

# **From Stylistics to Statistics: Computational Approaches to Latin Medieval Composition and Authorship**

Ghent, 27 November 2024

LW Research Day 2024 – From Source to Understanding

# From source to understanding

- Stylometry applied to medieval Latin: a **long bridge / distance?**
- **Source**
  - Dictation / self-dictation
  - Wax tablets /notes
  - Manuscripts
- **Understanding (hermeneutics)**
  - Edition
  - Digitized text (OCR)
  - Electronic text
  - Vectors ( = numerical representations of lexical properties)

# Stylometry

# History

“One class of words that we use has been called **function words**—the filler words of the language, such as *a*, *an*, *by*, *to*, and *that*. Generally they include prepositions, conjunctions, pronouns, and certain adverbs, adjectives, and auxiliary verbs. [...]

We like the function words rather well because many of them are not much influenced by the context of the writing [...].”



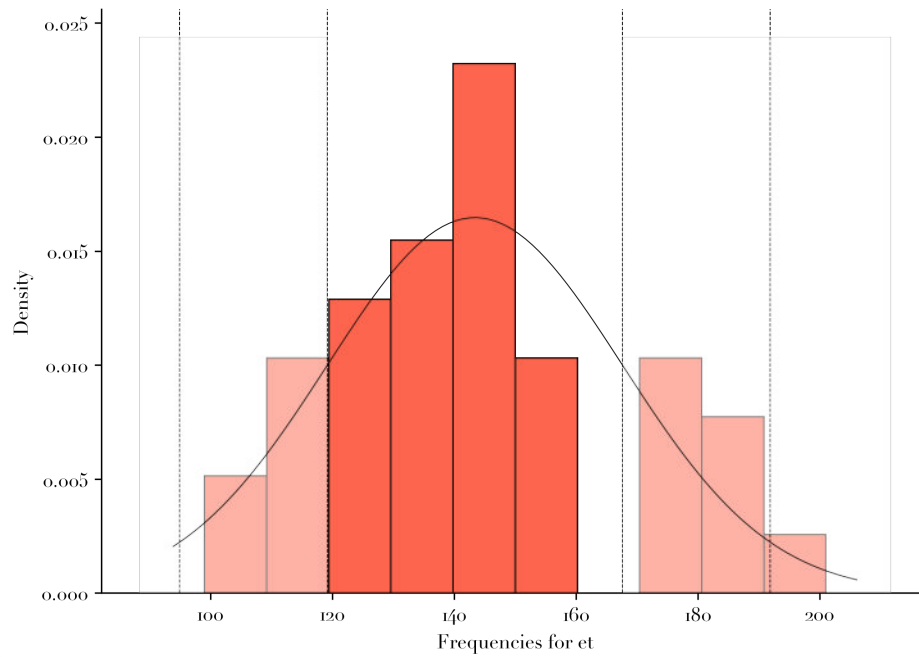
The Wallaces [left] with the Mostellers [right] at the University of Chicago in 1973.



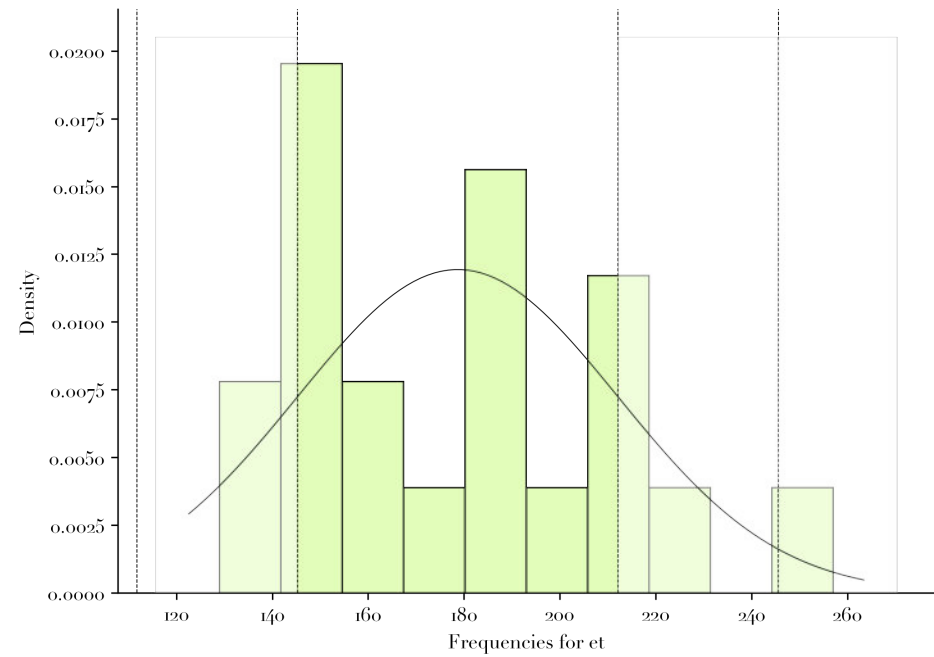
Example of IBM 7090.

