

Collation automatique

Panorama historique et exemple pratique



Contenu



Histoire de la collation automatique

Nury E., Spadini, E. « From giant despair to a new heaven: the early years of automatic collation », *Information Technology* 62, no 2, 61-73.

<https://archive-ouverte.unige.ch/unige:151057>



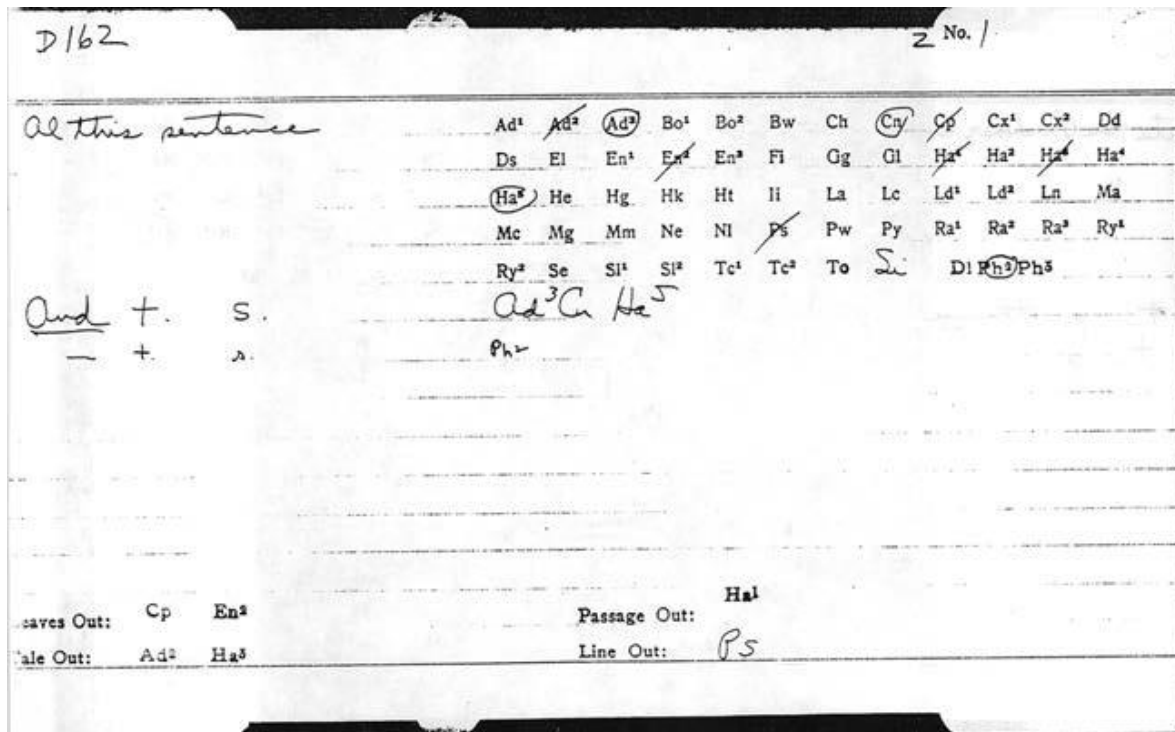
Exemple pratique

Calpurnius Flaccus – comment utiliser les résultats de la collation automatique pour analyser une tradition manuscrite ?

<https://github.com/enury/collation-viz>

1940-1960 – Préhiste

Notations formelles



Bordalejo 2021, Figure 1: Collation card by Manly and Rickert.

Collation optique



Hinman Collator. Folger Shakespeare Library (photographie de Julie Ainsworth) CC-BY-SA 4.0
Le voir en action: <https://vimeo.com/200751944>

Dearing 1962

Whereas in the past, then, textual editors, finding the way rough, got over the stile into By-Path Meadow and, benighted and storm-beaten, were captured by **Giant Despair**... – Dearing 1962, p. 3.

