

Molto Mag.^{co} et Rev.^{do} mio S.^{re}

Darmi hor mai tempo di rompere il silenzio si qui usato ed. V.S.M.R. che
da che mi partij di Roma, si per rinfrancarli nella memoria il desiderio
che ho di servirli, come ancora per darli occasione di soddisfare
al desiderio mio che è d'intender nuova di lei: et sentire il parere suo
circa alcune mie difficoltà, delle quali una è questa che co la presente
gli mando, intorno alla dimostrazione dell'infra scritto Lemma, la quale

Text Encoder and Annotator

an all-in-one editor for transcribing and
annotating manuscripts with RDF

Motivation

- [ClaviusOnTheWeb](#) Project
 - manuscripts in latin and italian language about Christopher Clavius
- Several kind of **users** such as historians, linguists, scholars and students deal with manuscripts
- Users often have to:
 - **transcribe** a manuscript from its digitized copy
 - **annotate** the text of a transcribed manuscript

Existing tools

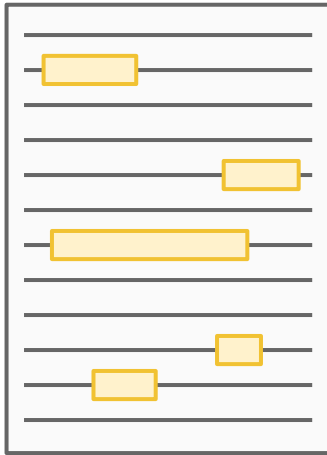
- Transcription and annotation are often treated as **distinct steps**
- Heavy weight languages are used for annotating text
- RDF-based annotation tools are often designed only for web page annotation

*None of them provide transcription and annotation features **together**, along with RDF support.*

Our Proposal

- Combine the transcription and annotation phases as they were **one joined process**
- Use a **lightweight** markup language
- Use **RDF** triples for specifying details of annotations

Inline vs Standoff



Inline:

- annotations are included **within** the text
- an annotation and its text are **kept close** each other
- annotations could make the text **heavier**
- text could become **hard to read**
- a **complex** annotation language could make the annotation process **difficult**



Standoff:

- annotations are **apart** from the text and store **offsets**
- no markup overloading
- standoff could raise problems if a **transcription error** is found during the subsequent annotation phase

