



K O N I N K L I J K E N E D E R L A N D S E  
A K A D E M I E V A N W E T E N S C H A P P E N

**Rapportage over de onderzoekspilot binnen Nederlab t.b.v de samenstelling van een corpus van de 15e en de 16e eeuw**

van der Sijs, Nicoline; Depuydt, Katrien; Rem, M.

2018

**document version**

Publisher's PDF, also known as Version of record

[Link to publication in KNAW Research Portal](#)

**citation for published version (APA)**

van der Sijs, N., Depuydt, K., & Rem, M. (2018). *Rapportage over de onderzoekspilot binnen Nederlab t.b.v de samenstelling van een corpus van de 15e en de 16e eeuw*. Meertens Instituut.

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the KNAW public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the KNAW public portal.

**Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

**E-mail address:**

[pure@knaw.nl](mailto:pure@knaw.nl)

## **Rapportage over de onderzoekspilot binnen Nederlab t.b.v de samenstelling van een corpus van de 15e en de 16e eeuw**

Katrien Depuydt (INT), Margit Rem (RU), Nicoline van der Sijs (RU)

Samenvatting: In de periode 1/12017 tot november 2017 is gewerkt aan de samenstelling van het corpus bestaande uit 15e- en 16e-eeuwse ambtelijke bescheiden. Het resultaat van de pilot is: CLVN, Corpus Laatmiddel- en Vroegnieuwnerlands, samengesteld door Ans van Kemenade, Margit Rem, Nicoline van der Sijs en INT

Projectbegeleiders: Nicoline van der Sijs, Margit Rem, Katrien Depuydt

Projectuitvoerders: Tamara van Seggelen, Niels Mulder, Sanne de Kroon en Emma Ringelberg (studentassistenten RU); Wil de Ruyter-Balt (dataproductiemedewerker INT); Jesse de Does (technische ondersteuning INT)

### **INLEIDING**

Het doel van dit project was het bij elkaar brengen, cureren en verrijken van 15e- en 16e-eeuws taalmateriaal om onderzoeksvragen mogelijk maken die momenteel niet kunnen worden beantwoord door een lacune in de data van Nederlab. Denk aan taalkundige onderzoeksvragen als “Hoe heeft het voorzetsel *na* ‘achter, naartoe, volgens’ zich sinds de 13e eeuw ontwikkeld tot de huidige situatie waarin twee verschillende voorzetsels bestaan met twee verschillende betekenissen: enerzijds *na* ‘achter’, en anderzijds *naar* ‘naartoe, volgens?’”.

Tot voor kort was een groot en onoverkomelijk probleem bij de beantwoording van een dergelijke onderzoeksvraag, het feit dat er een lacune in de data bestond voor de 15e en de 16e eeuw. Op basis van het Corpus Gysseling en het Corpus Van Reenen-Mulder (CRM) konden onderzoekers de talige ontwikkelingen in de 13e en de 14e eeuw volgen, voor de 17e eeuw was er weer veel materiaal, beschikbaar binnen Nederlab, maar voor de 15e en de 16e eeuw was er nauwelijks materiaal beschikbaar dat ook nog betrouwbaar gelokaliseerd en gedateerd is. Daardoor moest iedere uitspraak over de ontwikkeling van taalverschijnselen tussen het einde van het Middelnederlands en het begin van het Vroegnieuwnerlands speculatief blijven. Terwijl de 15e en 16e eeuw van essentieel belang zijn omdat juist in die periode de kiem van het Standaardnederlands ligt.

We zijn in het kader van dit pilotproject daarom aan de slag gegaan met het verzamelen, converteren en metadateren van tekstmateriaal uit de 15e en 16e eeuw, dat aan een aantal criteria moest voldoen:

1. De teksten zijn origineel (geen afschriften)
2. De teksten zijn gelokaliseerd en dit is niet op taalkundige gronden gebeurd;

3. De teksten zijn gedateerd;
4. De teksten zijn in proza gesteld;
5. De teksten zijn in het Nederlands gesteld. Dat betekent dat lange passages in bijvoorbeeld Latijn zijn weggehaald.

## **WERKWIJZE**

### Inventarisatie materiaal

Om te beginnen hebben we het DiT-corpus (Dutch in Transition) uit Nijmegen van Ans van Kemenade als basis genomen (inclusief de Drentse Etstoel), en gedeeltelijk door de student-assistenten met metadata laten verrijken. Vervolgens hebben we een van de student-assistenten op internet zoveel mogelijk archieven laten opzoeken en relevante digitale bestanden laten downloaden, en we hebben eigenaren van relevante data benaderd en gevraagd om een digitale kopie. Daarnaast hebben we vakgenoten om adviezen gevraagd. Tot slot hebben we relevante gegevens van [www.diachronie.nl/corpora](http://www.diachronie.nl/corpora) gehaald.

### Dataprocessing door INT

De tekstedities die we voor ons corpus gebruikt hebben, bestonden hoofdzakelijk uit pdf 's. Een kleiner aantal documenten waren in Word. We hebben alle pdf's via Adobe pro naar Word geconverteerd, en zijn met de Worddocumenten verder gegaan.

