

Automatic writer identification: the paleographer's new best friend?

Jinna Smit

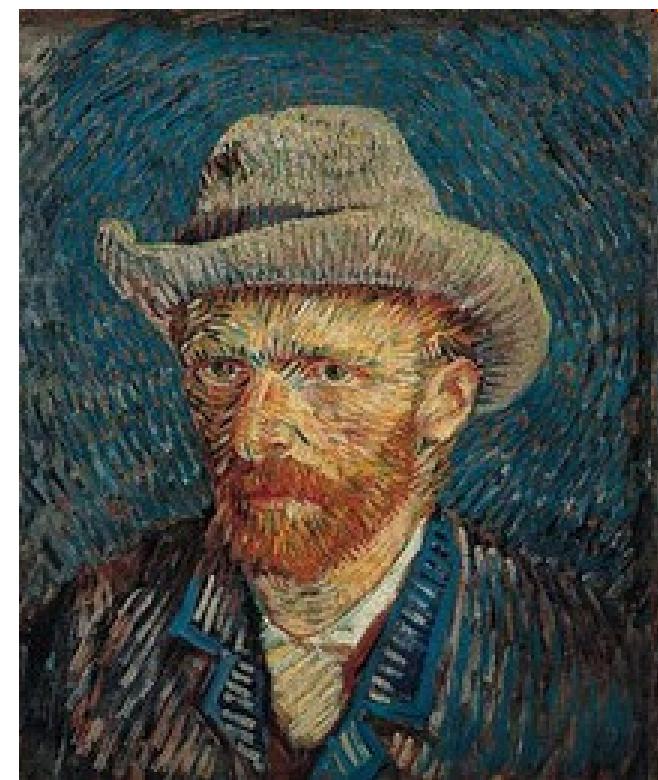
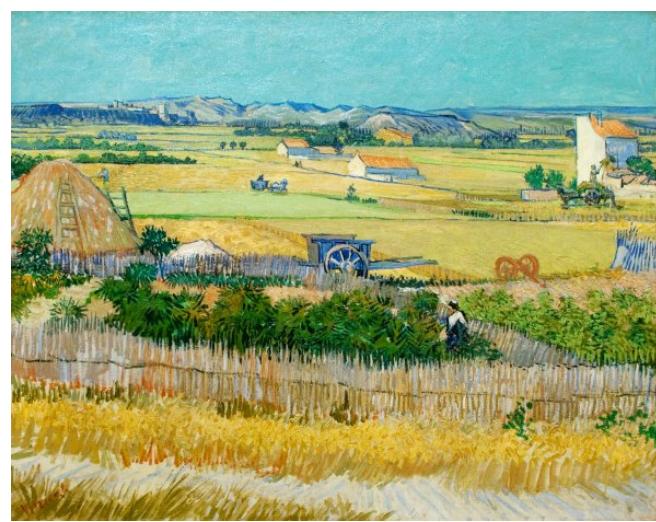
University of Amsterdam/Nationaal Archief

“Thinking of meat makes people less social and lonely”.

Article in Dutch newspapers on an experiment of three psychologists

Le Blute-Fin Mill





Axel Brink

Designer of GIWIS:

- writer identification program combining
- preprocessing steps and feature extraction methods
- requiring minimal user interaction
- visualizing the inner workings

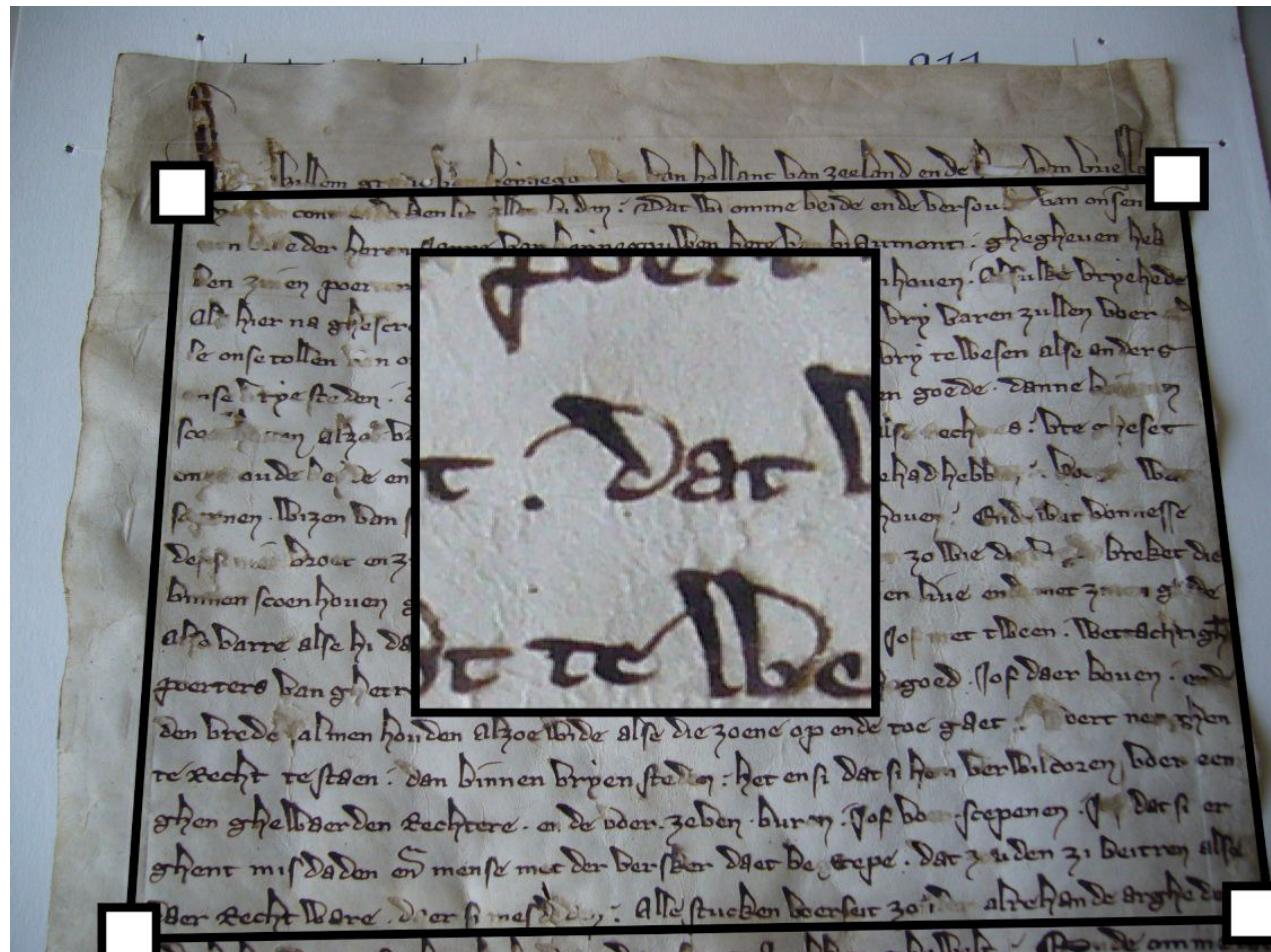
Designer of the Quill feature:

- one of the feature extraction methods of GIWIS
- measures width and direction of the ink trace
- performance rate of 92% and more

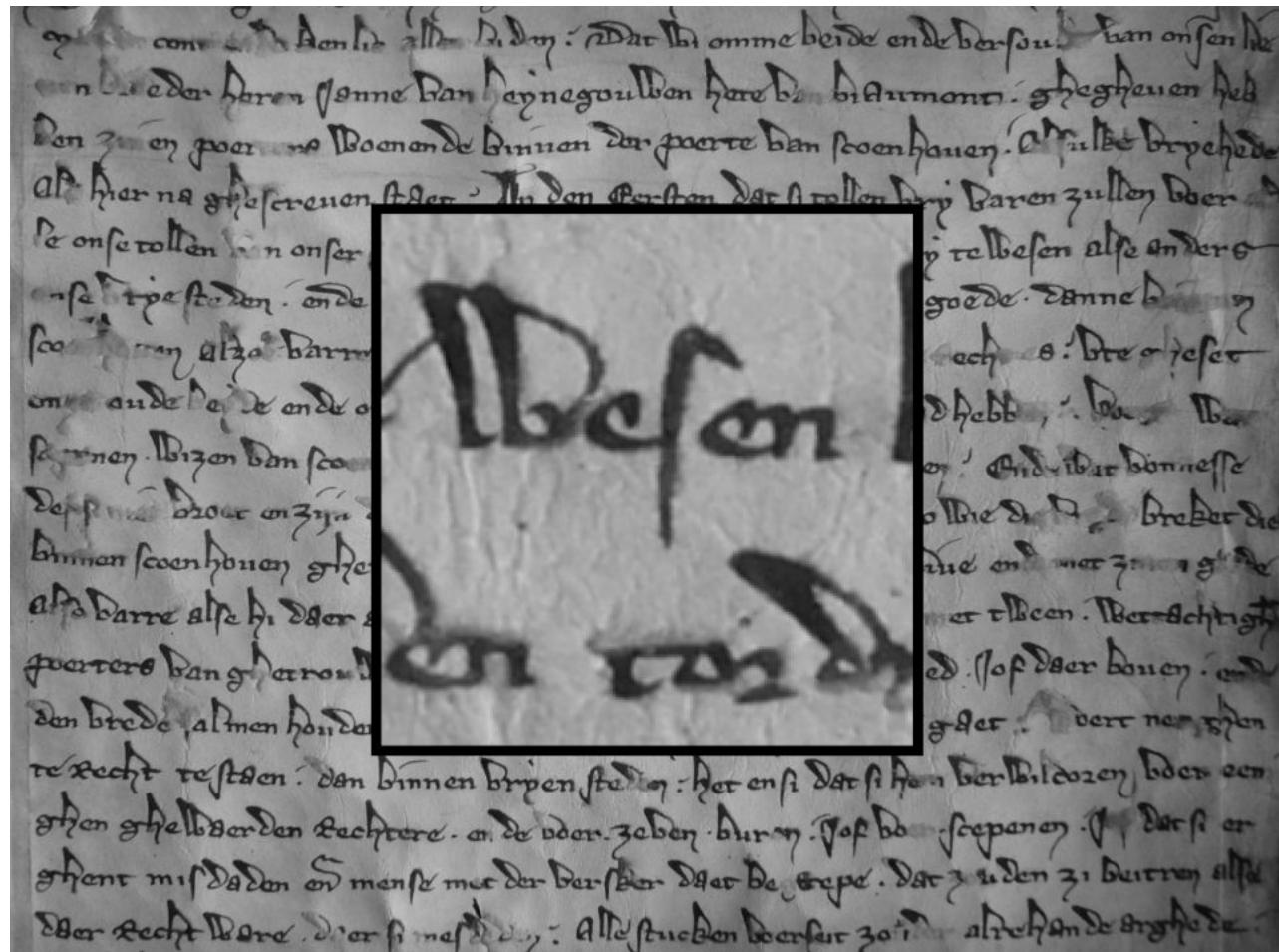
Research topic: **chancery of the count of Holland**

- Period 1299 until 1345
- County of Holland unified with county of Hainault
- Sources: output of the chancery
- Approach: paleographical method