# Features

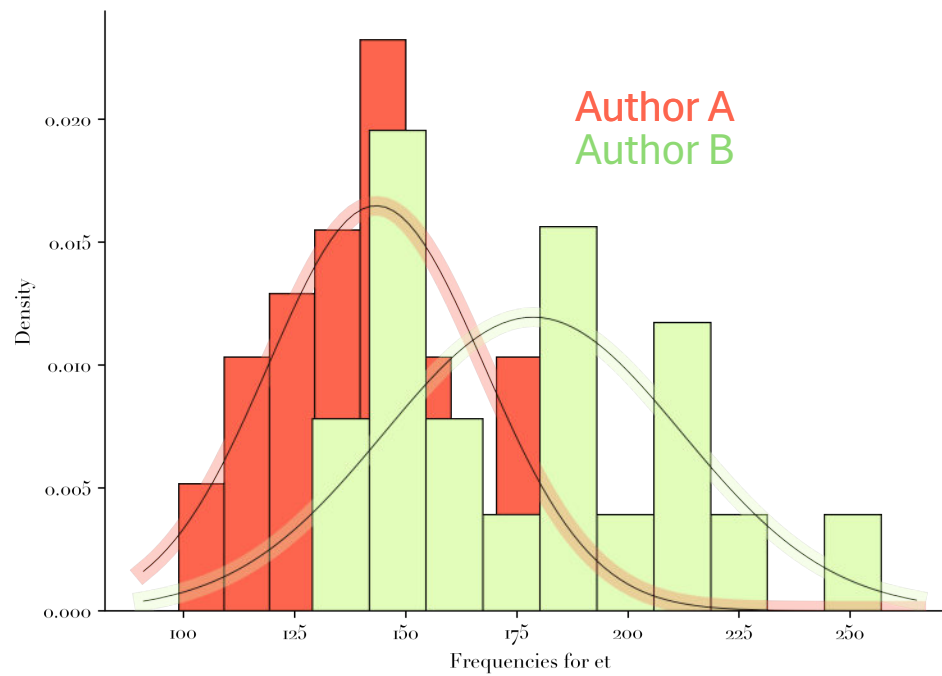
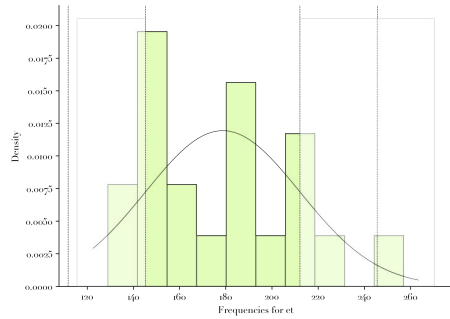
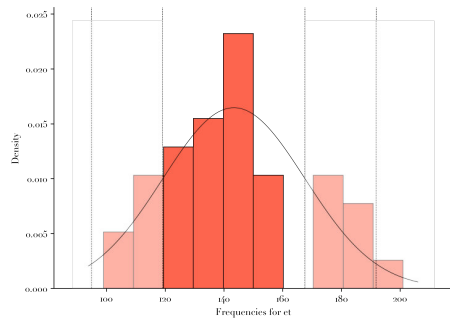
## Author A



## Author B



# Features



# Features and Vectors

	AUTHOR A			AUTHOR B		
	segment 1	segment 2	segment 3	segment 4	segment 5	segment 6
et	116	114	131	130	123	104
in	156	119	134	130	88	80
est	52	37	45	48	31	55
ad	46	29	43	63	43	28
qui	19	11	12	22	14	18
cum	30	62	58	82	45	66
ut	31	13	27	36	30	19
me	33	22	30	33	34	43
non	20	35	30	29	32	21
	41	45	23	32	22	41

# Features and Vectors

	AUTHOR A		AUTHOR B
	segment 1		segment 4
et	116		130
in	156		130
est	52		48
ad	46		63
qui	19		22
cum	30		82
ut	31		36
me	33		33
non	20		29
e	41		32



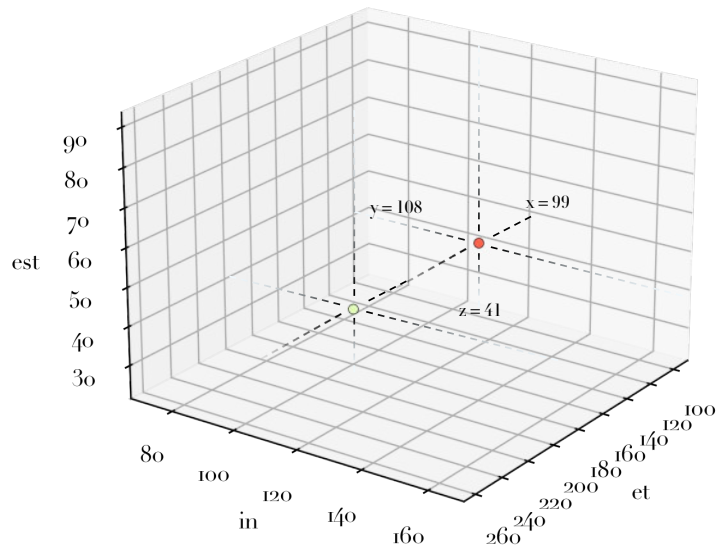
$n$  ROWS  
 $n$  variables  
= features  
= dimensions

