanae historiae principis*, scan 3/1148).

This image clearly shows the differences between Harvey's hands. The first part of the annotation is written in Harvey's secretary hand, whereas the second part, starting with 'At nullius apud Liuium' is written in his Italian hand.<sup>2</sup> For another example of Harvey's secretary hand, see the image on page 4.

<sup>2</sup> Virginia F. Stern, *Gabriel Harvey: His Life, Marginalia and Library* (Oxford, 1979), plate E.

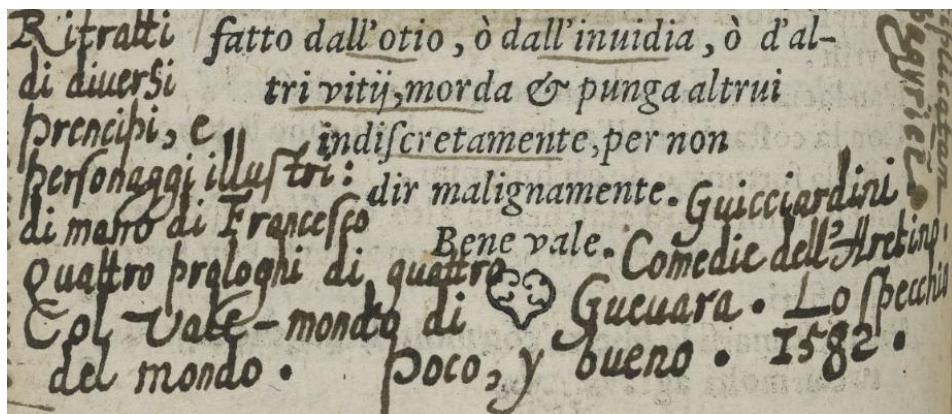
(Source: Walter Coleman, M.A., *Gabriel Harvey's Paper-Book. A critical edition of MS. Sloane 93 (British Library) with an introductory essay and explanatory notes*).

The date attribute contains the date at which the annotation was made: sometimes Harvey mentioned the date at which he read a book (on the title page, or in individual marginal notes), while it sometimes is possible to link an annotation to a specific year because of its content.<sup>3</sup>



(Source: AOR, Livy, *Romanae historiae principis*, 1r).<sup>4</sup>

One should take into account that some books were read more than once, something which can be betrayed by different hands and ink used. In these cases, the date mentioned on the cover page only provided a very limited degree of certainty about the date of individual marginal notes.



(Source: AOR, Giucciardini, *Detti, et fatti piacevoli*, 77v).

The reader attribute contains the name of another annotator (if any). The topic attribute specifies the topic of the marginal comment (which the transcriber can select from a predefined list of topics). By attaching a topic to the marginalia, we provide the users with a perspective through which the marginalia of different readers (i.e. also the readers included in the second phase of this project) can be studied. Although we started to generate a list of topics, we realised that attaching topics to marginal notes would drastically increase the editorial weight bearing upon the transcriptions. As this project aims to let the user do the interpretation, rather than creating transcriptions which have been thoroughly shaped by our own understanding of reading practices and strategies, **we do not make use of the topic tag**. However, other scholars, such as Philip Palmer, have fruitfully started to experiment with formulating a list of

<sup>3</sup> Anthony Grafton and Lisa Jardine, “”Studied for Action”: How Gabriel Harvey Read His Livy”, *Past & Present* 129 (1990), 37.

<sup>4</sup> The (folio) numbers refer to the number of the pages in the AOR viewer (rather than the original number of the books themselves).

possible topics (as well as typologies), and therefore the topic tag will remain part of our schema. Finally, the anchor\_text attribute can be used to capture the printed text to which the marginal comment refers. This attribute can be tricky, see sections 1.8—10. In the end, of the attributes of the marginalia element, only the hand attribute is required; the other attributes are optional, so, if you do not have the information one of these ‘implied’ attributes needs, do not use them.

```
<marginalia hand="Italian" date="1568" anchor_text="lex">
```

### 1.2.2 Language element

The marginalia element has two child elements, namely **language** and **translation**. Starting with the first of these two, the **language** element has one attribute, *ident*, which is the identifier of a particular language. The transcriber has to choose a language from this predefined list of languages.

EN = English

EL = Greek

FR = French

IT = Italian

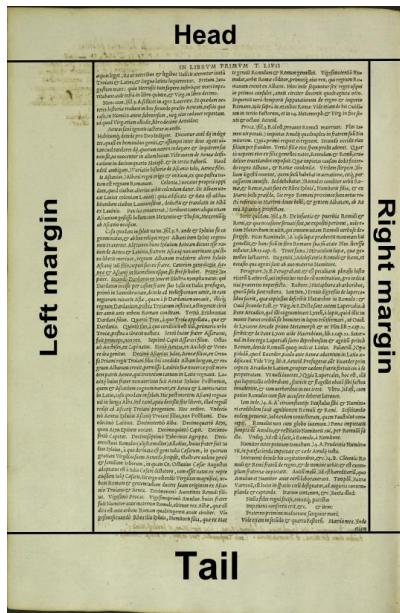
LA = Latin

ES = Spanish

The code for this element thus is: <language ident="XX">

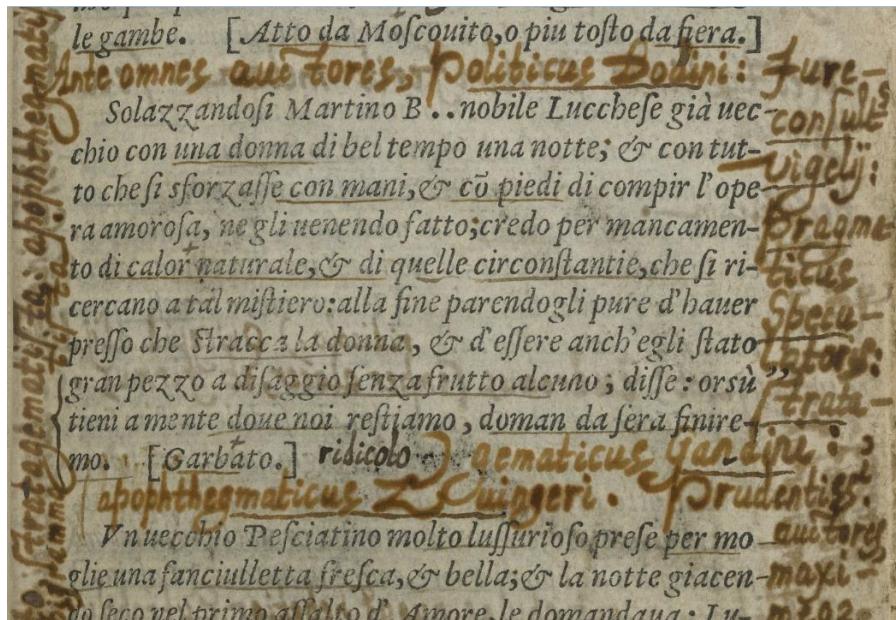
### 1.2.3 Position element

The **position** element is the only child element of the **language** element, and has two attributes which capture the place of the marginalia on the page as well as the orientation of the book relative to the reader. A page has four margins: left, right, head and tail. The left and right margins are the spaces on both sides of the printed text, while the head and tail are the margins above and below the printed text, respectively.



The remaining option of the place attribute is “full\_page”, which has to be used when Harvey writes marginal notes on a blank page (see the image on page 3 (scan 3/1148), for example).

The place attribute specifies the marginal space in which an annotation was written, but one has to take into account that many of Harvey marginal notes were spread over different marginal spaces.



(Source: AOR, Domenichi, *Facetie, motti, et burlì...*, 1r)

The marginal comment ‘Ante omnes auctores, Politicus Bodini: Jureconsultus Vigelij: pragmaticus speculatoris: stratagematicus ) Gandini:, apophthegmaticus Zuingeri.’ (the marginal note goes on even further), moves back and forward between two various marginal space, namely ‘intext’ and

‘right\_margin’. This should be captured in the XML transcriptions, and can be by closing and opening new instances of the position tag (other elements are left out of this example so as to focus on the role of the position element):

```
<marginalia hand="Italian">

    <language ident="LA">

        <position place="intext" book_orientation="0">
            <marginalia_text> Ante omnes auctores, politicus Bodini:
            </marginalia_text>
        </position>

        <position place="right_margin" book_orientation="0">
            <marginalia_text> Jureconsult[i]s Vigelij:
            <Pragmaticus Speculatoris:</marginalia_text>
            </position>

        <position place="intext" book_orientation="0">
            <marginalia_text> ) strategematicus Gandini: apophthegmaticus Zuingeri.
            </marginalia_text>
        </position>

    </marginalia>
```

It becomes clear that the position of the marginal note in the original source cannot exactly be captured because some words are spread across more than one marginal space, as is the case with the word “strategematicus” and “Gandini”. A rule of thumb can be to assign the entire word to the marginal space where the majority of its letters are written. Because of the many variations possible, this decision is left to the discretion of the transcriber.

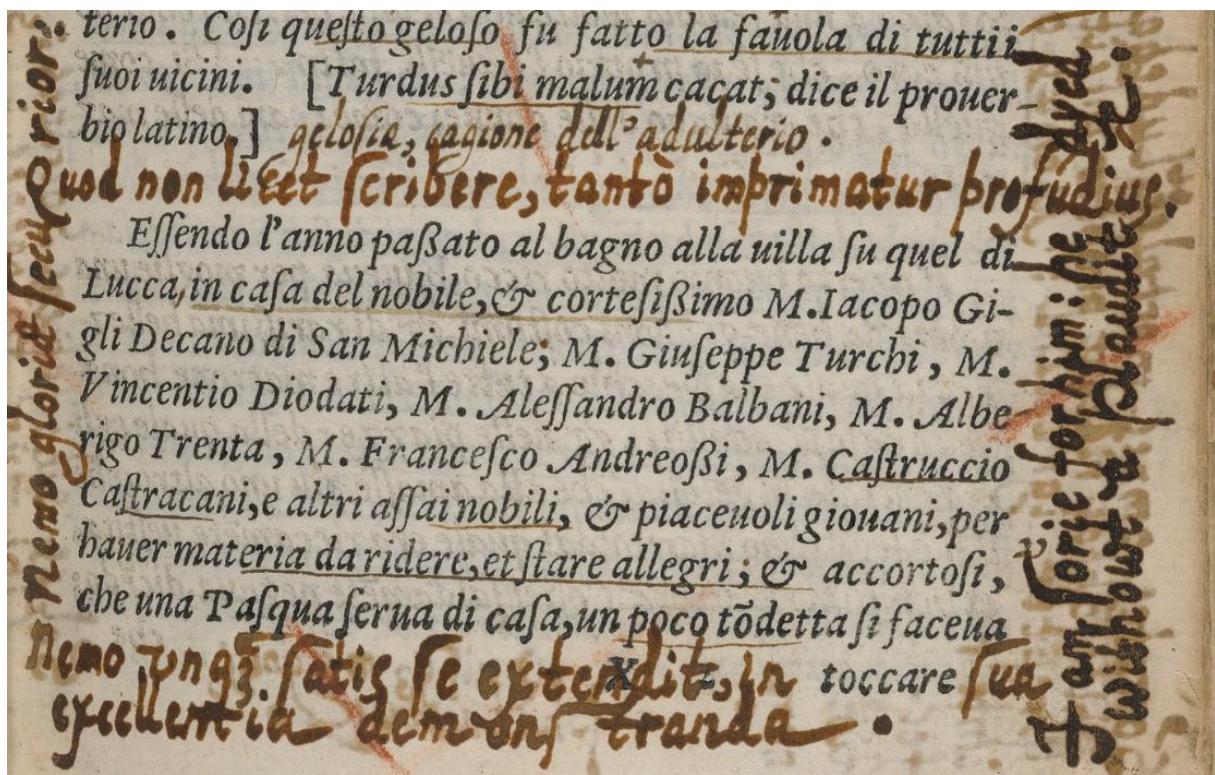
The book\_orientation attribute contains the position of the page relative to the reader. For instance, sometimes readers wrote marginalia in the left margin of the book, starting at the lower left corner and ending at the upper left corner (and this is but one of the variants which are possible).<sup>5</sup> In

---

<sup>5</sup> See, for instance, AOR, Domenichi, *Facetie, motti, et burli*, 2v.

order to be able to write an annotation in this manner, the reader had to rotate the book. To capture the various combinations, the book\_orientation attribute contains four different degrees in which the book could be rotated (in a clockwise direction). (This tag is based on the assumption that Harvey always wrote from left to right - and this tag might have to be revised when including readers such as Casaubon who wrote marginalia in right-to-left languages such as Arabic or Hebrew.)

[As a note on the side: transcribers should be aware of the interest scholars have in the use of capitals in early Italic hands. It is especially tricky to distinguish between a capital ‘S’ and ‘P’ and the lower case variants, since often their form (not always their size) is similar. In the example above, ‘Speculatoris’ is written with a capital ‘S’, but ‘stratagematicus’ is not (the so-called ‘long s’ never is a capital ‘S’). In case of uncertainty, highlight this by putting in a comment field the text ‘Uncertainty about capital’.]



(Source: AOR, Domenichi, *Facetie, motti, et burli*, 2r.)

Consider the marginal comment written in the lower part of the right margin. In order for Harvey to write this comment, he had to turn the book 90 degrees, so that the right margin effectively became the tail. The XML code would be:

```
<marginalia hand="Italian">
  <language ident="EN">
    <position place="right_margin" book_orientation="90">
```

```

<marginalia_text> I am sorie for him:  

    he dyed without a plaudite </marginalia_text>  

</position>  

</language>  

</marginalia>

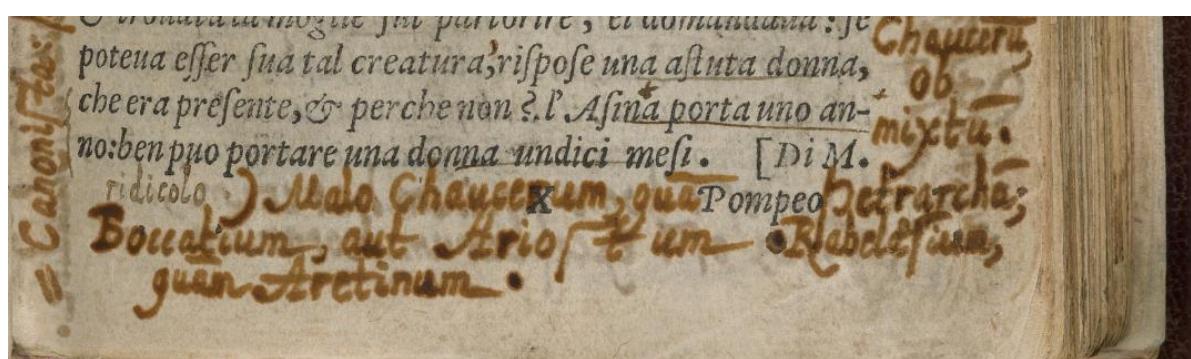
```

#### 1.2.4 Marginalia\_text element

The **marginalia\_text** element, a child element of the **position** element, captures the text of the marginalia written by Harvey. This element does not have any attributes, but just contains the text as it appears on the page (see the examples above).

#### 1.2.5 Person element

The second child element of the **position** element is the **person** element, which captures the persons mentioned in a particular marginalia. This empty element has one attribute, name, which is the standardized name of the person mentioned in the marginalia (the name attribute serves as the UID of a person in the corresponding spreadsheet). <person name="Augustine">. It is important that the names of the persons tagged are used in a coherent way, in order to prevent that tags are not recognized (by the database). The names can be transcribed as they appear in the text, but should always be tagged in a standardized way (in order to maximise coherence). As the names of the people Harvey refers to are stored in a spreadsheet, this spreadsheet should therefore be the first point of reference when looking for the standardized name of a person (or book or location). For the database, see chapter 4.



(Source: AOR, Domenichi, *Facetie, motti, et burlī*, 1r)

The marginal note at the bottom of the pages mentions a number of people, and the XML will be as follows:

```
<marginalia hand="Italian">

    <language ident="LA">

        <position place="tail" book_orientation="0">

            <marginalia_text> Malo Chaucerum, qua[m] Petrarcha[m]; Boccatium, aut
            Ariostum Rabelaesum, quam Aretinum.</marginalia_text>

            <person name="Francesco Petrarca" />
            <person name="Francois Rabelais" />
            <person name="Pietro Aretino" />
            <person name="Lodovico Ariosto" />
            <person name="Geoffrey Chaucer" />
            <person name="Giovanni Boccaccio" />

        </position>

    </language>

</marginalia>
```

Note here that the people mentioned in the marginal note will be only tagged after the **marginal\_text** element has been closed, thus observing the sequence of the child elements of the **position** element. In case books and locations are mentioned in the marginal note as well, first all the persons will be tagged, and only then the books and geographical locations mentioned will be tagged.

## 1.2.6 Book tag

Similar to the person element, the book element is empty, but contains the title of the book a reader refers to (in the title attribute) `<book title="De civitate Dei" />`. This tag encompasses all sorts of texts the reader refers to, and the title of the texts can be transcribed as it appears in the original text, but should always be tagged in a normalized way (in order to maximise coherence). The standardized titles of the books Harvey refers to are stored in the books spreadsheet.

Note: If Harvey mentions a book, its title should be put in between inverted commas in the translation of the marginal note. For example, <translation> Cicero's 'On the Orator' is the best book...</translation>.

### 1.2.7 Location tag

The fourth element contains the geographical locations mentioned by the reader in his annotations. This element is empty as well, and the standardized name of the location is given in the location\_name attribute (the name of the location referring to the standardized name in the corresponding column in the location spreadsheet). <location name= "London" />.

**Note: within one marginal note, the books, persons, locations should be tagged only once – i.e. if Harvey mentions Augustine twice in one marginal note, this person is only tagged once. However, every time a book/location/person is mentioned in a separate marginal note, this should be tagged (since we want to capture the frequency with which Harvey referred to them in his annotations).**

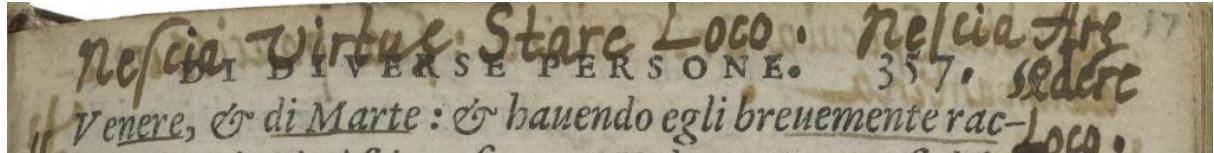
### 1.2.8 X-reference tag

Sometimes Harvey quoted from other sources, without mentioning the name of the author or the title of the book (for an example, see: AOR, Machiavelli, *The Art of Warre*, 2r: Imbelles Damae, quid nisi praeda sumus? This is a quote from Martial, XIII, 94). If a transcriber recognizes the source Harvey is quoting from, the name of the author can be captured in the person attribute of this element, while the title of the book can be given in its book\_title attribute. The text of the quote should be captured in the text attribute, and the language attribute should be used to specify the language in which the quote is written. Both the name of the author and the title of the book should be the standardized names also used in the **person** and **book** tag.

```
<X-ref person="Martial" book_title="Epigrams" language="LA" text=" Imbelles Damae, quid nisi praeda sumus?">
```

If you find such a 'hidden' reference, make sure to put the quote in between quotation marks in the translation, another sign for the end users that Harvey got this from another source. For example: <translation> "Helpless deer, what are we but prey?"</translation>

*Transcribers are not asked to check every bit of Harvey's writing to find such 'hidden' references*, for this would be too time consuming. This element is only to be used when the sources of phrases written by Harvey spring to mind. A google search can be pretty effective, especially since many classical text can be found online (e.g. at perseus.edu). For example, on the title page of Livius' *History of Rome* (see AOR, 1r.), Harvey quoted Ovid ('Acceptissima semper Munera sunt, Auctor quae preciosa facit.'), and google was quick to find the source. Even without a profound knowledge of classical literature, such references can be detected since they are often short, concise phrases:

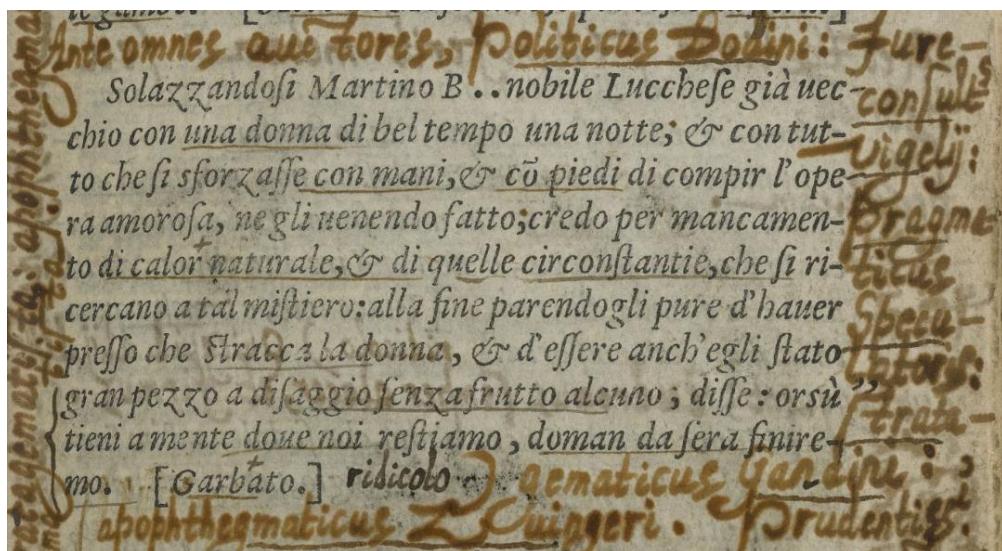


(Source: AOR, Domenichi, *Facetie, motti, et burli*, 17r)

In this example, the phrase 'Nescia virtus stare loco.' is taken from Lucan's *De bello Civili*, and also was found by using google.