Today, however, the textual editor sees a **new heaven** opening. [...] I tell you with absolute humility that I hear the morning stars sing together and all the sons of God shout for joy. – Dearing 1962, p. 34.



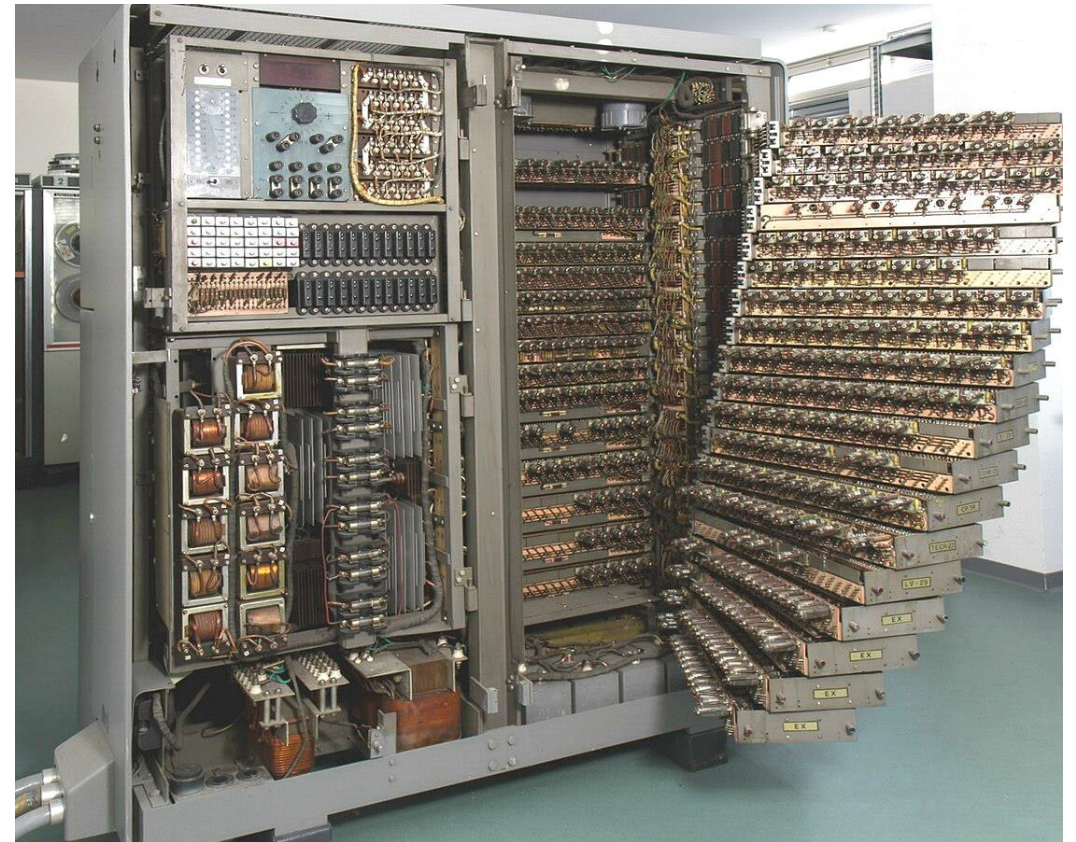
#4736

IBM 7090, modèle utilisé par Dearing.
Courtesy of International Business Machines Corporation, © (1961)
International Business Machines Corporation.

Froger 1968

Pour les collations l'emploi d'un ordinateur est dans la plupart de cas **irréalisable**. – Marichal, préface à *La critique des textes et son automatisation*, p.X

La collation des manuscrits est une opération qui relève en quelque sorte de la **comptabilité**: la machine peut fort bien l'exécuter. – Froger 1968, p. 230



Gamma 3, Bull. Technikum29 computer museum. CC-BY-SA 4.0 International.

Les années 1970

computer-aided critical editions, i. e., the use of the computer to **compare texts or manuscripts** and to indicate variants. – Gilbert 1973, p.139.

