

# Intro to automatic collation

Programme doctoral en études numériques.  
Formation **Textes et éditions numériques**, 25 et 26  
avril 2019, Université de Lausanne

Elena Spadini (Université de Lausanne)

1. What is collation
2. Where does it come from
3. The beginning of automatic collation
4. The present of automatic collation

What is collation

automatic  
machine assisted  
computer-supported  
semi-automatic

COLLATION

compare two or more texts,  
in order to spot their differences

compare two or more texts,  
in order to spot their differences

## Terminology

compare the *readings* of two or more *witnesses* of a work, in order to spot the *textual variants* between them

The sources to be collated, referred as **witnesses**, are the handwritten or printed copies of a textual work.

A **reading** is how a portion of the work is recorded in one witness. For example, considering the first verse of a poem,  
the reading of witness 1 (**W1**) is **the sun is high**  
the reading of witness 2 (**W2**) is **the star is high**  
the reading of witness 3 (**W3**) is **the sun is high**.

A **textual variation** occurs when the witnesses have different readings. In the example above, there is textual variation between **W1 and W2** and between **W2 and W3**. There is no textual variation between **W1 and W3**.

# Examples. How many witnesses?

- Julius Caesar's *De bello gallico*.
- Dante's *Commedia*.
- Samuel Beckett's *Stirring Still*.

# Examples. How many witnesses?

- Julius Caesar's *De bello gallico*. 255 (manuscripts)
- Dante's *Commedia*.
- Samuel Beckett's *Stirring Still*.

# Examples. How many witnesses?

- Julius Caesar's *De bello gallico*. 255 (manuscripts)
- Dante's *Commedia*. 600 (manuscripts and early prints)
- Samuel Beckett's *Stirring Still*.

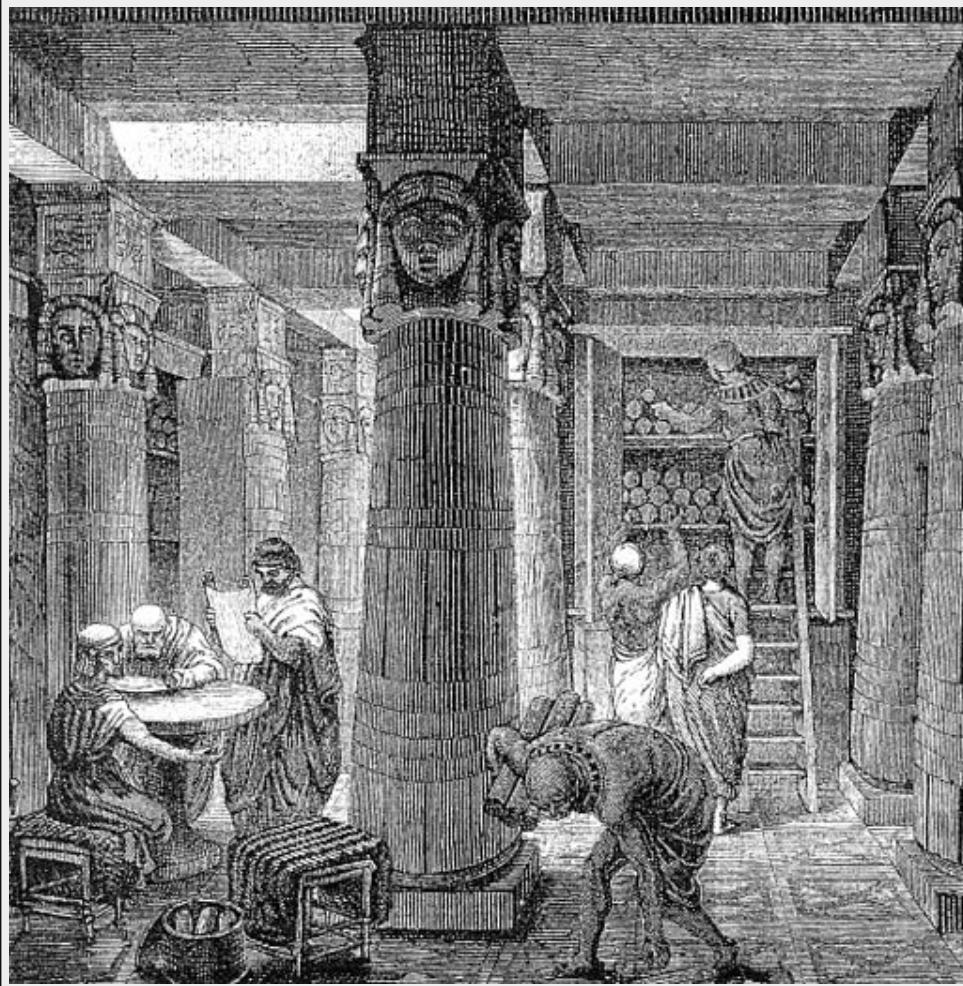
# Examples. How many witnesses?