### 1.2.9 Emphasis tag

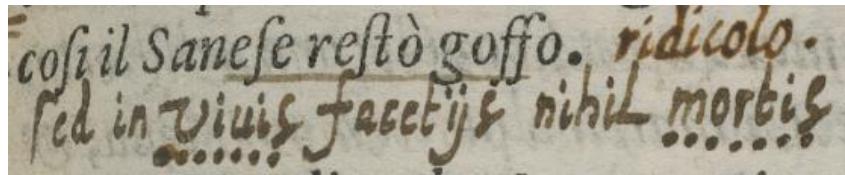
Sometimes, Harvey underlines (part of) the text of his own annotations. In order to capture this, the emphasis elements has a method attribute, as Harvey underlined by using a pen or chalk, a type attribute (as Harvey used various 'types' of underlining) and a text attribute, which contains the words underlined by Harvey.



(Source: AOR, Domenichi, *Facetie, motti, et burli*, 1r.)

```
<emphasis method="pen" type="straight" text="Jureconsulti[s] Vigelij: pragmaticus speculatoris" />
```

An example of words in his marginal note Harvey underscored with small dots:



(Source: ABO, Domenichi, *Facetie, motti, et burla...*, scan 19/224, right page)

### 1.2.10 Internal reference tag

Often Harvey uses words such as ‘supra’ and ‘infra’, sometimes accompanied by one or more page numbers, to refer to other places in the book he is reading and annotating. Such ‘internal references’, as we call them, need to be tagged separately so that we can create links and enable end users to quickly navigate to the place Harvey is referring to (NB: the actual creation of these links will be one of the main technical developments in phase II). The creating is the links in the XML files is done by these elements and attributes:

```
<!ELEMENT internal_ref (target+)>

<!ATTLIST internal_ref text CDATA #IMPLIED>

<!ATTLIST internal_ref anchor_text CDATA #IMPLIED>

<!ELEMENT target EMPTY>

  <!ATTLIST target filename CDATA #REQUIRED>
  <!ATTLIST target book_id CDATA #REQUIRED>
  <!ATTLIST target text CDATA #REQUIRED>
```

Basically Harvey makes use of two variants: supra/infra with or without page numbers, each of which should be tagged in a distinct way. In case of the second variant, say “supra, 50”, the XML will look like this:

```
<internal_ref text="s[upr]a" >
  <target filename="00000043.tif" bookid="PrincetonPA6452" text="50">
```

[Note: Other annotators, including John Dee, sometimes only mention a page number, without any preceding text such as supra or infra.<sup>6</sup> In these cases, the text attribute of the internal\_ref element is left empty, and the page number is inserted in the text attribute of the target element.]

The ‘internal\_ref text’ attribute captures the term (supra/infra) Harvey used (**without the page numbers he mentions!**), while the link is created through the attributes of a child element, ‘target’. Filename links to the name of the digital image of the page Harvey refers to, book\_id to the unique identifier of that book, the value of which can be found in the ‘Identifier column’ of the [corpus table](#) (these two attributes function as IDREFs, but since IDREF only can refer to an identifier declared within the same document, the values of these attributes is CDATA). The text attribute contains the page Harvey is referring to (which will, in AOR phase II, be transformed into an HTML clickable link in our viewer).

Note that we need to use a target element for every page number Harvey mentions: “s[upr]a, 50, 59” would become:

```
<internal_ref text="s[upr]a">
<target filename="00000043.tif" bookid="PrincetonPA6452" text="50">
<target filename="00000043.tif" bookid="PrincetonPA6452" text="59">
```

In case Harvey only writes down supra or infra, we need to attribute one or more page numbers in order to create the link. However, we need to show the end user that we inferred the page number, and this is why, in those cases in which Harvey did not supply references to specific pages himself, we need to **put the page number in between square brackets**. If he writes ‘infra’, we capture this as:

```
<internal_ref = "i[nfr]a">
<target filename="00000043.tif" bookid="PrincetonPA6452" text="[50]">
```

If our research shows that, while just using infra/supra, Harvey was referring to more than one page, we need to use several target elements to capture this. It’s important to realize that sometimes it is impossible to establish to what page Harvey was referring exactly, especially because he now and then wrote ‘s[upr]a, i[nfr]a’ in a single marginal note, the terms almost becoming the equivalent of ‘passim’. In case a link cannot be established with sufficient certainty, no link should be coded in XML.

---

<sup>6</sup> Thanks to Philip Palmer for pointing this out to me!

Harvey did not only use ‘supra’ or ‘infra’ to establish links within particular books. For example, Harvey writes ‘Liuius decadis 3. libro 3’. If we can find the exact page Harvey was referring to (since we can’t link to sections or parts of the book), this will be captured as follows:

```
<reference text="Liuius decadis 3. libro 3">
  <target filename="00000043.xml" bookid="PrincetonPA6452" text="[50]">
</reference>
```

The anchor\_text<sup>7</sup> attribute can be used to further specify the internal link by including words from the printed text or a marginal note. For example, on page 11 of Buchanan’s *De Regina Scotorum*, Harvey writes: “Morauius vt s[upr]a, i[nfr]a.” This is captured as follows:

```
<internal_ref text="s[upr]a" anchor_text="Comes Moraviae">
  <target filename="00000015.xml" book_id="PrincetonRB16th11" text="[3]">
</internal_ref>
```

In this case the value of the anchor\_text attribute, ‘Comes Moraviae’, are two words in the printed text which Harvey underlined.

There is one other variant of what we consider to be an internal reference, namely the cases in which Harvey refers to *another* book *within* our (digitized) corpus. In his Livy (p. 334), Harvey writes: “Frontinus libro I. cap. 6”. This reference is tagged as follows:

```
<internal_ref text="Frontinus libro I. cap. 6.">
  <target filename="earbm_stc_11402_0052.xml" book_id="HoughtonSTC11402"
text="[Bvii[r]]">
</internal_ref>
```

Keep in mind that such links can only be established for books *within* the AoR corpus. If Harvey refers to a book outside our corpus, the book will be captured by the book tag (since linking out to external object, i.e. objects external to our corpus, will be a feature of the second phase of AOR).

---

<sup>7</sup> I would like to thank Philip Palmer for his useful suggestions on the further development of the internal reference tag.

### 1.2.11 Translation tag

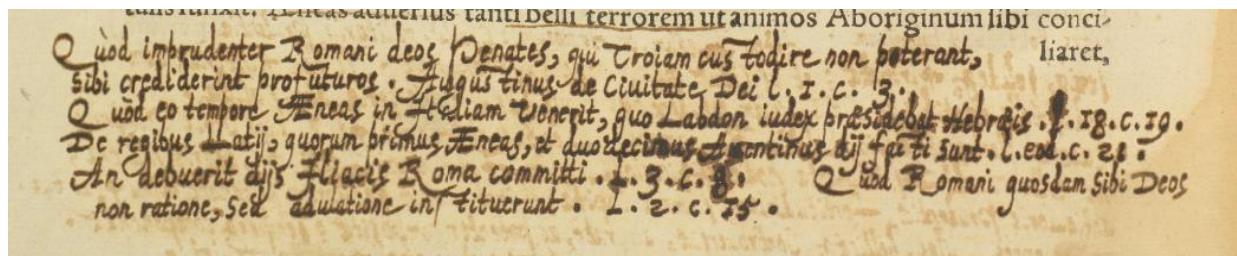
Once all the metadata is attached to the marginalia, the translation of the marginal note is put **translation** element, a child element of the **marginalia** element).

<translation>Skilfully and solidly </translation>

Those comments written in English will not be translated. All other languages will be translated into English, and the translation should include the standardized names of the people, books and places, among other things, Harvey mentions.

In some cases, for instance when Harvey copies a part from a (classical) text, translations can be found online. The Perseus database, for example, contains both editions and translations of a large number of classical texts.<sup>8</sup>

To give an example of all the elements we have covered so far:



(Source: AOR, Livy, *Romanae historiae principis*, 9v).

<marginalia hand="Italian">

<language ident="LA">

<position place="tail" book\_orientation="0">

<marginalia\_text> Quod imprudenter Romani deos Penates, qui Troiam custodire non poterant, sibi crediderint profuturos. Augustinus de Civitate Dei. 1. I. c. 3. Quod eo tempore Aeneas in Italiam venerit, quo Labdon iudex praesidebat Hebraeis. 1. 18. c. 19. De regibus Latii, quorum primus Aeneas, et duodecimus Aventinus dii facti sunt. 1. eod. c. 21. An debuerit diis Iliacis Roma committi. 1. 3. c. 8. Quod Romani quosdam sibi Deos non ratione, sed adulazione instituerunt. 1. 2. c. 15. </marginalia\_text>

<person name="Augustine" />

<person name="Aeneas" />

<person name="Labdon" />

<sup>8</sup> <http://www.perseus.tufts.edu/hopper/collection?collection=Perseus:collection:Greco-Roman>

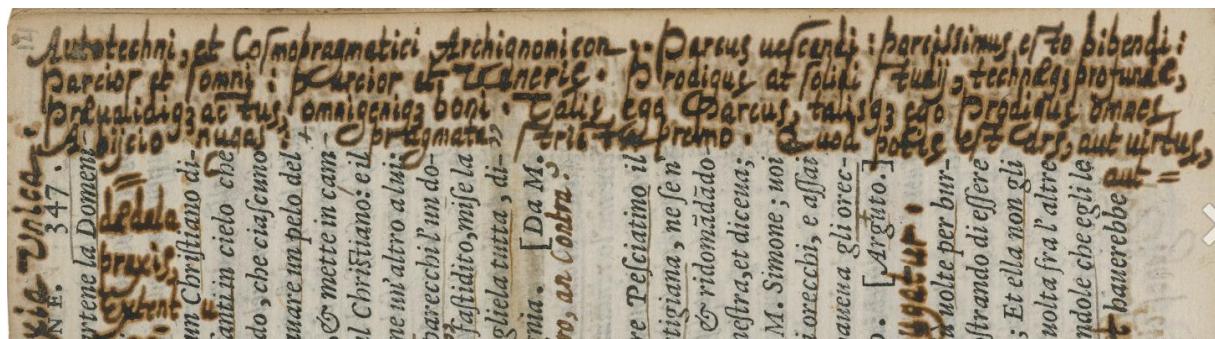
<person name="Aventinus" />  
 <book title="De Civitate Dei" />  
 <location name="Rome" />  
 <location name="Troy" />  
 </position>  
 </language>

<translation>How imprudent the Romans were in believing that they might derive any benefit from the Penates (household gods), who could not protect Troy: Augustine City of God, bk. 1, c. 3. That Aeneas came to Italy at the time when Labdon presided as judge over the Hebrews: bk 18, c. 19. Of the kings of Latium, among whom, Aeneas, the first, and Aventinus, the twelfth, were made gods. The same bk, c. 21. Whether Rome ought to be entrusted to the Trojan Gods. Bk. 3, c. 8. That the Romans established certain of their gods through flattery, not reason. bk. 2, c. 15. </translation>

</marginalia>

### 1.2.12 Links between marginal notes

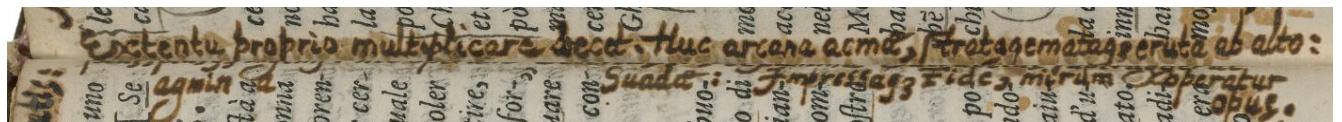
As Harvey's annotations activities were restricted by the available white space on the page, he was sometimes forced to continue his marginal notes on another page (see next section), or on another part of the same page. He used various marks in order to link the various parts of a marginal note, as the following example shows.



(Source: AOR, Domenichi, *Facetie, motti, et burlì*, 14r)

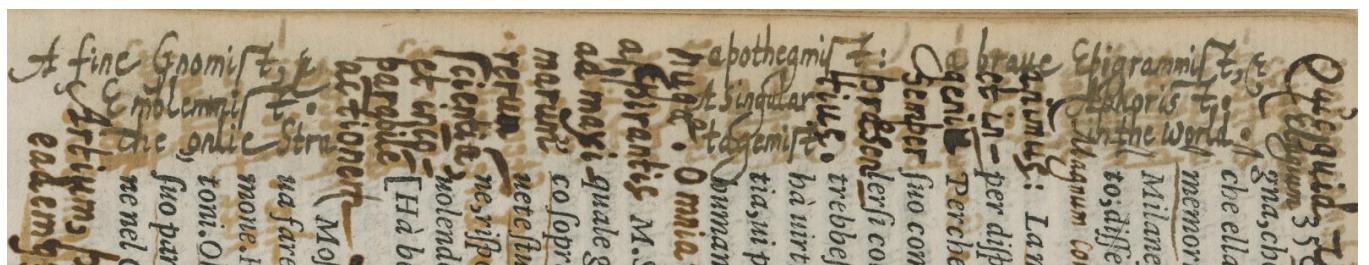
In the right margin of the page (the image has been turned to optimise the view), Harvey started with a long annotation that was not yet completed when he reached the lower end of the right margin. After the last word of this part of his marginal note, 'aut', Harvey also included an equal sign, and the last part of this marginal note that is written in between the printed text (the words 'da[e]dala praxis, extent) also starts with an equal sign.

Since Harvey virtually always ended his marginal notes with a dot, when he did not do so and instead ended his comment with another (punctuation) mark, this is a sign for the transcriber that the marginal note might have been continued elsewhere.



(Source: AOR, Domenichi, Domenichi, *Facetie, motti, et burli*, 14r)

In this image, the marginal note in the gutter of the page ends with a colon, and it is likely that the Harvey continued in the gutter of the previous page (the text of which is unfortunately not visible on the image. After having had a look at the original book, it turns out that the marginal note indeed continues in the gutter of the preceding page). It will not always be obvious where the marginal note continues, especially not when it is spread across pages, yet by looking at the marks used by Harvey, the ink (see the image below), and the meaning of the marginalia, it should be possible to connect the various parts of a marginal note.



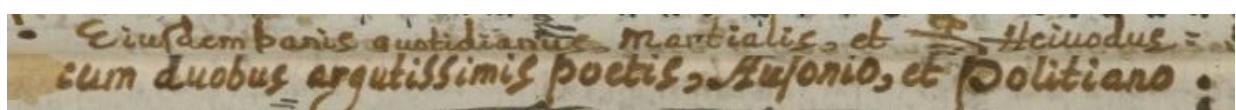
(Source: AOR, Domenichi, *Facetie, motti, et burli*, 15v)

### 1.2.13 Marginalia that run across pages

It happened that Harvey continued a marginal comment on another page, as the following examples show:



(Source: AOR, Domenichi, *Facetie, motti, et burli*, 3r-v)



(Source: AOR, Domenichi, *Facetie, motti, et burli*, 3v-4r)

In the first example, the marginal comment starts on the head of the left page, and continues on the head of the right page. The other example consists of a marginal note starting on the right\_margin of the left page, continuing on the left\_margin of the right page. This is tricky to capture, not in the least because we are transcribing individual pages. As a result, the marginal note is split and captured as two individual notes. However, by linking marginalia and the pages on which they are written via a set of attributes, a combination of ID's and IDREF's, we are able to connect these marginalia. These are the attributes:

<!ATTLIST marginalia marginalia\_id CDATA #IMPLIED> A unique ID, formed by the filename and a four digit number, always starting at 0001, separated by a underscore. The attribute is implied, since not every marginalia has an ID, but only those which continue on another page. The value of the attribute is CDATA rather than ID, since IDs only work within a XML file and hence cannot be used to refer to a unique value declared in another file.

These two attributes establish the link to the other part of the marginal note (written on either the preceding or following page, depending on which part of the marginal note is being transcribed). The attributes contain the unique ID of the other part of the marginal note (this value is CDATA rather than IDREF, since IDs and references to them only work within a single XML file). Note that only one of the two attributes should be used in the transcription of a single marginal note.

<!ATTLIST marginalia marginalia\_continues\_to CDATA#IMPLIED>

<!ATTLIST marginalia marginalia\_continues\_from CDATA#IMPLIED>

The following two attributes capture the ‘direction’ of the marginal note: the marginal note either is continued on the following page, or is a continuation of a marginal note on the preceding page. The value of the marginal note is the filename of the transcription onto which it continues or from which it is continued. The extension (or suffix) of the file is included as well (e.g. Ha2.099v.xml). We use the filename of the transcription and not of the digital image, since we need to derive the information which is part of the transcription.

<!ATTLIST marginalia marginalia\_to\_transcription CDATA #IMPLIED>

<!ATTLIST marginalia marginalia\_from\_transcription CDATA #IMPLIED>

The last marginal note contains the ID of the book in which the other part of the marginal note is written. The unique ID of the books in our corpus are the values in the ‘Identifier column’ of the corpus table, which can be found [here](#).

<!ATTLIST marginalia book\_id IDREF #IMPLIED>

Although most of the marginal notes which run across pages do not run across pages of different books, this attribute makes it possible to deal with the case in which this does occur.

Some examples:

On page 22 of the Dominichi, Harvey wrote in the head ‘{ Pro vno Amico singula:’, and continued on the head of the following page ‘{ In vnum Hostem omnia.’.

When transcribing the marginal note, starting on page 22, the transcribers should give the marginal note a unique ID. As mentioned, this ID is a combination of the filename (in this case Ha2.099v) and a four-digit number, starting at 0001. In this case, the ID becomes: Ha2.099v\_0001. Since we’re transcribing the first part of the marginal note, we need to use the ‘marginalia\_continues\_to’ attribute to link to the other part by inserting the ID of that part, HA2.100r\_0001 (note: this IDREF has to be an exact copy of that ID, otherwise a wrong link may be established). Having done that, we proceed with the ‘marginalia\_to\_filename’, Ha2.100r.tif (note: remember to include the extension here, since we’re linking to the digital image), followed by the ‘book\_id’, which is FolgersHa2.

```
<marginalia hand="Italian" marginalia_id="Ha2.099v_0001"
marginalia_continues_to="Ha2.100r_0001" marginalia_to_transcription="Ha2.100r.xml"
book_id="FolgersHa2">
```

The XML code for the other part of the marginal note will be:

```
<marginalia hand="Italian" marginalia_id="Ha2.100r_0001"
marginalia_continues_to="Ha2.099v_0001" marginalia_to_transcription="Ha2.099v.xml"
book_id="FolgersHa2">
```

Note: since this way of establishing links has been developed recently (August 2015), we should trace all the ‘running marginalia’ and insert the appropriate XML code. As we put a standard comment in the XML files when encountering a ‘running marginal note’, this should not be a problem:

“[The text from previous versions of the Manual.] We need to find a way to link the two marginalia that are now tagged individually on separate pages. A possible solution is to give the two marginalia IDs and link them via IDREF, but this has not been included in the DTD yet, as this will be discussed at the January workshop. In lieu of a final solution, it is important that transcribers insert a comment in the transcription, so that at a later stage it is easy to find the marginalia which are spread across pages.

For instance, when transcribing the marginal comment of the first example, the XML code would look like this.

```
<marginalia hand="Italian">
<language ident="LA">
<position place="head" book_orientation="0">
```

```

<marginalia_text>
    At non legit quotidi&egrave; rarissima mundi ingenia: -
</marginalia_text>

<!-- marginalia continues: Domenichi_008right, head, 'et spiritus omnium viuidissimos.'-->
</position>
</language>
</marginalia>

```

And the transcription of the next page includes:

```

<marginalia hand="Italian">
    <language ident="LA">
        <position place="head" book_orientation="0">
            <marginalia_text>
                - et spiritus omnium viuidissimos.
            </marginalia_text>
    </position>
    </language>
</marginalia>

```

<!-- Continued from page Domenichi\_008left, head, 'At non legit quotidi&egrave; rarissima mundi ingenia: -' -->

The information about the remainder of the marginal comment (the part on the other page) the transcriber should include in his or her comments are: **filename**, **place** of the marginalia on the page, and the **text** of the marginalia.