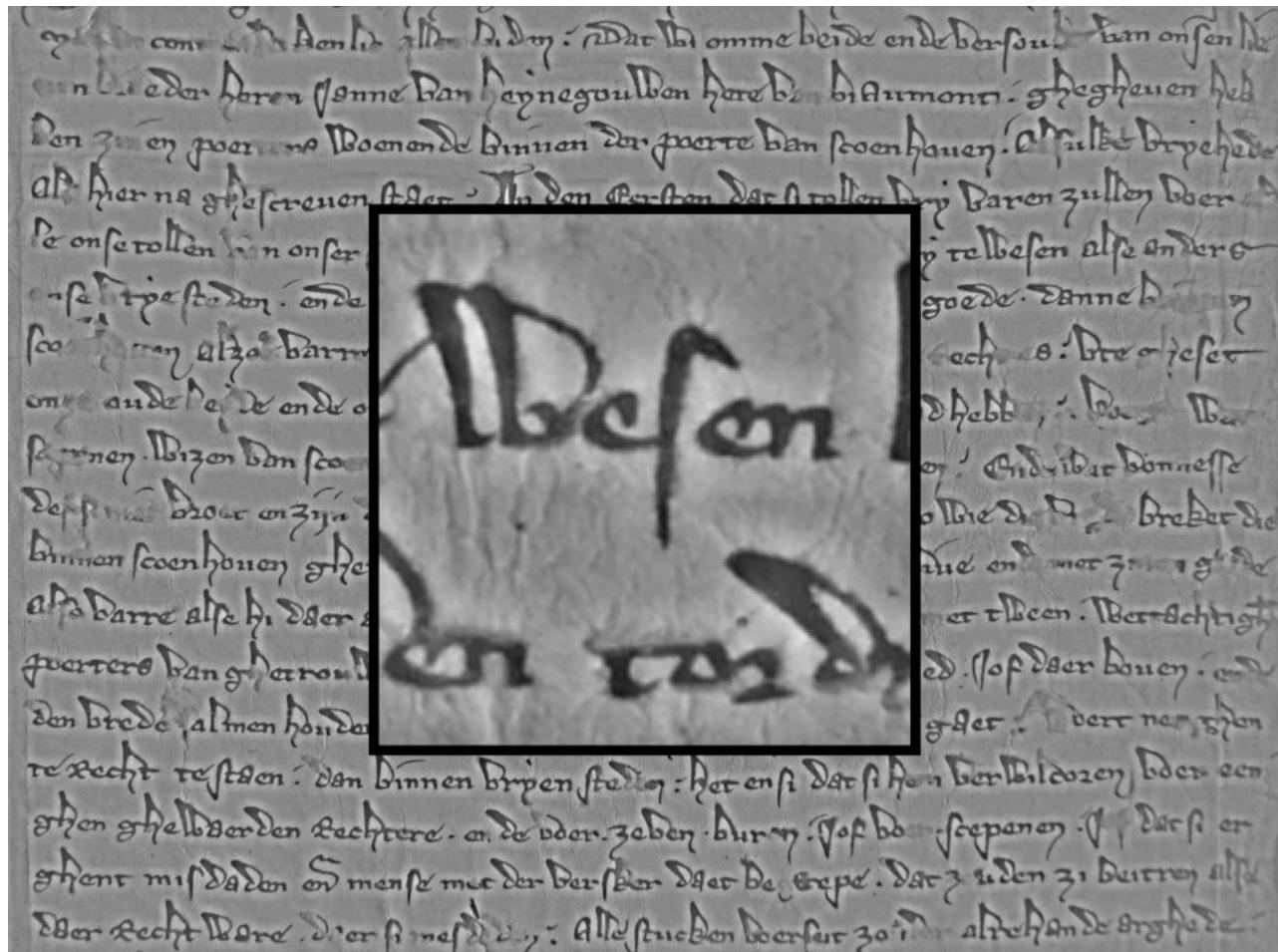
Manual text region selection:



Grayscale conversion, perspective correction and scaling:

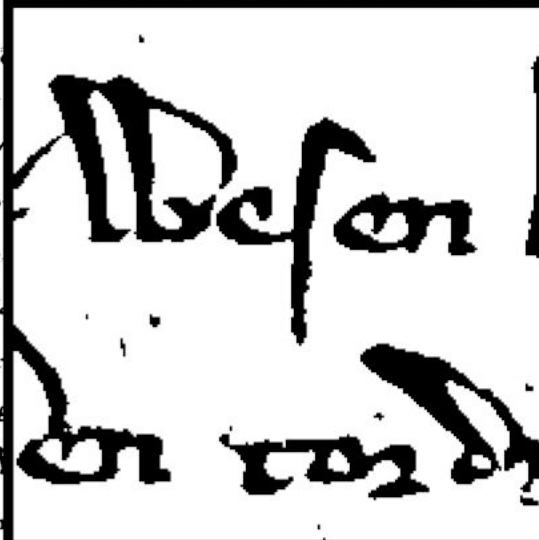


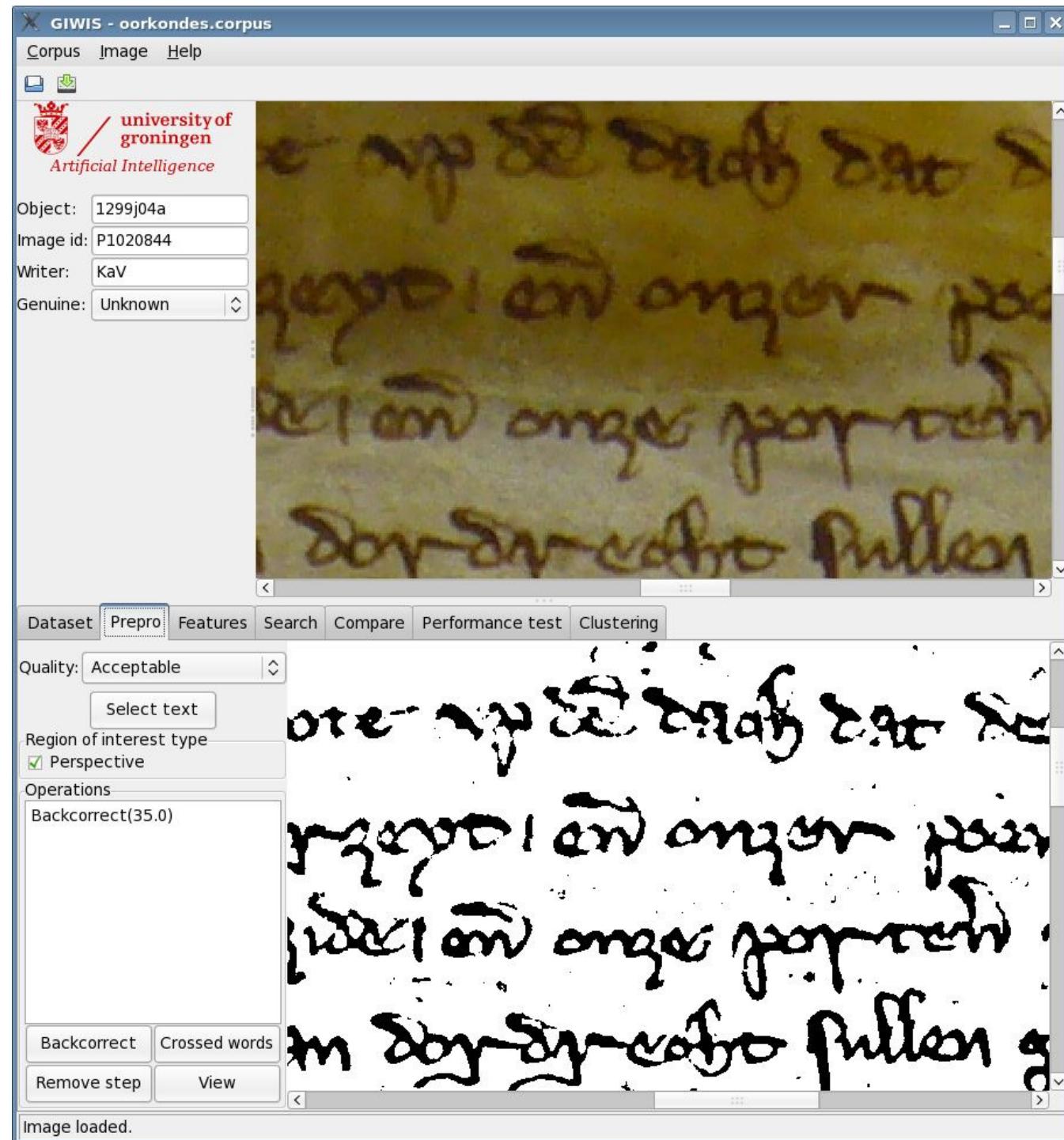
Removal of stains and equalizing of color variations by highpass filtering