$n$  COLUMNS  
 $n$  samples  
= observations  
= instances

Jeroen De Gussem

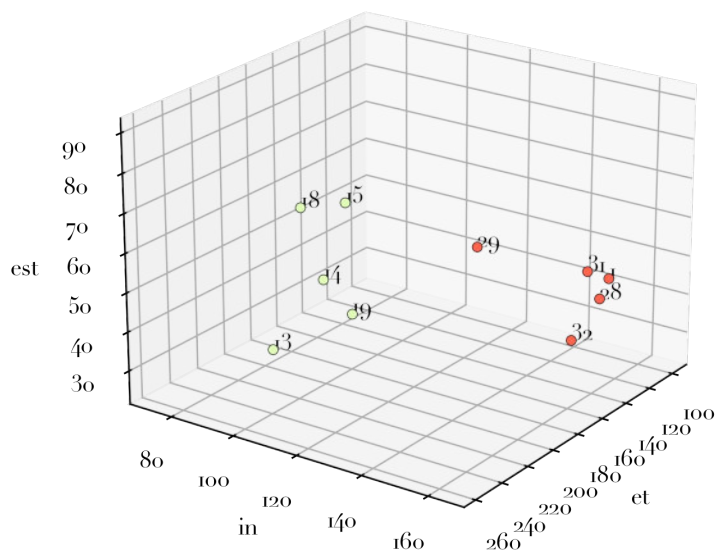


# Features



	AUTHORA		AUTHOR B	
	segment 1		segment 4	
et	116	<i>difference distance</i> 	130	 <i>n</i> ROWS <i>n</i> variables = features = dimensions
in	156		130	
est	52		48	
ad	46		63	
qui	19		22	
cum	30		82	
ut	31		36	
me	33		33	
non	20		29	
de	41		32	
			 <i>n</i> COLUMNS <i>n</i> samples = observations = instances	

# Features

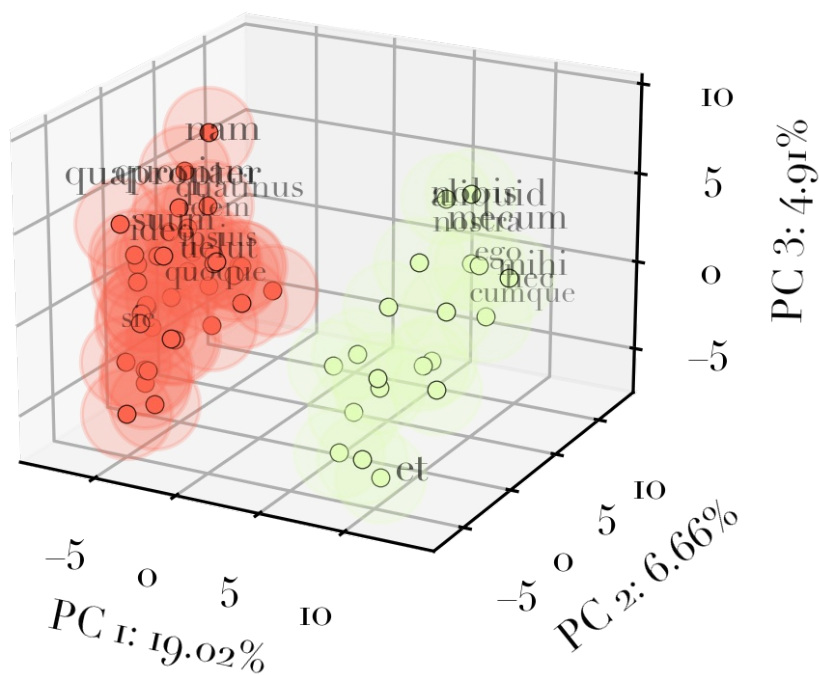


	AUTHOR A		AUTHOR B
	segment 1		segment 4
et	116	<i>difference distance</i> 	130
in	156		130
est	52		48
ad	46		63
qui	19		22
cum	30		82
ut	31		36
me	33		33
non	20		29
de	41		32

*n* ROWS  
*n* variables  
 = features  
 = dimensions

*n* COLUMNS  
*n* samples  
 = observations  
 = instances

# Principal Components Analysis (PCA)



	AUTHORA			AUTHORB		
	segment 1	segment 2	segment 3	segment 4	segment 5	segment 6
et	116	114	131	130	123	104
in	156	119	134	130	88	80
est	52	37	45	48	31	55
ad	46	29	43	63	43	28
qui	19	11	12	22	14	18
cum	30	62	58	82	45	66
ut	31	13	27	36	30	19
me	33	22	30	33	34	43
non	20	35	30	29	32	21
de	41	45	23	32	22	41

# Individual vs. Shared Authorship

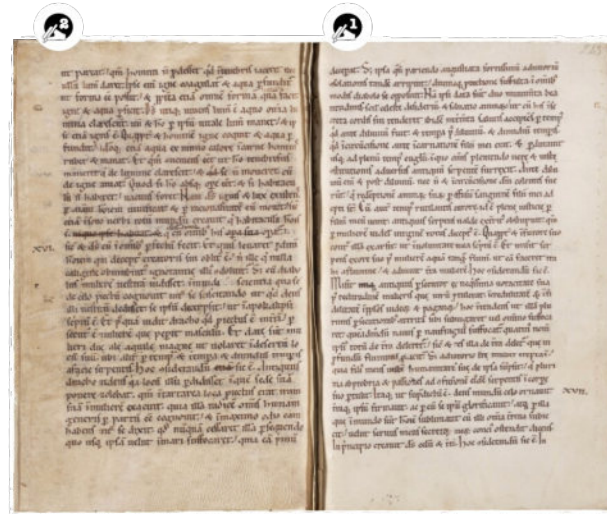
- **Individual authorship:** writing alone in the Latin Middle Ages?
  - difficult and technical
  - slow, physically demanding
  - expensive
  - autographs (copy where roles of author and scribe combined in one person) = very rare
- **Shared authorship:** contributors (secretaries) and scribes
  - Secretaries and scribes
  - Writing personnel attached to scriptoria / chanceries (official documents)
  - Involved in intellectual, spiritual, and administrative life
  - Proficiency in Latin, book learning and technical know-how in copying and manuscript production.
- **Every scribe a new author, every copy a new text?**