### Vaststellen metadataschema door INT

In navolging van het Corpus Gysseling en CRM, hebben we ook voor dit corpus de datering en de lokalisering van het materiaal zo precies mogelijk gemapt op de Nederlab-metadata. Er is een onderscheid gemaakt tussen de datering van de tekst (wanneer is de tekst geschreven/geconcipieerd) en de tekstgetuige (in welke legger is de tekst overgeleverd). De lokaliseringinformatie bestaat daar waar mogelijk uit plaats, regio, provincie en land, en is voorzien van Kloekecodes.

### Verwerking van de teksten

Er is per uitgangsdokument door de studenten vastgesteld of het een editie is waarin verschillende losse teksten achter elkaar zijn uitgegeven, of 1 dokument. Als het een verzameling van teksten bevat, dan hebben de studenten die teksten afzonderlijk gemarkeerd en van eigen metadata voorzien. Als een uitgangsdokument uit één tekst bestaat, betekende het niet altijd dat er slechts 1 set metadata aan die tekst moest worden gekoppeld. Documenten als bijvoorbeeld het *Hofboek Bredevoort*<sup>1</sup> bevatten de administratie van meerdere jaren van informatie over de horigen in de heerlijkheid in de vorm van korte notities. De datering van de tekstgetuige van een dergelijk dokument is een lappendeken aan dateringen bij stukjes tekst, gekoppeld aan de dag, maand en jaar waarop die notities zijn gemaakt.

---

<sup>1</sup> <https://www.heerlijkheidbredevoort.nl/horigheid/>

De collecties zijn door de student-assistenten verder bewerkt volgens de instructies in de Appendix: ze hebben de metadata geëxtraheerd, gecureerd en aangevuld. Per uitgangsdokument is vervolgens bekeken wat de editieprincipes waren. Historische teksten bevatten afkortingen, bevatten plaatsen waarin de tekst niet leesbaar is, etc. Het was belangrijk om vast te stellen hoe e.e.a. in de editie was weergegeven omdat deze informatie in Nederlab uiteindelijk overal op dezelfde manier weergegeven moet worden. Waar nodig is er een eerste vertaalslag gemaakt naar een meer uniforme codering, die in het Worddocument is aangebracht, om de verdere processing voor Nederlab te vereenvoudigen.

In de meeste bestanden hebben de studenten het begin en einde van een document aangegeven met een code, en die code plus aanvullende informatie hebben ze tevens in een begeleidend excel-bestand opgenomen, met de volgende kolommen:

- Record\_ID (dit id vormt de link tussen de metadata in het Excel en de tekst of het tekstgedeelte)
- Titel document
- Periode of jaar
- Dag (indien bekend)
- Plaats (Kloekcode)
- Provincie
- Genre: bv. akte
- Auteur (indien bekend)
- Bron (verwijzing naar website)
- Aanvullende Informatie: b.v. Transcribent: Chris Picard. Nagekeken door: Kristiaan Magnus. Afkortingen zijn opgelost tussen ronde haken (). Superscript staat tussen vierkante haken []. Doorgehaalde tekst staat tussen <del></del>. Opmerkingen of onzekerheden transcribent staan tussen apenstaartjes @@.

### Nabewerking door INT

Teksten en tekstgedeeltes zijn door het INT in de Worddocumenten gescheiden door een XML-codering op basis van de id's. De metadata van de studenten en de id-codes zijn door Wil de Ruyter-Balt van INT gecontroleerd en zij heeft de meest complexe teksten aangepakt. De verdere dataprocessing hiervan is door Jesse de Does gedaan.

### **RESULTAAT**

De afgesproken deliverables zijn gerealiseerd:

*Deliverable 1:* toevoegen van gestructureerde metadata aan beschikbare ambtelijke documenten en egodocumenten en kwaliteitsbeoordeling transcripties (zie boven);

*Deliverable 2:* verzamelen van relevante egodocumenten op basis van [www.egodocument.net](http://www.egodocument.net) en kwaliteitsbeoordeling transcripties (zie overzicht hieronder);

Aan de derde afgesproken *deliverable 3*: “verzamelen van relevante letterkundige edities die door individuele onderzoekers zijn vervaardigd” hebben we een iets andere invulling

gegeven. We hebben op advies van Prof. dr. Jan Burgers gebruik gemaakt van kronieken en niet van literaire werken.

### Overzicht uiteindelijk gebruikte bronnen.

#### **Seyger**

Gerard Seyger , *De geboorte en bloei van het twents als schrijftaal in de late middeleeuwen. Corpus van 14e en 15e eeuwse oorkonden in de volkstaal, uit de steden Almelo, Enschede, Oldenzaal en Ootmarsum en van ambtman en rentmeester in Twenthe. Bijdrage tot een bronnenboek van Twenthe.* Oldenzaal 1999-2011

#### **Heemkundekring de Kommanderij Gemert**

[www.heemkundekringgemert.nl](http://www.heemkundekringgemert.nl)

- Erfpachtregister Gemert.
- Gemert, Oorkonden 1401-1485.

#### **Henk Ruessink en Yvette Hoitink**

<https://www.heerlijkheidbredevoort.nl>

Henk Ruessink en Yvette Hoitink, Hofboek Bredevoort 1506-1596. Akten en notities. [z.p.], 2000-2009.

#### **Hoogheemraadschap van Delfland**

<https://www.hhdelfland.nl>

Dingtalen- of presentatieboek, Historisch archief hoogheemraadschap van Delfland 2014. [z.p.], Hoogheemraadschap van Delfland, 2014.