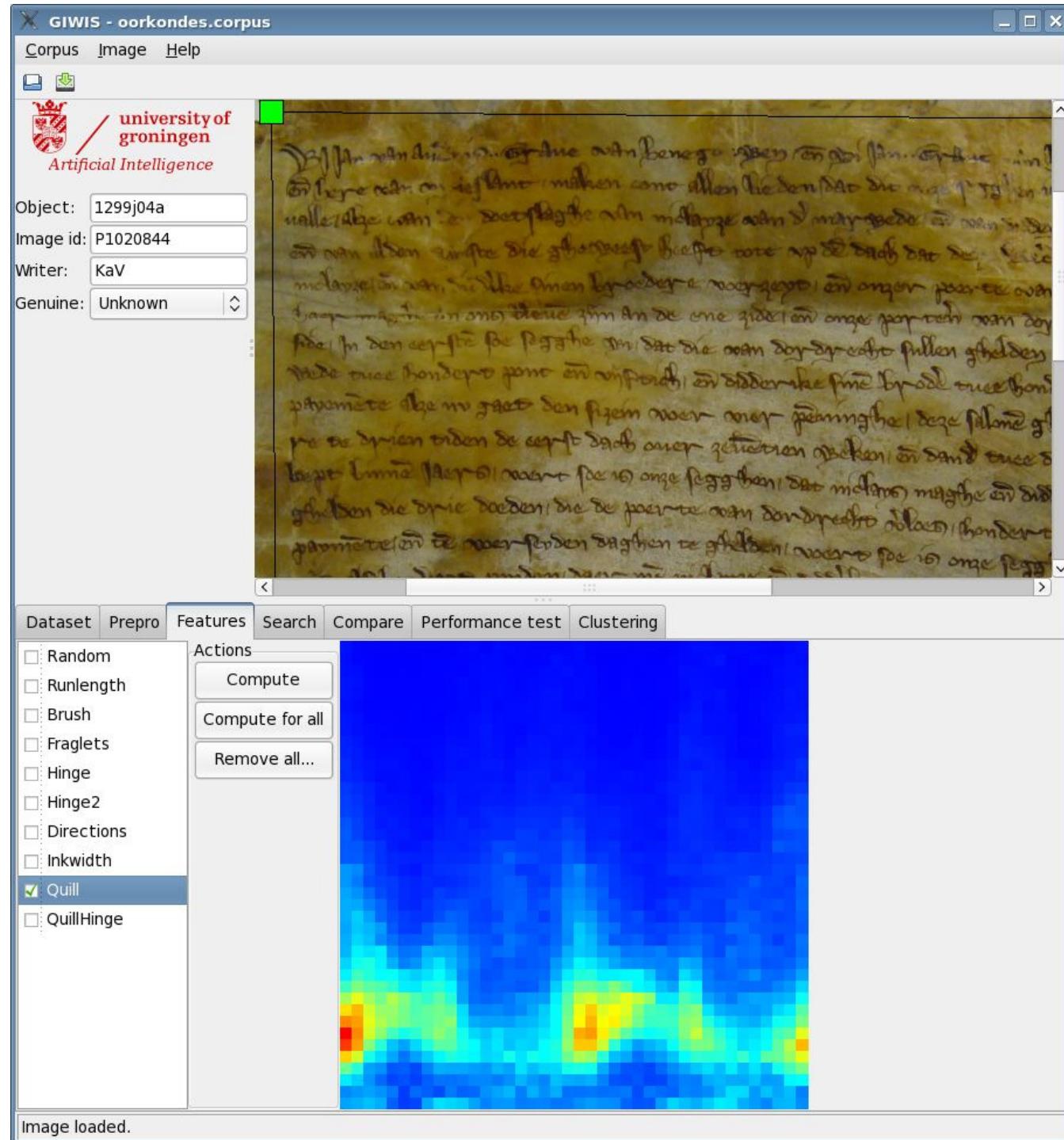


Binarization by Otsu thresholding:

...cont. En den drie h. d. d. dat omme beide ende versonen van onser lie
vaders eder doren gonne van Leynegouw van here bi blaimont. ghedeghenen heb
Den z. en poer. en woonende binnen der gaerte van scoenhoven. en dat brychede
al hier na gescreuen staet. In den eersten dat scrollen dat daren zullen doen
te onser tollen in onser
infestie staden. ende
soo gheven aldaar
ons oude leide ende o
fijnen. Wijzen van sco
den p. a droet en zyn
binnen scoenhoven ghe
ghabanne alse hi doren
gaerters hangt harrou
den brede valmen hondes
re recht resten. den binnen bryen stell. dat en si dat sij hen verbildoren. sder een
ghen gheldderden bechtere. en de vader. geben. blurn. soff bo. scopenen. v. dat si er
ghent misdaden en mensc met der verster. dat de grepe. dat z. uden z. beurt en
ver recht ware. dat si mes. v. alle stukken doersur zo. alrehande brughe.







X GIWIS - oorkondes.corpus

Corpus Image Help

university of
groningen
Artificial Intelligence

Object: 1299j04a
Image id: P1020844
Writer: KaV
Genuine: Unknown

Dataset Prepro Features Search Compare Performance test Clustering

Feature(s) used: Quill

Search

1 1299k24a Writer: KaV Distance: 0.03

2 1301h05a Writer: KaV Distance: 0.04

3 1313e14a Writer: KaV Distance: 0.04

4 1303k11a Writer: KaV Distance: 0.04

5 1299k24a Writer: KaV Distance: 0.04

Image loaded.