Définitions de la collation automatique :

<https://lexiconse.uantwerpen.be/lexicon/collationAutomatic.html>

- Dearing, *Computer Aids to Editing the text of Dryden*.
- Gibson and Petty, *Project OCCULT*.
- Cabaniss, *Using the Computer for Text Collation*.
- Widmann, *The computer in historical collation*.
- *TUSTEP*
- Gilbert, *Automatic Collation*.
- Cannon, *OPCOL*.
- Shillingsburg, *Computer Assistance to Scholarly Editing*.

1980-2000 – collation interactive

‘provided that there is no doubt about the readings in the manuscripts, the collation of manuscripts is a **purely mechanical process**, consisting of merely identifying places where two or more texts do not match’ – Hockey 1980, 144-145

‘It was initially thought that collation was a fairly mechanical process which could be simulated by computer. However, the kinds of **comparisons needed for complex text are much more sophisticated** than those provided by standard software for comparing computer files’ – Hockey 2000, 125.

- Stringer and Vilberg, The Donne Variorum Textual Collation Program.
- Robinson, The Collation and Textual Criticism of Icelandic Manuscripts
- Cannon and Oakman, Interactive Collation on a Micro-Computer: the URICA! Approach.
- Marín and García, Requisitos para la edición crítica informatizada: UNITE.
- Horton, Sequence Comparison and Old-Spelling Texts.

2000-2024



- ChrysoCollate
- CollateX
- Classical Text Editor (CTE)
- Ecdosis
- iAligner
- LERA / LAKomp
- Saktumiva
- TEI Comparator
- TRAViz

Limités à une tradition:

- Capitularia
- Prahbed



- Juxta
- Variance viewer

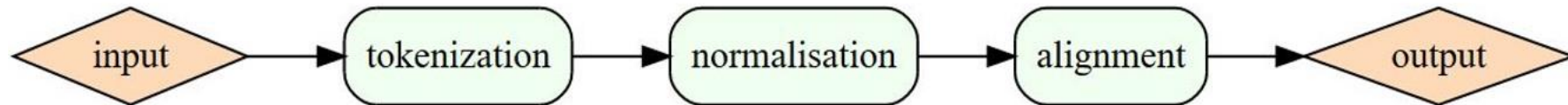
Thèses / non publié :

- Collator
- Ncritic
- Roelli & Bachmann
- Salemans

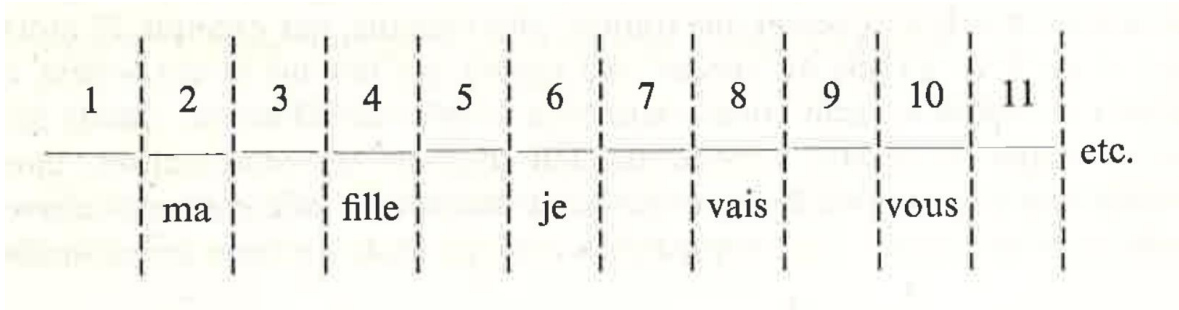
Principes de base

Le modèle de Gothenburg (2009)

- Tokenisation
- (Normalisation)
- Alignement
- (Analyse)
- Visualisation / output



Tokenisation



Diviser le textes en unites de comparaison

- Caractères
- Mots
- Phrases

Limitation: gestion d'éléments non-textuels.