- Julius Caesar's *De bello gallico*. 255 (manuscripts)
- Dante's *Commedia*. 600 (manuscripts and early prints)
- Samuel Beckett's *Stirring Still*. 40 (copybooks, holograph manuscripts, typescripts, carbon copies)

# Collate ... why?

- establishing a critical text, as for a critical edition
- tracking the writing process, as for a genetic edition
- studying the textual transmission, for purposes other than the edition

Where does it come from



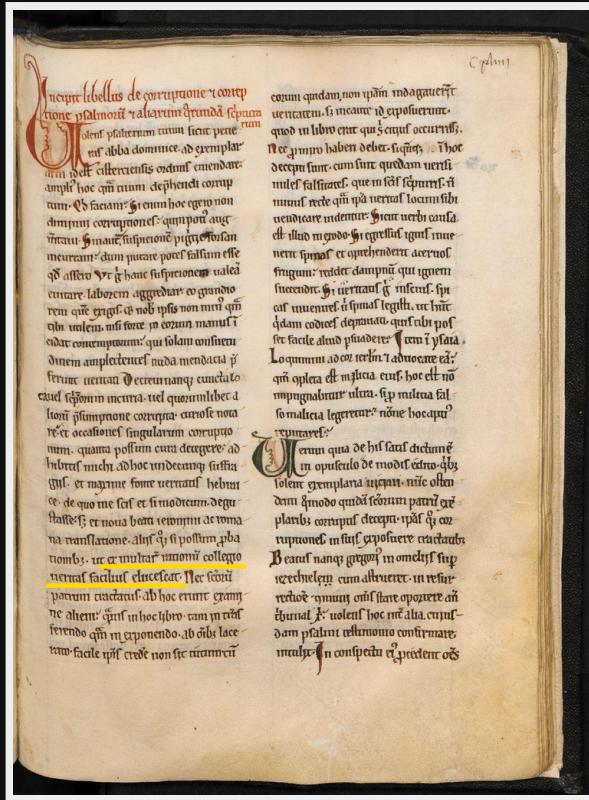
# Library of Alexandria

Public Domain, [Link](#)



Antonio da Fabriano II, *Saint Jerome*

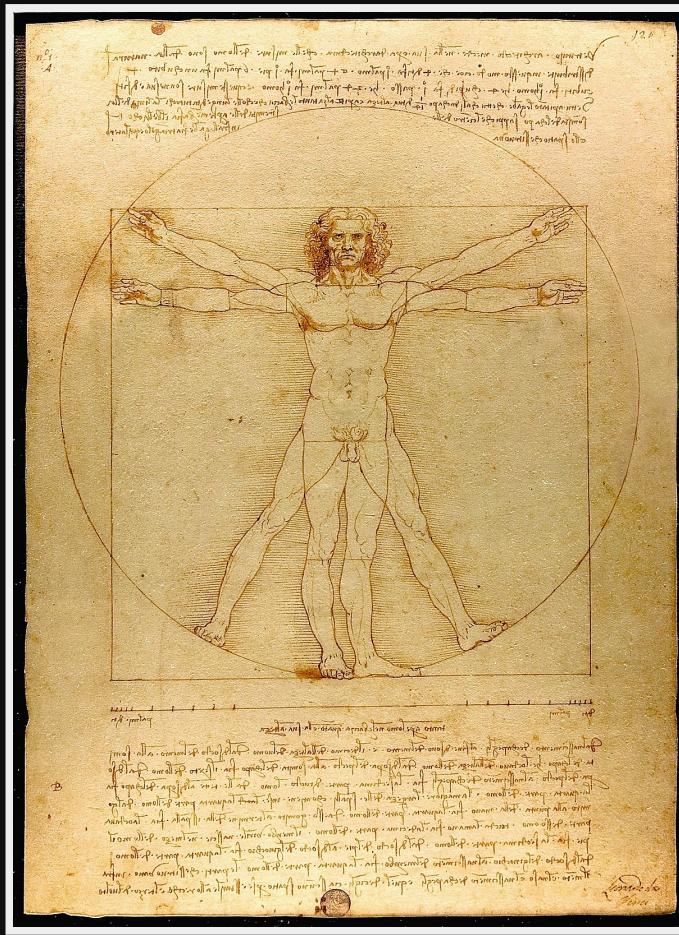
Public Domain, [Link](#)