# Hildegard of Bingen & Ghent MS 241

# Hildegard of Bingen (1098–1179) and Ghent MS 241

irrationale aut ayal nichil aliud scit. n qd semu...  
**E**t in hōie id ē iuuenti scientia hōis uim & forti-  
 tudine pspicue mltitue creat. ita qd nec i seipso nec i aliis  
 iquirari nullatenē prauē cedat. qd qsi uirile ē. In eo qd erat  
 qd p donū diuine grē hōi i peccati uoluntate cū mīa parcat.  
 & ut miserial ipse ira adtendat. <sup>quatuor</sup> uimū penitentię ei ifundat  
 ipsiq; oleo mię unguat. sic ne ipse hō ppter excedentē mīsurā &  
 irremissibile penitentiā cadat. & ne repidus i uanitatē prauē  
 opū iuoluat. qd qsi femineū ē. Et istud dō benedicit qā h u

ut sit ei i cibū refectio nūc aīe. ita ut i omī uiuentes uirtu-  
 tes p humilitatē dō subiecte. & uirtutes i celesti milicia uolan-  
 tes. & uirtutes que hōie a terrenis reb' ad celestia remouēt.  
 atq; omī uirtutes i quib' uiuentes uirtes de plantatione spe-  
 sci sūt. cū eo i aīa ipse pascant. & ipse i cū illis i hīs omībus  
 pascant. Et hec omīa i hōie illo ita fiūt. q' sic i dō studet.  
 Viditq; dō cuncta que fecerat. & erat ualde bona. Tunc  
 i hōie dō qd omīa que spe sēs tribuit ualde bona sūt. qā  
 i plenitudine omīū uirtutū pfecta sūt. cū pri' una queq; uir-  
 tutū i semetipsa tantū unū bonū fiūt. hic aut' omīa parit  
 bona sūt. qā hōe omīa bona i hōie illo simul apparendo im-



1  
 peccatoꝝ

2  
 peccatoꝝ



Hildegard of Bingen receiving inspiration from the Holy Ghost, and engraving her ideas on a wax tablet. She is assisted by a male monk transcribing her words to parchment (left) and a female nun (right). MS Lucca 1942, Biblioteca Statale, fol. 1v.

# Hildegard of Bingen (1098–1179) and Ghent MS 241

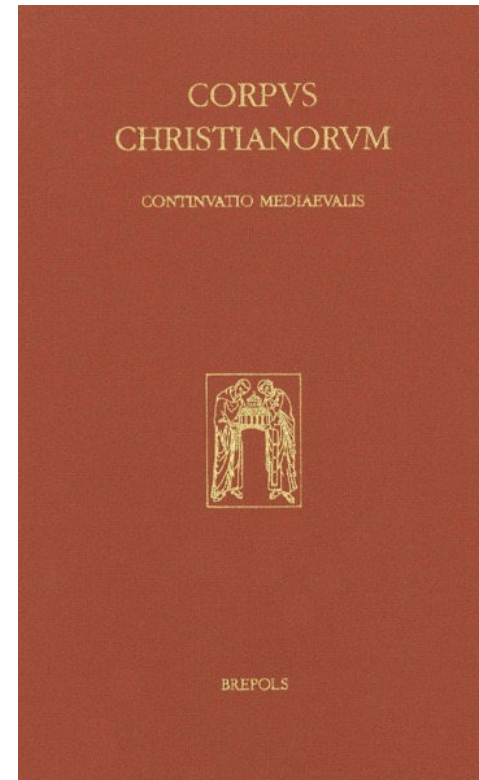
#	<i>ante corr.</i>	<i>post corr.</i>	<i>freq.</i>	#	<i>ante corr.</i>	<i>post corr.</i>	<i>freq.</i>
1.	quod	ut	392	11.	cum	-	22
2.	ac	et	207	12.	prefato	-	22
3.	que	-	73	13.	autem	sententie	21
4.	quod	-	70	14.	quod	quia	19
5.	-	et	69	15.	prefati	-	19
6.	etiam	-	34	16.	quod	quatenus	14
7.	ut	-	26	17.	que	quoque	14
8.	qui	-	26	18.	est	sit	13
9.	ad intellectum	intellectui	24	19.	predicto	-	13
10.	quoque	-	24	20.	-	que	12

1		11	
2		12	
3		13	
4		14	
5		15	
6		16	
7		17	
8		18	
9		19	
10		20	

# Two recensions of the *LDO*



ante correctionem



post correctionem



# Reconstruction to *ante correctionem*

*Sed et ante faciem eiusdem imaginis quasi tabula ut cristallus perlucida appareret, in qua scriptum erat: Pulchram formam argentei coloris manifestabo, quia diuinitas, quae initio caret, magnam claritatem habet; sed unumquodque quod initium habet ambiguum in terroribus est nec secreta Dei in plena scientia capere potest.*