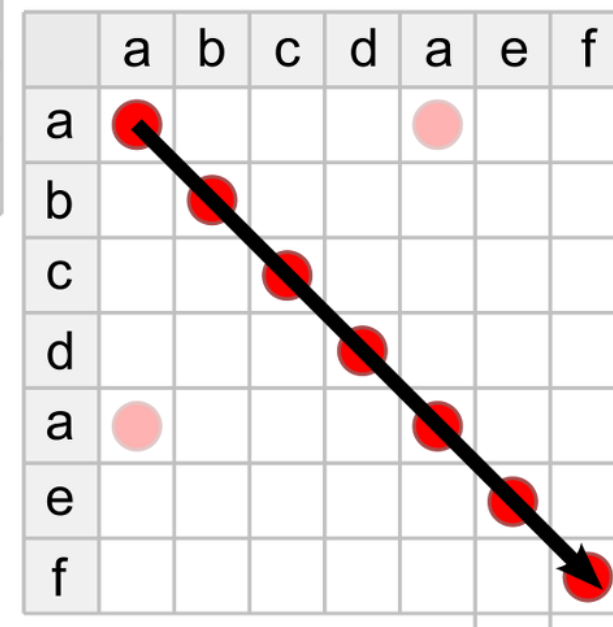
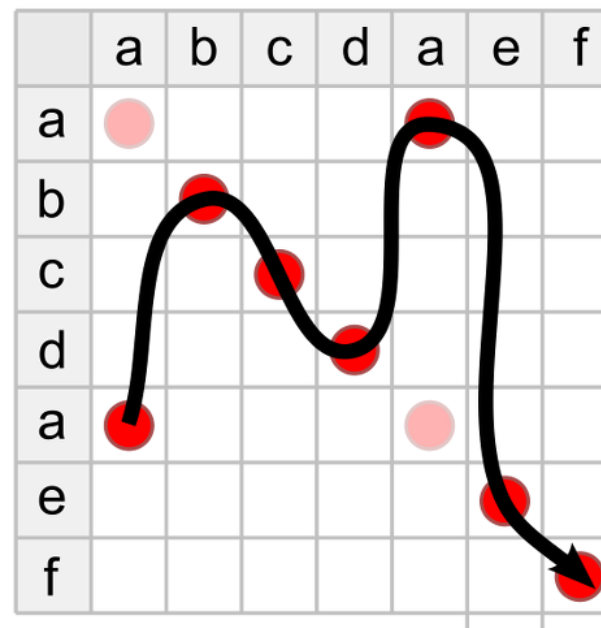
Alignement

N° de lieu	14	15	16	17	18	19
Texte n° 1	une		étrange		nouvelle	...
Texte n° 2	une		nouvelle		où	...

Diagram illustrating word alignment between two texts. Text 1 (Texte n° 1) and Text 2 (Texte n° 2) are shown with their respective words and positions (N° de lieu). The alignment is indicated by arrows: a vertical arrow between 'une' at position 14, and a diagonal arrow between 'étrange' at position 16 and 'où' at position 18.

