



In nieuwe tekens, die overigens de taal niet vervangen. De totstandkoming van een digitale onderzoeksdocumentatie bij de *Volledige Werken* van Willem Frederik Hermans

Peter Kegel en Bert Van Elsacker

NEERLANDISTIEK.NL 04.04; GEPUBLICEERD: 22 SEPTEMBER 2004

'Voor de wetenschap houdt de taal op te bestaan. Zij denkt in nieuwe tekens, die overigens de taal niet vervangen, maar een aparte denkvorm uitmaken, de enige denkvorm die zekerheid oplevert, omdat de onzekerheid er bij afspraak uit verwijderd gehouden wordt, voor zover men daarin is geslaagd.'

– Willem Frederik Hermans, *Paranoia*, zestiende druk 1993, p. 13.

Samenvatting

Het Constantijn Huygens Instituut bereidt momenteel de uitgave voor van de *Volledige Werken* van Willem Frederik Hermans. Het omvangrijke onderzoek resulteert uiteindelijk in een leeseditie van circa drieëntwintig verzamelbanden, die in een reeks van jaren zal verschijnen bij de Bezige Bij. Zowel het onderzoek als de voorbereiding voor publicatie zijn ondenkbaar zonder uitgebreide toepassing van digitaliseringstechnieken. Voor de collatie van alle voor het tekstueel onderzoek relevante drukken wordt gebruik gemaakt van het softwareprogramma *Collate*. Alle teksten worden na collatie ondergebracht in een full-text XML-TEI-bestandsformaat met geïntegreerd variantenoverzicht. De digitale onderzoeksomgeving die zo ontstaat vormt de basis voor vervolgonderzoek en onderzoeksverslaglegging door de editoren. De elektronische bestanden dienen bovendien als directe kopij voor de boekuitgave en worden gebruikt voor digitale publicatie van onderzoeksmateriaal. De digitale duurzaamheid van de XML-TEI-coderingen maakt dat teksten en onderzoeksgegevens beschikbaar blijven voor nieuwe publicaties en digitaal vervolgonderzoek.

Summary

At present the Constantijn Huygens Institute is preparing an edition of the Complete Works of the Dutch 20th century author Willem Frederik Hermans. The extensive research on this project will eventually result in the publication of a reading edition. This will comprise a set of about 23 volumes, that are to be published by Dutch publishing house "de Bezige Bij" in the course of several years. It became clear that it would be inconceivable that research on and preparations for this publication would be possible without extensive use of several techniques of digitisation.

For the collation of all of the printed editions of Hermans' work that are relevant for textual research the Constantijn Huygens Institute uses the software programme Collate. After collation all texts are digitally stored in a full text XMLfile format, in compliance with rules set by TEI. Integral part of the corpus is a survey of textual variants.

The digital environment that is thus created for research will serve as a sound base for further scholarly research and for a protocol for administration and reports on the work in progress by the editors during the realisation of the project. Also the digital files will serve as direct (soft) copy for the printed edition and besides will be used for digital publication of research material. Preserving the texts and research in XML format, in compliance with the rules set by TEI, is meant to facilitate the lasting (digital) availability of the texts for new publications and further computer aided research.

1 Inleiding

In februari 2000 zijn bij het Constantijn Huygens Instituut de voorbereidingen begonnen voor de uitgave van de *Volledige Werken* van Willem Frederik Hermans, die vanaf najaar 2005 bij uitgeverij De Bezige Bij zullen verschijnen. Behalve al het zelfstandig verschenen werk van Hermans bevat de leeseditie honderden niet eerder gebundelde teksten, waaronder veel artikelen en recensies uit de vroege naoorlogse jaren. De *Volledige Werken* worden gepresenteerd in verzamelbanden, onderverdeeld naar genre, met binnen de genres een chronologische opzet. Uitgangspunt voor de editie zijn de 'ultima manus'-uitgaven van de afzonderlijke teksten. Een voorlopige indeling gaat uit van circa drieëntwintig banden van elk 800 à 1000 pagina's. De uitgave van de *Volledige Werken* wordt gerealiseerd in samenwerking met het Willem Frederik Hermans Instituut (WFHI).¹

Bij de voorbereidingen en realisatie van de editie wordt op grote schaal gebruik gemaakt van nieuwe ontwikkelingen op ICT-gebied. Ten behoeve van het tekstkritisch onderzoek worden per titel alle voor het onderzoek relevante drukken gecollationeerd met behulp van het softwareprogramma Collate. Het editeuronderzoek dat op basis van de door Collate geproduceerde gegevens plaats vindt, wordt systematisch digitaal gedocumenteerd. Hierbij wordt gebruik gemaakt van XML-TEI-standaarden. Ook de kopij voor de gedrukte editie wordt in XML-TEI-formaat bij de uitgeverij aangeleverd. Naast de boekuitgave worden onderzoeksresultaten en digitale publicaties op het internet gepresenteerd. Ten slotte moet langdurige toegang tot de elektronische documenten verzekerd zijn: zowel de teksten als het overig onderzoeksmateriaal moeten in de toekomst gebruikt kunnen worden voor nieuwe uitgaven en voor nieuw onderzoek met behulp van digitale technieken.