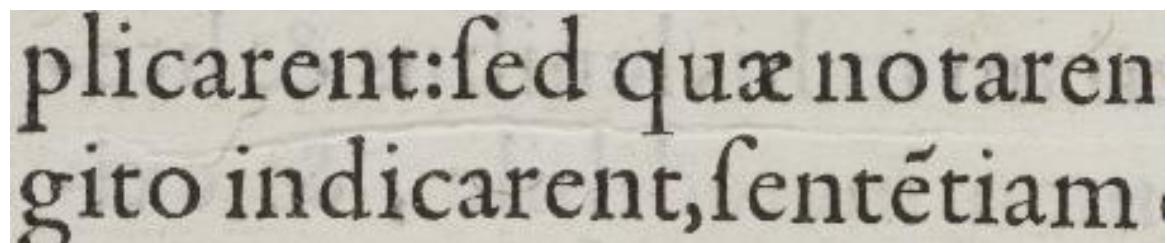
It was not only the lack of whitespace on a page that caused Harvey to write marginal notes that were linked to each other on different pages. For Harvey was telling a story, if you like, or a larger argument, which consisted of various marginal notes that could be spread throughout a book or even was dispersed in various books. On page 34 of Livy's *History of Rome* (AOR, 25v), for example, Harvey referred to a number of authors and books which attacked and defended the rights of monarchs, such as the notorious *Vindiciae contra tyrannos*, and he ended his comment with: 'Nisi quod magna etiam

controversia inter iurisconsultos Imperiales, et Pontificios; de qua alias accuratius' (Except that there is also a great disagreement between imperial jurists and pontifical ones, about which elsewhere in more detail). Since we do not yet have a mechanism to link various marginalia, for the moment we should stick to referring to other marginalia by making use of the comment field."

### 1.3 Tagging and transcribing reader's interventions: **underline**

The **underline** element incorporates all underlining of the *printed text* on a page. This is an empty element with four attributes: method (the method of underlining, chalk, pen or scoring), type (straight, curved or dotted) text (the words that are underlined by the reader) and language (the language of the text that is underlined).

The 'scoring' option of the method attributes captures the instances in which Harvey underlined words by making a physical mark in this page, as is visible on the next image:

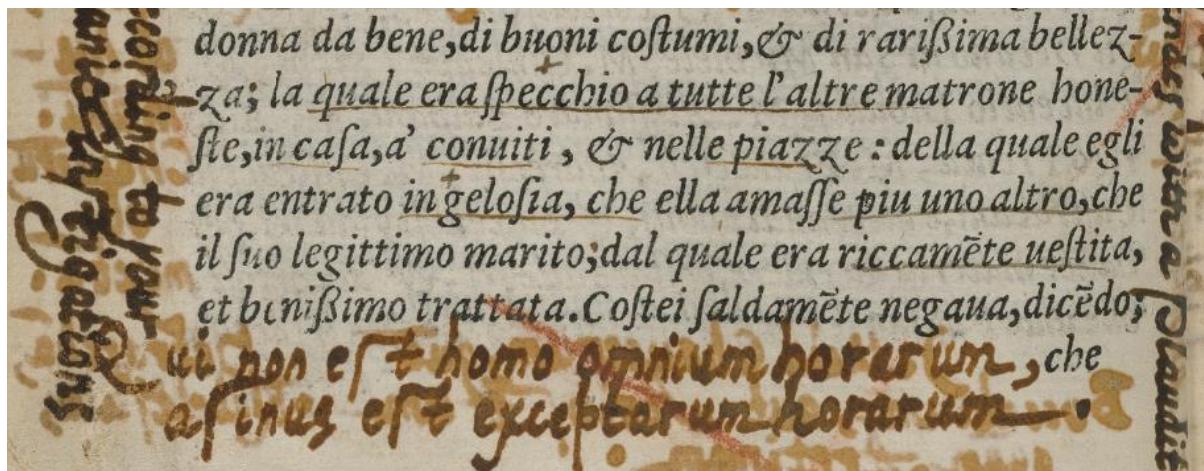


(Source: AOR, T. Smith, *De linguae Anglice*, 47v).

This underling should be captured as follows:

```
<underline method="scoring" type="straight" language="LA" text="explicarent: sed quae notarent" />
```

As has already become clear, the text attribute contains the word (or words) underlined by Harvey or other readers. It is important to note that this attribute only contains a word or a sequence of words which are underlined. However, if in one line two words are underlined, but separated by a word that is not underscored, both of the underlined words have to be captured individually.



(Source: AOR Domenichi, *Facetie, motti, et burlī*, 1v).

<underline method="pen" type="straight" language="IT" text="quala era specchio a tutte l'altre matrone" />

Here *matrone* is included in the text attribute, for part of the word is underlined, and it would not make sense only to include the letters that are underlined.

<underline method="pen" type="straight" language="IT" text="in casa" />

<underline method="pen" type="straight" language="IT" text="conuiti" />

Although the underlined words ‘in casa’ and ‘conuiti’ appear on the same line (and in the same sentence), they are tagged individually, for they are separated by printed text which is not underlined. Also remark that Harvey used chalk to highlight a part of the printed text (as well as his marginal note on the bottom of the page). In this case, the underlining cannot be linked to specific words; hence the text attribute of this element is optional. As this underlining both ‘touches’ the printed text as well as the marginal note, this reader’s intervention should be capture twice.

<underline method="pen" type="straight" />

Within the marginalia tag, the emphasize tag should be used to capture the red line in chalk.

<marginalia hand="Italian">

<language ident="LA">

<position place="tail" book\_orientation="0">

<marginalia\_text> ui non est homo omnium horarum, asinus est exceptorum horarum.  
 </marginalia\_text>

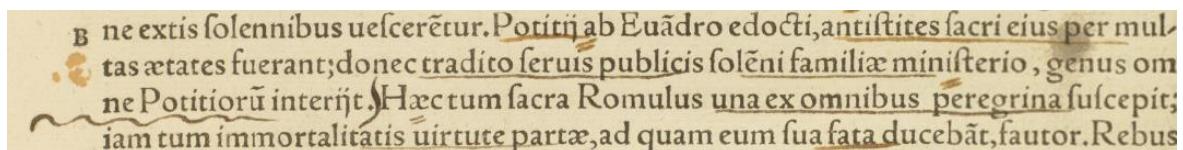
<emphasis method="chalk" type="straight" />

</position>

</language>

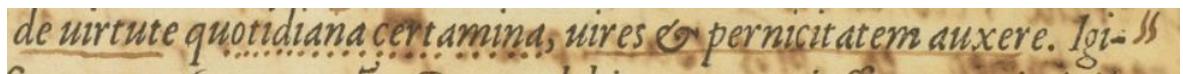
</marginalia>

It might seem strange to say that in this example the type of underlining/emphasis is straight, but remember that the line itself is straight not curved. More importantly, we do not record the direction (or angle) of the line, which would cause us to work with coordinates, but we only capture that Harvey used chalk to ‘mark’ some printed text and one of his marginalia. Curved underlining is clearly discernible as such, and looks like this:



(Source: AOR, Livy, *Romanae historiae principis*, 11r).

The last type of underlining employed by Harvey is small dots. It seems that he did not do this very frequently, yet we do capture this variant by using a unique value of the type attribute.



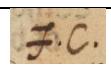
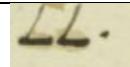
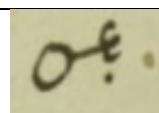
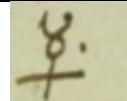
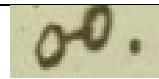
(Source: AOR, Livy, *Romanae historiae principis*, 8r).

## 1.4 Tagging and transcribing reader's interventions: symbols

Harvey used a set of astrological symbols which referred to larger abstract topics – for example, the Mars symbol refers to passages on war and warfare, while the sun symbol represents kings or kingship – to mark up passages in the printed text. The **symbol** tag incorporates all the symbols used by Harvey, but within this group each symbol has a fixed ID or type, enabling users to search all symbols as well as specific symbols. Such a system also makes it possible to incorporate the symbols used by other readers, without affecting the general structure of the XML schema. The **symbol** element has four attributes: name (the class or ID of the symbol), place (the place of the symbol on the page), language (the language of the text captured in the text attribute) and text (if the symbol clearly refers to printed text or the text written by the annotator, the text to which it refers is put in this attribute). <symbol name= "Mars" place= "left\_margin" language="LA" text="bellum" />. Note that the text attribute is not required, as many symbols, as well as other interventions by the reader are standing on their own (or are difficult to link to a specific passage or word in the printed text). The language attribute specifies

the language of the text captured in the text attribute, and these two attributes should only be used in combination with one another (i.e. the use of the text attribute necessitates the use of the language attribute).

The transcriber can chose a symbol from this predefined list of symbols:

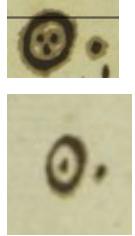
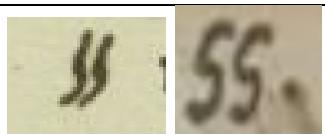
Symbol	Meaning	Figure
Asterisk *	Astronomy	
Biseected circle	The biseected circle covers topics about the earth or natural history. <sup>9</sup>	
Crown	Appears in Machiavelli's <i>The Arte of Warre</i> (ABO, scan 187). Deals with warfare and the role of generals.	
HT	The sign 'may be a '3' joined to a version of the sign for Mercury. If so, it might stand for Hermes Trimegistus and its reference would be appropriate enough in the context here cited'. <sup>10</sup>	
J.C.	Stands for <i>juris consultus</i>	
LL	Stands for <i>leges</i> or <i>legibus</i>	
Mars	Denotes war, warfare	
Mercury	Harvey uses this sign to signify eloquence, but sometimes trickery as well. 'Harvey used the sign to mark titles or passages in which he perceived any obvious relation to the commonly assigned characteristics or jurisdiction of Hermes or Mercury, in mythology or astrology'.	
Moon	Used to signify references to the moon. According to Wilson, '[i]t is used in Simlerus in conjunction with the sign for Sol...'. <sup>11</sup> This is also the case in the Newberry's Castiglione, f. Ppii r.	
Opposite planets	Used to mark controversy or opposition of any kind <sup>12</sup>	
Saturn		

<sup>9</sup> Stern, Harvey, 141. Wilson, 'Gabriel Harvey's Method of Annotating', 356.

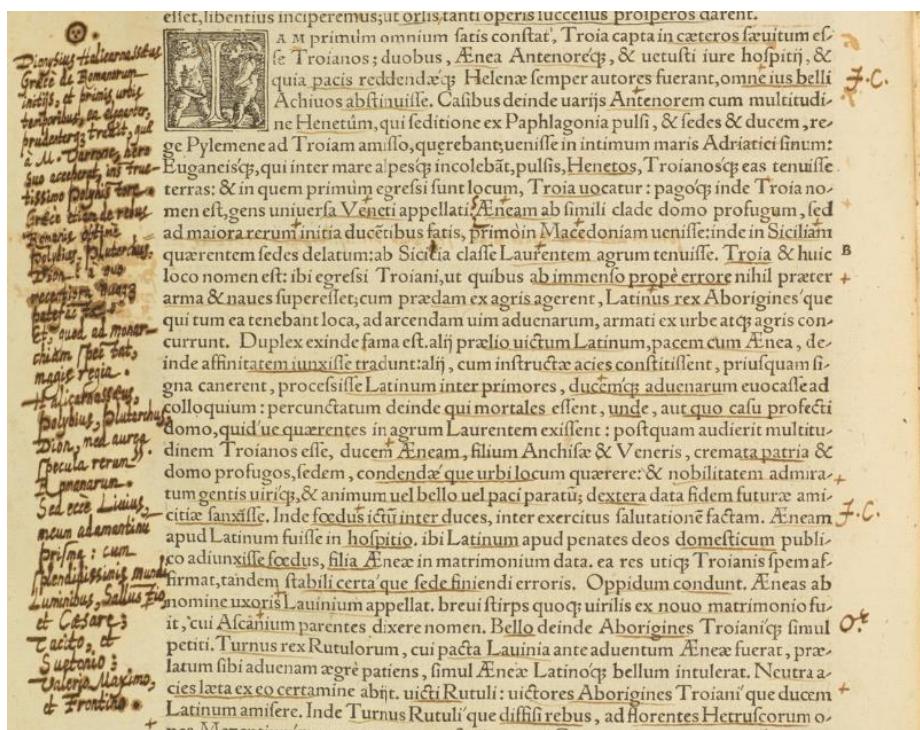
<sup>10</sup> Ibid., 358.

<sup>11</sup> Wilson, 'Gabriel Harvey's Method of Annotating', 356.

<sup>12</sup> Wilson, 'Gabriel Harvey's Method of Annotating', 355.

Square	According to Wilson, this symbol was only used in <i>Simlerus</i> to ‘distinguish writers in the humanities, and especially writers on theology’ <sup>13</sup>	
Sun	Sign for Sol, occurs ‘in Simlerus in its common astrological reference to emperors, kings, and lords’. Was also associated by Harvey to mortality and to medical professions, as well as to land and property. <sup>14</sup>	
SS	Scilicet; also ‘senatus sententiae’ (see Harvey’s copy of Machiavelli’s <i>Art of Warre</i> ).	
Venus		

This list covers all the symbols used by Harvey, and working with them is relatively easy. Take the following example, for instance.



(Source: AOR, Livy, *Romanae historiae principis*, 9v).

<sup>13</sup> Wilson, ‘Gabriel Harvey’s Method of Annotating’, 358.

<sup>14</sup> Wilson, ‘Gabriel Harvey’s Method of Annotating’, 356–78. Stern, *Harvey*, 141: The sun denotes kingship.

The correct code for tagging the symbols on the image above would be:

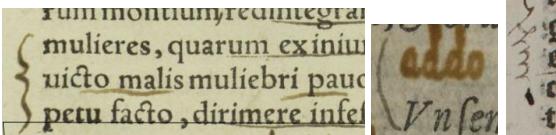
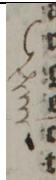
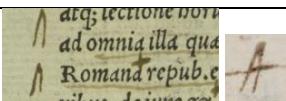
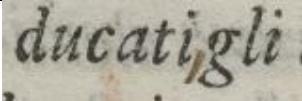
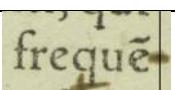
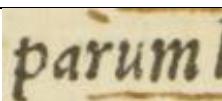
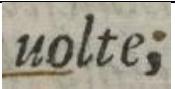
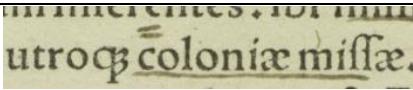
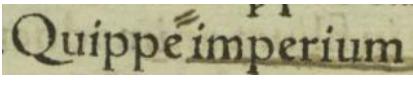
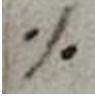
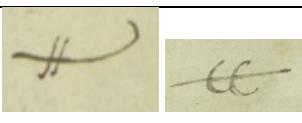
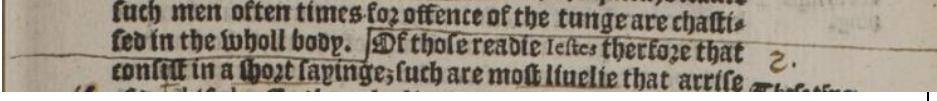
```
<symbol name="sun place="left_margin" />  
<symbol name="JC" place="right_margin" language="LA" text="jus belli"/>  
  
<symbol name="JC" place="right_margin" />  
  
<symbol name="Mars" place="right_margin" language="LA" text="bello" />
```

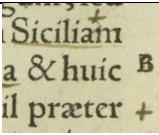
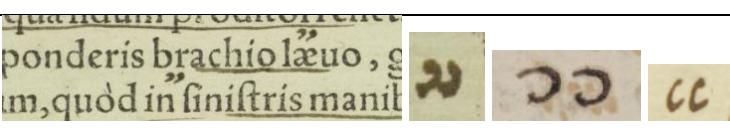
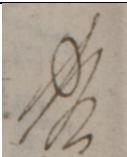
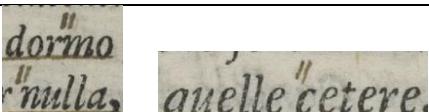
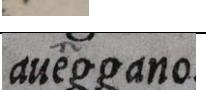
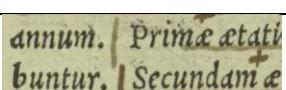
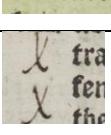
In the last tag, the Mars symbol is linked to the printed word ‘bello’ (by using the text attribute), for in this case between the meaning of the symbol and the printed text (in this case the word ‘bello’, which Harvey underlined as well).

## 1.5 Tagging and transcribing reader's interventions: marks

The **mark** tag incorporates all the marks used by Harvey, and within this group each mark has a fixed ID or type, enabling users to search all marks as well as specific marks. Such a system also makes it possible to incorporate the marks used by other readers, without affecting the general structure of the XML schema. Like the **symbol** element, the **mark** element has four attributes: name (the class or ID of the symbol), place (the place of the mark on the page), language (the language of the text captured in the text attribute) and text (if the mark clearly refers to printed text or text written by the annotator, the text to which it refers is put in this attribute). Note that the text attribute is not required, as many marks, as well as other interventions by the reader are standing on their own (or are difficult to link to a specific passage or word in the printed text). The language attribute specifies the language of the text captured in the text attribute, and these two attributes should only be used in combination with one another (i.e. the use of the text attribute necessitates the use of the language attribute).

Apostrophe	
Box	Harvey sometimes drew a box around a word or passage: 

Bracket		
Circumflex		
Colon		
Comma		
Dash		
Diacritic		
Dot		
Double_vertical_bar		
Equal sign		 utroq; coloniæ missæ.
		
Est mark		(see below, pp. 27-8).
Hash		
Horizontal bar		
Page break		
Pen trial		

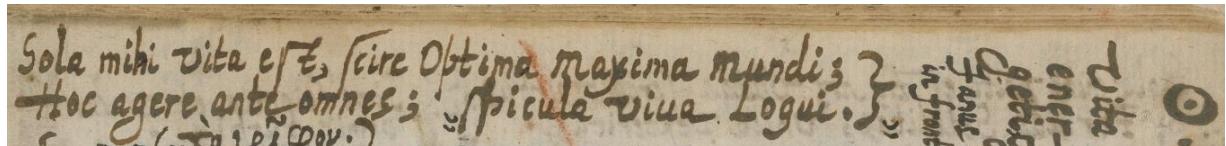
Plus sign	
Quotation mark	
Scribble	
Section_sign	 <sup>15</sup>
Semicolon	
Slash	
Straight quotation mark	
Tick	
Tilde	
Triple dash	
Vertical bar	
X-sign	

Due to the large variety of marks, the place on the page at which they appear, and the proximity of other reader's interventions such as symbol, underlining, and marginalia, can make it somewhat difficult to

<sup>15</sup> Also known as 'pilcrow', 'capitulum', and 'paragraph'. Thanks to Claire M. L. Bourne for pointing this out. For examples, see: Machiavelli, *Arte of Warre*, images \_136; \_137; \_138.

work with them. The following examples aim to shed some light on how the mark element should be used.

### 1.5.1 Brackets

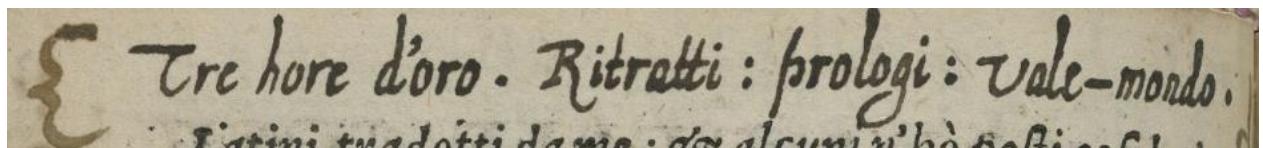


(Source: AOR, Domenichi, *Facetie, motti, et burli*, 7v)

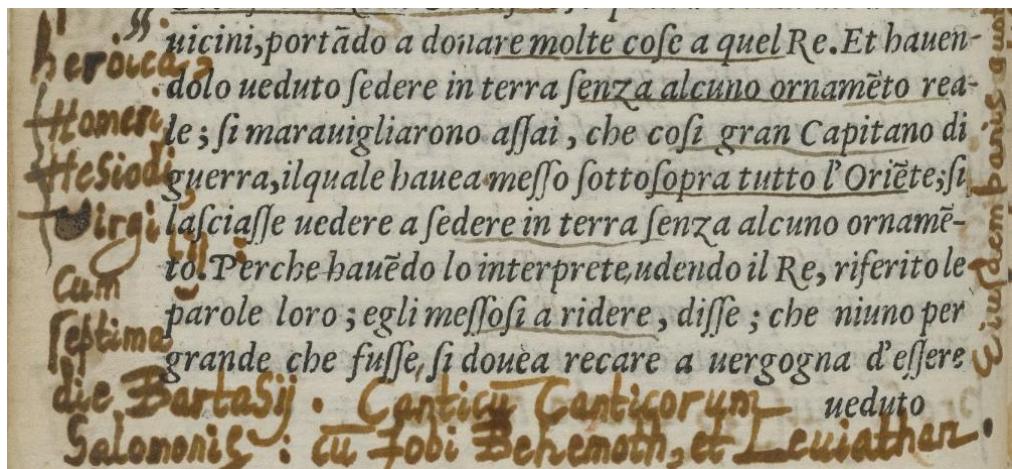
On this page, the bracket } clearly is part of the marginalia (the same ink is used, for example), and thus should be part of the marginalia\_text element.

<marginalia\_text>Sola mihi vita est, scire optima maxima mundi; Hoc agree ante omnes; spicula viua Loqui. } </marginalia text>

The other brackets in this image are part of the marginal annotation Harvey made in Greek.



(Source: AOR, Guiccardini, *Detti, et fatti piacevoli*, 77v.)