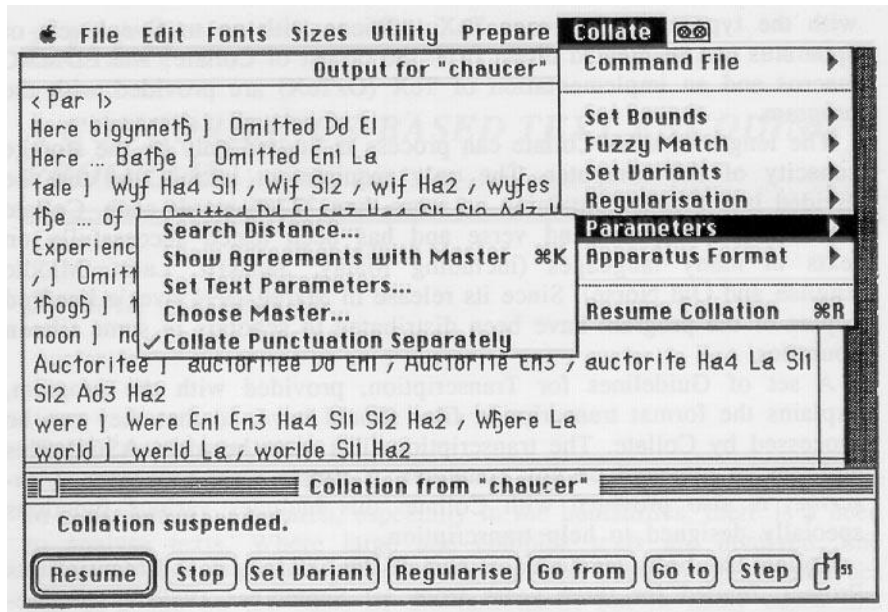
N° de lieu	14	15	16	17	18
Texte n° 1	une		étrange		nouvelle
Texte n° 2	une		()		nouvelle

Diagram illustrating word alignment between two texts. Text 1 (Texte n° 1) and Text 2 (Texte n° 2) are shown with their respective words and positions (N° de lieu). The alignment is indicated by a vertical arrow between 'une' at position 14 and a diagonal arrow between 'étrange' at position 16 and '()' at position 17.



Normalisation

Régularisation des tokens

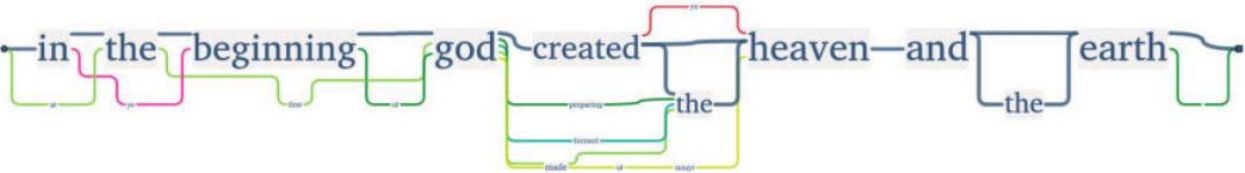
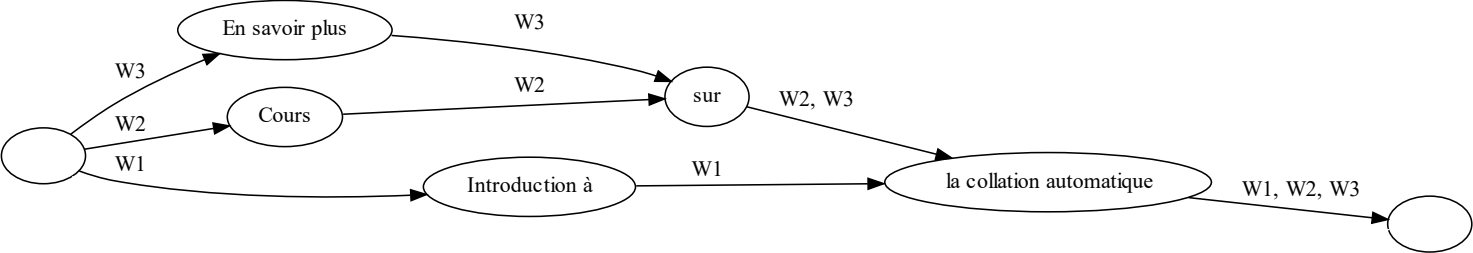
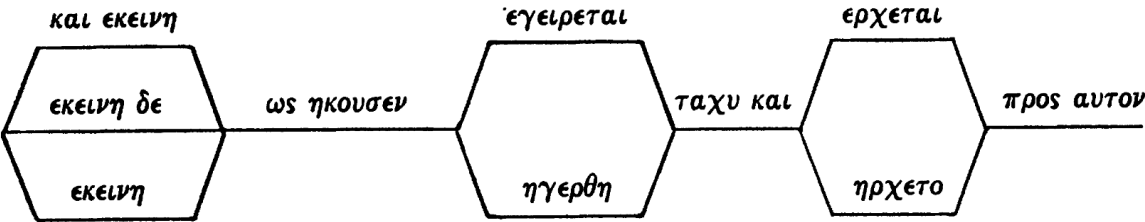