*ex multarum rationum collegio veritas facilius elucescat*  
from the gathering of many proofs the truth comes to light more easily

Nicolò Maniacutia, *Libellus de corruptione et correptione Psalmorum  
et aliarum quarundam scripturarum*

Bibliothèque interuniversitaire Montpellier, ms 294 f 144r. [Link](#)



# Leonardo da Vinci, Vitruvian man

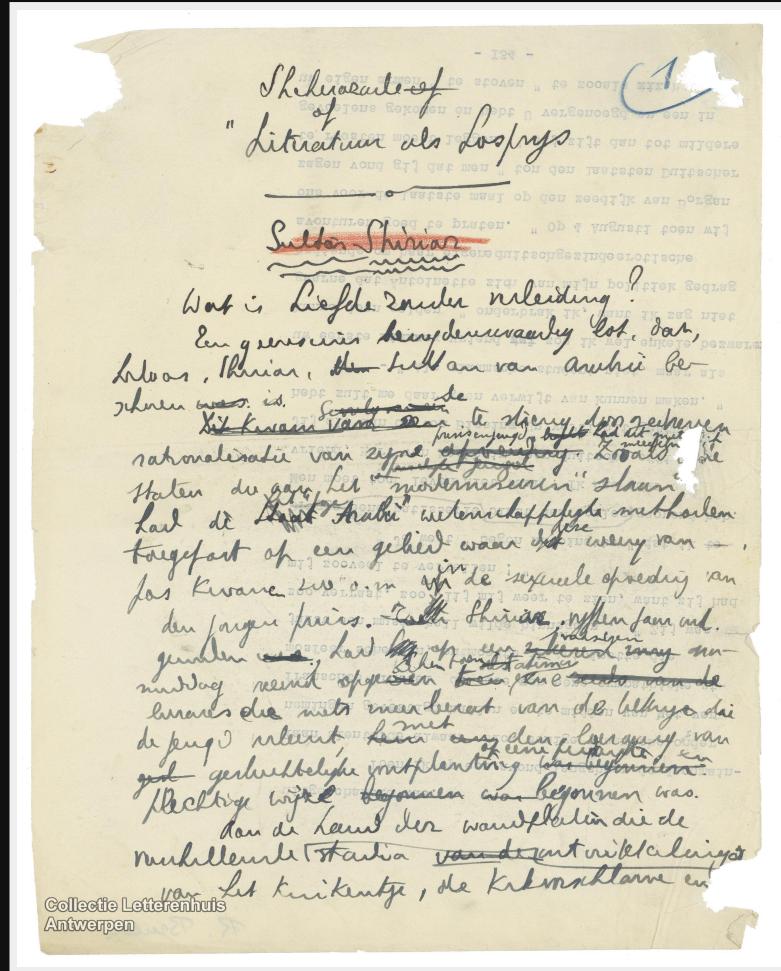
Public domain. [Link](#)