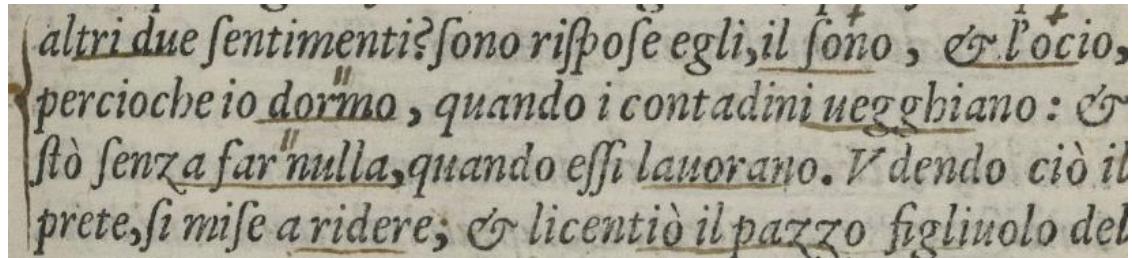
(Source: AOR, Domenichi, *Facetie, motti, et burli*, 3v)

In these two images, the ink of the brackets is different than that of the marginal notes, hence these brackets should be tagged as a mark.

<mark name="bracket" method="pen" place="head" language="IT" text="Tre hore d'oro. Ritratti: prologi: vale-mondo."/>

<mark name="bracket" method="pen" place="left\_margin" language="LA" text="Homeri, Hesiodi, Virgilij"/>

It is likely that these brackets were added at a later stage, and the use was similar to that of putting a bracket in front of printed text, as Harvey did here:

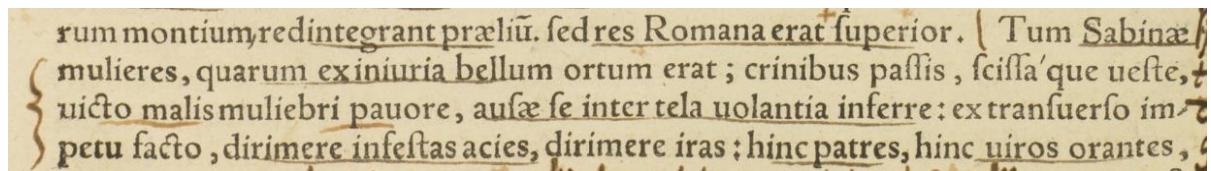


(Source: AOR, Domenichi, *Facetie, motti, et burli*, 4r)

Whereas in the previous example, the bracket could be clearly linked to the (written) text, in this case it's much more difficult to use the text attribute of the mark element, for we do not know exactly which sentences (or parts of the sentences) Harvey captures. Hence the best way to capture this tag is:

<mark name="bracket" method="pen" place="left\_margin" />

Sometimes, though, it is possible to link the bracket to a part of the printed text.



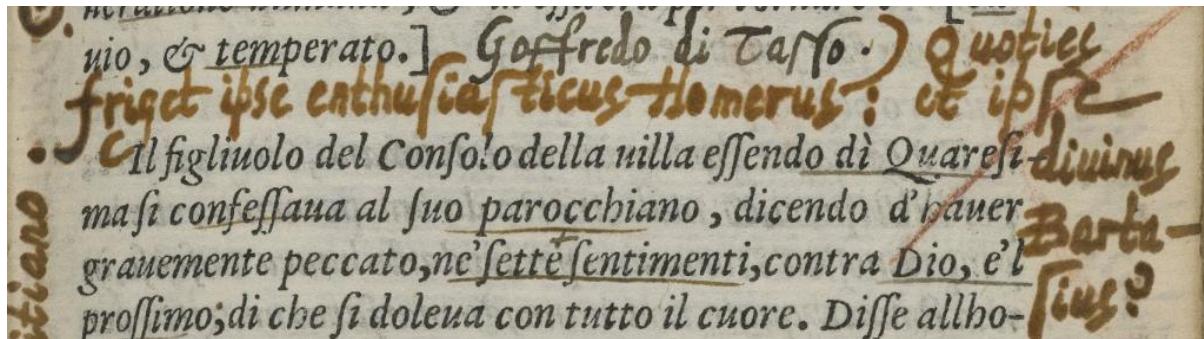
(Source: AOR, Livy, *Romanae historiae principis*, 12r).

Here, for instance, we can use the fact that Harvey put a vertical bar in the text (before 'Tum'), thereby isolating the sentence, as it were. Moreover, this sentence also is the last sentence on the page. Therefore we can capture this bracket as follows.

<mark name="bracket" method="pen" place="left\_margin" language="LA" text="Tum Sabinae mulieres, quarum ex iniuria bellum ortum erat; crinibus passis, scissa que ueste, victo malis muliebri pauore, ausae se inter tela volantia inferre; ex transuerso impetu facto, dirimere infestas acies, dirimere iras: hinc patres, hinc uiros orantes, ne se"/>

So even though we tag all interventions Harvey made in the text separately, we can use a combination of interventions (here the ‘|’ and the ‘{‘ mark to establish the link between the bracket and the printed text.

Harvey also used brackets to distinguish between and to separate different marginalia. This was in particular necessary on pages with a limited amount of white space.



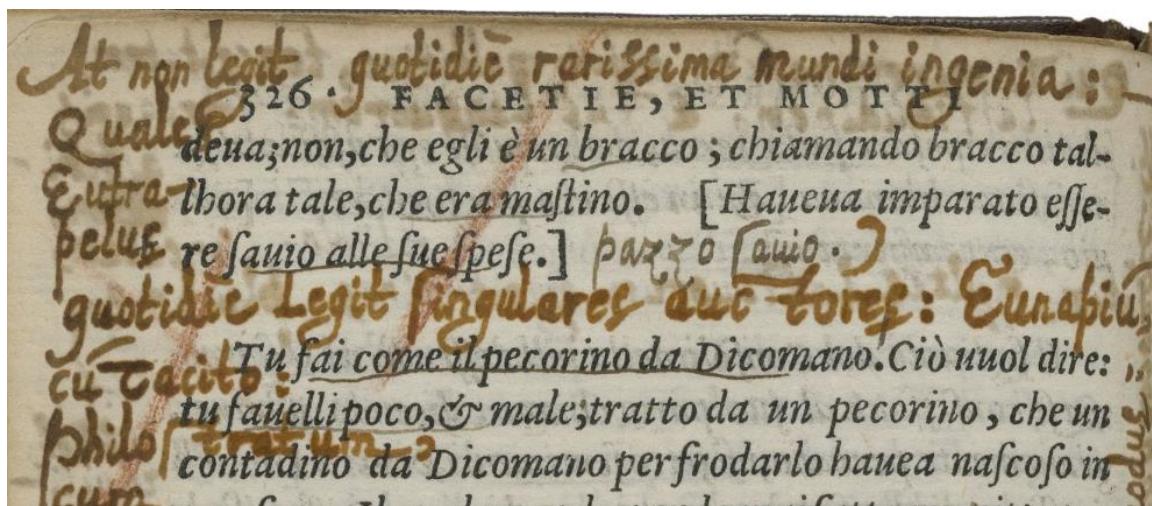
(Source: AOR, Domenichi, *Facetie, motti, et burli*, 4r)

Probably Harvey first wrote ‘Goffredo di Tarso.’ and later wrote ‘ ) Quoties friget ipse enthusiasticus Homerus: et ipse diuinus Bartasius?’ This can easily be captured in two marginalia\_text elements (note: which are part of two marginalia elements, as the ‘Goffredo di Tarso’ is treated as a separate marginal comment).

<marginalia\_text> Goffredo di Tarso </marginalia\_text>

<marginalia\_text> ) Quoties friget ipse enthusiasticus Homerus: et ipse diuinus Bartasius?  
</marginalia\_text>

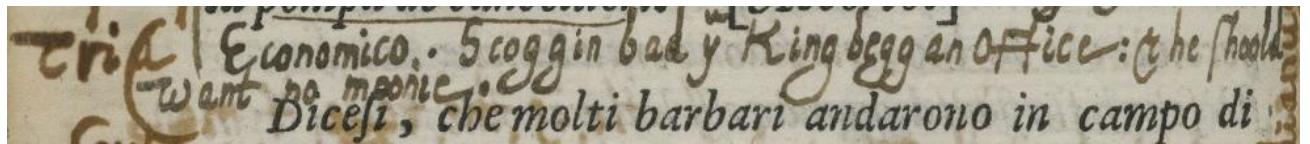
Sometimes, however, it is more difficult to apply this method. Consider this example:



(Source: AOR, Domenichi, *Facetie, motti, et burli*, 3v).

Here, the ‘)’ is written as part of the marginalia that starts in the left upper corner (‘At non legit...’) and that covers most of the page (again, note the similarities between the ink of the bracket and that of this

marginal comment). The bracket can be included in this marginalia, but in such a way that it does not split words (in this case, the most logical transcription would be: ‘...Eutrapelus ) quotidie legit...’.<sup>16</sup>



(Source: AOR, Domenichi, *Facetie, motti, et burla*, 3v).

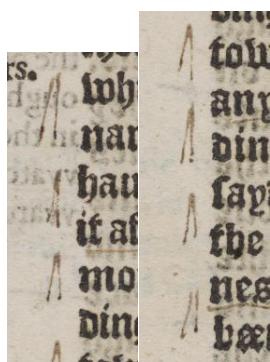
This example is even more complex. Harvey first seems to have written a ‘|’ before ‘Economico’, and although it is not entirely clear whether this vertical bar belongs to the marginalia, it is not standing next to or in the printed text (the way in which these vertical bars are used most frequently), hence we can assume that Harvey used the ‘|’ in order to demarcate this marginal note. Apparently this was not enough, for later Harvey added a ‘(‘, which splits the word ‘Tria’ in two. Transcribing this as ‘Tri(a’ would not make sense, and therefore it is best to put the ‘(‘ behind ‘Tria’ and include the ‘|’ in the marginalia starting with ‘Economico’. The XML would be:

```
<marginalia_text> | Economico. </marginalia_text>
```

```
<marginalia_text> Tria ( </marginalia_text>
```

### 1.5.2 Circumflex

Sometimes Harvey gets a bit sloppy when writing these marks, and he not always properly connects the end of the lines. Yet these marks should be tagged as circumflexes.



(Source: AOR: Machiavelli, *Art of Warre*, 105v).

<sup>16</sup> In order to improve the readability, we also could decide to put these brackets at the beginning of the text of the marginal notes. If we decide to do so, we should explain our convention on the website.

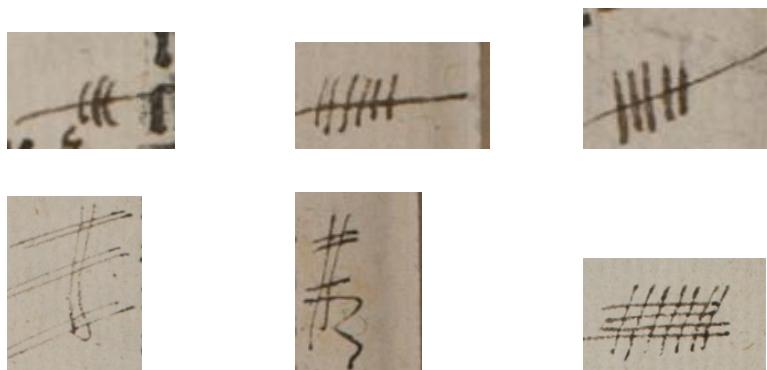
### 1.5.3 Est mark

There is a large variety of similar marks which denote ‘est’, ‘hoc est’, or ‘id est’. So far we have only encountered one mark Harvey used to denote this, but in his *Dizionario di Abbreviature latine ed italiane* Capelli lists a number of other varieties of this mark.<sup>17</sup>

—	?	7,3,3,β
ꝑ	ꝑote	
ꝑ	ꝑ	
ꝑ	ꝑꝑ	
ꝑ	ꝑ	
ꝑꝑ	ꝑꝑ	
ꝑ	ꝑꝑꝑ	

### 1.5.4 Hash

There are many varieties of the hash tag but we all tag these instances with the hash tag, rather than devising separate tags, which only would lead to a proliferation of tags with little hermeneutical value (instead, it would only confuse the transcribers as well as the end users).



(Source: AOR, Thomas Hobby (transl.), *The book of the Courtier*, passim).

Harvey also seems to have created a mark which hovers between a circumflex and a hash:

---

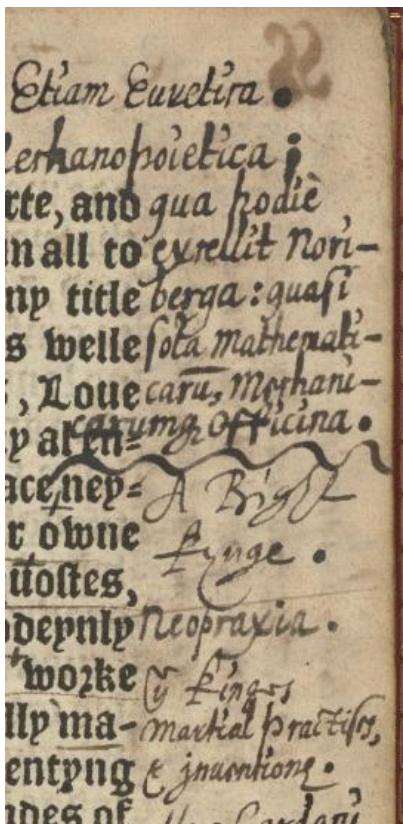
<sup>17</sup> <http://www.hist.msu.ru/Departments/Medieval/Cappelli/CPLLI406.HTM>;  
<http://www.hist.msu.ru/Departments/Medieval/Cappelli/CPLLI408.HTM>



It is possible that Harvey first made a ‘regular’ circumflex and that he later, during another reading of the book (or page), tallied the mark, perhaps to register the fact that he copied something (likely the printed text standing next to the mark) into a commonplace or paper book. It is also possible that Harvey was sloppy or hasty when writing down a hash. Therefore, it’s probably best to look at the surrounding marks on the page: if there are other hash marks, this particular mark could be best tagged as a hash mark, whereas if this mark is part of a sequence of circumflexes, it should be tagged as a circumflex.

### 1.5.5 Horizontal bar

The horizontal bar used bar Harvey have different shapes – they can be straight over curved, as the following example shows:

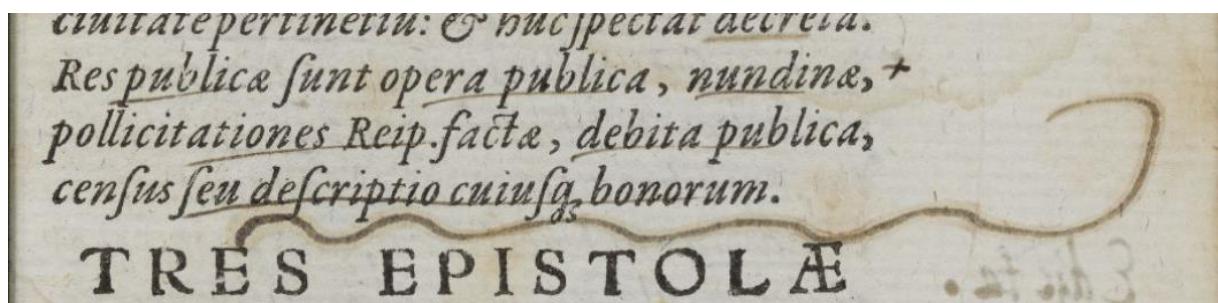


(Source: AOR, Frontinus, *Stratagems*, 2r).

The curved horizontal bar underneath ‘mechanicarumq[ue] officina’ and the straight horizontal bar underneath ‘kynge’ fulfil the same function, namely to distinguish between the various marginal notes. Hence we do not distinguish between the various shapes of a horizontal bar, but just tag them by using the horizontal\_bar tag.

### 1.5.6 Underlining vs horizontal bars

Because Harvey underlined words in the printed text and in his own marginal notes, it can sometimes be tricky to decide whether to use the underline tag or the horizontal\_bar mark. In the image below, Harvey underlined some of the printed text, but there is also a curved line beneath it. In this case, the curved line should be tagged as a horizontal bar, since Harvey occasionally included such a bar to signal the end of a book or chapter (sometimes he used a page break, see the images in section 1.5.7.).



(Source: AOR, Freigius, *Paratitla*, 8r).

```
<underline method="pen" type="straight" language="LA" text="seu descriptio cuiusq[ue]  
bonorum"/>
```

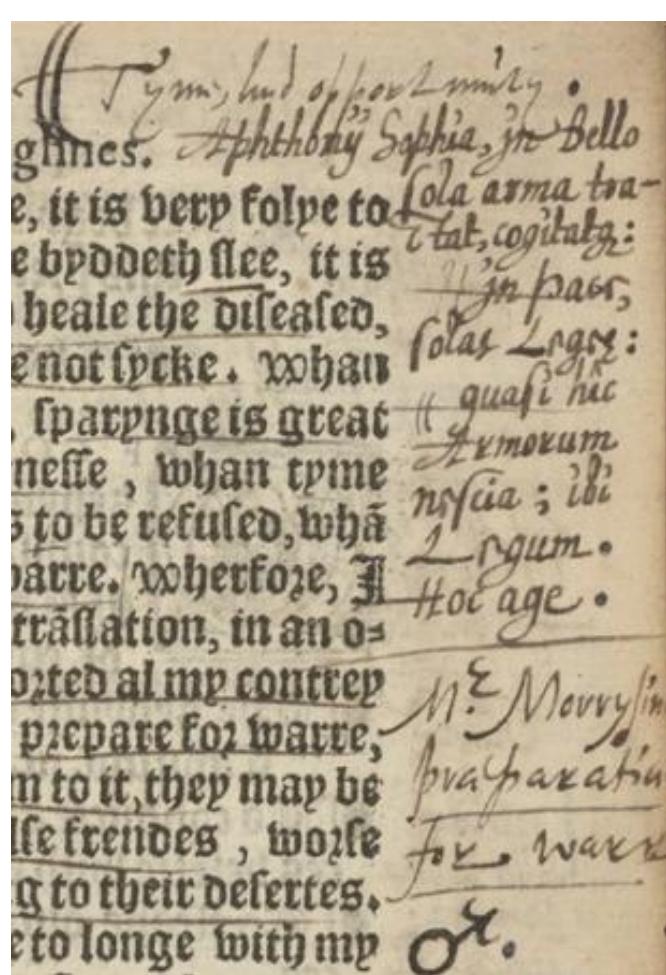
```
<mark name="horizontal_bar" method="pen" place="intext" />
```

In the image below, Harvey used a horizontal bar (below ‘Hoc age’) to distinguish that marginal note from the one below, in which Harvey highlighted some words by underlining them (captured by the emphasis tag, since he underlines text in his own marginal notes). The correct way to capture this in XML thus would be:

```
<marginalia hand="Italian">  
  <language ident="LA">  
    <position place="right_margin" book_orientation="0">  
      <marginalia_text>
```

Mr. Morrys in[s] praeparatio[n] for warr  
 </marginalia\_text>  
 <emphasis method="pen" type="straight" text="praeparatio[n] for warr"/>  
 </position>  
 </language>  
 </marginalia>

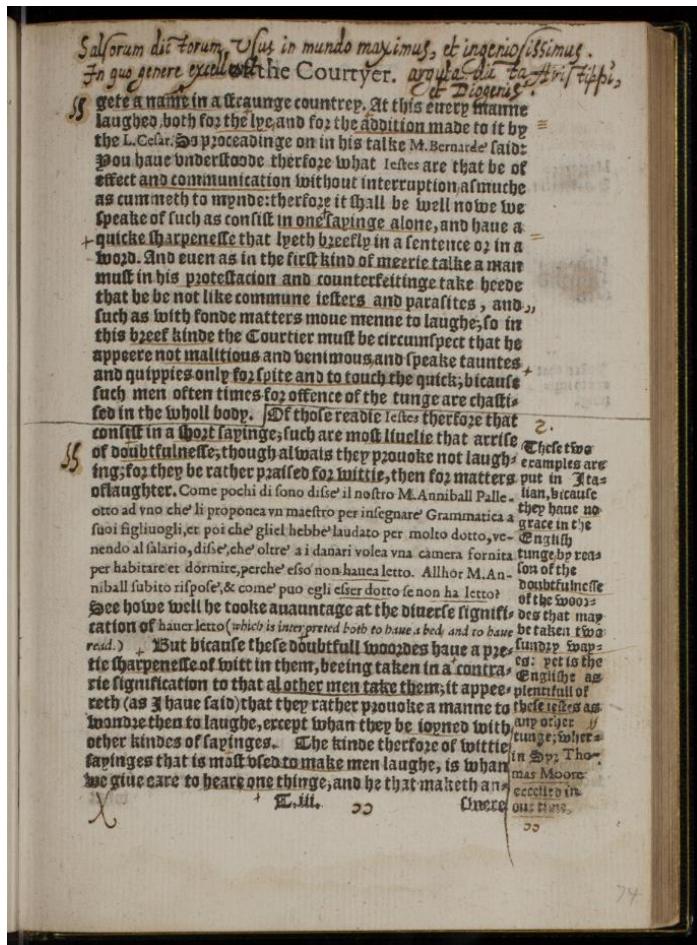
<mark name="horizontal\_bar" method="pen" place="right\_margin" />



(Source: AOR, Frontinus, *Stratagems*, 5r).

### 1.5.7 Page break

Sometimes Harvey made use of horizontal lines to divide the printed text into two distinct parts:

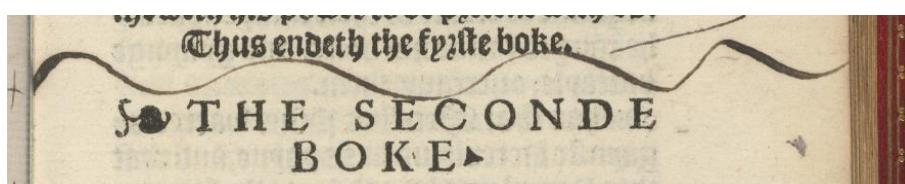


(Source: AOR, Thomas Hobby (transl.), *The book of the Courtier*, 85r).