Correspondance approximative

W1	le	petit	chien	noir
W2	le	chiot		noir

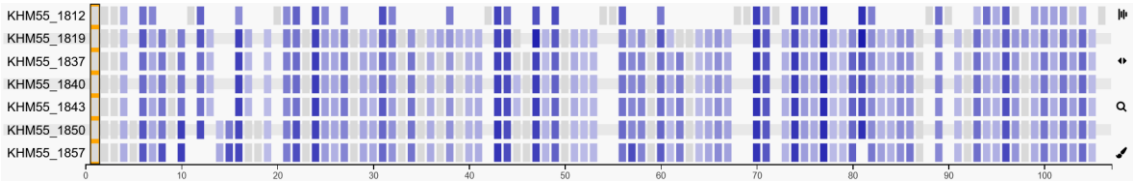
W1	le	petit	chien	noir
W2	le		chiot	noir

Visualisations



2179 And this ditty after mee, Sing; and daunce it trippingly.
q2 Ditty me, Sing dance
f1 Ditty me, sing dance trippinglie.
f2 Ditty me, sing dance trippinglye.
f3 Ditty me, sing dance
f4 Ditty me, sing dance trippingly.> w
r1 Ditty me, Sing Dance
r2 Ditty me, Sing Dance
r3 Ditty me>w
p1 ditty me>w
p2 ditty me>w
t1 ditty me>w
t2 ditty me>w
h1 ditty me>w
wa Sing, and dance it trippingly.> w
j1 ditty me>w
ca ditty, me,> w
h3 ditty, me,> w
ma ditty, me,> w

Fletus passeris Lesbie
Passer delicie mee puelle
Qui cum ludere que in sinu tene
Qui primu digitum dare at petenti
Et acris solet incitare morsus
Cum desiderio meo nitenti
Carum nescio quid libet iocari
Et solaciolum sui doloris
Credo ut tu gravis adquiescet ardor
Tecum ludere sicut ipsa possem
Et tristis animi levare curas
Tam gratum est michi q ferut puelle
Pernici aureolum fuisse malum
Qd conam soluit diu negatam



Exemple pratique

Calpurnius Flaccus & PyCoviz

Méthode de Lachmann

Erreurs communes

- Leçons partagées par un groupe de témoins
- ET absentes des autres témoins
- ET n'ont pas d'autorité manuscrite

Erreurs uniques

- Leçons uniques à un témoin
- ET n'auraient pas pu être corrigées aisément

Paratexte

The collation of a manuscript frequently requires quite **detailed notes**. It cannot be reduced to the bare recording in a register of the presence or absence of precisely defined readings – Whittaker 1991.

Ex: localisation dans le manuscrit (folio/page), passage difficile, hésitations, résolutions d'abréviations.

Les besoins d'un éditeur

Comparaisons

- Voir leçons uniques d'un témoin
- Voir leçons différentes entre deux témoins
- Voir leçons communes entre 2+ témoins et différentes chez les autres témoins

Autres

- Visualiser une grande quantité de données, dont des éléments paratextuels
- Partager le résultat
- permettre une évaluation et/ou reproduction par les pairs

Mise en pratique: PyCoviz

PyCoviz: Visualizing Collation Results

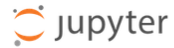
DOI [10.5281/zenodo.1186178](https://doi.org/10.5281/zenodo.1186178)

launch [binder](#)

Visualization and interactive manipulation of the collation results from [CollateX](#) with a jupyter notebook.