*Et imago haec praedictam tabulam inspiciebat. Vnde et linea in qua sedebat mouebatur; et mox ubi eadem linea praedictae rotae in sinistra eius parte coniuncta uidebatur, exterior pars ipsius rotae per breue spatium aliquantum aquosa, et deinde aliquantum ultra medietatem medietatis eiusdem rotae, quae sub praefata linea in transversum ducebatur, rubicunda, et postmodum pura et lucida, et tandem uelut turbida et procelliosa tempestas efficiebatur, scilicet prope finem eiusdem medietatis ubi praefata linea eidem rotae affixa erat.*

Audiuique uocem de celo michi dicentem:

II. O homo, audi et intellige uerba illius qui erat et qui est absque officio mutabilitatis temporum; quia in ipso hoc antiquum consilium fuit, quod diuersa opera facere uoluit, et haec quasi radium solis ante antiquitatem dierum inspexit, quoniam futura erant. Nam Deus unus est et huic unitati nichil se conferre potest; sed ipse praesciuit quia quoddam opus, quod facturus erat, similitudinem huius unitatis sibi usurpare temptaret. Vnde et illi obstaculum repercussionis opposuit, quoniam ipse illa unitas est, quae nullum sibi similem habet: alioquin ipse unitas nominari non posset. Et ideo etiam illum a se abiecit qui similitudinem hanc peruerse appetit; et sic quelibet rationalis anima in homine, quae de illo qui uerus Deus est existit, ut quod sibi placet eligat et quod sibi displicet abiciat, quia quid bonum seu quid nocuum sibi sit cognoscit. Sed quamuis Deus unus sit, in ui tamen cordis sui opus quoddam praesciuit, quod magnifice

24 tabula ut cristallus] cfr Hildeg., *ibidem*: tabula eburnea

25 appa[-ruit\rebat] C2 31/33 rotę in sinistra - aquosa: in ras. C1 35  
1<sup>o</sup>: ducta [-fuit\erat] C2; 2<sup>o</sup>: [-ducta erat\ducebatur] C2 [ac>et] postmodum  
36 tempestas \efficiebatur/ C2

2. 5/6 [ac>et] huic unitati [-nulla similitudo adest\nichil se conferre potest]  
C2 6 [-quod\quia] quoddam C2 6/8 quod facturus - temptaret: *part. in*  
*ras., part. in marg.* 11 [-habere cupit\peruerse appetit] C2 [ac>et] sic 12  
qu[i>e] de illo 12/13 \ut/ quod ... eligi\at] et ... abici\at

33 aliquantum: aliquantulum LM 37 eadem rotę: eiusdem r. LM 39  
Audiuique: *hic incipit cap. 2 in T*

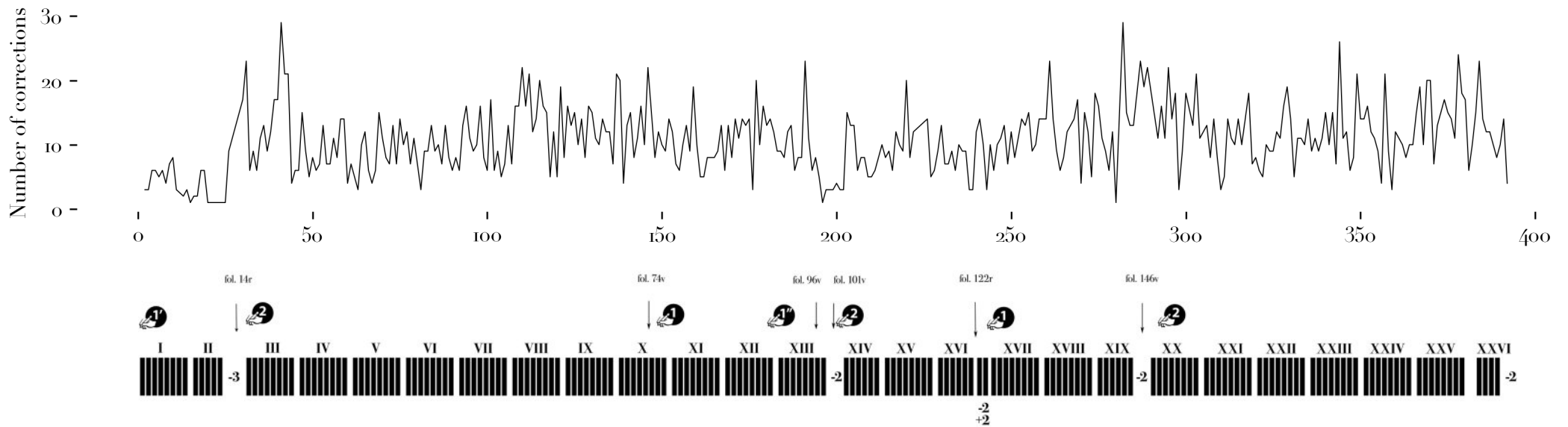
24 tabula ut cristallus] cfr Hildeg., *ibidem*: tabula eburnea

25 appa[-ruit\rebat] C2 31/33 rotę in sinistra - aquosa: in ras. C1 35  
1<sup>o</sup>: ducta [-fuit\erat] C2; 2<sup>o</sup>: [-ducta erat\ducebatur] C2 [ac>et] postmodum  
36 tempestas \efficiebatur/ C2