Although this mark runs across three different spaces on the page (namely left margin, intext, and right margin – see 1.2.3), it clearly was an intervention in the printed text, hence when using this mark the place attribute should be ‘intext’.

<mark name="page\_break" method="pen" place="intext"/>

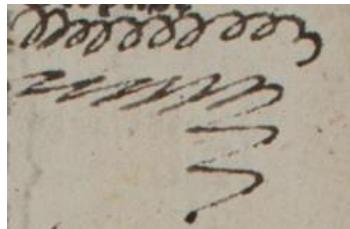
There are curved page breaks as well, but since their function is similar, we tag them as page breaks (without making a distinction between straight and curved lines).



(Source: AOR, Frontinus, *Stratagems*, 035r).

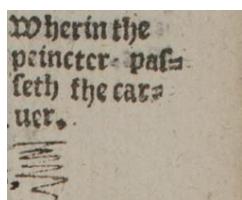
### 1.5.8 Pen trial and scribble

These two marks can be seen as leftover categories, tags which can be used to capture Harvey's interventions which are difficult to categorize, partly because they do not have a recurrent shape and usage. This is especially the case with pen trials:



(Source: AOR, Thomas Hobby (transl.), *The book of the Courtier*, 52r).

Some scribbles do occur more often and might have had a function, such as:



(Source: AOR, Thomas Hobby (transl.), *The book of the Courtier*, 47r).



(Source: AOR, Freigius, *Paratitla*, 15r).

Other examples of scribbles:

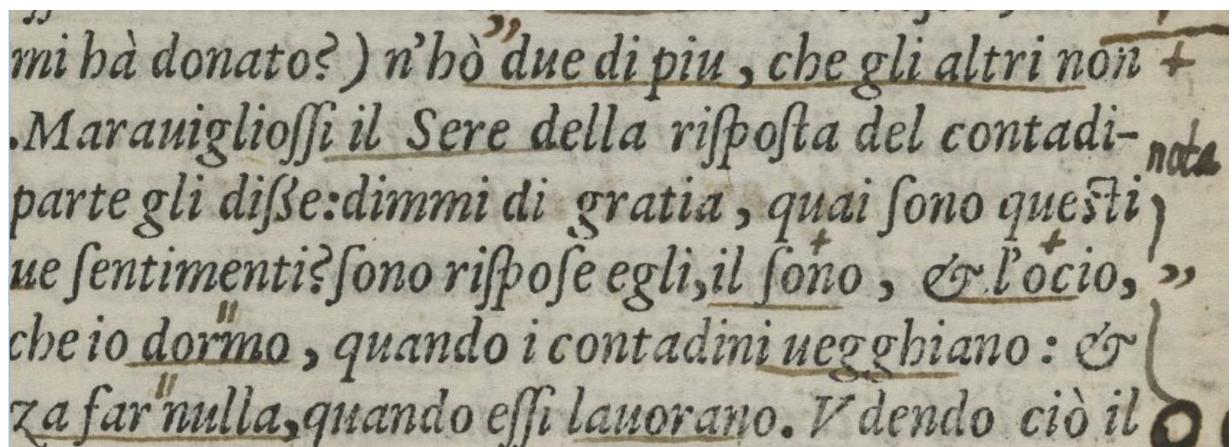


(Source: AOR, Thomas Hobby (transl.), *The book of the Courtier*, 194v).

It is left to the transcriber to decide whether a mark should be tagged as a scribble or a pen trial. As a guideline, when the transcriber feels that a mark was purely the result of a reader testing his or her pen, the mark should be tagged as a pen trial.

### 1.5.9 Straight quotation marks vs quotation marks

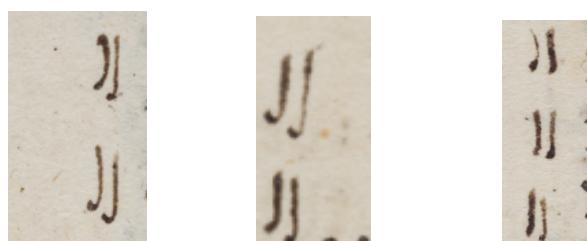
Harvey made use of two types of quotations marks, namely the straight quotation marks  and (curved) quotation marks . The use of these marks is not entirely clear, although it is likely that Harvey used quotation marks to highlight text or phrases he wanted to copy into one of his notebooks, the straight quotations marks he probably used as reading marks – these marks often appear in the Italian books he owned. Since these marks are clearly distinguishable, as the picture below shows, both of the marks have a unique name, enabling users to search for a specific variant of the mark.



(Source: AOR, Domenichi, *Facetie, motti, et burli*, 4r).

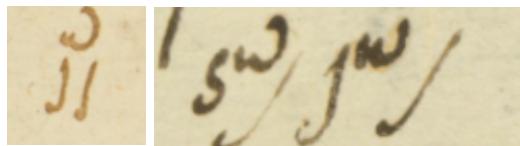
### 1.5.10 Unknown marks

In his copy of Machiavelli's *Art of War*, Harvey sometimes uses a mark which has the shape of a double vertical bar, yet its endings are slightly bended, creating the impression of two 'I's.



(Sources: AOR, Machiavelli, *Art of warre*, 14v; 15r; 41v).

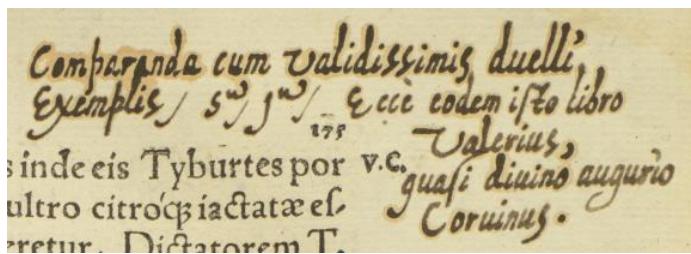
Compare the shape of this mark with the letter or line that follows Harvey's abbreviation of the word 'infra', or the letter/line after infra and between supra and infra (in the second image):



(Source: AOR: Livy, *Romanae historiae principis*, 8r; 62r).

These images suggest that, rather than being a letter, it was a line, some sort of slash, which divided the two abbreviations or had the function to distinguish between the marginal note and the printed text. It even could have been some sort of extended dot (since Harvey always ends his marginal notes with a dot) or a comma (when standing in between supra and infra, as above and below). The marginal note on the second image therefore is transcribed as follows:

<marginalia\_text>s[upr]a, i[nfr]a. </marginalia\_text>



(Source: AOR: Livy, *Romanae historiae principis*, 96r).

However, the mark also seems to appear in what seems to be a succession of SS symbols, as is visible on the following images:



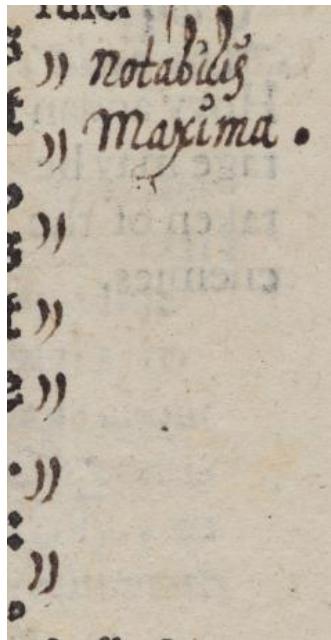
(Sources: AOR: Livy, *Romanae historiae principis*, 94r; 107v; 232v;

Thomas Hobby (transl.), *The book of the Courtier*, 179v; 180v; 181c).<sup>18</sup>

<sup>18</sup> For other clear examples, see Livy, 231v-232r.

In these examples, the marks seem to be somewhat sloppy variants of the SS symbol – as if Harvey wrote them hastily - and they can be tagged as such. Possibly these marks in the examples derived from Harvey's annotations in Machiavelli's *The Art of War* are also variants of the SS symbol.

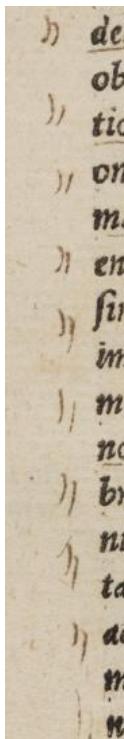
However, in other cases this mark seems more closely related to a quotation mark:



(Source: AOR: Machiavelli, *Art of warre*, 66r).

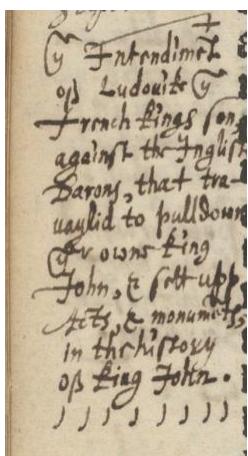
In combination with the marginal note 'Notabilis Maxima' (most notable), these marks seems to highlight a passage in the printed text. Although Harvey uses symbols to do this as well, as we saw earlier, because there is not a 'clear' SS symbol written on this page and because these marks resemble a quotation mark more closely, they should be tagged as such.

The same applies to the marks in the Melanchton:



(Source: AOR, Melanchton, *Selectarum*, 340v).

It seems, however, that this ‘unknown mark’, which, as it turns out, can be known in some occasions, was not only used in pairs:

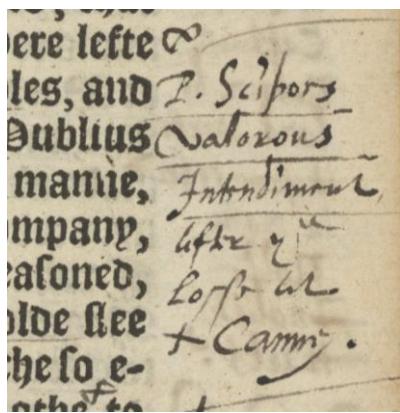


(Source: AOR: Frontinus, *Stratagems*, 98v).

In this particular example, it is not entirely clear how to tag the mark, and in case of persistent uncertainty about particular instances of this annotation, just put in the comment field in the XML files: <!--Unknown mark--> so that we can get back to them at a later point.

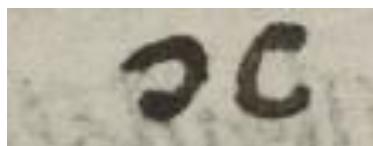
There are some other examples of unknown marks (and possibly symbols):

The mark above the marginal note and above a word in the printed text:



(Source: AOR, Frontinus, *Stratagems*, 108r).

The ')(‘ mark or symbol (appears in various books):

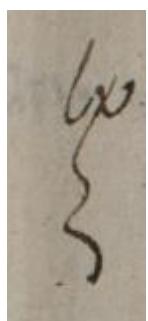


(Source: AOR: Domenichi, *Facetie, motti, et burli*, 129v).

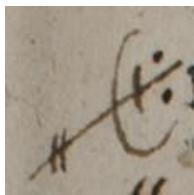
Several others:



(Source: AOR, Thomas Hobby (transl.), *The book of the Courtier*, 17r).



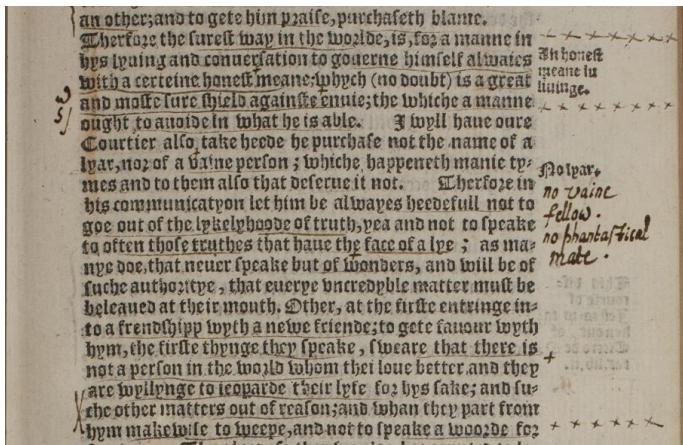
(Source: AOR, Thomas Hobby (transl.), *The book of the Courtier*, 52r).



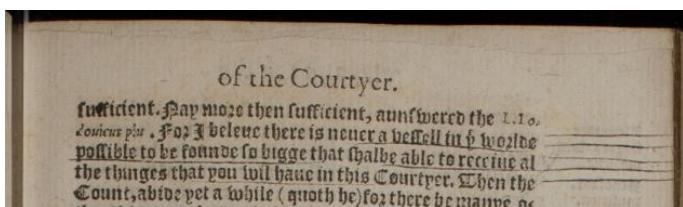
(Source: AOR, Thomas Hobby (transl.), *The book of the Courtier*, 97r).

### 1.5.11 Rows of marks

Harvey sometimes put a number of the same mark next to or below each other:



(Source: AOR, Thomas Hobby (transl.), *The book of the Courtier*, 76r).



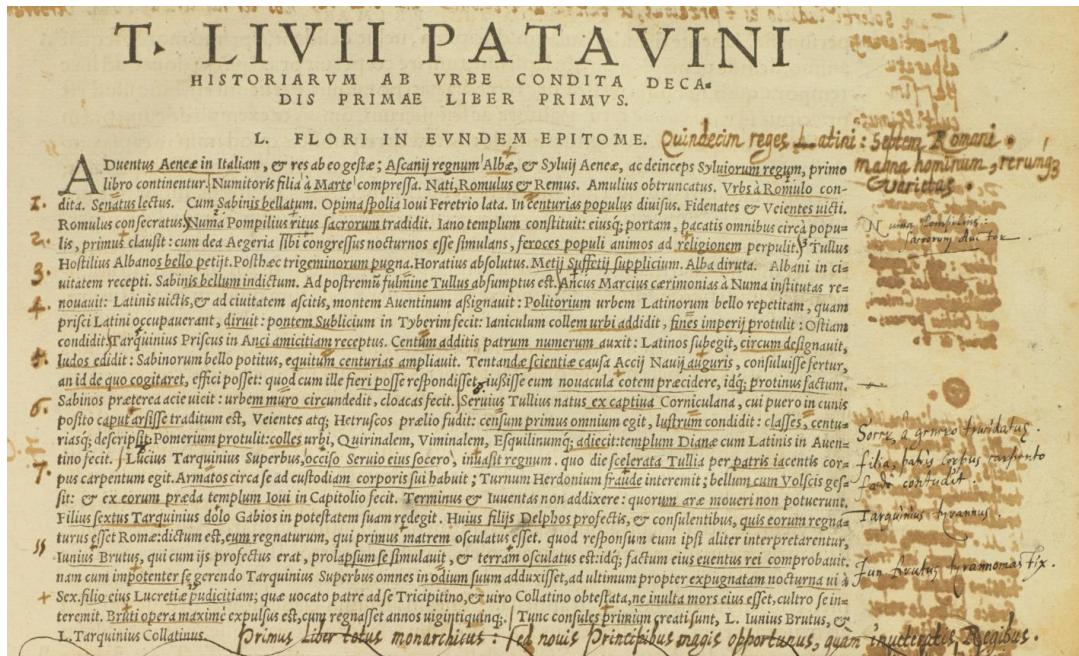
(Source: AOR, Thomas Hobby (transl.), *The book of the Courtier*, 44r).

These and other similar forms of annotations are not going to be captured by separate tags. Instead, we will use the existing tags to tag all the marks individually (in case of the two examples above, using the marks X\_sign and horizontal\_bar). It nevertheless is a possibility that, say, one cross in a margin had a different function than several crosses in the margin. Therefore, we should give the user the opportunity not only to search for pages where a particular mark (or a number of this mark) appears, but also to search for a sequence of marks (e.g. give the pages of book X in which five horizontal bars are tagged subsequently). This will allow the user to find passages of the printed text that are marked by a sequence of marks (for another example, see the images in section 1.5.2).

## 1.6. Numerals

Numbers that are part of a marginal note, for instance when Harvey is referring to a book chapter of mentions a date, are captured in the text of the marginal note. However, when Harvey only writes down a number in order to mark up (part) of the printed text without any additional verbal annotation (i.e. his own manuscript annotation), this is captured by the numeral element.

For example:



(Sources: AOR: Livy, *Romanae historiae principis*, 9r).

```

<numeral place="left_margin">1.</numeral>
<numeral place="left_margin">2.</numeral>
<numeral place="left_margin">3.</numeral>
<numeral place="left_margin">4.</numeral>
<numeral place="left_margin">4.</numeral>
<numeral place="left_margin">6.</numeral>
<numeral place="left_margin">7.</numeral>

```

In case a number inserted by Harvey can be linked to one or more words in the printed text or in a marginal note, this can be captured by using the text attribute and the language attribute (specifying the language of the printed text/marginal note).

## 1.7 Drawings

Although so far it appears that Harvey's annotations do not include many drawings, we have encountered a couple:



(Source: AOR: Thomas Hobby (transl.), *The book of the Courtier*, 194r; 190r).

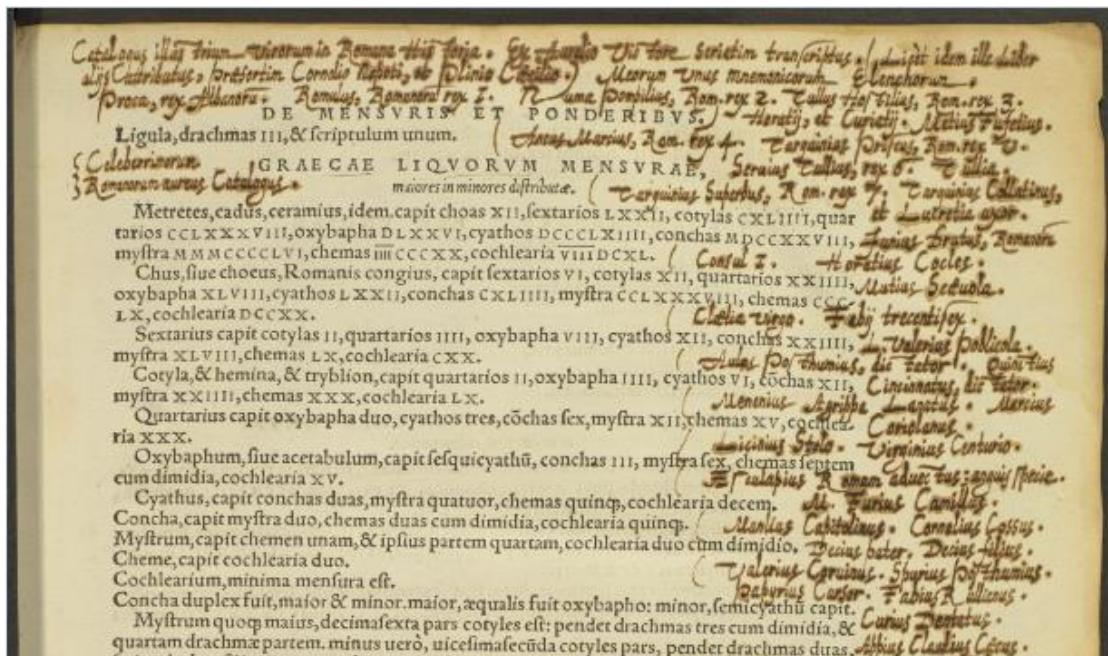
These and possible other drawing are captured by the drawing element, which has a name attribute (a predefined list of various types of drawing), method and place attributes, and text and language attributes (in case the drawing can be lined to printed text, the language attribute specifying the language of the printed text). For example:

```
<drawing name="manicule" method="pen" place="left_margin"/>
```

## 1.8 Linking reader's interventions to the printed text I: marginalia

As mentioned, the various elements which capture the interventions Harvey made in his books (the marginalia, symbol, mark, and underline elements) all have a text – or in case of marginalia anchor\_text – attribute, by which the annotations can be linked to the printed text. These links, however, are not always obvious or straight-forward; on the contrary, establishing the relationship between the printed text and the annotations often depends on the interpretation of the transcriber. Before giving some observations which might help transcribers in establishing such links, it is important to stress that the text attributes of the various elements are optional: the transcriber does not have to establish such relationships, and in case it is unclear whether there is a relationship between printed text and annotation, it is better not to use the text attribute rather than to invent a relationship that might not have existed. Relating the interventions to the printed text primarily is a user's activity.

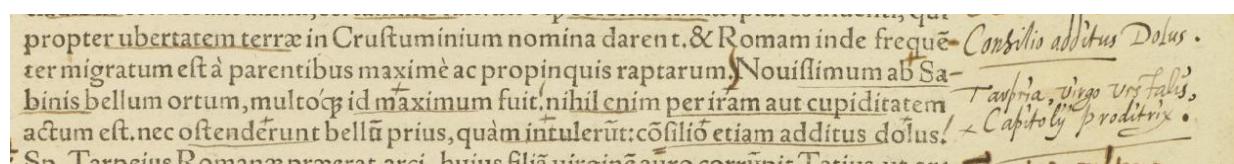
Marginal notes, for example, do not necessarily have to refer to the printed text next to which they are written. The clearest example is this page of Livy's History of Rome:



(Source: AOR: Livy, *Romanae historiae principis*, 4r).

Whereas the printed text list various measurements used in Ancient Greece, Harvey compiled a ‘catalogue of famous men in Roman history’, which was ‘one of my lists for memory’. In this case, there is no direct relationship between the printed text and Harvey’s annotation, and the anchor\_text attribute of the marginalia element should therefore not be used.