*recensere sine interpretatione et debemus et possemus*

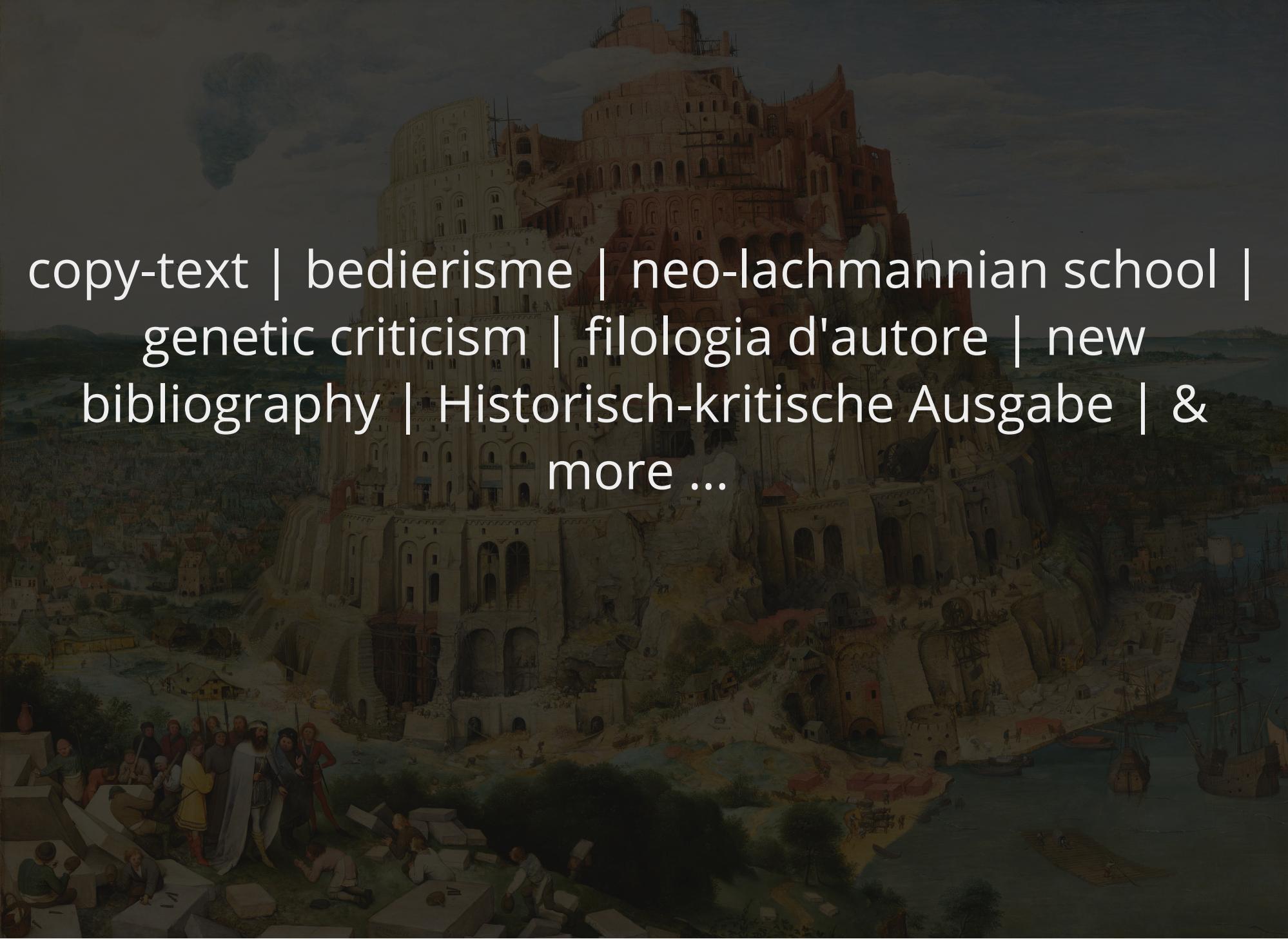
Karl Lachmann

Public Domain, Link

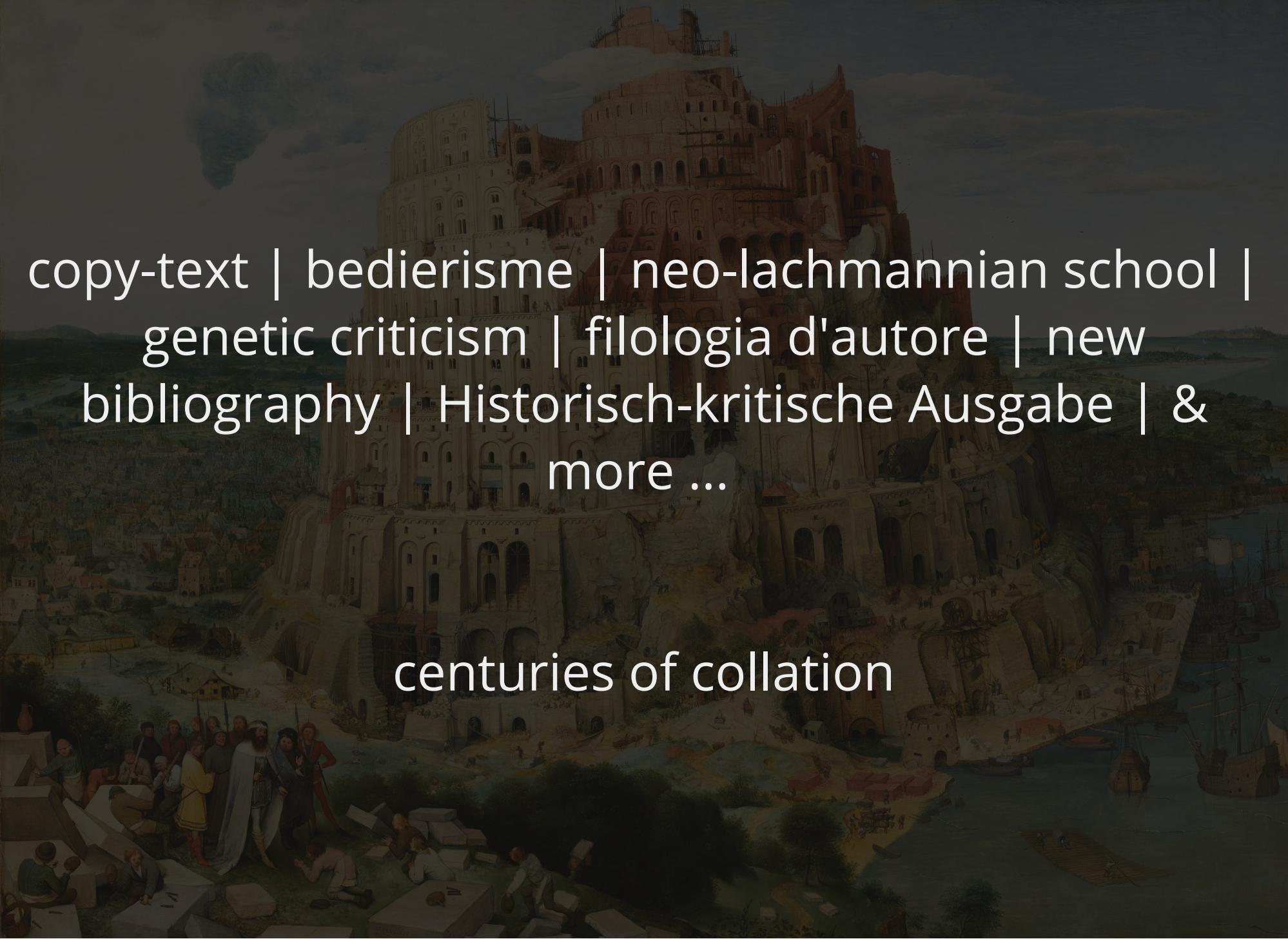


# Raymond Brulez, ms of *Sheherazade* (1932)

Reproduced only for didactic purposes



copy-text | bedierisme | neo-lachmannian school |  
genetic criticism | filologia d'autore | new  
bibliography | Historisch-kritische Ausgabe | &  
more ...

The background of the slide features a detailed painting by Pieter Bruegel the Elder, depicting the Tower of Babel. In the center, a massive multi-tiered tower reaches towards the clouds. The tower is surrounded by a bustling scene of people, animals, and various structures, illustrating the chaotic and diverse nature of the babel mentioned in the Bible.

copy-text | bedierisme | neo-lachmannian school |  
genetic criticism | filologia d'autore | new  
bibliography | Historisch-kritische Ausgabe | &  
more ...

centuries of collation

# How manual collation works

1. choose a base witness
2. compare one by one the other witnesses to the base one (silent or aloud)
3. register important variants

# How manual collation works

1. choose a base witness
2. compare one by one the other witnesses to the base one (silent or aloud)
3. register important variants

tiring, error-prone, difficult to record all data

The beginning of automatic collation



Hinman Collator

See in action!

link and licence on the image



*"Pour les collations  
l'emploi d'un ordinateur  
est dans la plupart de cas  
irréalisable."*

R. Marichal, Préface a *La critique des  
textes et son automatisation*

*"La collation des manuscrits est une opération qui  
relève en quelque sorte de la comptabilité: la machine  
peut fort bien l'exécuter. "*

Froger, Dom Jacques. *La critique des textes et son automatisation*. Dunod, 1968.

Voir article Froger 1965.