2. 5/6 [ac>et] huic unitati [-nulla similitudo adest\nichil se conferre potest]  
C2 6 [-quod\quia] quoddam C2 6/8 quod facturus - temptaret: *part. in*  
*ras., part. in marg.* 11 [-habere cupit\peruerse appetit] C2 [ac>et] sic 12  
qu[i>e] de illo 12/13 \ut/ quod ... eligi\at] et ... abici\at

33 aliquantum: aliquantulum LM 37 eadem rotę: eiusdem r. LM 39  
Audiuique: *hic incipit cap. 2 in T*

# Hildegard of Bingen (1098–1179) and Ghent MS 241



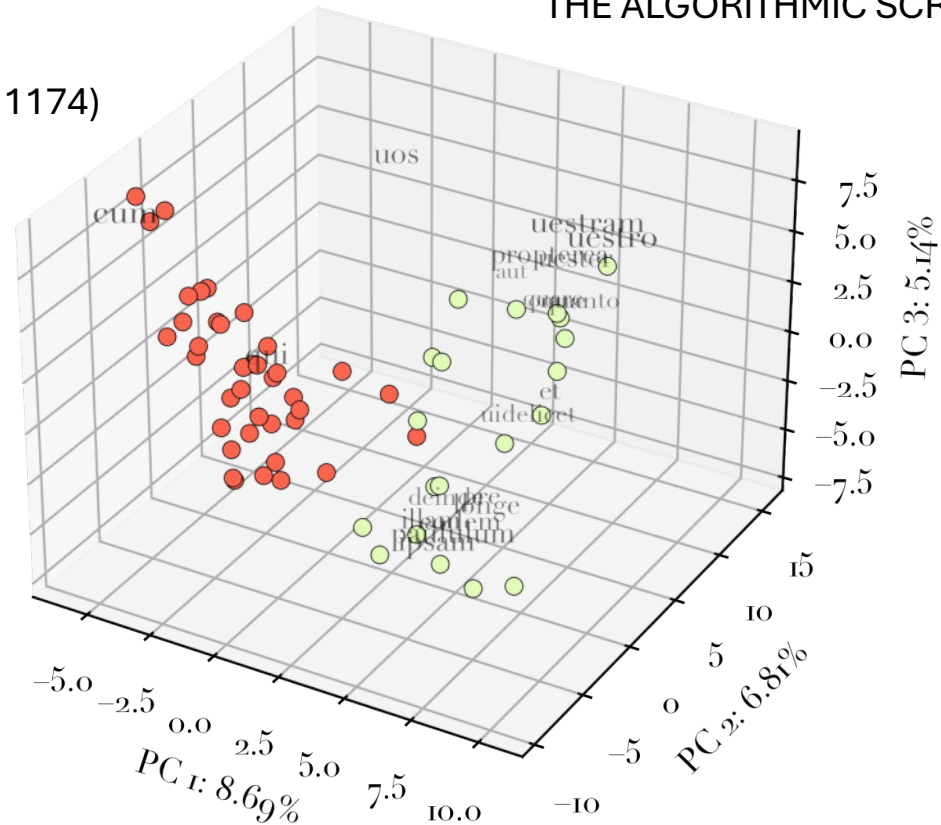
# The Algorithmic Scribe

# Hildegard of Bingen (1098–1179) and Ghent MS 241

- Hildegard of Bingen  
*Liber diuinorum operum* (finished c. 1174)
- Elisabeth of Schönau

0 corrections

THE ALGORITHMIC SCRIBE

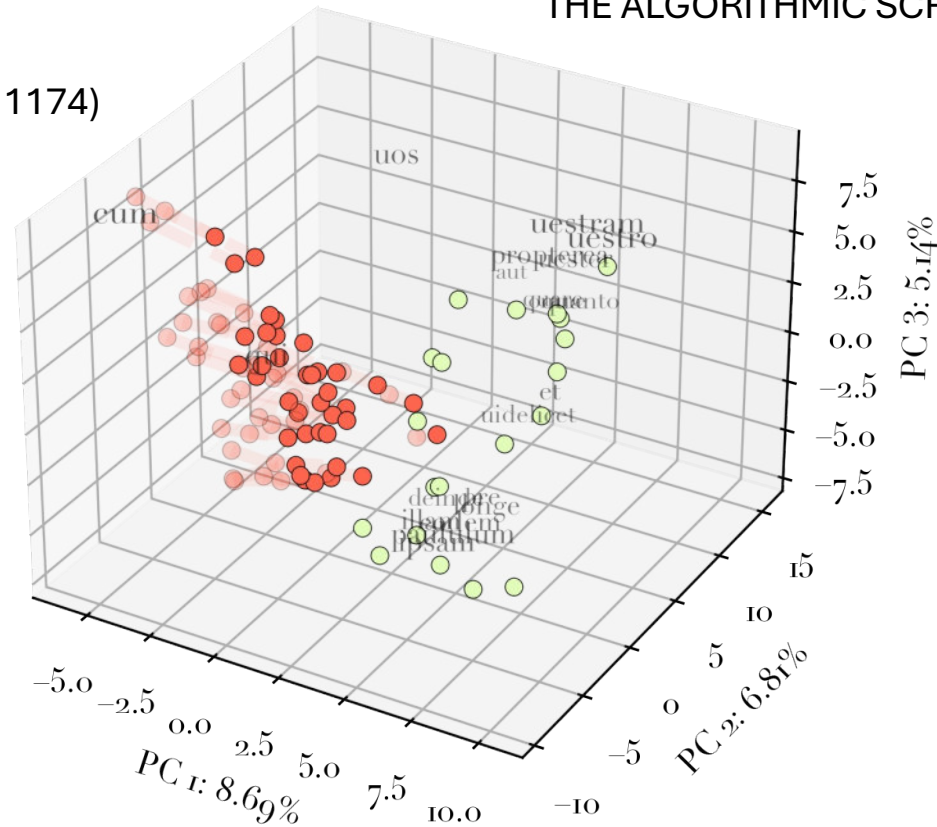


# Hildegard of Bingen (1098–1179) and Ghent MS 241

- Hildegard of Bingen  
*Liber diuinorum operum* (finished c. 1174)
- Elisabeth of Schönau