Sometimes, there were links between marginal notes and the printed text, for instance when Harvey summarized a passage from the printed text or when he copied part of the printed text:



(Source: Princeton, Livy, *Romanae historiae principis*, 12r).

One of Harvey’s annotations in the right margin is the phrase ‘Consilio additus Dolus’, a partial copy of the printed text (which Harvey underlined) ‘co[n]ſilio etiam additus dolus’. In this case, the link between marginal note and printed text is clear, and would be captured as follows:

<marginalia hand="English secretary" anchor\_text="co[n]ſilio etiam additus Dolus">

<language ident="LA">

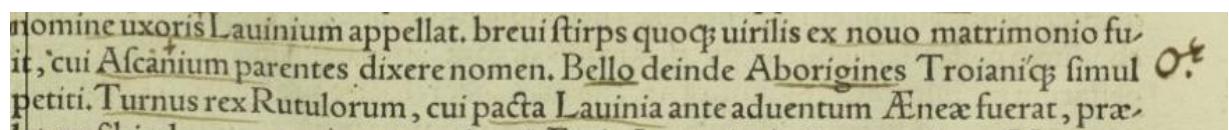
<position place="right\_margin" book\_orientation="0">

<marginalia\_text> Consilio additus Dolus </marginalia\_text>  
 </position>  
 </language>  
 <translation>Deliberation coupled with deceit<translation/>  
 </marginalia>

In this example, Harvey did not only copy the words of the printed text, but he also underlined these words and put a couple of marks (plus signs) above them, other indications of Harvey's interest in this passage. Transcribers thus should look at the content of the marginal note, and see whether this content is mirrored in the printed text, and whether other interventions made by Harvey in the printed text point towards the specific passage. Again, these observations may help the transcribers to establish a link, but due to the sheer variety of interventions and marginalia, no formal rules are created to discern such links. Ultimately, this is left to the interpretation of the transcriber.

## 1.9 Linking reader's interventions to the printed text II: symbols

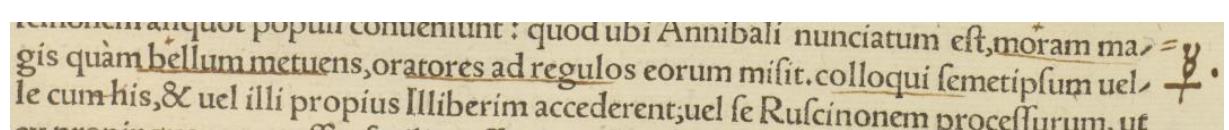
Whereas marginal notes could serve various purposes, some of them functioning as mnemonic or heuristic tools, symbols were often used to index a certain passage of the printed text; a symbol was a concise summary of the printed text which enabled Harvey to quickly find a passage about a certain topic in his books. Because Harvey used his symbols in this way, it can be easier to establish a link between the symbol and the printed text.



(Source: AOR: Livy, *Romanae historiae principis*, 10r).

Harvey uses the Mars symbol to signify war or warfare, and in this case the Mars symbol can be linked to the word *bello* (which is also underlined). <symbol name="Mars" place="right\_margin" text="bello" />. Often it is sufficient to only capture a keyword ('*bello*' in this case), for it is up to users to research the exact topics (or historical events) Harvey was interested in when using the Mars or other symbols.

Another example:

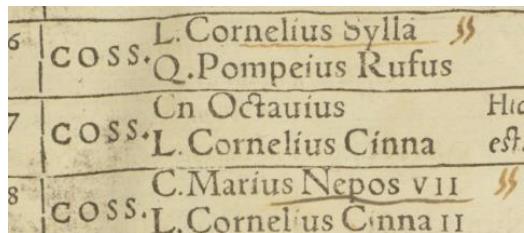


(Source: AOR: Livy, *Romanae historiae principis*, 147r).

Since Harvey often used the Mercury symbol to denote eloquence (and sometimes trickery), one can capture the link to the printed text by tagging various keywords, separated by a semicolon:

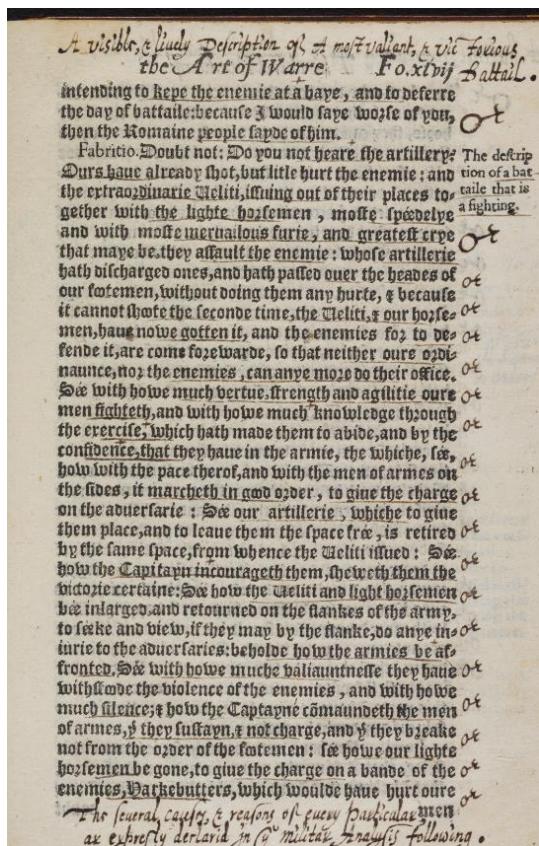
```
<symbol name="Mercury" place="right_margin" language="LA" text="oratores; colloqui"/>
```

In some cases, the link between a symbol and the (underlined) printed text is even clearer (since the symbol is not standing next to a passage of text, but above or next to a couple of words):



(Source: AOR: Livy, *Romanae historiae principis*, 533v).

However, Harvey also could have used a symbol to highlight the content of an entire page (symbols that were written in the left upper corner of a page, for instance), and in this case the link with a specific passage is much harder to discern. Because through the tagging of symbols, we already offer the end user a way to search through Harvey's annotations, we should not worry too much to establish a link between symbols and printed text, and if such links are not obvious or clear, the text attribute is best left empty. Moreover, sometimes Harvey used the same symbol in succession to mark up a large passage (see the image below), and in such cases using the text attribute is not necessary: the fact that a symbol appears several times on one page is a sufficient indication for the end user that Harvey was extremely interested in the printed text:



(Source: AOR, Machiavelli, *The Art of Warre*, 51r).

### 1.10 Linking reader's interventions to the printed text III: marks

Marks can appear anywhere on the page, but quite often Harvey puts marks such as plus signs, equal signs and quotation marks in the printed text. Tagging these marks is fairly straight-forward; more difficult is when to link these marks to one or more words in the printed text. One thing to remember is that the text attribute of the mark element is optional; a transcriber does not have to link the mark to the printed text. On the contrary, if the transcriber is not sure about the existence of a link between a reader's intervention and the printed text, it is better not to use the text attribute. There are, however, some signs which suggest a link between a mark and one or more words. For instance, if a mark is directly above a word, and this word is underlined as well, this suggests that Harvey had a clear interest in this particular word.

*at, docebatq; patrum id superbia factum, qui cōnubium finitimis negassent, illas ta-*

(Source: AOR: Livy, *Romanae historiae principis*, scan 11v).

Sometimes, more than one words were underlined, and the meaning or function of the words ‘betray’ whether one or more words should be included in the text attribute.

è id Romana pubes passa, & haud dubiè ad vim spectare rescoepit, cui tēpus

(Source: AOR: Livy, *Romanae historiae principis*, scan 11v).

It is unlikely that Harvey thought the preposition ‘ad’ was of special interest to Harvey, so in this case a transcriber can put ‘ad vim’ (‘towards’ or ‘by strength’) in the text attribute.

uictores pelli à portis urbis? Reclamantibus uniuersis, primus ipse scuto super caput elato pergit ad portā; seuti alij testudine facta in urbēm perrumpunt: deturbatisq; Samnitib. quæ circa portam erant muri occupauere: penetrare interiora urbis, quia pauci admodum erant, non audent. Hæc priuò ignorare cōsul, & intētus recipiēdo exercitui esse. iam enim præceps in occasum sol erat: & appetens nox periculosa & lūspecta omnia etiā uictorib. faciebat. progressus lōgius ab dextra, cāstra capta uidet: ab læua clamorē in ūrbe, mistum pugnantium ac pauentium fremitu esse. & tum forte certamen ad portam erat. aduectus deinde equo proprius, ut suos in muris uidet, nec iam integri quicquam esse; quoniam temeritate paucorum magnē rei parta occasio esset: acciri quas receperit copias, signaç in ūrbem inferri iussit. ingressi proxima ex parte, quia nox appropinquabat, quieuere. nocte oppidum ab hostibus desertum est. Cæsa illo die ad Aquiloniam Samnitium millia triginta trecenti quadraginta: capta tria millia octingēti & septuaginta: signa militaria nonaginta septem. Cæterū illud memoriæ traditur, non fermè alium ducē lāetiore in acie uisum: seu suopte ingenio, seu fiducia bene gerundæ rei. Ab eodem roboře animi, necq; cōtrouerso auspicio, reuocari à prælio potuit: & in ipso discrimine, quo templa dñs immortalibus uoueri

(Source: AOR: Livy, *Romanae historiae principis*, 140v).

When Harvey repeatedly put marks above the same word (see the previous image, above various declensions of the word *urbs*), this can be a signifier for the special interest he had in a word, rather than in the couple of words he underlined. In these cases the transcriber can link the mark to that specific word.

ogn' uno famia più, ch'io non ami. Piacque molto questo giuoco,  
et già ogn' uno si preparaua di parlar sopra tal mater-

(Source: AOR: Castiglione, *Il Cortegiano*, 20v)

In this example, the link between mark, marginal note and printed text is crystal clear. Harvey used what we have christened as the ‘est\_mark’ to link a word in the printed text to the marginal note. The text of the marginal note can be put in the text attribute of the mark in the left margin, whereas the printed word can be included in the text attribute of the mark in the text.

Sometimes, though, the use of marks and their relation to the text is a bit less clear. For instance, it occurred that Harvey wrote a mark in between two words, reducing the certainty with which we can establish a link between the mark and the printed text. Based on the meaning of the text and the use of other reader’s interventions, the transcriber should decide whether or not to use the text attribute and which words to include in this attribute.

che se io proprio l'hauessi fatto.

(Source: AOR, Domenichi, *Facetie, motti, et burli*, 1r).

siete grauida, soggiunse il Vinitiano; oh costei cominciò Galeazzo, il quale mentì

(Source: AOR, Domenichi, *Facetie, motti, et burli*, 2v).

Link between marks that appear in the margins and the printed text are even more difficult to establish, so often it is best leave out the text attribute when dealing with these marks.

set; pergit ad proximam speluncam, si forte eō uestigia ferrent. quæ, ubi omnia foras  
uerfa uidit, nec in partem aliam ferre: confusus atq; incertus animi ex loco infesto age  
re porro armentum occœpit. Inde cum actæ boues quædā ad desiderium (ut fit) re  
lictarum mugissent; redditæ inclusarum ex spelunca bouum uox Herculem conuertit.  
quem cum uidentem ad speluncam Cacus ui prohibere conatus esset: iactus claua, fi  
dem pastorum ne quicquam inuocans, morte occubuit. Euāder tum ea, profugus ex  
Peloponneso, autoritate magis quam imperio regebat loca: uenerabilis uir miraculo  
literarum, rei nouæ inter rudes artium homines; uenerabilior diuinitate credita Car  
mentæ matris, quam fatiloquam ante Sibyllæ in Italiam aduentum, miratæ hæ gëtes

(Source: AOR: Livy, *Romanae historiae principis*, 11r).

Especially when dealing with pages that are heavily annotated, it is very difficult to attach marks that are written in the margins to (parts of) the printed text, since often these marks are accompanied with marks in the text, as the image above shows. However, in the case of less-densely annotated pages, it might very well be possible to attach a mark in the margin to the printed text.

gli parea: ma tanto era la reverenza, che si portava al  
uoler della Sig. Duchessa, che la medesima libertà era  
grandissimo freno: ne era alcuno che non estimasse per  
lo maggior piacere, che al mondo hauer potesse, il com-

(Source: AOR: Castiglione, *Il cortegiano*, 16r)

In this case, it is fairly easy to put the underlined text in the attribute of the mark. In the following two examples, using the text attribute is already a bit trickier, since it is less clear than in the previous example above which text to include in the text attribute (due to the occurrence of more underlining as well as the proximity of other marks).

nia. Iudicij publici processu continentur Ci  
tatio seu requisitio, questio, pæna, restitu  
tio, cadauera. Atq; hæc de prima instatia;  
sequitur secunda seu appellatio.

(Source: AOR, Freigius, *Paratitla*, 7v).

Behald, the stryng sodenly brake , and  
downe with a great noyse fell Dame  
Kerese, a woman very heauy baith by  
vnweldy age & massy substance . But  
+ sche an auld beaten soldiar , nothing  
+ dismayed with the darkenes of the  
night, the heighth of the wall, nor with  
the sodennesse of the fall , vp sche get-  
teth, & winneth into Bothwels cham-  
ber, sche gyt the dore open , and out of

(Source: AOR: Buchanan, *De Maria scotorum*, 5v).

+ uolse, et disse, S. Vnico , non è alcun di noi qui , che non  
ui ceda in ogni cosa , ma molto più nel conoscer l'animo  
della S. Duchessa, et così come più che gli altri lo cono-  
sciete per l'ingegno uostro diuino , l'amate anchor più

(Source, AOR: Castiglione, *Il Cortegiano*, 9v).

Besides capturing the underlining with the underline tag, the combination of plus signs in the image above and underlining can be captured as follows:

```
<mark name="plus_sign" method="pen" place="left_margin" language="IT" text="ui ceda in ogni cosa"/>  
<mark name="plus_sign" method="pen" place="intext" language="IT" text="l'ingegno uostro diuino"/>
```

deprive vs therot. They aledge the Romaynes, who by  
meane of their owne powers, lost their libertie. They  
aledge the Venicians, and the Frenche king; which Ve-

(Source: AOR: Machiavelli, *The Arte of Warre*, 17r).

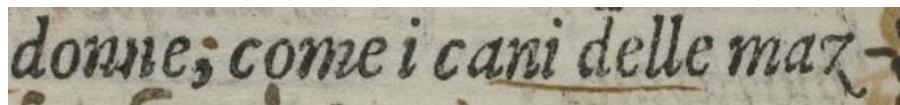
This image from Harvey's copy of Machiavelli's book shows that sometimes there is a clear link between a mark (a plus sign, in this case) and a single word, which should be tagged as follows:

```
<mark name="plus_sign" method="pen" place="intext" language="EN" text="their owne powers"/>  
<mark name="plus_sign" method="pen" place="intext" language="EN" text="Venicians"/>  
<mark name="plus_sign" method="pen" place="intext" language="EN" text="Frenche"/>
```

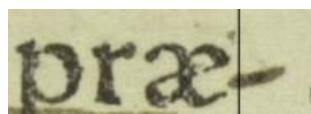
Because of the large variety of combinations possible, it is undoable and undesirable to formulate strict rules which should be applied to the text. As suggested, other reader's interventions, as well as the meanings of the word(s) and the syntactic function can be of help when determining the link between a mark and text. However, the text attribute does not have to be used, and in case of being uncertain about the link between marks and words, it is better not to use it. In the end, it is important to remember that the primary goal of the project is to capture all the reader's interventions and not the relationships between these interventions and the printed text. Put differently: do not spend too much time in trying to establish these links.

## 1.11 Changing the punctuation

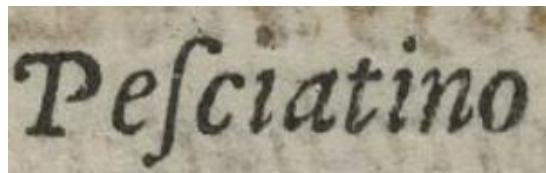
Harvey could be pretty pedantic when annotating his books, and one of the things he did was changing the punctuation, for instance by turning a comma into a semicolon, adding dashes to indicate word breaks, or even putting dots above i's if the printer had failed to do so.



(Source: AOR, Domenichi, *Facetie, motti, et burli*, 1v)



(Source: AOR: Livy, *Romanae historiae principis*, 12r).

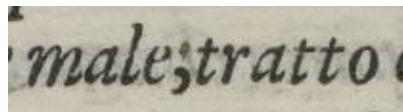


(Source: AOR, Domenichi, *Facetie, motti, et burli*, 1v).

Although the use of marks to change (and, according to Harvey to improve) the punctuation is clear, we are not recording why Harvey used certain mark, but just the fact that he used them. Therefore, if Harvey changed a comma into a semicolon, we capture this as follows.

```
<mark name="dot" method="pen" place="intext" />
```

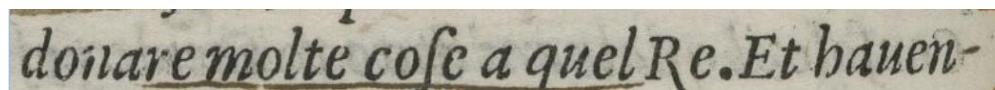
The text attribute should be empty, for the use of this mark cannot be linked to a specific word (or words) in the printed text. However, it is helpful if the transcriber can include a comment where in the text to find such an intervention, for they can be difficult to spot for the person checking the transcriptions. This can be done in such a way:



(Source: AOR: Domenichi, *Facetie, motti, et burli*, 3v)

```
<mark name="dot" method="pen" place="intext"/><!-- male, tratto -->
```

In a similar fashion a dash (used as a word break) can be indicated.

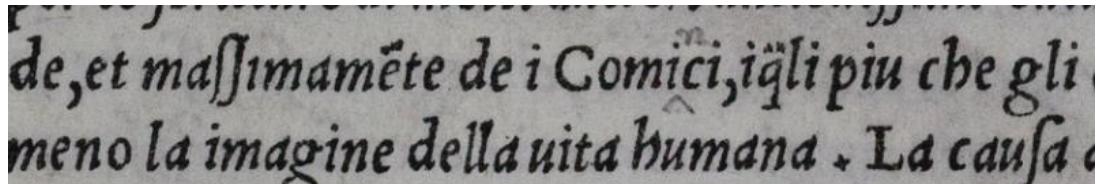


(Source: AOR: Domenichi, *Facetie, motti, et burli*, 3v)

```
<mark name="dash" method="pen" place="right_margin"/><!-- hauen -->
```

## 1.12 Changing the spelling: the errata tag

The errata tag captures the instances in which the readers makes amendments to the printed text, which include corrections but also instances in which the readers aims to preserve the text. In case of Harvey, his pedantry also included changing the spelling of a word in the printed text:

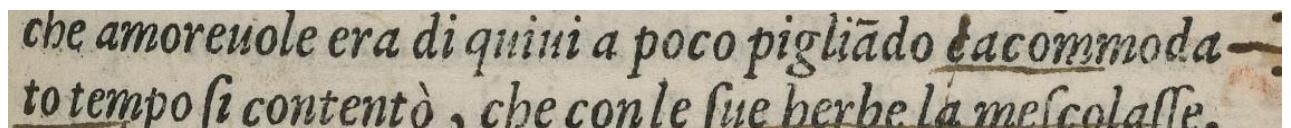


(Source: AOR: Castiglione, *Il Cortegiane*, 56r)

This should be captured by the errata tag. The errata element consists of two attributes, namely *copytext* and *amendedtext*; the former captures the printed text whereas the latter consists of the word as corrected by Harvey.

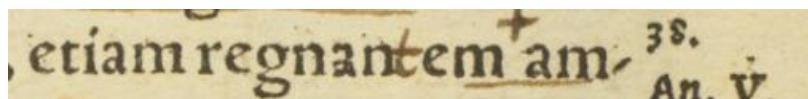
```
<errata copytext="Comici" amendedtext="Cominici"/>
```

Two other examples:



(Source: AOR: Domenichi, *Facetie, motti, et burli*, 36r)

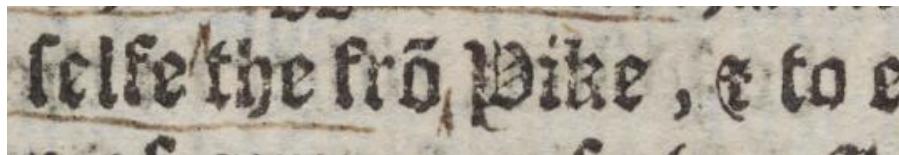
```
<errata copytext="caccommodato" amendedtext="acommodato"/>
```



(Source: AOR: Livy, *Romanae historiae principis*, 18r)

```
<errata copytext="regnan em" amendedtext="regnantem"/>
```

The errata tag can also be used to capture the instances in which Harvey corrected the word order of the printed text.



(Source: AOR: Machiavelli, *The Arte of Warre*, 28r).

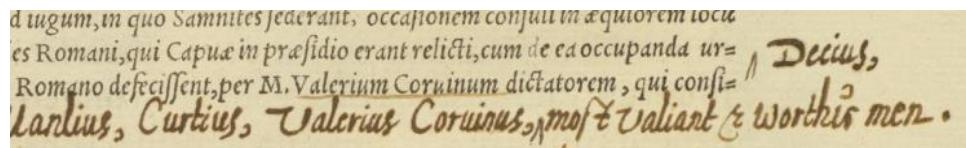
Harvey uses two circumflexes to indicate that the word order should be changed. The circumflexes can be tagged individually, and the errata tag can capture the changes Harvey made:

```
<errata copytext="the fro[m]" amendedtext="fro[m] the"/>
```

The marks are tagged as follows (the word in the text attribute indicating the word that has been ‘moved’):

```
<mark name="circumflex" method="pen" place="intext" language="EN" text="the"/>
<mark name="circumflex" method="pen" place="intext" language="LA" text="the"/>
```

Another example of this practice:



(Source: AOR: Livy, *Romanae historiae principis*, 94r).