8

II | I MA FILLE JE VAIS VOUS DIRE UNE ETRANGE NOUVELLE

卷之三

ପ୍ରକାଶନ ମାଟେ

1

2

3

4

5

6

7

8

9

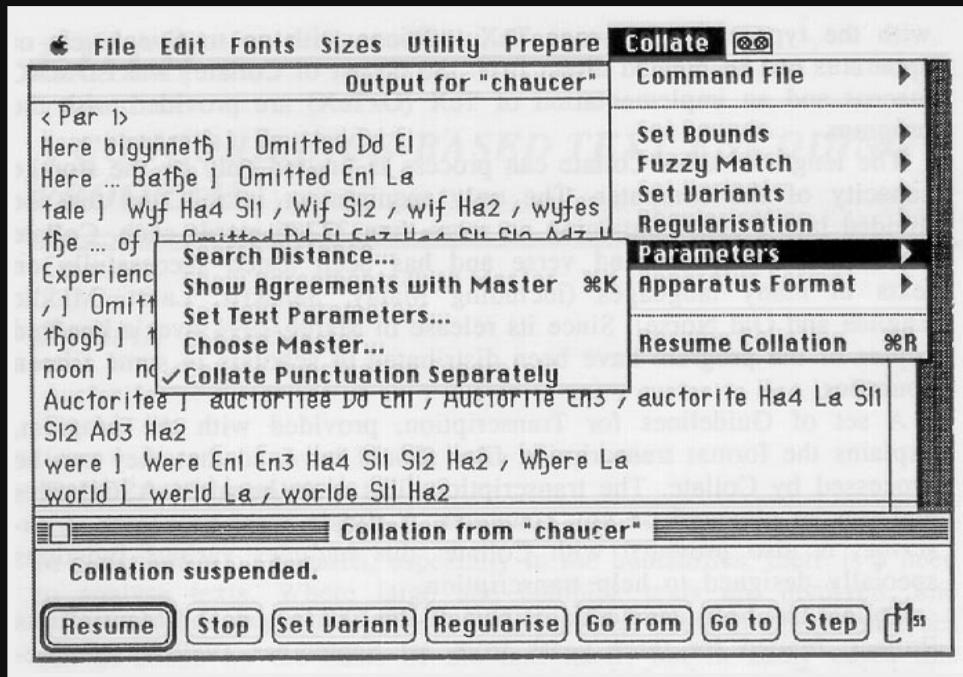
1

1

etc.

Nº de lieu	14	15	16	17	18	19
Texte n° 1	une		étrange		nouvelle	...
Texte n° 2	une		nouvelle		où	...

Nº de lieu	14	15	16	17	18	
Texte n° 1	une		étrange		nouvelle	...
Texte n° 2	une		( )		nouvelle	...