#### **Huygens-ING**

A.M. van Lith-Droogleever Fortuijn, J.G.M. Sanders en G.A.M. Van Synghel (eds.), *Kroniek van Peter van Os tot 1523.* 's-Gravenhage, 1997. Rijks Geschiedkundige Publicatiën uitgegeven in opdracht van de minister van Onderwijs en Wetenschappen onder toezicht van de Rijkscommissie voor Vaderlandse Geschiedenis. Kleine Serie 87

<http://resources.huygens.knaw.nl/retroboeken/vanos/#page=0&accessor=toc&view=homePane>

**Paul Bouw**

<https://www.paulbouw-geschiedenispagina.nl>

- Album amicorum van Johanna Bentinck. Hoge Raad van Adel, Coll. Van Spaen, nr. 87c.
- Gelders Archief toegang 0583. Huis Laar.
- Gelders Archief, toegang 024. Memories van Successie, Kantoor Doesburg.
- Gelders Archief, toegang 0243. Charterverzameling.
- Gelders Archief, toegang 0405. Huis ter Horst.
- Gelders Archief, toegang 0409. Huis Keppel.
- Gelders Archief, toegang 0418. Huizen Nettelhorst en De Heest.

### **Radboud Universiteit Nijmegen**

- DiT-corpora, Ans van Kemenade

<http://mjtangelder.ruhosting.nl/DITC/DITCORPUS.html>

### **Regionaal Archief Alkmaar**

<https://www.regionaalarchiefalkmaar.nl>

- Archief van de gemeente Alkmaar, 1325-1815, inventarisnummer 2076. Stukken betreffende de opstand in het Noorderkwartier in het algemeen en het Beleg en Ontzet van Alkmaar in het bijzonder. 1571-1578.
- Cort verhael van die gheschiedenisse ende belegeringhe der stat Alcmaer anno 1573.
- Nanning van Foreest: Kort verhaal van het beleg van Alkmaar. Een cort verhael van de strenghe belegheringhe ende aftreck der Spangiaerden van de stadt Alcmaer gheleghen in Hollandt, die welcke de Spaengiaerts nae dat sij se int jaer 1573 seven weecken lanck belegert hadden (...) uut het Latyn int Duytsch overgeset.
- Oorlog in het Noorderkwartier, de Niedorpen, 1570-1590. Archief van de gemeente Oude en Nieuwe Niedorp, inventarisnummer 139. Stukken betreffende requisitiën, wachtdiensten, sauvegardes e.d., 1570-1590.
- Oorlog in het Noorderkwartier, Graft - De Rijk, 1573-1588. Archief van de Banne Graft-De Rijk, 1348-1607, inventarisnummer 66. Stukken betreffende de krijgshandelingen in het Noorderkwartier, leveranties en requisities., 1573-1593.

- Resoluties van de Alkmaarse vroedschap, 1549-1565. Archief van de stad Alkmaar, 1254 (1325) - 1815 (1816), inventarisnummer 38.
- Resoluties van de Alkmaarse vroedschap, 1565-1682. Archief van de stad Alkmaar, 1254 (1325) – 1815 (1816), inventarisnummer 39.
- Resoluties van de vroedschap van Alkmaar, december 1571 - januari 1575. Archief van de gemeente Alkmaar, 1325-1815, inventarisnummer 39.
- Resoluties van de Alkmaarse vroedschap, 1582-1597. Archief van de stad Alkmaar, 1254 (1325) – 1815 (1816), inventarisnummer 40. Register van resoluties van de vroedschap, 11 augustus 1582 – 30 december 1597.

### **Regionaal Archief Zutphen**

<http://www.regionaalarchiefzutphen.nl>

- Brieven 1460, 1461, 1571-1580.
- Regionaal Archief Zutphen. Transcriptie Brummen, Mark van Empe, notulenboek 1540 - 1708. RAZ, archiefnummer 2008, inv. Nr. 1
- Archief van het Vrouwenconvent Het Oude Convent te Zutphen (1340-1553) Toegang 355, inventarisnummers 1-12
- Oud Rechterlijk Archief Zutphen. Transcriptie Vredeboek, 1573 - 1587, Inventarisnummer 16

### **Stadsarchief Enschede**

<http://stadsarchief.enschede.nl>

Processen stadgericht Enschede, nummer 73, boek 01

### **Stadsarchief van Leuven – Itinera Nova**

<http://itineranova.be/in/home>

- Archief Leuven SAL7303
- Archief Leuven SAL7304
- Archief Leuven SAL7305
- Archief Leuven SAL7306
- Archief Leuven SAL7307
- Archief Leuven SAL7308
- Archief Leuven SAL7310
- Archief Leuven SAL7311
- Archief Leuven SAL7312-7313
- Archief Leuven SAL7314