The screenshot shows the GIWIS software interface for a corpus of medieval manuscripts. At the top, there's a menu bar with 'GIWIS - oorkondes.corpus', 'Corpus', 'Image', and 'Help'. Below the menu is a logo for 'university of groningen Artificial Intelligence' and a sidebar with object details: Object: 1299j04a, Image id: P1020844, Writer: KaV, and Genuine: Unknown. The main window has tabs for 'Dataset', 'Prepro', 'Features', 'Search' (which is selected), 'Compare', 'Performance test', and 'Clustering'. A sub-menu under 'Search' says 'Feature(s) used: Quill' with a 'Search' button. The search results list five entries, each with a number, manuscript ID, writer, distance, and a thumbnail image. The thumbnails show fragments of medieval Latin text in Gothic script. The first result is 1299k24a by KaV with a distance of 0.03. The second is 1301h05a by KaV with a distance of 0.04. The third is 1313e14a by KaV with a distance of 0.04. The fourth is 1303k11a by KaV with a distance of 0.04. The fifth is 1299k24a by KaV with a distance of 0.04. The bottom status bar says 'Image loaded.'

Image id: P1020844

Writer: KaV

Genuine: Unknown

en van allen enste die ghebrughe heeft tot op de dagh dat se, sive
mclanze en van, sive van broederen voorgegaen en ongen geven
van haren, in den ons denee sijn an de ene zide en onge geven van den
sider. In den scriptie toe ghegaen dat die oom dor drie ollen children

Dataset

Prepro

Features

Search

Compare

Performance test

Clustering

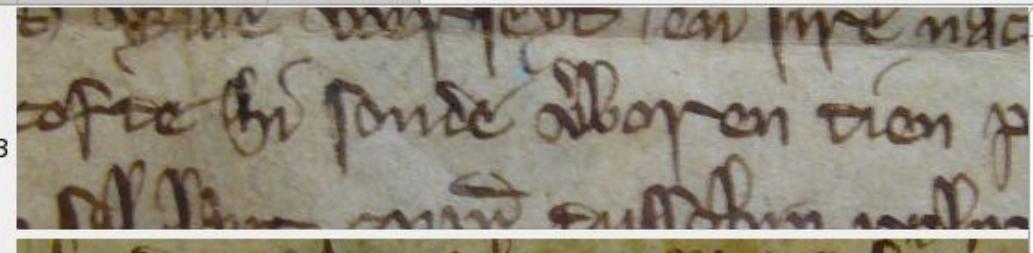
Feature(s) used: Quill

Search

1299k24a

1 Writer: KaV

Distance: 0.03

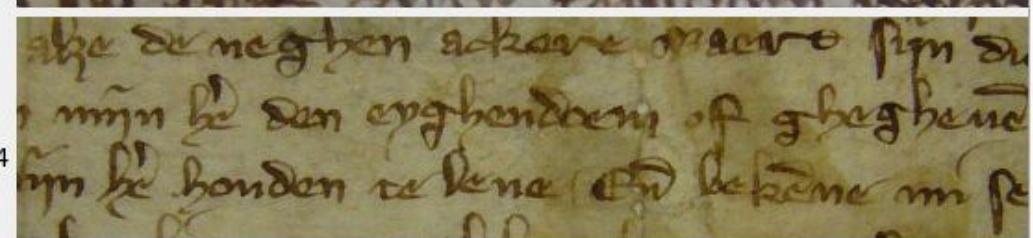


sofie hi sonde alboren dien p
al si... mit auf ihm wille

1301h05a

2 Writer: KaV

Distance: 0.04

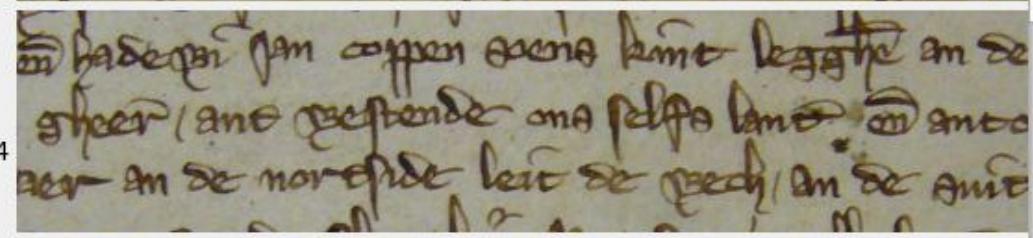


ake de neghen ackere macro sijn du
mijn he den eyghendem of gheghene
ijn he honden te leue. En bekene mi se

1313e14a

3 Writer: KaV

Distance: 0.04

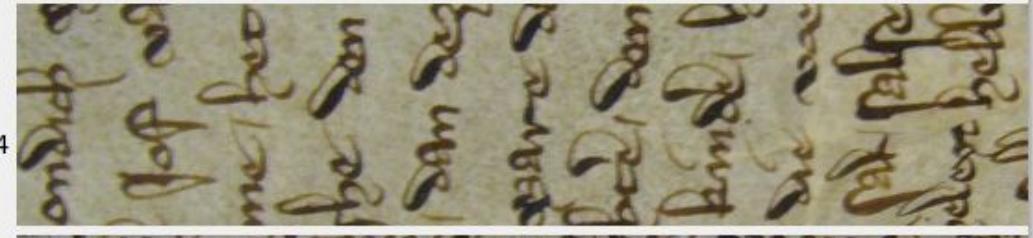


en hadegh jan copper sene kint legghet am de
gheer/ant gestende ma selfs land. en anto
her am de noordse leit de gecij am de ame