Lightweight Markup Language

_____<text>(id1)_____

_____<other text>(id2)_____

+++

(id1) predicate object

(id1) predicate object

(id2) predicate object

+++

_____<another
portion of text>(id3)_____

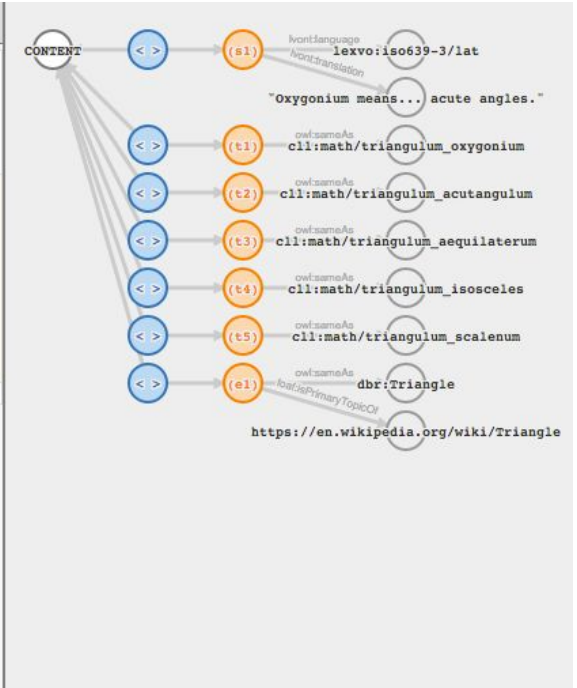
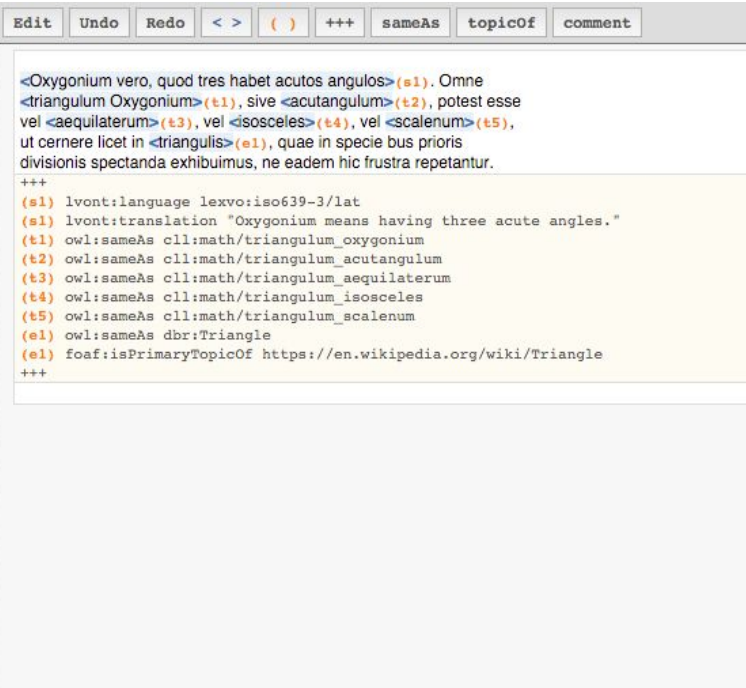
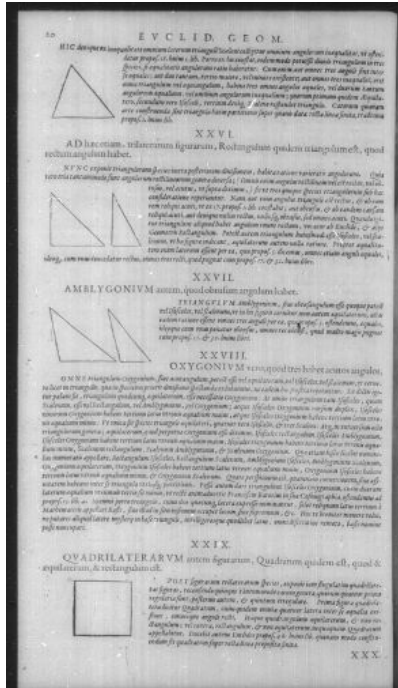
Inline lightweight syntax for:

- **marking portions** of text
- labeling them with an **identifier**

“Standoff-like” syntax for:

- creating a **directive section** in which the annotation details can be specified using **RDF**

Text Encoder and Annotator (TEA)