- Archief Leuven SAL7315
- Archief Leuven SAL7319
- Archief Leuven SAL731B 1423
- Archief Leuven SAL7320
- Archief Leuven SAL7321
- Archief Leuven SAL7322
- Archief Leuven SAL7323
- Archief Leuven SAL7324
- Archief Leuven SAL7325
- Archief Leuven SAL7326
- Archief Leuven SAL7327
- Archief Leuven SAL7328
- Archief Leuven SAL7329
- Archief Leuven SAL7330
- Archief Leuven SAL7331
- Archief Leuven SAL7332
- Archief Leuven SAL7333
- Archief Leuven SAL7334
- Archief Leuven SAL7335
- Archief Leuven SAL7336
- Archief Leuven SAL7337
- Archief Leuven SAL7343
- Archief Leuven SAL7344
- Archief Leuven SAL7345
- Archief Leuven SAL7346
- Archief Leuven SAL7347
- Archief Leuven SAL7349
- Archief Leuven SAL7350
- Archief Leuven SAL7351
- Archief Leuven SAL7352
- Archief Leuven SAL7353
- Archief Leuven SAL7354
- Archief Leuven SAL7355
- Archief Leuven SAL7356
- Archief Leuven SAL7357
- Archief Leuven SAL7358
- Archief Leuven SAL7359

**Streekarchivariaat Noordwest-Veluwe**

[www.streekarchivariaat.nl](http://www.streekarchivariaat.nl)



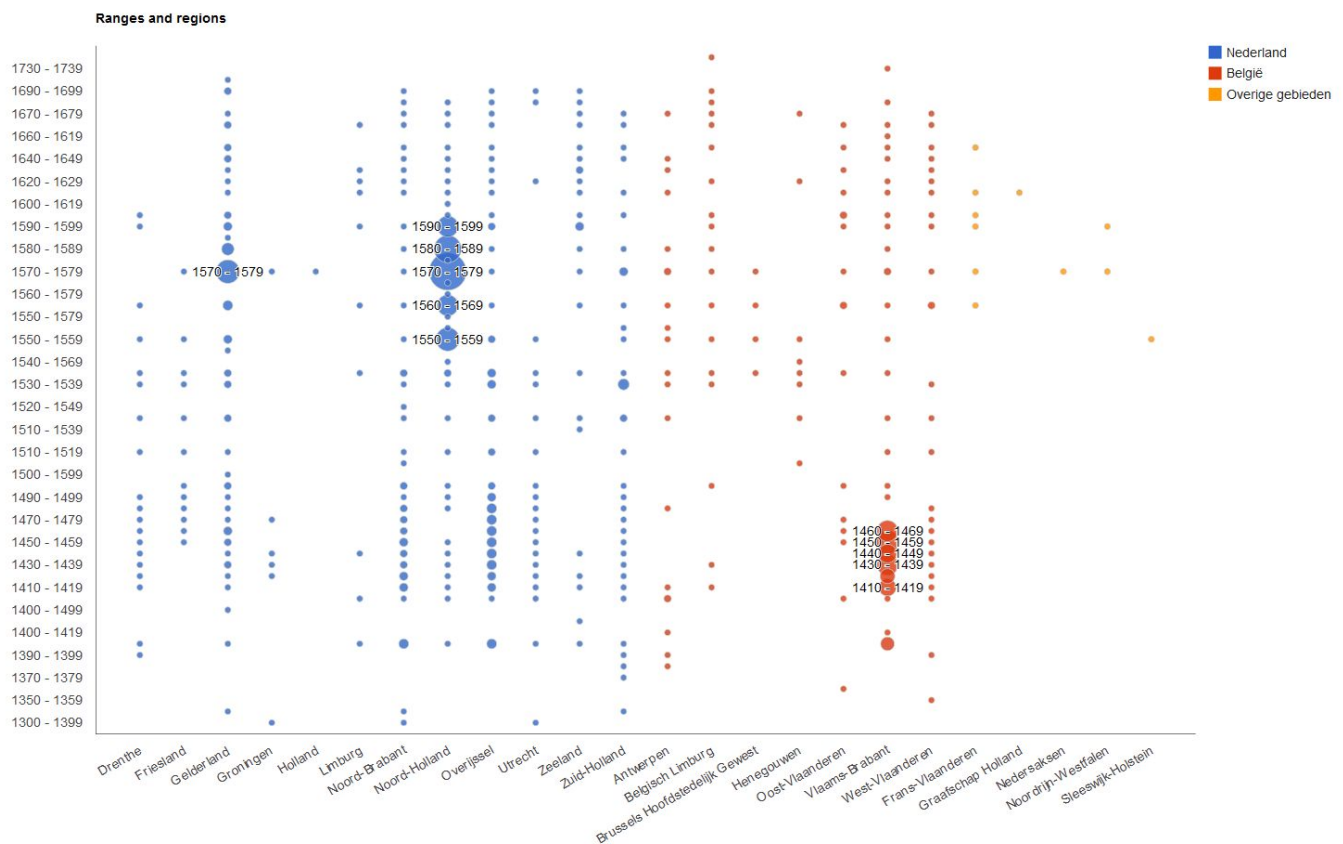
Transcriptie op de criminele rechtspraak 1563-1596 Oorvedenboek Elburg E.  
 Kranenburg-van der Beek, 2000, via de website van Streekarchivariaat Noordwest-Veluwe