In this case the errata tag is not used (since this tag is only used for corrections made to the printed text), but the text of the marginal note is captured as follows (following the word order Harvey intended):

```
<marginalia_text> Manlius, Curtius, Valerius Corvinus, Decius, most valiant & worthiest men.
</marginalia_text>
```

The circumflexes are tagged as well:

```
<mark name="circumflex" method="pen" place="right_margin" language="EN" text="Decius"/>
<mark name="circumflex" method="pen" place="intext" language="EN" text="Decius"/>
```

It seems that Harvey also wanted to preserve the integrity of the printed text by ‘restoring’ letters that had partly faded or that were only printed partly.

A photograph of a manuscript fragment. The text is written in a Gothic script. The first line reads 'vnstable she is: and therfore,'. The word 'she' is underlined.

A photograph of a manuscript fragment. The text is written in a Gothic script. The lines read: 'b that I desir', 'b getherto, you', and 'thoſe I hould'. The letter 'b' in the first line and 'b' in the second line are underlined.

(Source: AOR: Machiavelli, *The Arte of Warre*, 60v).

The second image in particular shows that Harvey carefully restored letters that had partly faded and inserted letters that were not printed or that had faded completely (there is no way of telling this). We use the errata tag to capture these inventions as well. The partly faded letters are included in the *copytext* attribute, whereas the letters which have completely faded (or have not been printed) are not included in this attribute. For example:

```
<errata language="LA" copytext="she" amendedtext="she"/>  
<errata language="LA" copytext="he rto" amendedtext="hetherto"/>
```

A final use of the errata tag is to capture the instances in which Harvey deleted one or more words of the printed text (i.e. we do not use a separate strikethrough tag to capture this practice).

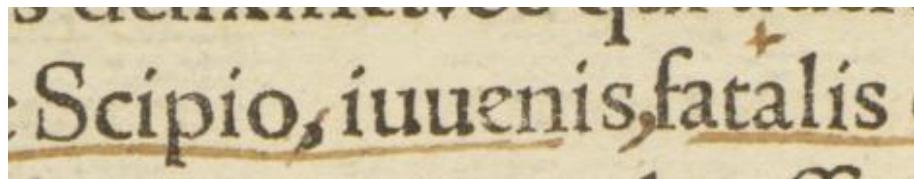
A photograph of a manuscript fragment. The text is written in a Gothic script. The line reads 'LIBER TERTIVS.'. The word 'TERTIVS.' is crossed out with a horizontal line.

(Source: AOR: Livy, *Romanae historiae principis*, 404r).

```
<errata language="LA" copytext="TERTIVS" amendedtext="[deleted]"/>
```

The deleted word is put in the *copytext* attribute, whereas the *amendedtext* attribute contains the text [deleted].

In the following example Harvey deleted the comma after Scipio (and instead inserted a comma after *iuuenis*).

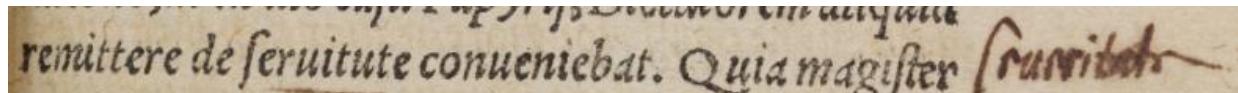


(Source: AOR: Livy, *Romanae historiae principis*, 167v).

This is captured as follows:

```
<errata language="LA" copytext="," amendedtext="[deleted]" />
```

There is one exception to the use of the errata tag, namely when Harvey writes the corrected version of a misspelled word in the margin:



(Source: AOR: Melanchton, *Selectarum*, 130r).<sup>19</sup>

We could use the errata tag to capture such a modification, but from the perspective of the end user this would be weird, for this technically is a marginal note, and the end user expects to have more information than the errata tag can provide (e.g. the language of the marginal comment, its translation). Therefore we use the marginalia rather than the errata tag to capture these specific interventions made by Harvey:

```
<marginalia hand="Italian" anchor_text="seruitute">
  <language ident="LA">
    <position place="right_margin" book_orientation="0">
      <marginalia_text>
        seueritate
      </marginalia_text>
    </position>
  </language>
</marginalia>
```

---

<sup>19</sup> For another example see the same book, 119r.

```

</marginalia_text>
</position>
</language>
<translation>
    With severity.
</translation>
</marginalia>

```

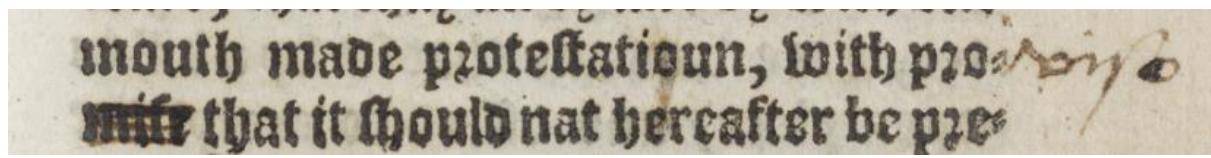
The anchor text attribute is used to link the marginal note to the word in the printed text Harvey is correction. As usual, an underline tag is used to capture the underlined word in the printed text.

For a similar example:



(Source: AOR: Buchanan, *Ane Detection*, 6v).

In this case the marginalia tag should be used, but when Harvey writes part of the word in the margin in order to correct a word in the printed text (see the example below), this should be captured by the errata tag (it wouldn't make sense to capture part of a word by the marginalia tag, the use of which in this case does not have added value).



(Source: AOR: Buchanan, *Ane Detection*, 6v).

The corresponding example for this example is:

```
<errata language="EN" copytext="promise" amendedtext="proviso"/>
```

### 1.13 Missing text & Uncertainty

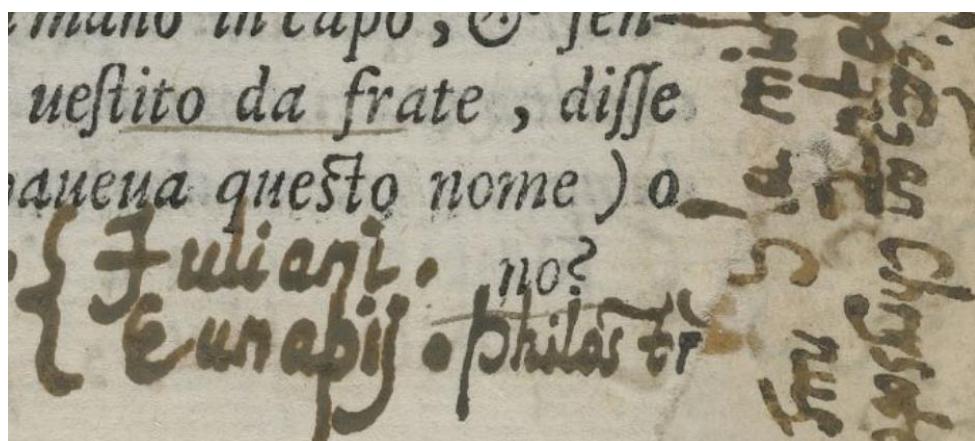
It happens that part of the written text is not legible due to the fact that the page is damaged, or that the exact text cannot be deciphered because of a difficult handwriting. This needs to be made clear in the transcription, and will be done as follows.

[ab]ove = only the letters ‘ove’ are clearly legible, but based on the context the transcriber made an educated guess that the missing letters are ‘ab’. The brackets here denote the uncertainty about the letters they contain. A similar use is: a[bo]ve. Note that the use of brackets is similar to expanding a contraction, such as q[ue].

[-]ove = the text preceding ‘ove’ is missing, and the transcriber does not have a clue about the missing text, nor about the numbers of letters that are missing.

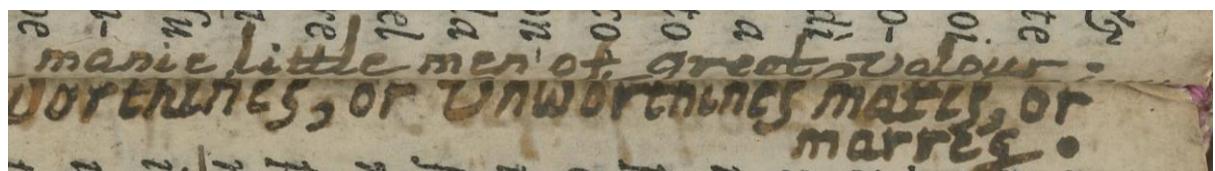
a[...]e = three letters of this word cannot be recognized (every dot represents one letter).

In the following example, the letters after ‘Philastr’ have disappeared (the page is torn). Since the line above the ‘r’ seems to signify a contraction, the missing letter could very well be an ‘i’, making ‘Philastri’, the genitive form of Philastrus, which matches with the preceding names (i.e. Juliani, Eunapij) which are also in the genitive. This particular example could be transcribed as ‘Philastr[i]’.



(Source: AOR: Domenichi, *Facetie, motti, et burli*, 8r).

Even though the images we are using for this project are of an amazing quality, sometimes, due to the bent of the paper, text is partly illegible or missing, as the image below shows (e.g. the text below ‘valour’).



(Source: AOR: Domenichi, *Facetie, motti, et burli*, 9r).

This requires someone to have a look at the original sources, and in order to quickly find all such instances, a standardized comment field in the XML should be used, namely:

<!-- MISSING TEXT: ILLEGIBLE DUE TO BENT PAGE. ORIGINAL SOURCE NEEDED-->

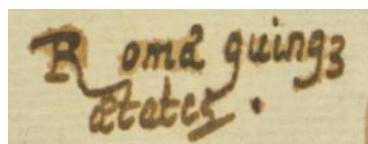
Once the original source is consulted, and the transcription of such marginal notes has been completed, a comment field should be used to highlight these formerly missing marginalia in order to make them easy to find for the people checking these new transcriptions.

<!--FORMERLY MISSING TEXT-->

## 1.14 Contractions and Abbreviations

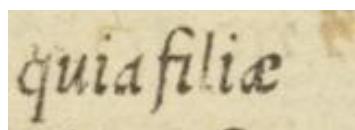
Contractions were used often, both in the printed text as in the marginal notes of the readers. **The general rule is that all contractions should be fully expanded, while indicating that they are contractions by placing the added letters between brackets. The only example to this rule is the ‘ae’ ligature in the printed text, which is written out without using brackets.**

For example, it often happens that the final “e” of a word is replaced by a sort of apostrophe (often happens with the genitive of female nouns, e.g. Romae).



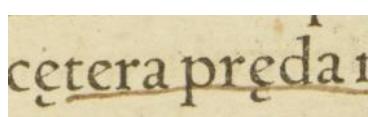
(Source: AOR: Livy, *Romanae historiae principis*, 478v).

In case Harvey uses this contraction, the ‘missing’ letter should be put in between brackets: Roma[e]. However, since in the printed text the ‘ae’ ligature is pretty clear, as the example below shows, we do not put the ‘e’ between brackets, but we write it out instead: *filiae*.



(Source: AOR: Livy, *Romanae historiae principis*, 479r).

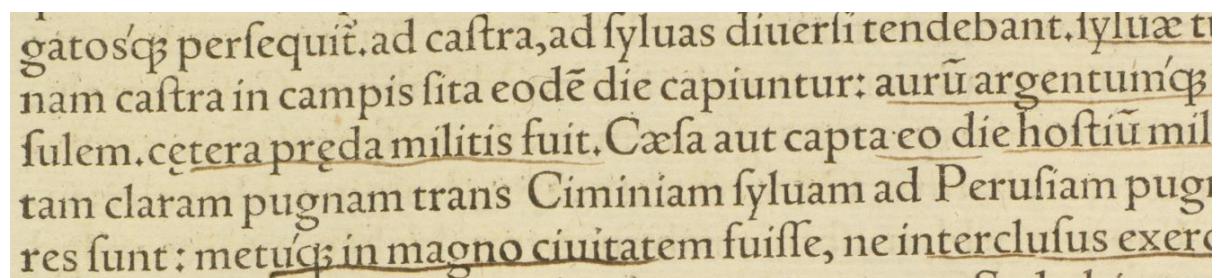
In the printed text ‘ae’ is sometimes rendered as ‘ę’:



(Source: AOR: Livy, *Romanae historiae principis*, 126v).

Because in this case the letter ‘a’ is completely missing, we put it between brackets: c[a]etera pr[a]eda.

Other contractions which are commonly used both by Harvey and in the printed text are placing a line above a word which replaces an ‘m’ or ‘n’ (e.g romanoru + line, which forms the genitive plural, romanorum; or: adamantinu + line, which is adamantinum).<sup>20</sup> The missing letters should always be put between brackets. Another common contraction which appears in the printed text as well as in Harvey’s marginal notes is omitting the “ue” in que (see the image on the previous page). Following the general rule, this should be transcribed as: q[ue].



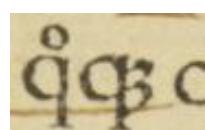
gatosq; persequit̄ ad castra, ad syluas diuerlī tendebant. Iyluæ t̄  
nam castra in campis sita eodē die capiuntur: aurū argentumq;  
sulem. c̄etera pr̄da militis fuit. C̄esa aut capta eo die hostiū mil  
tam claram pugnam trans Ciminiam syluam ad Perusiam pugi  
res fuit: metuq; in magno ciuitatem fuisse, ne interclusus exere

(Source: AOR: Livy, *Romanae historiae principis*, 126v).

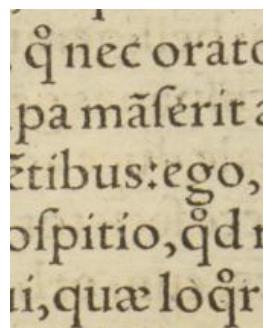
For example, the various contractions on this page should be transcribed as:

auru[m] argentumq[ue]; c[a]etera pr[a]eda; hostiu[m]; metuq[ue].

A couple of other examples of contractions in the printed text:



(Source: AOR: Livy, *Romanae historiae principis*, 143v).



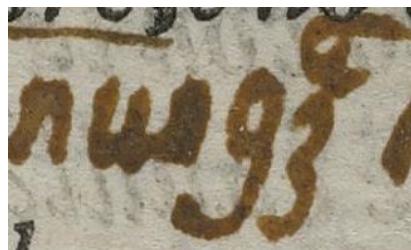
<sup>20</sup> ABO, Livy, 24.

(Source: AOR: Livy, *Romanae historiae principis*, 145r).

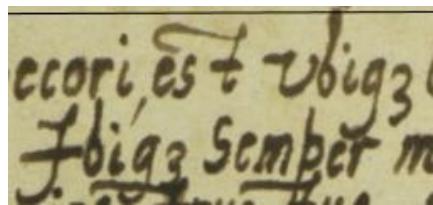
These are the abbreviations of the word: q[uo]q[ue], q[uo], q[uo]d, and loq[uo]r.

Another example is the contractions of the words supra and infra, rendered as s<sup>a</sup> and i<sup>a</sup> respectively. Like que, these word should be transcribed as s[upr]a and i[nfr]a. Other contractions, such as w<sup>th</sup> or w<sup>ch</sup> will be expanded by putting the omitted letter(s) between brackets: w[i]th, w[hi]ch.

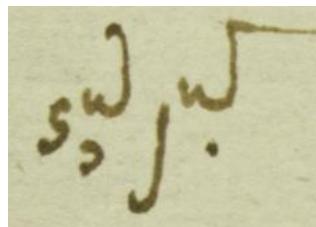
Some examples of contractions frequently used by Harvey: tamq[uam], numq[uam]; ubiq[ue]; ibiq[ue]; s[upr]a; i[nfr]a; M[aste]r.



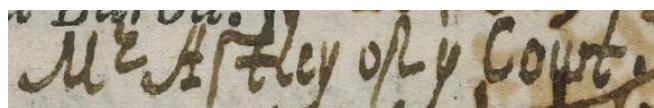
(Source: AOR: Domenichi, *Facetie, motti, et burli*, 2r.).



(Source: AOR: Livy, *Romanae historiae principis*, 9v).



(Source: AOR: Livy, *Romanae historiae principis*, 386r).

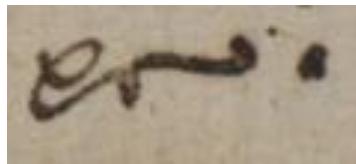


(Source: AOR: Domenichi, *Facetie, motti, et burli*, 24r).

Another abbreviation Harvey uses, albeit less frequently than those mentioned above, is ‘xc’ or ‘&cc’ which stands for et cetera and should be transcribed as [et cetera].



(Source: AOR: Castiglione, *Book of the Courtier*, 179v).

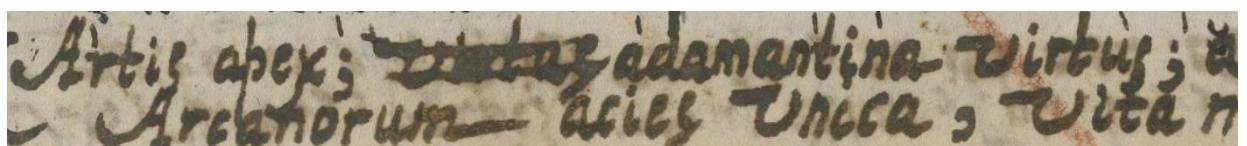


(Source: AOR: Castiglione, *Book of the Courtier*, 4v).

In case Harvey used more exotic contractions, it is wise to consult A. Cappelli, *Dizionario di Abbreviature latine ed italiane* (6<sup>th</sup> edition, Milan), an older edition of which can be accessed at: <http://www.hist.msu.ru/Departments/Medieval/Cappelli/>

## 1.15 Strikethrough

Sometimes Harvey crossed out words in his own marginal notes, as the following image shows:



(Source: AOR: Domenichi, *Facetie, motti, et burli*, 11r).

In this case, the word crossed out by Harvey is still visible, and should be transcribed as follows [virtus deleted]. However, if the text is not legible anymore, this can be transcribed as follows [deleted].

## **2. Transcribing in XML**

### **2.1 Using an XML editor**

When transcribing a page of one of the books annotated by Harvey, the transcriber should use the XML editor Oxygen.<sup>21</sup> One of the main benefits of using an XML editor is that it speeds up the process of transcribing and it validates the XML transcriptions against the schema (also see 2.2). If, for whatever reason, the XML editor is not working or available, use text editors such as Notepad+ (on PC) or TextWrangler (on MAC), and refrain from using Microsoft Word, since word uses curved quotation marks (“”) which are not recognized by XML, whereas the text editors use straight quotations marks (") which XML does recognize. Because of this, one should not copy text directly from a digital source into the XML editor, for it is likely that the source text contains curved quotation marks as well as special characters which are not recognized as such by XML (e.g. an ‘&’, which has to be declared in XML by using the code & - to mention but one example).

### **2.2 Validation**

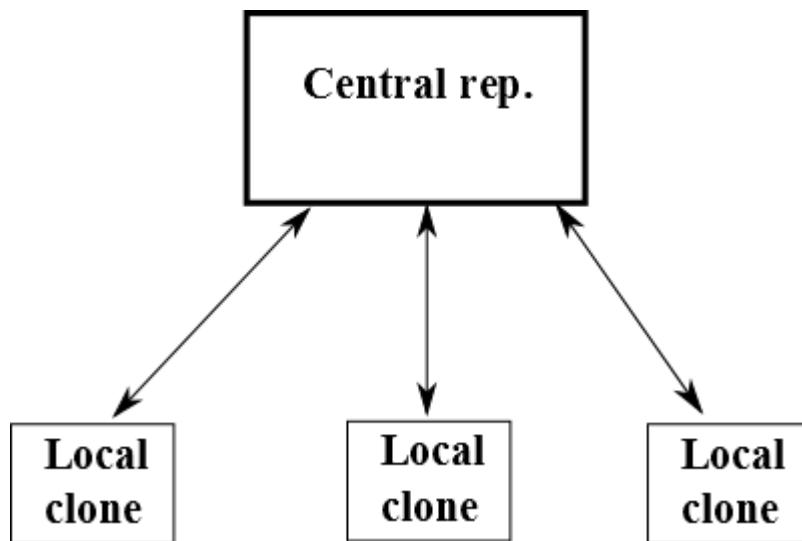
XML editors have other advantages, one of which is mistakes in the XML will be highlighted and that the XML files will be validated against the schema, thus ensuring that the transcription respects the hierarchy of the schema (and the DTD, on which the schema is based), that all the required attributes are used, and so on. As explained in section 1.1., all transcriptions should be linked to the external DTD and schema. In case the software is not working, the XML can also be validated online, using a site as <http://www.xmlvalidation.com/>. This site also provides the possibility to validate the XML against the schema.