1303k11a

4 Writer: KaV

Distance: 0.04

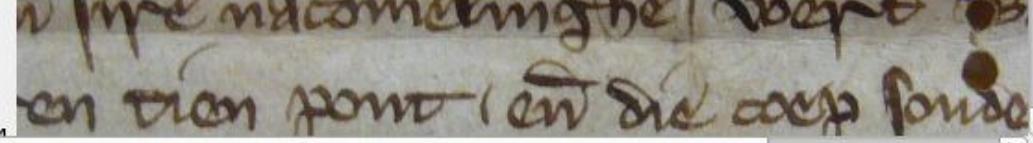


en ijre naconinghe werpt ob
en dien punt (en die coep sonde

1299k24a

5 Writer: KaV

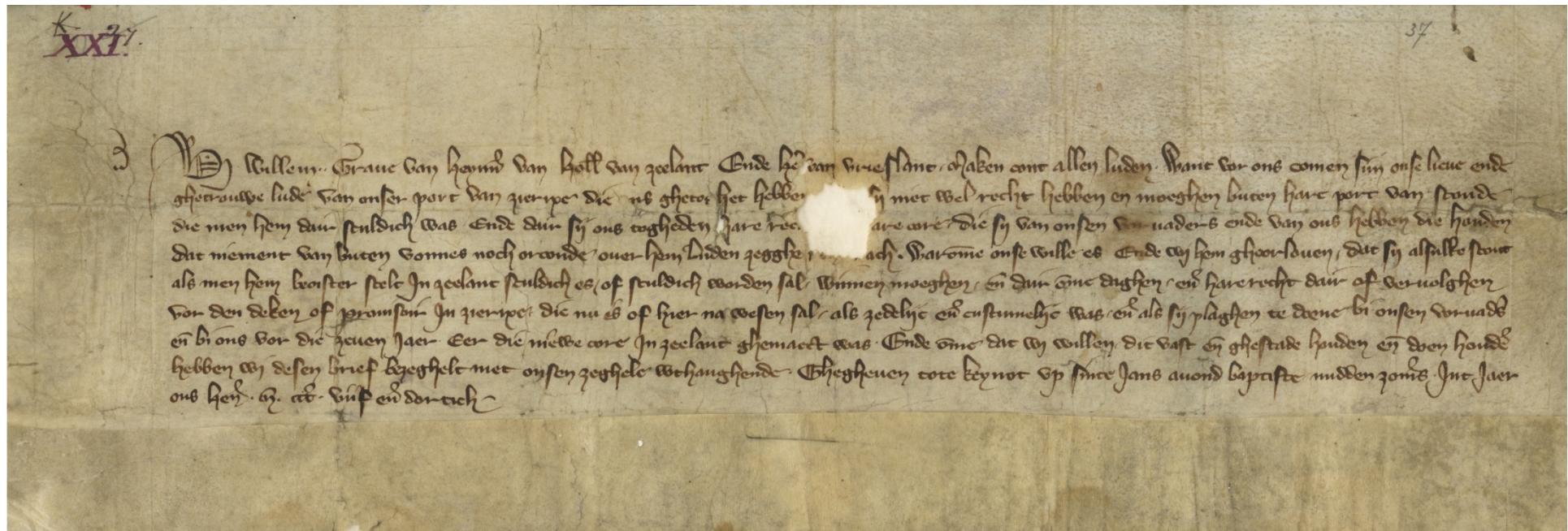
Distance: 0.04



en dien punt (en die coep sonde

Image loaded.

Charter written by chancery clerk Richard Vlec



Et auant la mort de Renart. Monseigneur le pape fit faire un bref auquel il auroit ordonne que nul autre pape ne pecteroit par maniere de sacrement. Et li aultres papes furent de lenu que de liege
et toutz l'acord entierement en la chose estoit fait en une cedula escripte par la
dictes papes. Discours que de Cologne et d'Arras et Mayenne dessus dis laquelle nos autres
papes d'auant part a la See de Bressuire de liege et les foyers d'autre. En tenuoingage de ces

Willem Graue van Heym van hoff van zeeland en heire van vriesland
ghetrouw porten van delff ghemengelijker alle priuilegien en alle hantuerster
van onsen vorstaren die Grauen in hoff hebben ghewesen vaste houd
onsen naomen den hem en huerey naamekinghen vaste ende ghestade et
autre

Willem en dene van heymeghender van hollant van geeltnie
quatre en dix ans ouys midme coene van gafspend en dreech o
fmeister Renart die hi van ons hele ten leene obteust tien pe
en leueen philips nadregale en sene naamekinghen van

RV: identification nos. 1-11 by m.i. & GIWIS

Testcharter

Identification writing hand

GIWIS top 3											
		m.i.	1	m.i.		2	m.i.		3	m.i.	
1	1334f18	RV	1338i10	RV	0.09	1343a12	RV	0.09	1342l20a	RV	0.10
2	1334h27	RV	1337h27a	RV	0.10	1338i10	RV	0.10	1334f18	RV	0.10
3	1335d02a1	RV	B1338g31	RV	0.08	1335d02a3	RV	0.09	1338l09a	RV	0.09
4	1335d02a3	RV	1338l09a	RV	0.08	1340i29	RV	0.08	1335d02a1	RV	0.09
5	1335f23a	RV	1340i29	RV	0.10	1335d02a1	RV	0.10	1337k10	RV	0.10
6	1335k19	RV	1342a12	RV	0.08	1342l20a	RV	0.10	1338i10	RV	0.10
7	1335l03	RV	1342a12	RV	0.06	1338j09a	RV	0.09	1335k19	RV	0.11
8	1336b2a	RV	1337h22a	RV	0.10	1337k30a	RV	0.10	1342a12	RV	0.10
9	1336d08	RV	1342l20a	RV	0.10	1335k19	RV	0.11	1337h27a	RV	0.11
10	1336e26	RV	1337h27a	RV	0.14	1342a12	RV	0.14	1335k19	RV	0.15
11	1337a25	RV	1335d02a3	RV	0.09	1338l09a	RV	0.10	1335d02a1	RV	0.11

RV: identification nos. 12-23 by m.i. & GIWIS

Testcharter

Identification writing hand

GIWIS top 3											
		m.i.	1	m.i.		2	m.i.		3	m.i.	
12	1337e16	RV	1338e06	RV	0.07	1335d02a4	RV	0.11	B1338g31	RV	0.11
13	1337gx	RV	B1338g31	RV	0.09	1338l09a	RV	0.12	1338e06	RV	0.12
14	1337h22a	RV	1336b2a	RV	0.10	1338l09	RV	0.10	1337k30a	RV	0.10
15	1337h27a	RV	1338i10	RV	0.09	1334h27	RV	0.10	1342l20a	RV	0.10
16	1337i04a	RV	1337h27a	RV	0.09	1335d02a2	n.i.	0.11	1326j05	JvL	0.12
17	1337k10	RV	1338e06	RV	0.08	1340j06a	RV	0.09	B1338g31	RV	0.10
18	1337k30a	RV	1342a12	RV	0.08	1342l20a	RV	0.09	1338a12	RV	0.09
19	1338a12	RV	1342a12	RV	0.07	1337k30a	RV	0.09	1338j09a	RV	0.10
20	1338c17	RV	1338j09a	RV	0.09	1338l09	RV	0.09	1342a12	RV	0.10
21	1338e06	RV	1337e16	RV	0.07	1337k10	RV	0.08	1344d22a	RV	0.09
22	1338f19	RV	1338j09a	RV	0.10	1342l20a	RV	0.11	1342a12	RV	0.11
23	1338g08a	RV	1303k06a	T	0.11	1337h27a	RV	0.12	1337e16	RV	0.12