## RESULTAAT

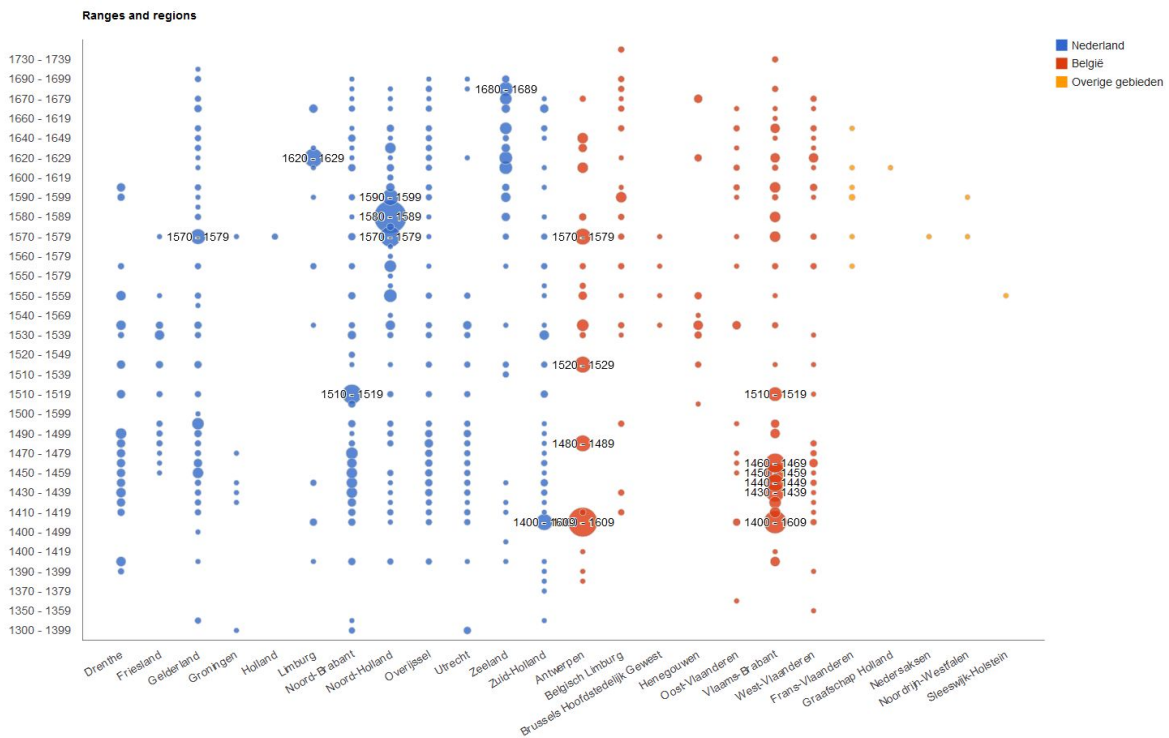
Het uitgangscorpus, waarop het CLVN-corpus een uitbreiding bevat, bestrijkt een iets ruimere range (met viertiende-eeuws tot begin achttiende-eeuws materiaal). We hebben ervoor gekozen om dit niet uit het corpus te verwijderen. Omdat de data precies gedateerd is, kan de gebruiker besluiten om dit specifieke materiaal uit te sluiten.

Het corpus bevat in totaal 5.748.762 woorden uit 4475 documenten. Het vijftiende-eeuwse materiaal bestaat uit 1.843.604 woorden. Het zestiende-eeuws materiaal bestaat uit 2.022.142 woorden. Materiaal na de zestiende eeuw bestaat uit 998.062 woorden, het materiaal voor de vijftiende eeuw uit 37.625 woorden.

Hieronder een view op de spreiding van de documenten per locatie en per decade.



Hieronder een view op de spreiding van het aantal woorden per locatie en er decade



## EVALUATIE

Het CLVN zoals opgeleverd voor Nederlab bestond uit getagde Worddocumenten en 80 Excelbestanden met metadata.

Hoe goed we ook geprobeerd hebben om zorgvuldig te werken, toch bleken er meer issues te zijn met de opgeleverde data dan verwacht.

- Metadatatavelden waren niet in elk Exceldocument hetzelfde
- De waarden waren niet overall consistent ingevuld. Een eenvoudig voorbeeld zijn dateringen: maanden waren aangegeven met 1 of twee cijfers, of met een afkorting; maar in de uiteindelijke metadatadatabase van Nederlab moet de dateringsinformatie geparseerd op dag, maand en jaar erin komen.
- Door tikfouten konden id's in het Excel niet gematcht worden met id's in de tekst
- Door voortschrijdend inzicht waren de laatst verwerkte teksten qua aanwezige codering in het Worddocument anders dan de teksten die in het begin verwerkt waren; binnen het tijdsbestek van het kleine pilotproject was het onmogelijk om dat nog te herzien; dat werk is in de postprocessing voor Nederlab moeten gebeuren
- Door verschillende studenten op eenzelfde deelcollectie te zetten waren er soms kleine verschillen van aanpak binnen één deelcollectie, hetgeen het verder automatisch verwerken van de bestanden complexer maakte.

Conclusie is dat voor een corpus van deze omvang, de werkwijze anders moet. Er moet meer tijd bij om

- a. voor alle geselecteerde bestanden een zorgvuldige analyse te maken van de editieprincipes en dan vast te stellen hoe e.e.a. verwerkt moet worden
- b. om een infrastructuur op te bouwen die ervoor zorgt dat
  - er gewerkt wordt met een metadataschema waar niet vanaf geweken kan worden
  - er gewerkt wordt op een manier dat er valide XML wordt opgeleverd (en niet meer een Worddocument met alvast wat XML-codering erin)
  - er gewerkt wordt in een omgeving waarin id's nooit handmatig in tekst en metadatarecord moeten worden gezet zodat er geen mismatches meer zijn tussen tekst en metadata.