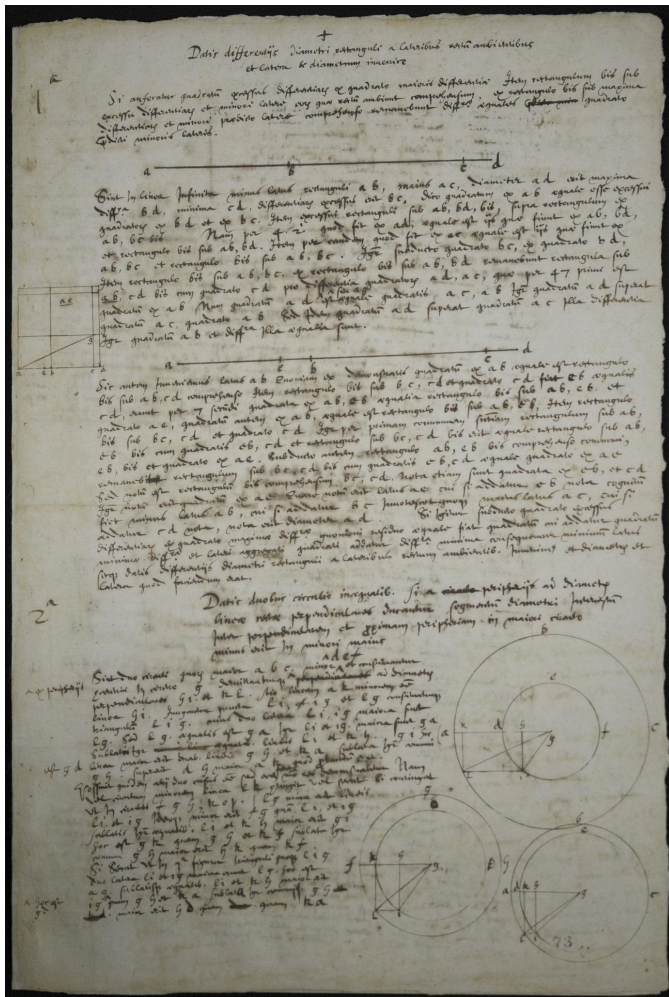


Manuscript
image

Editor

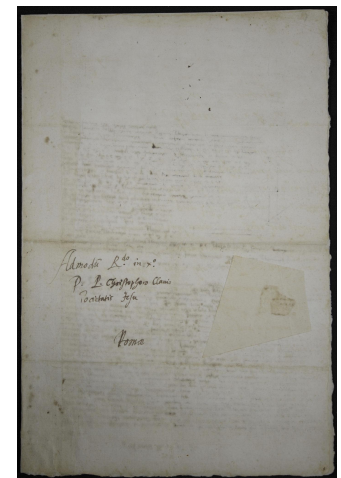
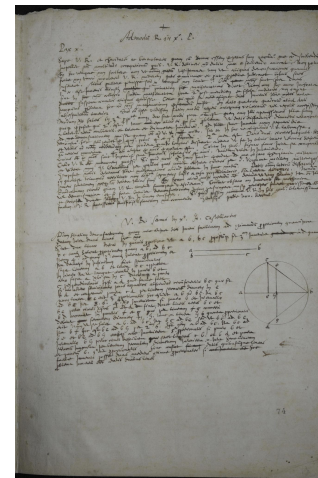
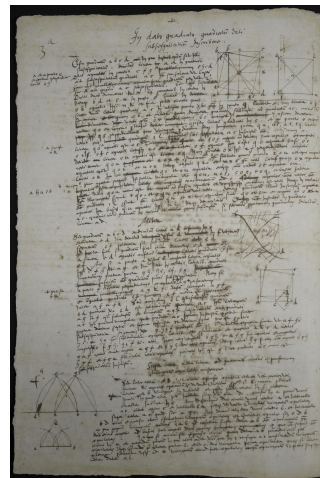
Summary Diagram

Example of annotations



Letter written
by Jean Chastellier
to Christoph Clavius
Paris 4 dicembre 1594

(Historical Archives of the Pontifical Gregorian
University, APUG 530 cc. 73r-74v)



Example of annotations

- **Annotation about the whole document:**

Jean Chastellier a Christoph Clavius in Roma

Paris 4 dicembre [1594] (1)

APUG 530 cc. 73r-74v

Autografa

- **Editor's reconstructions:**

... solutionem neque tamen debitor illius sum [quoniam replicavi] aliud mihi agendum esse deinde ut v[olo] [solvimus] ita ...

- **Editor's comments:**

... hoc nostro collegio studia (3) mathematica (4) hora 2a post meridiem et schola ...

- **Author's additions:**

... erit ig {hoc est gd} quam gh et ...