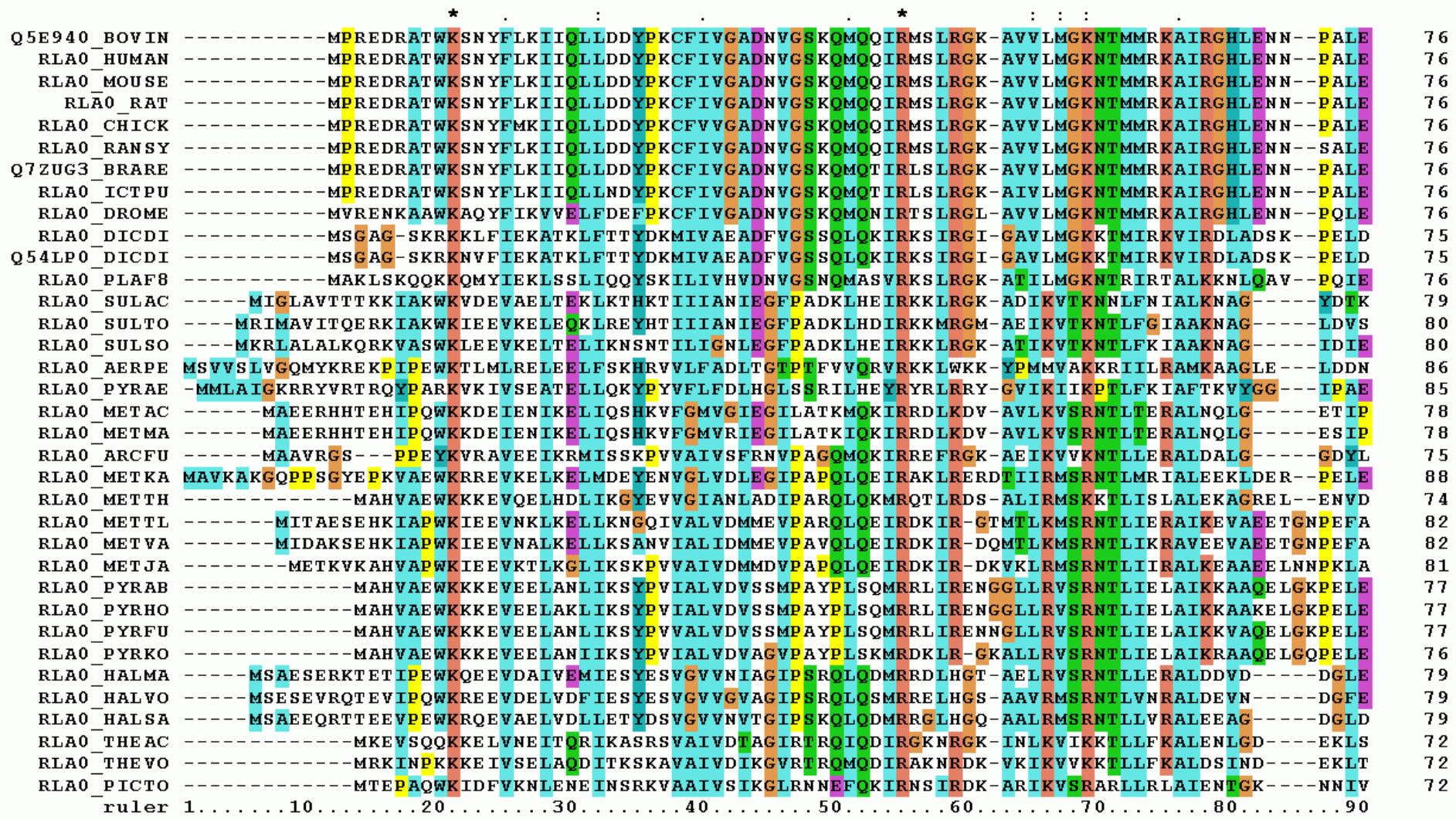


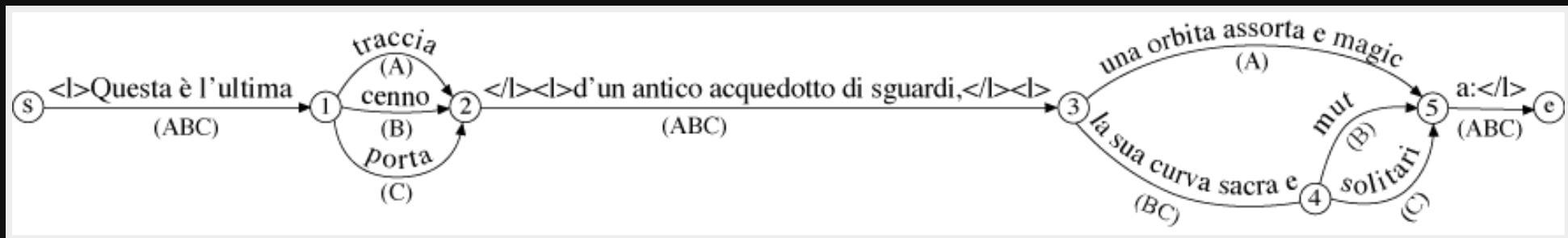
# Collate

Robinson, P. M. W. 1989. "The Collation and Textual Criticism of Icelandic Manuscripts (1): Collation." *Literary and Linguistic Computing* 4 (2): 99–105.

1991, annunciato nella lista di diffusione Humanist  
([dhhumanist.org/Archives/Virginia/v04/1240.html](http://dhhumanist.org/Archives/Virginia/v04/1240.html))

The present of automatic collation





Schmidt, D. and Colomb, R., 2009. "A data structure for representing multi-version texts online". *International Journal of Human-Computer Studies*, 67.6, 497-514.

Schmidt, D., 2009. "Merging Multi-Version Texts: a Generic Solution to the Overlap Problem". In Proceedings of *Balisage: The Markup Conference 2009*. Balisage Series on Markup Technologies, vol. 3 (2009).

# Most used software nowadays

- Juxta  
Commons
- CollateX

These slides reuse materials from

- <http://collatex.net/>
- <https://github.com/DiXiT-eu/collatex-tutorial>
- <https://github.com/interedition/collatex>

Elena Spadini, Intro to automatic collation  
CC-BY-NC-SA