142 corrections  
/ 3,000 words

THE ALGORITHMIC SCRIBE

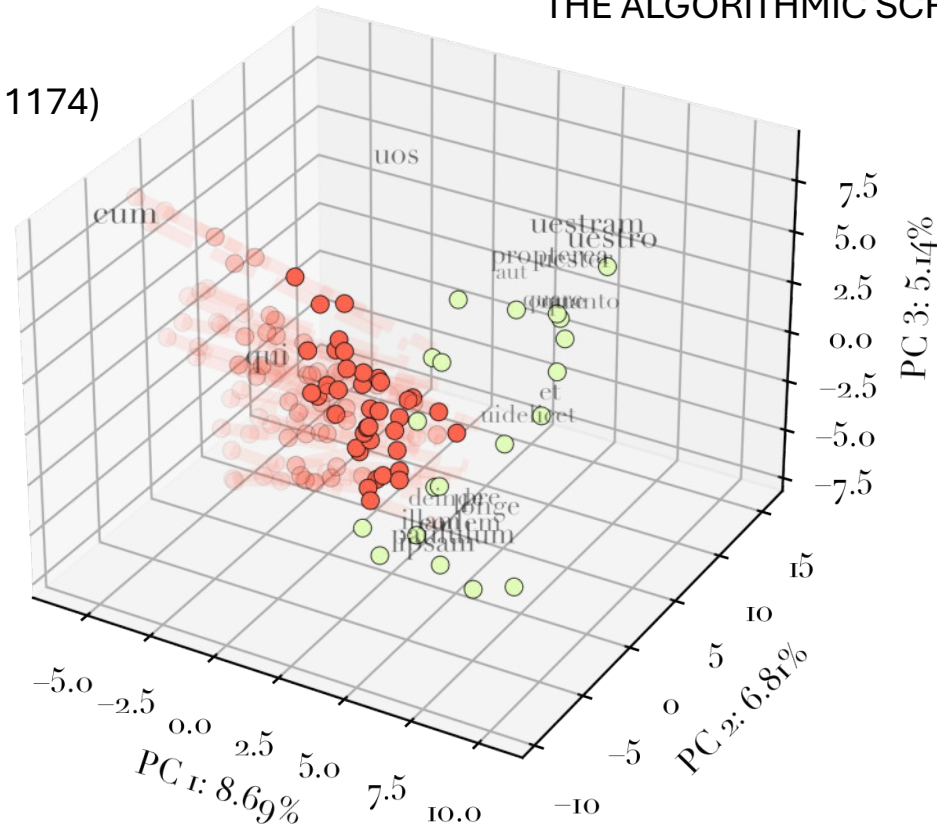


# Hildegard of Bingen (1098–1179) and Ghent MS 241

- Hildegard of Bingen  
*Liber diuinorum operum* (finished c. 1174)
- Elisabeth of Schönau