<https://github.com/enury/collation-viz/>

Jupyter Notebook (mode app)

[Edit App](#)[Logout](#)

Possible interactions are grouped by category:

1. **Extract:** view a collation extract.
2. **Modifications:** update the collation by adding/deleting rows, moving tokens up or down, or adding/deleting notes.
3. **Agreements:** select groups of witnesses to see what readings they have in common, against other witnesses.
Use cases: (1) selecting only witnesses from the first dropdown menu will show by default the readings they have in common against all others; (2) as a result, select only one witness in the first menu to see its unique readings. (3) Compare groups of witnesses: for instance, what do B, C and P1594 have in common against M and N? What do M, N and P1594 have in common against B and C?
4. **Search:** search the collation for a single reading, for which you can see more information in the last page.
5. **Clarify:** see more information about one reading, i.e. one cell in the table (defined by a witness and a specific row).

Extract

Modifications

Find Agreements

Search

Clarify

From: 6

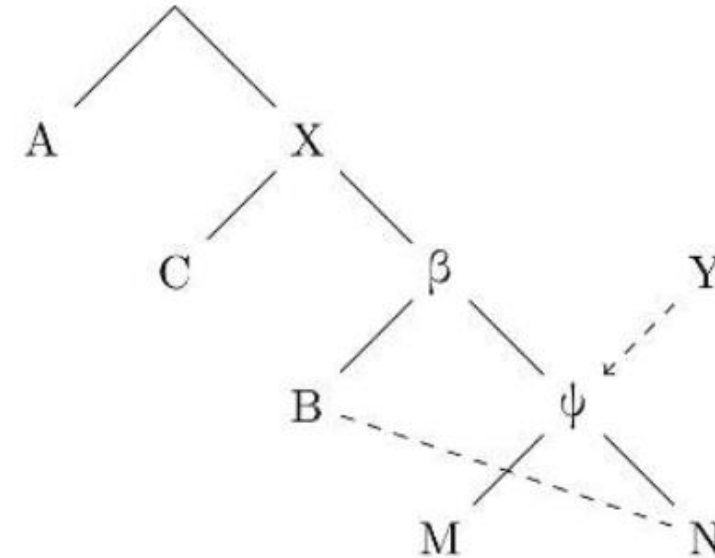
To: 11

B1	B2	C1	C2	LH	M1	M2	N1	N2	P1594	ID
quinque cum tyranno	quinque cum tyranno	quinque cum tyranno	quinque cum tyranno	quinque cum tyranno	quinque cum tyranno	quinque cum tyranno	quinque cum tyranno	quinque cum tyranno	Quinque cum tyranno	6
							familiae			7
proxime	proxime	proxime	proxime	proximi	proximae	proximae	proximae	proximae	proximae	8
familię	familię	familię	familię	familiae	familiae	familiae		familiae	familiae	9
puniantur	puniantur	puniantur	puniantur	puniantur	puniantur	puniantur	puniantur	puniantur	puniantur	10
quędam	quędam				quaedam	quaedam	quaedam	quaedam		11

<https://github.com/enury/collation-viz/>

Déclmations de Calpurnius Flaccus

- 5 manuscrits, prose en latin classique
 - Un manuscrit perdu (X)
 - *Editio princeps* de Pierre Pithou 1580/1594
 - Håkanson 1978
-
- Quelles sont les relations entre l'*editio princeps* et les manuscrits ?
 - Quel témoin est la source des corrections de N2 ?
 - Qu'est-ce qui sépare C des autres manuscrits ?



Comparaisons

Leçons uniques de Pithou

Extract

Modifications

Find Agreements

Search

Clarify

Agreements:

M1

M2

N1

N2

P1594

Against:

B1

B2

C1

C2

LH

No witness selected.