---

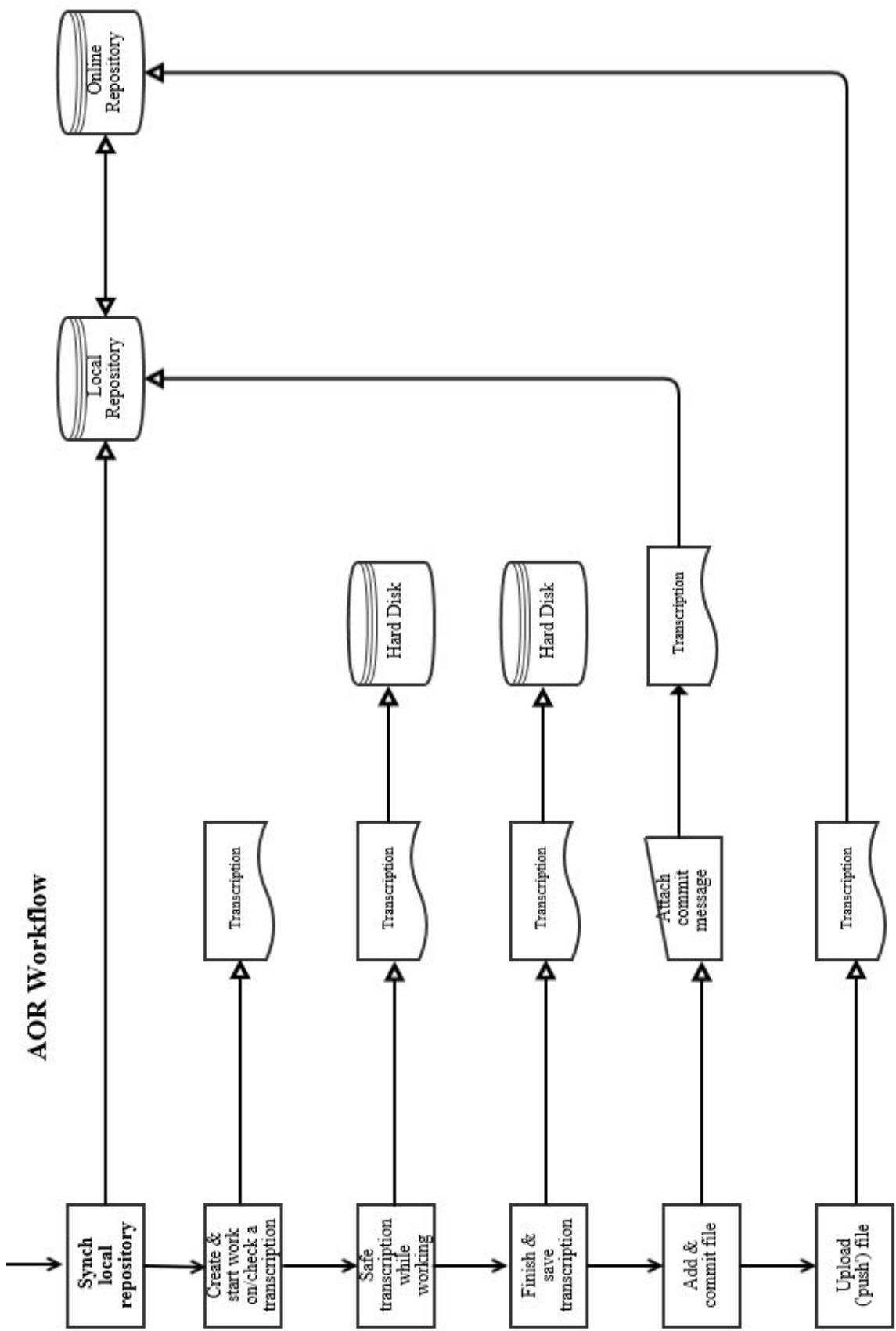
<sup>21</sup> <http://www.oxygenxml.com/>

### 3. Workflow

The process of transcribing consists of three phases or stages: in the first stage, the transcriber is working on the transcription. Once this work has been finished, the transcription will be checked by a member of the project team, the second phase. If all the people involved are satisfied with the transcription, the transcription is finalised (the third stage). In order to coordinate the work on the files, work which is incremental in nature and is done by various people, this project makes use of Git and GitHub. Git is a Version Control Systems, whereas GitHub is Git's online platform, making it possible for various people to contribute to a project. The project's central GitHub repository can be found at: <https://github.com/livesandletters/aor>. It is necessary to install GitHub (<http://git-scm.com/downloads>).



Every transcriber has a clone of the central repository. A clone is set up easily: open Git Shell, which takes you to the GitHub directory on the hard drive. Type git clone [url of the repository] [name of the clone]. For example: c:\Users\User1\Documents\GitHub> git clone https://github.com/livesandletters/aor AORCLONE, which will create a new folder (c:\Users\User1\Documents\GitHub\AORCLONE) where the local repository is created. Transcribers will save their work in the local repository, but also upload their files to the central repository, so that files are saved online as well.



### 3.1 First phase: transcribing

The process of transcribing broadly consists of three phases, the first of which is the ‘transcribing phase’. In this phase mostly consists of one person working on the transcription of a single page. As the image depicting the workflow clearly shows, **prior** to starting to work on a new transcription, or to continuing working on an already existing one, **the transcriber should always synchronise his or her local copy of the repository (i.e. the clone) with the online repository in order to avoid the emergence of different versions of the same file**. When working on older versions of files, all kinds of errors can occur due to the merger of clashing versions (when synchronising the repositories after having done work on the transcriptions). Although it is possible to solve these errors, this is likely to be very time consuming, hence we should try to avoid this from happening and the easiest way is to always synchronise the repositories before commencing with work.

Once this is done, the work on the actual transcriptions can start, and throughout the process of transcribing, the transcriber should make sure that the files are saved while working on it, lest any work be lost. According to the infrastructure describe above, transcribers have their local clone of project’s central GitHub repository, where they can store their transcriptions. The transcriptions should be saved (and uploaded) as .xml files, the filename reflecting the part of the name of the Tiff file (i.e. if the Tiff file is name Ha2.001v.tif, the transcription of this image should be named Ha2.001v.xml). Moreover, the files should be saved in a subfolder of the local repository. The names of these folders are the surnames of the author of the book that is being transcribed (e.g. Domenichi). The use of these subfolders is mandatory, for otherwise the database will become a mess of transcriptions from different books. In order to store a file in the local database (or repository), the file should first be ‘added’ (moving the file to the staging area), after which it can be ‘committed’, a term (and command) in Git which moves the file from the staging area to the repository. Comments can be attached to the file every time it is committed, and this, together with the availability of the different ‘commits’ of the file, make Git and GitHub such a useful tool to work with. For example, after finishing a transcription (the first stage of the process), the transcriber adds and then commits the file. After giving the commit command, a text file opens, and comments can be attached to the ‘commit’ of this file:



In the **first stage** of the process, during which the transcriber is working on the file, this note can contain information to the transcriber her or himself, such as a quick word where on the page the transcribers should continue their work the following day. All the transcriptions, also those which are still under construction, should not only be saved in the local repository, but should also be uploaded to the online repository so that every transcriber always works with the most up-to-date repository. Uploading files can be done through the ‘push’ command – for information on GitHub and Git commands see the Source section of this document.

### **3.2 Second phase: checking**

Once the transcriber has finished a transcription, the transcription has to be checked (the **second stage** of the process). After the transcriber has finished the work on a transcription, he or she should commit it to the local repository, and the comments (the ‘commit message’) attached to the message have to be standardized to a certain extent, so that the people who are checking the transcriptions of a particular book, can easily find the transcriptions they are supposed to check. **Therefore, the first line of the ‘commit message’ should read: ‘First pass: to be checked’**, for the GitHub overview will give the name of the file plus the first bit of the commit message attached to it.

When going to the relevant folder (i.e. that of a particular book) of the online repository, one can click on the button ‘view latest commits’, while this also can be seen in this GitHub software on the transcriber’s local computer. In this way, it is easy to spot the transcriptions which have moved to the second phase of the process.

In order for the comment field of Git/GitHub to be a useful overview, the comments added to the commit message should therefore be concise and follow the standard explained above. For example, when someone has finished checking a transcription, the mission that should be attached to the commit should include: [filename] [action undertaken], e.g. Domenichi\_009left: checked by JG, comments added in XML. It is highly recommended that the transcribers insert comments in the XML file to highlight uncertainties or raise questions for the person checking the transcriptions.

The transcribers will check each other’s work. Checking someone’s work is an important part of the project and it should be done thoroughly. Every aspect of a transcription should be checked, such as the scholarly content but also more mundane facets such as typos. It is recommended that the person who is checking the transcription only changes errors such as typos or includes marginal annotations that have been overlooked. However, when uncertainties about ‘scholarly content’ or varying ideas about correct interpretations or translations occur, a message should be left in the XML file to highlight this. In this way, a discussion can take place between members of the team, all of which will be

documented by the version control offered by Git. Transcribers are encouraged to highlight uncertainties themselves in the comment fields to attract the attention of the person checking the transcriptions.

Because transcriptions are checked and have to be updated by the transcribers, this will cause that transcriptions will be bounced back and forth between the original transcriber and the person that is checking the transcription. However, because all of this happens in Git/GitHub it is possible to keep track of all the changes. Every time a concise commit message should be attached, including the filename, the action undertaken, and the action required (and by which person). For example: Ha2.015r: comments added in XML, to be checked by Chris. Or: Ha2.015r: transcriptions checked: missing text remaining, which signifies the problem of missing text due to page bents. The exact wording is not that important, as long as the work that still needs to be done on a transcription is clear.

Undoubtedly there will be questions or uncertainties that will require the expertise of other project members such as Earle Havens, Matt Symonds, or Tony Grafton. If a transcription is checked and approved, except for one or more uncertainties, it should have a commit message which reads: **All checked; expert advice needed**. Moreover, transcribers should indicate instances of persisting uncertainty in the transcription itself by using a comment that starts with '**Expert advice:**' followed by **a short explanation of the problem**. After such transcriptions have been uploaded to the central repository, Jaap Geraerts will liaise with the various experts who are involved in the project and coordinate this last stage of the checking process.

### 3.3 Third phase: finalised transcriptions

After all the work on a transcription has been finished, Jaap Geraerts will upload it to a separate folder called 'finalised transcriptions'. This is a sub-folder or directory of the folder of a specific book, and the structure of the GitHub thus looks like this:

**AOR**

XMLschema

Livy

Finalised transcriptions

Domenichi

Finalised transcriptions

In this way, everyone can easily locate the finalised transcriptions, without having to do a survey of the commit messages in order to find this out.

### **3.4 How to deal with errors**

In case that, despite the precautions mentioned above, one encounters errors such as files being corrupted due to a merger of conflicting versions, an email should be sent to all the transcribers. Moreover, all work on these files should be stopped and for the time being no files should be uploaded to the online repository. Once the corrupted files are cordoned off, I (Jaap Geraerts) will try to solve the problems and, if necessary, will divide the tasks that need to be done to restore the files. Thereafter, once all the problems are solved, I'll notify the transcribers that work can be resumed.

As mentioned, conflicting versions can emerge when different people are working on the same file. Normally, because the work on the transcriptions is divided in different phases, and because each phase only involves one person working on a file, this problem is largely avoided. However, problems can arise when people are going to update files based on the weekly error reports (the results of the XML validation) sent by John Abrahams or Mark Patton. Transcribers should therefore refrain from trying to resolve these errors on their own accord. Instead, I (Jaap Geraerts), will coordinate the work that needs to be done or solve the errors myself.

## 4. Spreadsheets

Three Excel spreadsheets contain information about the people, books and locations mentioned by Harvey. These spreadsheets can be found on CELL's google drive (you need permission to access it). If a transcriber encounters a person, book title, or location mentioned by Harvey, the first point of reference is these spreadsheets. If a person, book title, or location is already mentioned in one of the spreadsheets, this name (and exactly this name) has to be used to tag this person/book title/location. If however, this person/book title/location is not listed in one of the spreadsheets, then the transcriber has to add this to the spreadsheet.

When adding the name of a new person or location it is important to find the modern English name, commonly used by academics and non-academics alike. A possible source for finding the modern English name is the Oxford Dictionary of National Biography (ODNB - <http://www.oxforddnb.com/>). In case of persons, it is best to give the first name followed by the surname (e.g. Guillaume Du Bartas), rather than Du Bartas, Guillaume. Sticking to the rule as far as possible enlarges the consistency, and prevents transcribers from making mistakes. Moreover, it is better to use someone's name rather than his title (e.g. 'Robert of Dudley' rather than 'Earl of Leicester'), although in some cases it is better to use the commonly-known and used name (e.g. Queen Elizabeth, Philip II, Charles V). There are more exceptions though, for of some persons only their surname is commonly used (e.g. Aristotle, Augustine, Livy, Cicero). Moreover, sometimes only the name of a family is mentioned (e.g. Borgia), and in this and other cases only the family name should be tagged. Alternative spellings of a name are recorded in the database as well, so that end users can find the person they are looking for even when using the non-standardized variant of that person's name. Geographical locations can be handled in the same way as people, and the modern English names of towns, cities, areas, and countries, should be used.

Just as persons and locations, books are tagged by using one standardized name. This standardized name is used across editions and languages, i.e. if Harvey is referring to a translation of a particular book, will still tag this instance by using the standardized name. In this way all the references to a particular book and its various editions and translations are captured in one tag. Since the end users have the images and the transcriptions, they can work out for themselves if and when Harvey is referring to a translation or a particular edition of the book. For instance, the standardized tag of Augustine's *De Civitate Dei* could be De civitate Dei, and this tag is used also when Harvey mentioned the whole title of this book (*De Civitate Dei contra Paganos*). If Harvey refers to *The city of God* it should still be tagged as *De Civitate Dei*. A good source for finding the standardized title is the Universal Short Title Catalogue (USTC - <http://www.ustc.ac.uk/>).

The book spreadsheet also has a column called 'bibl. information: title', in which extra information about the book can be stored, such as the complete title (in case the book is quite obscure,

or if the title is similar to that of another book) and the link(s) to digital copies of that book. **The name of the author of the book should also be included in the column ‘bibl. information: author’.** Moreover, if the name of the author is not yet recorded in the people’s spreadsheet, this should be added.

As agreed in the meeting of Wednesday, October 1 (2014), Jaap Geraerts, will regularly check the spreadsheets for updates. If someone makes a mistake when updating the spreadsheets, the proposed change should be emailed to all the transcribers and, once agreed, I will update the spreadsheets. Furthermore, I’ll also update the XML files in which the superseded name appears (I reckon it is best when one person is responsible for such an update, rather than dividing this task among a number of people). DRCC will regularly run checks to validate the names in the XML transcriptions against those in the Excel spreadsheets.

## 5. Sources

### Biographical and bibliographical information:

- Oxford Dictionary of National Biography: <http://www.oxforddnb.com/>
- Universal Short Title Catalogue: <http://www.ustc.ac.uk/>

### Git & GitHub:

- Git site: <http://git-scm.com/>
- ‘Try Git’ introduction: <https://try.github.io/levels/1/challenges/1>
- GitHub help: <https://help.github.com/>
- Full reference book: <http://git-scm.com/book>
- Some tutorials: <https://www.atlassian.com/git/tutorials>
- Git Wiki: [http://en.wikipedia.org/wiki/Git\\_%28software%29](http://en.wikipedia.org/wiki/Git_%28software%29)

### Dictionaries

- A. Cappelli, *Dizionario di Abbreviature latine ed italiane* (6<sup>th</sup> edition, Milan); an older edition can be accessed at: <http://www.hist.msu.ru/Departments/Medieval/Cappelli/>

## Appendix A: DTD

```
<?xml version="1.0" encoding="UTF-8"?>

<!ELEMENT transcription (page, annotation)>
<!ELEMENT page EMPTY>
    <!ATTLIST page filename CDATA #REQUIRED>
    <!ATTLIST page pagination CDATA #IMPLIED>
    <!ATTLIST page signature CDATA #IMPLIED>
    <!ATTLIST page reader CDATA #REQUIRED>

<!ELEMENT annotation (marginalia*, underline*, symbol*, mark*, numeral*, errata*, drawing*)>
    <!ELEMENT marginalia (language*, translation?)>
    <!ATTLIST marginalia hand (English_secretary|Italian) #REQUIRED>
    <!ATTLIST marginalia marginalia_id CDATA #IMPLIED>
    <!ATTLIST marginalia marginalia_continues_to CDATA #IMPLIED>
    <!ATTLIST marginalia marginalia_continues_from CDATA #IMPLIED>
    <!ATTLIST marginalia marginalia_to_transcription CDATA #IMPLIED>
    <!ATTLIST marginalia marginalia_from_transcription CDATA #IMPLIED>
    <!ATTLIST marginalia book_id CDATA #IMPLIED>
    <!ATTLIST marginalia date CDATA #IMPLIED>
    <!ATTLIST marginalia other_reader CDATA #IMPLIED>
    <!ATTLIST marginalia topic (Law|Astronomy|Warfare|Rhetoric) #IMPLIED>
    <!ATTLIST marginalia anchor_text CDATA #IMPLIED>
        <!ELEMENT language (position*)>
            <!ATTLIST language ident (EN|EL|FR|IT|LA|ES) #REQUIRED>
                <!ELEMENT position (marginalia_text*, person*, book*, location*, X-ref*, emphasis*, internal_ref*)>
                    <!ATTLIST position place
(head|tail|left_margin|right_margin|intext|full_page) #REQUIRED>
                    <!ATTLIST position book_orientation
(0|90|180|270) #REQUIRED>
                        <!ELEMENT marginalia_text (#PCDATA)>
                        <!ELEMENT person EMPTY>
                            <!ATTLIST person name CDATA
#REQUIRED>
                        <!ELEMENT book EMPTY>
                            <!ATTLIST book title CDATA
#REQUIRED>
                        <!ELEMENT location EMPTY>
                            <!ATTLIST location name CDATA
#REQUIRED>
                        <!ELEMENT X-ref EMPTY>
```

```

        <!ATTLIST X-ref person CDATA
#REQUIRED>

        <!ATTLIST X-ref book_title CDATA
#REQUIRED>

        <!ATTLIST X-ref language
(EN|EL|FR|IT|LA|ES) #IMPLIED>

        <!ATTLIST X-ref text CDATA
#IMPLIED>

        <!ELEMENT emphasis EMPTY>
        <!ATTLIST emphasis method
(chalk|pen) #REQUIRED>

        <!ATTLIST emphasis type
(straight|curved|dotted) #REQUIRED>

        <!ATTLIST emphasis text CDATA
#IMPLIED>

        <!ELEMENT internal_ref (target+)>
        <!ATTLIST internal_ref text CDATA
#IMPLIED>

        <!ATTLIST internal_ref anchor_text CDATA
#IMPLIED>

        <!ELEMENT target EMPTY>
        <!ATTLIST target filename CDATA
#REQUIRED>

        <!ATTLIST target book_id CDATA
#REQUIRED>

        <!ATTLIST target text CDATA
#REQUIRED>

        <!ELEMENT translation (#PCDATA)>
<!ELEMENT underline EMPTY>
        <!ATTLIST underline method (chalk|pen|scoring) #REQUIRED>
        <!ATTLIST underline type (straight|curved|dotted) #REQUIRED>
        <!ATTLIST underline language (EN|EL|FR|IT|LA|ES) #IMPLIED>
        <!ATTLIST underline text CDATA #IMPLIED>

        <!ELEMENT symbol EMPTY>
        <!ATTLIST symbol name
(Asterisk|Bisected_circle|Crown|JC|HT|LL|Mars|Mercury|Moon|Opposite_planets|Saturn|Squ
are|SS|Sun|Venus) #REQUIRED>

        <!ATTLIST symbol place (head|tail|left_margin|right_margin|intext|full_page)
#REQUIRED>

        <!ATTLIST symbol language (EN|EL|FR|IT|LA|ES) #IMPLIED>
        <!ATTLIST symbol text CDATA #IMPLIED>

        <!ELEMENT mark EMPTY>
        <!ATTLIST mark name
(apostrophe|box|bracket|circumflex|colon|comma|dash|diacritic|dot|double_vertical_ba

```

```

r|equal_sign|est_mark|hash|horizontal_bar|page_break|pen_trial|plus_sign|quotation_mark|scri
bble|section_sign|semicolon|slash|straight_quotation_mark|tick|tilde|triple_dash|vertical_bar|
X_sign) #REQUIRED>
    <!ATTLIST mark method (chalk|pen) #REQUIRED>
    <!ATTLIST mark place (head|tail|left_margin|right_margin|intext|full_page)
#REQUIRED>
        <!ATTLIST mark language (EN|EL|FR|IT|LA|ES) #IMPLIED>
        <!ATTLIST mark text CDATA #IMPLIED>
<!ELEMENT numeral (#PCDATA)>
    <!ATTLIST numeral place (head|tail|left_margin|right_margin|intext|full_page)
#REQUIRED>
        <!ATTLIST numeral language (EN|EL|FR|IT|LA|ES) #IMPLIED>
        <!ATTLIST numeral text CDATA #IMPLIED>
<!ELEMENT errata EMPTY>
    <!ATTLIST errata language (EN|EL|FR|IT|LA|ES) #REQUIRED>
    <!ATTLIST errata copytext CDATA #REQUIRED>
    <!ATTLIST errata amendedtext CDATA #REQUIRED>
<!ELEMENT drawing EMPTY>
    <!ATTLIST drawing name (face|manicule|map|florilegium) #REQUIRED>
    <!ATTLIST drawing method (chalk|pen) #REQUIRED>
    <!ATTLIST drawing place (head|tail|left_margin|right_margin|intext|full_page)
#REQUIRED>
        <!ATTLIST drawing language (EN|EL|FR|IT|LA|ES) #IMPLIED>
        <!ATTLIST drawing text CDATA #IMPLIED>

```