284 corrections  
/ 3,000 words

THE ALGORITHMIC SCRIBE

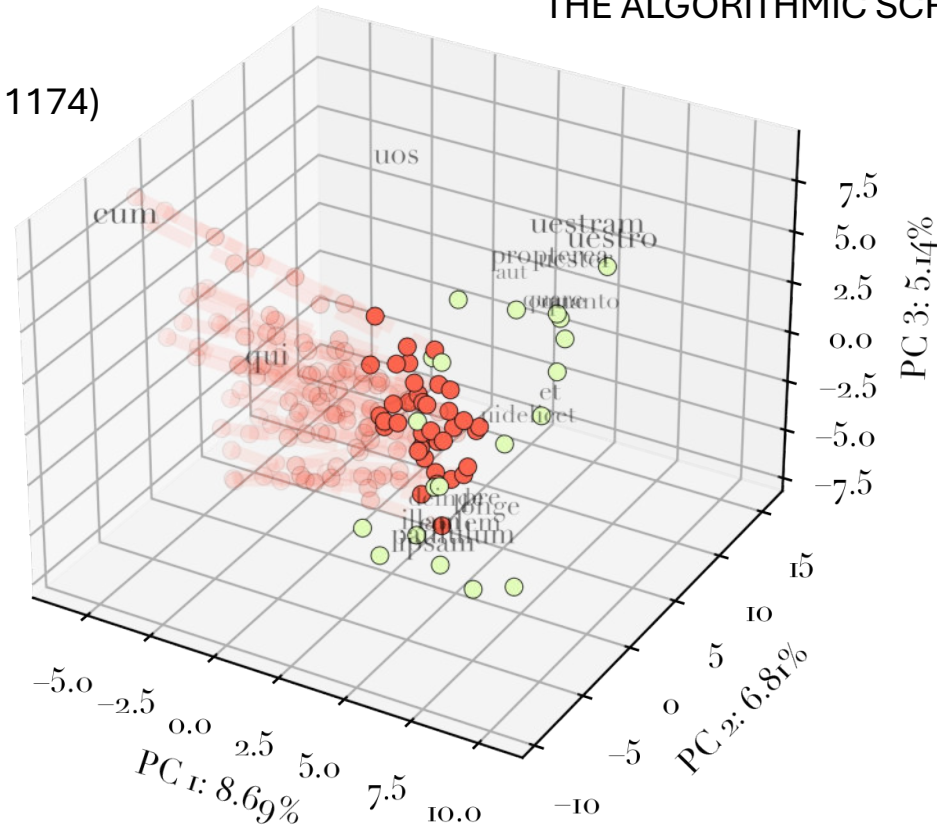


# Hildegard of Bingen (1098–1179) and Ghent MS 241

- Hildegard of Bingen  
*Liber diuinorum operum* (finished c. 1174)
- Elisabeth of Schönau

426 corrections  
/ 3,000 words

THE ALGORITHMIC SCRIBE

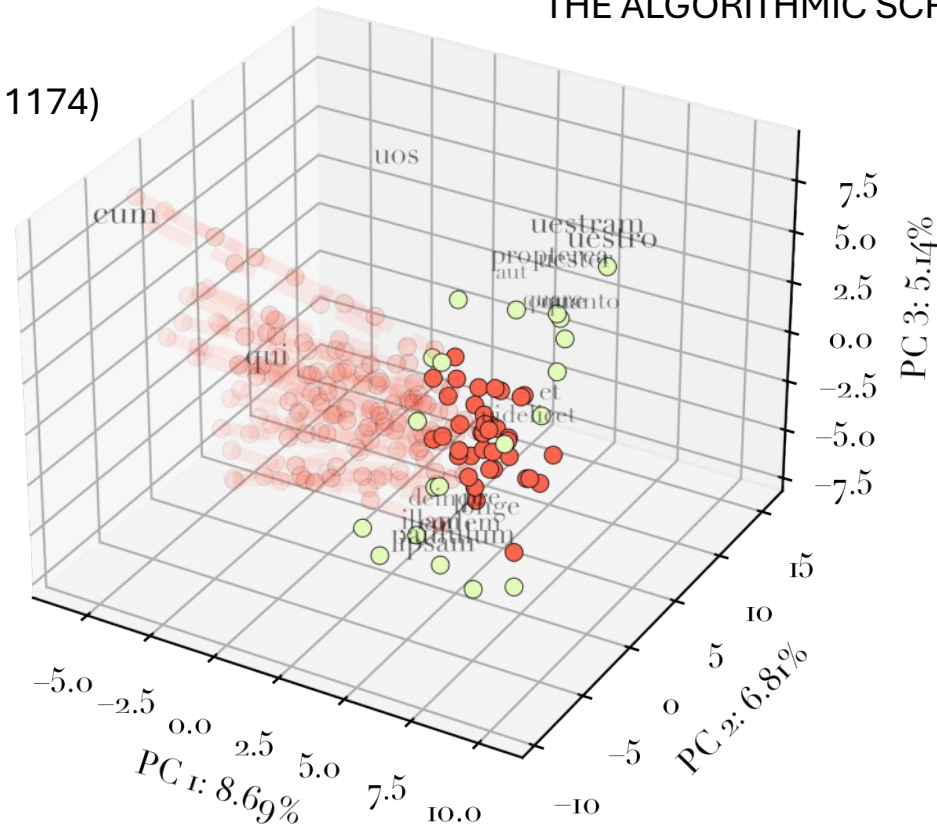


# Hildegard of Bingen (1098–1179) and Ghent MS 241

- Hildegard of Bingen  
*Liber diuinorum operum* (finished c. 1174)
- Elisabeth of Schönau

568 corrections  
/ 3,000 words

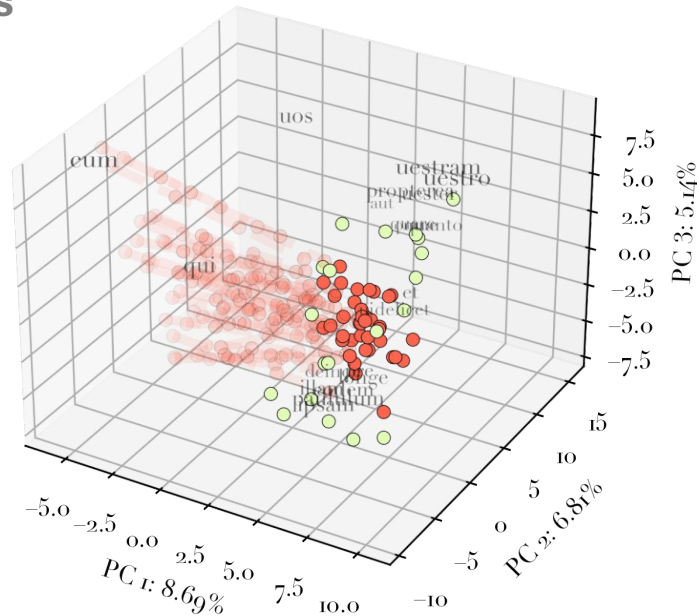
THE ALGORITHMIC SCRIBE





# Algorithmic Scribe

**18.9%**  
of all words



# Ghent MS 241

**2%**  
of all words



# Conclusions