Tot slot merken wij op dat het uiteindelijke doel, namelijk om een lacune te dichten in de data van Nederlab, maar zeer beperkt is geslaagd. Weliswaar zijn er extra data aan Nederlab toegevoegd, maar die data kunnen we niet goed bevragen, de data kunnen niet worden gesorteerd op plaats, en er kunnen niet automatisch kaarten van de data worden gemaakt, waardoor er geen mogelijkheid bestaat zicht te krijgen op de geografische verbreiding van taalkenmerken, zoals de verbreiding van *naar*.

## Appendix.

### **Instructies voor de student-assistenten met betrekking tot opgeloste verkortingen en andere redactionele toevoegingen**

#### *1. Vooraf*

Alle teksten die jullie bewerken, zijn door redacteuren getranscribeerd. Bij dergelijke transcripties moeten onvermijdelijk keuzes worden gemaakt, en die keuzes worden op verschillende typografische manieren aangegeven. In principe moeten jullie die keuzes respecteren, maar jullie moeten wel inventariseren welke keuzes er zijn gemaakt. Die inventarisatie moeten jullie vermelden in de metadata bij de opmerkingen. Kijk om te beginnen even in de inleiding van de teksten wat er precies is veranderd door de redacteuren, en neem dat op in een opmerking in de metadata. Soms staat er bijvoorbeeld dat er voor de leesbaarheid hoofdletters aan het begin van een zin zijn toegevoegd, of dat men interpunctie heeft toegevoegd. Maak daar altijd even een opmerking over in de metadata.

Verder moet je in de transcripties bekijken welke conventies er zijn gebruikt met betrekking tot het oplossen van verkortingen en het toevoegen van redactionele opmerkingen. Zie de volgende drie punten. De meest gebruikte conventie is tegenwoordig: verkortingen cursief of tussen rondhaken en redactionele toevoegingen tussen vierkante haken. In de gevallen dat je een keuze moet maken (zie punt 4), zou je op deze conventie kunnen terugvallen. Belangrijk is te weten dat in deze periode eigenlijk geen ronde of vierkante haken in de originele tekst voorkwamen, haken zijn van latere datum.

In het algemeen: zodra je niet weet wat je moet doen, neem dan geen zelfstandige beslissingen maar overleg met Margit en mij, zodat we voorkomen dat werk over moet worden gedaan. Als het om een enkel los geval gaat waar je twijfel over hebt, zet daar dan in de tekst via de functie Opmerking toevoegen een opmerking bij. Als je de tekst dan inlevert, moet je ons even wijzen op deze opmerkingen en dan kijken wij er even naar.

#### *2. Verkortingen*

In de 15<sup>e</sup> en 16<sup>e</sup> eeuw werden nogal eens verkortingen gebruikt, afkortingen. Het is gebruikelijk dat redacteuren die verkortingen voluit schrijven, maar over het algemeen geven ze wel aan dat het gaat om een uitwerking door de redacteur. Dat doen ze op verschillende manieren: soms cursiveren ze de letters die ze hebben toegevoegd – *en* in het origineel wordt getranscribeerd als *ende* –, soms plaatsen ze die tussen ronde haken – *en(de)* – en soms tussen vierkante haken – *en[de]*.

In de teksten die jullie bewerken, komen jullie deze uitgewerkte verkortingen ook tegen; in punt 4 van dit document staat een aantal voorbeelden van uitwerkingen uit de Bijbelvertaling, waar we cursief gebruikt hebben, en uit de gekaapte brieven, waar ronde haken zijn gebruikt. Die voorbeelden laten zien wat voor zaken allemaal in het verleden verkort werden.

De instructie is dat jullie in de metadata als algemene opmerking vermelden welke conventie er is gebruikt, dus, afhankelijk van de situatie:

- verkortingen zijn gecursiveerd,
- verkortingen zijn tussen ronde haken geplaatst,

- verkortingen zijn tussen vierkante haken geplaatst
- verkortingen zijn stilzwijgend opgelost.

Uiteraard moeten uiteindelijk overal in alle bestanden verkortingen op dezelfde manier worden weergegeven, maar het rechte trekken van de verschillende notatiewijzen gebeurt in Leiden op het INT; daarom is het belangrijk dat men in Leiden via de metadata weet hoe de verkortingen in ieder bestand zijn behandeld.

Jullie hoeven verder in principe niets te doen aan de verkortingen. Dus geen letters of haken weghalen.

### *3. Redactionele opmerkingen*

Behalve uitgewerkte verkortingen zijn er soms door de redacteurs ook andere opmerkingen toegevoegd. Bijvoorbeeld [onleesbaar] of (onleesbaar). Opnieuw zijn hier door de redacteurs soms verschillende haken gebruikt, namelijk vierkant (de meest gebruikte conventie tegenwoordig) of rond, bijvoorbeeld: [marge] of [?] of [onleesbaar], maar ook: (marge) etc.

Kijk even of dit soort informatie in een bestand voorkomt en zo ja: plaats een opmerking in de metadata hierover.