B1	B2	C1	C2	LH	M1	M2	N1	N2	P1594	ID
at	at	at	at	et	at	at	at	at		13
voluit	voluit	voluit	voluit	voluit	voluit	voluit	voluit	voluit	noluit	53
alte	alte	alte	alte	alte	alte	alte	alte	alte	ita	57
iam	iam	iam	iam	iam	iam	iam	iam	iam		69

Corrections de N2

Extract

Modifications

Find Agreements

Search

Clarify

Agreements:

M1

M2

N1

N2

P1594

Against:

M1

M2

N1

N2

P1594

B1	B2	C1	C2	LH	M1	M2	N1	N2	P1594	ID
							familiae			7
familię	familię	familię	familię	familiae	familiae	familiae		familiae	familiae	9
ethiopus	ethiopus	ęthiopus	ęthiopus	Aethiops	aethiops	aethiopus	aethiops	aethiopus	Aethiops	22

Comparaisons

Pithou et C

B1	B2	C1	C2	LH	M1	M2	N1	N2	P1594	ID
decem	decem	x	x						x	2
quēdam	quēdam				quaedam	quaedam	quaedam	quaedam		11
si	si	sic	sic	sic	si	si	si	si	sic	203
oculum	oculum	oculorum	oculorum	oculorum	oculum	oculum	oculum	oculum	oculorum	1145

Pithou et N

B1	B2	C1	C2	LH	M1	M2	N1	N2	P1594	ID
erant	erant	erant	erant	erant	erant	erant				205
							erant	erant	erant	207
me	me	me	me	me	me	me				614
							me	me	me	616
dixit	dixit	dixit	dixit	dicit	dixit	dixit	dicit	dicit	dicit	810
eiusmodi	eiusmodi	eiusmodi	eiusmodi	eiusmodi	eiusmodi	eiusmodi	huiusmodi	huiusmodi	huiusmodi	971
misere	misere	misere	misere	miserae	misere	misere	miserae	miserae	miserae	1017
reddidit	reddidit	reddidit	reddidit	reddidit	reddidit	reddidit	reddit	reddit	reddit	1334
explicitē ex calpurnio flacco excerptē	explicitē ex calpurnio flacco excerptē	explicitē ex calpurnio flacco excerptē	explicitē ex calpurnio flacco excerptē		explicitae ex calpurnio flacco excerptae	explicitae ex calpurnio flacco excerptae				1356

Total: 9 rows in the table.

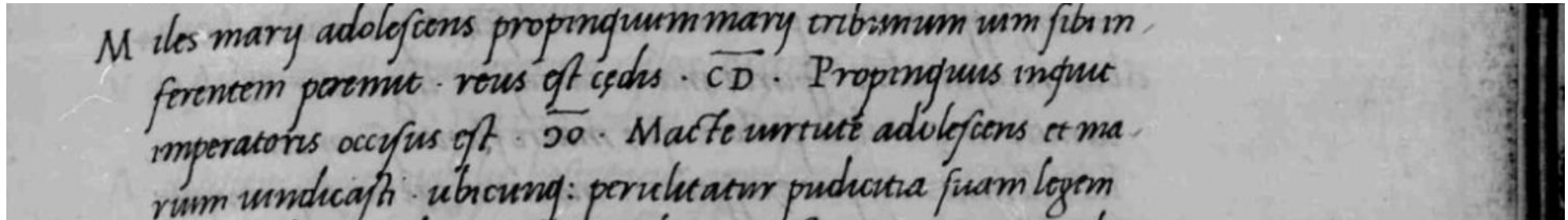
Paratexte

Extract	Modifications	Find Agreements	Search	Clarify
---------	---------------	-----------------	--------	---------

ID:

Witness:

Ø : 3, Unknown abbreviation. Normalized form supplied by Lehnert., 148r:8, 30, contra, http://daten.digital-sammlungen.de/bsb00090859/image_296





UNIVERSITÉ
DE GENÈVE



Fonds national
suisse

grammateus

Merci de votre attention !

elisa.nury@unige.ch