RV: identification nos. 24-33 by m.i. & GIWIS

Testcharter

Identification writing hand

GIWIS top 3											
		m.i.	1	m.i.		2	m.i.		3	m.i.	
24	B1338g31	RV	1335d02a1	RV	0.08	1338l09a	RV	0.09	1337gx	RV	0.09
25	1338h22a	RV	1338i10	RV	0.09	1342a12	RV	0.09	1340f07a	RV	0.10
26	1338i10	RV	1342a12	RV	0.08	1338h22a	RV	0.09	1334f18	RV	0.09
27	1338j09a	RV	1342a12	RV	0.07	1338c17	RV	0.09	1335l03	RV	0.09
28	1338l09	RV	1338c17	RV	0.09	1342a12	RV	0.10	1337h22a	RV	0.10
29	1338l09a	RV	1335d02a3	RV	0.08	B1338g31	RV	0.09	1339c10	RV	0.09
30	1339c10	RV	1340j06a	RV	0.06	1340i29	RV	0.07	1338l09a	RV	0.09
31	1339c20	RV	1335d02a3	RV	0.09	1338l09a	RV	0.11	1340i29	RV	0.11
32	1339f17	RV	1335d02a3	RV	0.11	1338l09a	RV	0.11	1339c10	RV	0.12
33	1339g20	RV	1340j06a	RV	0.09	1340i29	RV	0.09	1308k05a	V	0.11

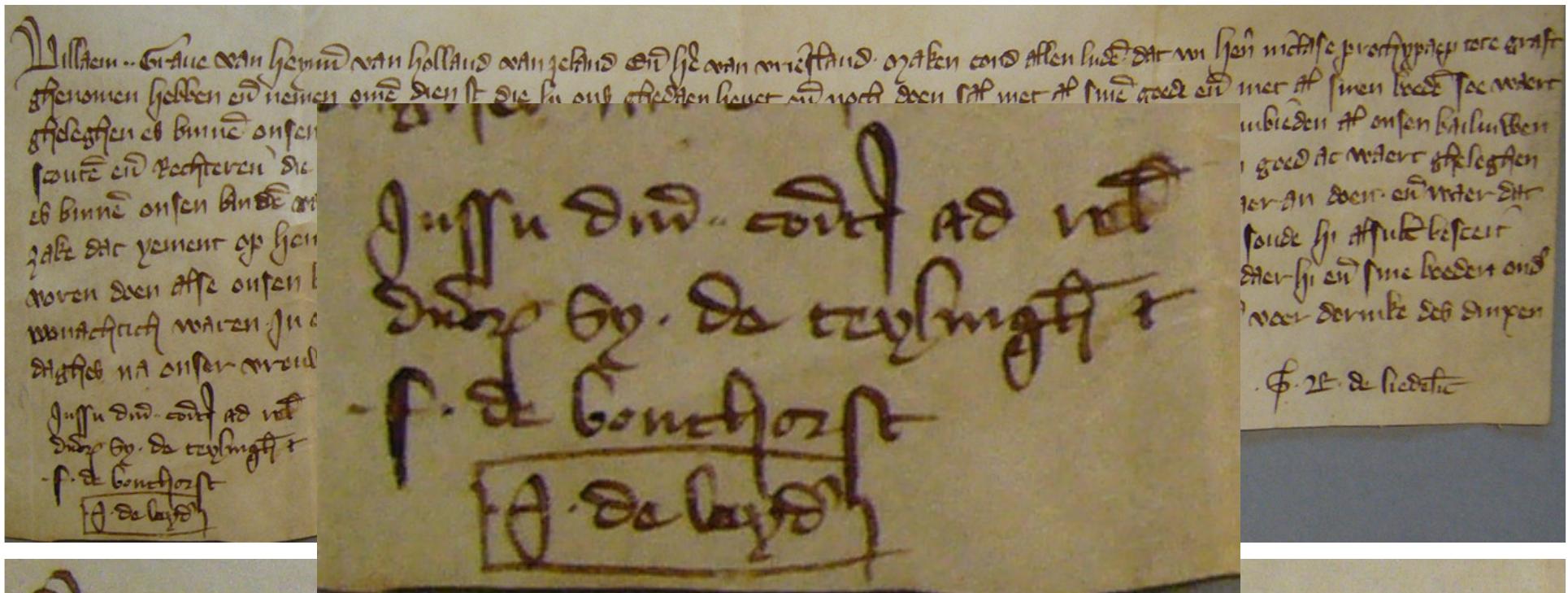
RV: identification nos. 34-43 by m.i. & GIWIS

Testcharter

Identification writing hand

GIWIS top 3											
		m.i.	1	m.i.		2	m.i.		3	m.i.	
34	B340e12	RV	1340i29	RV	0.13	1338l09a	RV	0.14	1309a05a	n.i.	0.14
35	1340f07a	RV	1342a12	RV	0.07	1310g08f	S	0.09	1342l20a	RV	0.09
36	1340f09a	RV	1339g20	RV	0.12	1340j06a	RV	0.13	1340i29	RV	0.14
37	1340i29	RV	1340j06a	RV	0.05	1339c10	RV	0.07	1335d02a3	RV	0.08
38	1340j06a	RV	1340i29	RV	0.05	1339c10	RV	0.06	1339g20	RV	0.09
39	1341c29a	RV	1340i29	RV	0.12	1339c10	RV	0.12	1301c08a	n.i.	0.12
40	1342a12	RV	1335l03	RV	0.06	1338a12	RV	0.07	1340f07a	RV	0.07
41	1342a26	RV	1342a12	RV	0.09	1337h27a	RV	0.12	1338a12	RV	0.12
42	1342l20a	RV	1342a12	RV	0.08	1337k30a	RV	0.09	1338i10	RV	0.09
43	1344d22a	RV	1338e06	RV	0.09	1337e16	RV	0.11	1335d02a4	RV	0.11