Daarnaast worden er zo af en toe redactionele opmerkingen of moderne vertalingen in de vorm van noten toegevoegd aan de tekst. Die noten moeten jullie weghalen uit het bestand of coderen. Maar opnieuw wel een opmerking toevoegen in de metadata dat je dat hebt gedaan.

### *4. Dezelfde soort haken voor meerdere functies*

Jullie moeten goed op het volgende letten: als er ter aanduiding van redactionele opmerkingen en uitgewerkte verkortingen dezelfde soort haken (rond dan wel vierkant) zijn gebruikt, hebben we een probleem. Het liefste zou je willen dat redactionele opmerkingen tussen vierkante haken staan en uitgewerkte kortingen tussen ronde of cursief. Dan wordt er namelijk een typografisch verschil gemaakt tussen de twee soorten gegevens. De redactionele opmerkingen zijn moderne opmerkingen die je uit de tekst wilt houden, de uitgewerkte verkortingen zijn weliswaar interpretaties maar ze zijn wel onderdeel van de oorspronkelijke tekst, het is eigenlijk een spellingskwestie: we gebruiken een wat andere spelling dan de oorspronkelijke tekst voor de leesbaarheid en doorzoekbaarheid.

In het geval dat beide soorten opmerkingen met dezelfde soort haken zijn aangeduid, overleg dan even met Margit en mij over wat te doen. Als bijvoorbeeld als enige redactionele opmerkingen voorkomt (onleesbaar), dan zou mijn voorstel zijn om te zoeken op dat woord en het overal tussen vierkante haken te zetten. Het beste is even per document te bekijken hoe de situatie is. Maar ga niet heel veel veranderingen aanbrengen zonder eerst te overleggen met Margit en mij.

De volgende oplossing is in dit soort gevallen wat ons betreft prima: als de opgeloste afkortingen tussen ronde haken staan (en(de)) en er redactionele opmerkingen zijn tussen vierkante haken [onleesbaar], en er daarnaast ook nog in de tekst stukken tekst staan die niet een afkorting zijn maar wel tussen ronde haken staat, zet die dan tussen @@. (Dergelijke ronde haken zijn ook door de redacteurs toegevoegd, waarschijnlijk ontbrak deze in het

origineel, maar sommige redacteuren hebben voor de leesbaarheid interpunctie toegevoegd). Dus bijvoorbeeld (met de tekst waar het om gaat vetgezet):

"...en[de] op Arnhem an d[en] lantrentm[eeste]r **(die van de schattinge id ontfanck hefft)** gelievert sijn word[en]..."

wordt, om te voorkomen dat het clasht met de opgeloste afkortingen (en(de)):

"...en(de) op Arnhem an d(en) lantrentm(eeste)r **@die van de schattinge id ontfanck hefft@** gelievert sijn word(en)..."

Doe dit consequent en zet een opmerking in de metadata.

### 5. Doorhalingen

Soms is tekst in de oorspronkelijke documenten doorgehaald, en loopt dus een streep doorheen in het Worddocument. Bijvoorbeeld: cleyner. Dit moeten jullie vervangen door: <del>cleyner</del>

Let op: jullie moeten het doorhalen dus ook weghalen (selecteer de bewuste tekst, kies Lettertype en haal het vionkje weg bij doorhalen).

### 6. Voorbeelden van verkortingen uit de Bijbelvertaling en de gekaapte brieven

Bijbelvertaling:

Breviaturen moeten in de bijbeltekst worden opgelost, en de oplossingen moeten cursief worden gezet. Het gaat alleen om wat afgekort is met tekenjes waar slechts één of een beperkt aantal oplossingen voor zijn (= de oorspronkelijke, specifieke afkortingstekens van de middeleeuwse kopiist). Dit zijn er slechts enkele:

- de tilde boven een letter die alleen een n, m of -de kan representeren.
- de komma of slash door een v die alleen -er- of -ers representeert (afkorting voor -er- in de ver- van vergelijk of voor -er- of -ers in vers); het gaat om de v.'s in de capittelinleidingen, waar een v met een komma (of een punt) erboven-en-erdoor staat. De punt achter de v geeft nogmaals aan dat de v een afkorting voorstelt. Zo'n v wordt opgelost als *vers*, waarin *ers* cursief is. Hiermee vervalt ook de punt achter de v, dus: v. 21 wordt *vers* 21 (met *ers* cursief). De breviatuur vss. lossen we op als *versen*. **NB.** Het oplossen van *vers* en *versen* gebeurt alléén in de (inleidende stukken van de) broodtekst en de registers; in de kanttekeningen handhaven we de afkorting V. (zie de instructie voor de kanttekeningen hieronder)
- het speciale teken dat aan de c. voorafgaat in etc. en dat et- representeert (vaak &c., soms nog een ander teken), bijvoorbeeld: ende troost der genadiger verhooringe Godes, 55, *etc.*

Enkele voorbeelden van opgeloste breviaturen: dē HEERE wordt uitgewerkt als: *den* HEERE, eē wordt uitgewerkt als *een*. Let daarbij goed op; er staat bijvoorbeeld bij 2 Makkabeeën 15: 32 een klein streepje boven de n van het eerste woord 'en', en verderop staat 'tege'. Dit wordt dus:

*Ende* haer toonende het hoofd des schelmachtigen Nicanors, ende de handt deses Godts-lasteraers, welcken hy uytgestreckt hadde *tegen* het heylige huys des Almachtigen, ende den hals hadde opgesteken:

Gekaapte brieven:

Afkortingen kunnen vaak meerdere betekenissen en dus meerdere uitwerkingen hebben, afhankelijk van de context. Sommige uitwerkingen kunnen op verschillende manieren worden gespeld; kijk daarvoor naar de context en neem de uitwerking die het meest lijkt op de verder gebruikte spelling. Zo kan oudh. dus worden uitgewerkt als oudh(eijdt), oudh(eijt) of oudh(eydt).