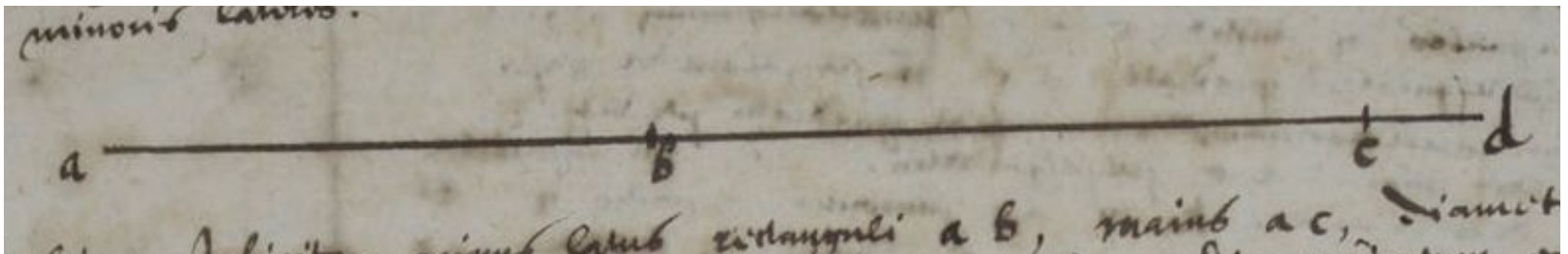
- **Sheet identifiers:**

73v

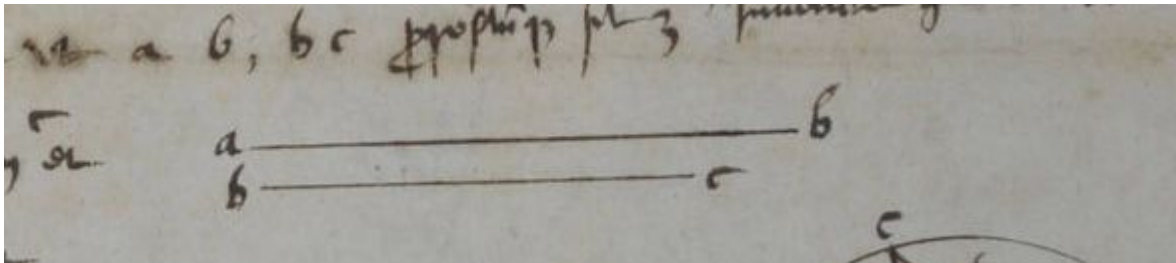
74r

Example of annotations

Drawing of figures by typing characters



a_____b_____c_____d

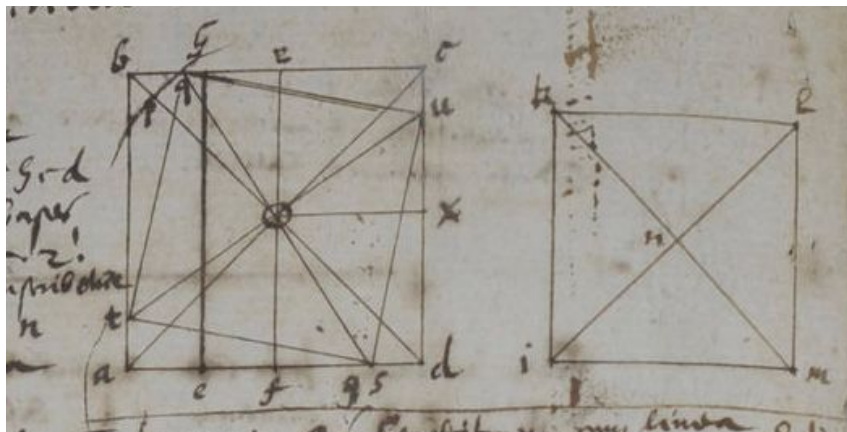


a_____b

b_____c

Example of annotations

Annotating the presence of an image

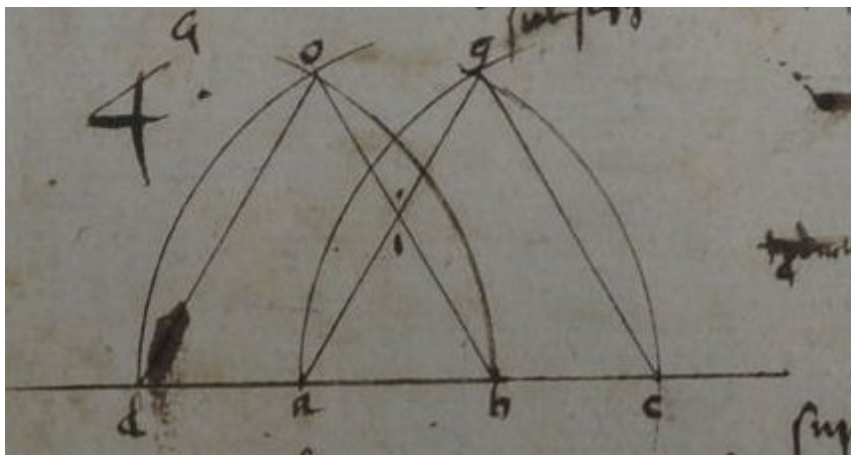


[FIGURE]

[FIGURE 2]

[FIGURE 3a-3b-3c]

...