Usual procedure: subscription hand ≠ charter hand



Willaim... Graue van heym van hollant aangeleid en he van nietland oocken con allen huet dat wi hem nietse presteypaer ree graet
ghenomen hebben en neinen omc den st die hi ons gheghen heeft en noch doen dat met al sijn gode en met sijn leed toe medert
ghedeghen es binne onsen
soule en pechters die
es binne onsen lande en
gakte dat yment op hem
overen deen alte onsen l
wantschafft waren jn e
digtes na onsen land
Jussu dom. cont ad not
dowc. sy. de tylmgh t
f. de Goudorst
H. de Lijst

Charter written and signed by Richard Vlech

Mittem.. Graue van Regem van hoff van ze
dat ghi onder zonet en verheid als en
toe ghehouden soude hebben. dat hi hem schrey
en e hen Ianne hen Simon van borchem en
hen denste als si van ons ghehouden hebben ha
tich van onser weghen met desen breue bege
neue charre magdalene dach. gne jore
p dny
Enck. snt de bellonete. 2 falconete / n alns de qsilu
.. S. Vlech.

postm. En vullen den c
ans zone en hefft dat p
n te bedraghen. Want o
vullen En van hefft van s
heffen. En des maken wi
e ghehoude ghe pmaende
S an de saure

*Testcharter**Identification writing hand*

GIWIS top 3														
1	1337h27a	RV	1	1338i10	rv	0.09	2	1334h27	rv	0.10	3	1342l20a	RV	0.10
2	1337k10	RV	1338eo6	RV	0.08		1340j06a	RV	0.09		B1338g31	rv	0.10	
3	1337k30a	RV	1342a12	rv	0.08	1342l20a	RV	0.09	1338a12	RV	0.09			
4	1338a12	RV	1342a12	rv	0.07	1337k30a	RV	0.09	1338j09a	RV	0.10			
5	1338c17	RV	1338j09a	RV	0.09	1338l09	RV	0.09	1342a12	rv	0.10			
6	1338e06	RV	1337e16	rv	0.07	1337k10	RV	0.08	1344d22a	rv	0.09			
7	1338f19	RV	1338j09a	RV	0.10	1342l20a	RV	0.11	1342a12	rv	0.11			
8	1338h22a	RV	1338i10	rv	0.09	1342a12	rv	0.09	1340f07a	RV	0.10			
9	1338j09a	RV	1342a12	rv	0.07	1338c17	RV	0.09	1335l03	rv	0.09			
10	1338l09	RV	1338c17	RV	0.09	1342a12	rv	0.10	1337h22a	rv	0.10			
11	1338l09a	RV	1335d02a3	rv	0.08	B1338g31	rv	0.09	1339c10	RV	0.09			
12	1339c10	RV	1340j06a	RV	0.06	1340i29	RV	0.07	1338l09a	RV	0.09			
13	1339f17	RV	1335d02a3	rv	0.11	1338l09a	RV	0.11	1339c10	RV	0.12			
14	1339g20	RV	1340j06a	RV	0.09	1340i29	RV	0.09	1308k05a	V	0.11			
15	1340f07a	RV	1342a12	rv	0.07	1310g08f	S	0.09	1342l20a	RV	0.09			
16	1340f09a	RV	1339g20	RV	0.12	1340j06a	RV	0.13	1340i29	RV	0.14			
17	1340i29	RV	1340j06a	RV	0.05	1339c10	RV	0.07	1335d02a3	rv	0.08			
18	1340j06a	RV	1340i29	RV	0.05	1339c10	RV	0.06	1339g20	RV	0.09			
19	1342l20a	RV	1335d02a3	rv	0.08	1337k30a	RV	0.09	1335d02a1	rv	0.09			

Writing samples Richard Vlec

GIWIS: distance = 0.05

Willen.. Gracie van Heyn van Holland van gelant en he dan
stadeloet te quitenre onsen henen en ghettulden. Excep Baed-
ghelosten als zy dor ons ghedoeft hebben den vullen van de
polamen smeren newt Jannic Kliere. Vullen kerremay en
zestien stelinghe en neghen penninghe groter dan 12 al-

.. Willen.. Gracie van Heyn van Holland van gelant en he dan
drecht - alsulke gracie ghedaen hebben om dienst die hi on-
daer toe behore en die wissene te drouwen - also groot al-
moers die ghedeghen sijn westwaerts van loememert ge-
die en gaertie daer op hebben - wort om houden sal van
myn zegheli - Ghelijcende te

1st = RV: 1339 July 20

2nd = V: 1308 November 5

GIWIS: distance = 0.11

Mittem... Graue van henev van helle om zet en h^e gey
dat ghi onde zonet en wer heid als em dat justi
toe ghehouden soude hebben. dat hi hem schryf siet
en e hen Janne /hen Simon van banchem en vullen den tu
ben dienst als sy van ons gheuert hebben. Hoe dat sy he
trog van onser ewen en dan hene van helle.

Bi Willaem Graue van henev van helle van gelame en here van
ponge onze baelin van middelborch dat ghi der abdissen van
van den goede daer haer de abde van middelborch onbeleent of
haer in besitt of ghekeest heeft (is dat sake dat de aker
die hene hem op ghesent hebben) en dat mes on dat gheghue

Writing samples Richard Vlec

GIWIS: distance = 0.14

• Willen.. Gracie van Heym van Acht van gelant en hec van
goden riden van enster porten van roedenane - dat hebbe en ope
pac. • Van elden dat sic begheven sulken en ghescreet sun op hi
hem van dene hadden op ~ zegghen en op ~ gredsminken - dat
en ons malle is dat sic hebbe in den maneren Gasthuis

• Willen.. Gracie van Heym van Holland van gelant en hec van
drecht - assulke gracie ghedien hadden om dienst die hi or
dan toe behoort en die wistenie te droughen - also groot al
maers die gedeghen sun westwaerds van Loenenere ge
die en gaertin dan op hebbe - wort om houden sal van

More about GIWIS and the Quill feature:

- Axel Brink, *Robust and applicable handwriting biometrics* [PhD thesis University of Groningen, December 2011].
- A. Brink, J. Smit, M. Bulacu, and L. Schomaker, “Writer identification using directional ink-trace width measurements.” *Pattern Recognition*, (45):162-171, 2012.
doi: 10.1016/j.patcog.2011.07.005
- J. Smit, “The death of the palaeographer? Experiences with the Groningen Intelligent Writer Identification System (GIWIS)” [*Archiv für Diplomatik*]