Voorts:

@ = a(nno), soms ook a(men)

a:p: = a(nno) p(assato)

ady, adi = a(nno) d(omin)i; ady, adi kan ook zijn: a(d) di(em) (Latijn voor: op de dag)

Ao. overnemen als A(nn)o [zonder punt erachter, omdat de afkorting is uitgewerkt]

ap (waarbij de p wordt geschreven met een slinger omhoog, zoals de afkorting voor pre):

a(nno) p(assato)

b̄ = b(er)

c = c(ito); c(arolusgulden)

Comp = Comp(agnie)

compt. = compt(oir) of comp(aran)t

Curat: = Curat(or)

d = d(ito); d(enier)

DGG = D(ie) G(od) G(eleide)

D:H: = D(en) H(eer)

d'E Comp = d'E(dele) Comp(agnie)

decemb. = decemb(ris)

den 11<sup>o</sup> dezēs – nemen we over als: den 11. dezēs

dn = d(e)n

do = do(minee)

dto = d(a)to

Dublic: = Dublic(aat)

DW = D(ienst)W(illige)

Edlgr Agtb<sup>e</sup> = EdelGr(oot) Agtb(ar)e

em = em(mer)

f = f(acit)

G:t = G(roo)t (een Wissel G(roo)t f3514:-:-)

geauth = geauth(oriseerd)

Geb<sup>e</sup> = Geb(oor)e (titulatuur: Wel Edele Geboore Heer)

Gem<sup>e</sup> = Geme(lde) (gemelde capitein ...)

Geste = Gest(reng)e

geteek. = geteek(end)

gld = g(u)ld(en)

Gouv<sup>r</sup> Gen<sup>l</sup> = Gouv(erneu)r Gen(eraa)l

h<sup>e</sup> k<sup>y</sup> = he(le) k(off)y

Het teken voor florijn overnemen als in de gewone f

k = k(an)

Lb [ligatuur met slinger erdoor] = L(i)b(ra); NB. pond stond behalve als waarde aanduiding (geld) ook voor een gewichtsaanduiding (pond), en daarbij maakten de meesten een consequent onderscheid tussen Lb (met of zonder punt of slangetje) = geld, lb (met of zonder punt of slangetje) = gewicht. Meestal worden geldponden Lb met hoofdletter, en gewichtsponden lb met kleine letter; denk aan het Engelse pondteken, dat dus een nog sterker tot een symbool geworden hoofdletter L als basis heeft.

ld = l(oo)d [is 1/32 pond]

ll = l(est)l(eden)

LZelve (L met een schuin streepje door de stok) = (des)Zelve

oxe = ox(hoofd)e

pic = pic(ol)

rds = r(ijks)d(aalder)s

m = m(uts), m(orgen) als landmaat, m(erk), m(aakt)

md. = m(aan)d of m(aan)d(en)

ODW = O(nderdanige) en D(ienst)W(illige)

p [lees: p met een horizontale streep door de stok, zie Van der Gouw] = p(er)

p [lees: p met haaltje links door door de stok, zie Van der Gouw] = p(ro)

p = p(enningen)

þ = p(re)

pass = pass(ato)

plant<sup>e</sup> = plant(ag)e

pochie [lees: : p met een horizontale streep door de stok] = p(ar)ochie

pt = p(ar)t

qq = q(ualitate) q(ua)

rb = r(oe)b(el)

rd = r(ijks)d(aalder)

S = S(eigneur) of in Italiaanse context S(ignore), on Spaanse context S(eñor)

sa = sa(lige)

sch = sch(ellingen)

Sr = S(eigneu)r

st = st(ui)vers

stv = st(ui)v(ers)

Trip: = Triplic(aat)

U L = U L(iefhebbende), UL(ieve); let op: in de betekenis Ulieden of Ulieder laten we het verkort staan (met punt als die in het origineel staat)

UEdg = UEd(el)g(estrenge)

uwelEdgs = uwelEd(el)g(e)s(trengen) (*al of niet met meervouds n*)

uwesEdgs = uwesEd(el)g(e)s(trengen) *of ook 2e naamval* uwesEd(el)g(estrengen)s

v [lees: v met / erdoor] = v(er)

wed. = wed(uwe)



xbris = (decem)bris

z (klein of groot, aan het eind van een word) = h(eid) [of h(eyt), h(eit), kies spelling conform rest tekst]; in namen staat het voor zoon, dat we niet oplossen (zie boven); ook de vadersnaam is soms afgekort, dat zullen we meestal ook onopgelost laten omdat we niet weten wat het betekent, bijvoorbeeld Chz; dan lossen we op Ch(ristiaans)z. of Ch(ristoffels)z.  
Z = Z(eer)