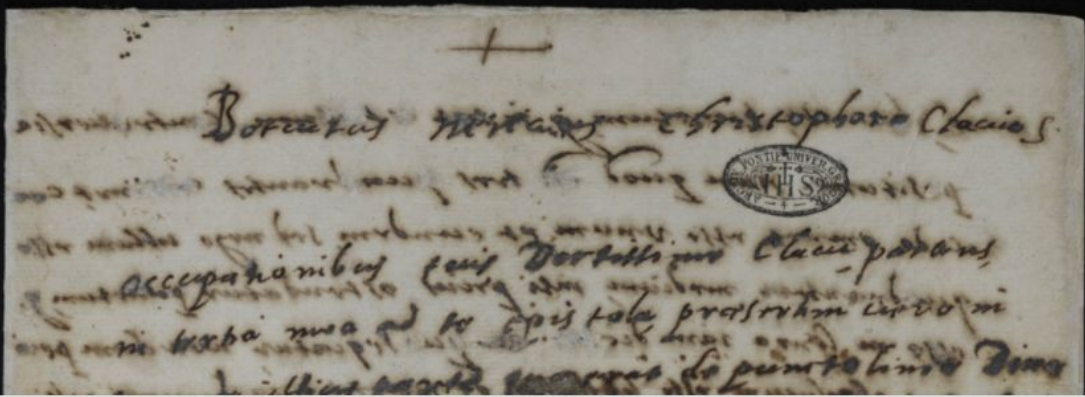
Future Works

- New syntax allowing **overlapping** and **nested** annotations
- New interface designed for hosting **multiple** views (manuscript image, editor, summary)
- Triple **export** using different **formats** (XML, Turtle, JSON-LD) in compliance with different **models** (Open Annotation and NIF2)
- Annotation on **images**. Even big ones like digitizations of maps (Matteo Ricci, World map of 1602)

Text Encoder and Annotator (TEA)

Previous

Next



Undo Redo << >> +++ triple ▾

Botvitus Nericius <clavius<Christophoro Clavio>clavius> S.]]s[[
Occupationibus tuis Doctissime <clavius<Clavi>clavius> parcens,
in tertia mea ad te epistola, praesertim vero in
secunda illius parte, quae erat de puncto <d<lineae Dino
strati>d> ultimo abs te invento; brevior fortasse fui,
quam oportuit, ideoq. hac scriptione rem totam
paullo uberius, si potero, explicabo.]]s[[Principio
igitur lineae conchoidis accidens certissimum est,
nimirum, si per lineam Dinostrati transeat, eam
per lineas rectas dividere in partes supra infraq.
aequales, propter motum lineae a summo, velut polo,
basim permeantis; qui quidem motus familiam in
hoc negotio, vel in primis ducit.]]s[[Certum quoque est
punctum, quod aequidistantiae dicitur nempe g esse
in linea conchoide, utpote ipsam describens.]]s[[

+++
clavius is wd:Q76728
clavius type foaf:Person
clavius page http://claviusontheweb.it
clavius comment "Our project is named after him."

All ok.

clavius	Christophoro Clavio	is	wd:Q76728
		type	foaf:Person
		page	http://claviusontheweb.it
		comment	"Our project is named after him."

clavius	Clavi	is	wd:Q76728
		type	foaf:Person
		page	http://claviusontheweb.it
		comment	"Our project is named after him."

Id lineae Dino←strati

mp	mp	comment	"Manu propria"
mp	mp	comment	"Manu propria"
mp	mp	comment	"Manu propria"

Thank you!

Questions?

fabio.valsecchi@iit.cnr.it