In dit artikel geven we een overzicht van de gevolgde werkwijze bij de digitalisering voor de *Volledige Werken* van Willem Frederik Hermans. Na een korte uiteenzetting van de eerste fase van de digitalisering, staan we stil bij de technische voorwaarden en toepassingsmogelijkheden van Collate; we demonstreren het gebruik van dit programma aan de hand van een concreet voorbeeld. Bijzondere aandacht gaat daarna uit naar de conversie van de collatiebestanden naar een XML-TEI-opslagformaat, de belangrijkste standaard voor het blijvend beschikbaar houden en het verrijken van digitale documenten. De in XML-TEI gegenereerde collatiebestanden vormen het uitgangspunt voor het creëren van een digitale onderzoeksomgeving voor de editoren van de *Volledige Werken*. We beschrijven, opnieuw aan de hand van voorbeelden, enkele mogelijkheden van een verregaand gedigitaliseerde editiepraktijk en gaan in op de voordelen van digitalisering bij onderzoek en publicatie. Ten slotte kijken we vooruit naar het toekomstig gebruik van elektronische onderzoeksdocumentaties door vakgenoten en andere letterenonderzoekers.

2 Microverfilming, digitalisering en codering

Het onderzoek bij de voorbereidingen van de *Volledige Werken* is in belangrijke mate afhankelijk van Collate. Het gebruik van het collatieprogramma veronderstelt dat al het te vergelijken materiaal in

¹ Zie voor algemene informatie over het project: <http://www.chi.knaw.nl/nieuwetijd/hermans/project.html>

elektronische vorm beschikbaar is. Daarom zijn in een eerdere fase van het project alle voor het onderzoek benodigde teksten in twee fases gedigitaliseerd.² In eerste instantie werd, in nauwe samenwerking met specialisten van het [Bureau Metamorfoze](#) van de Koninklijke Bibliotheek, al het onderzoeksmateriaal op microfilm gezet. Bij de verfilming waren extreem hoge kwaliteitsstandaarden noodzakelijk: elke vorm van vertekening, elk gebrek of teveel aan contrast op de films, elke onvolledigheid of verdubbeling op het filmbeeld immers zou, bij de digitalisering of in een latere fase van het project, onherroepelijk tot grote problemen leiden. Alle films werden vervolgens gescand, waarna de gescande afbeeldingen met behulp van Optical Character Recognition (OCR) zijn omgezet naar computerleesbare tekst. Ook hier moesten, hoewel het digitaliseren van twintigste-eeuws materiaal in het algemeen goede resultaten oplevert, talrijke detailproblemen worden opgelost.³ Bij het digitaliseren werd gebruik gemaakt van de expertise van het [Nederlands Instituut voor Wetenschappelijke Informatiediensten](#) (NIWI).

Na de digitalisering van het materiaal volgde een laatste fase in het voorbereidende traject. Om elektronisch collationeren mogelijk te maken, was het noodzakelijk de OCR-bestanden op systematische wijze te voorzien van eenduidige codes: structuurkenmerken en typografische informatie met inhoudelijke betekenis moesten expliciet worden gecodeerd, en er werden commandocodes voor Collate toegevoegd.⁴ Op basis van talrijke praktische tests werden voor dit codeertraject, dat voor het overgrote deel is geautomatiseerd, gedetailleerde procedures vastgesteld, die zijn beschreven in een uitgebreide handleiding voor codeurs. De handleiding is toegespitst op het tekstcorpus van het Hermansproject, maar kan algemeen worden gebruikt.⁵ De codering van de genres 'Romans' en 'Verhalen en novellen' is inmiddels grotendeels voltooid. Het gehele codeertraject zal naar verwachting medio 2005 worden afgerond.⁶

3 Collate

Collate kan grote hoeveelheden tekst in zeer korte tijd automatisch met elkaar vergelijken; de onderzoeker kan voor een groot deel zelf bepalen hoe en in welke volgorde tekstdelen met elkaar

² Bij het vaststellen van de bronnen diende de bibliografie *Het bibliografische universum van Willem Frederik Hermans* (Janssen en Van Stek 2000) als uitgangspunt. Voor de reconstructie van de tekstgeschiedenis worden - naast alle overgeleverde typoscripten en manuscripten, drukproeven en correctie-exemplaren - van elke tekst de eerste en de laatste bij leven verschenen druk, alle herziene en vermeerderde drukken, alle 'opnieuw gezette' drukken en alle overige drukken met varianten onderzocht. Grofweg is dat een derde van het in *Het bibliografische universum* verzamelde materiaal, in totaal goed voor circa 50.000 pagina's.

³ Op de websites van het [Electronic Text Center Virginia](#) en het [Leeds Electronic Text Centre](#) is relevante informatie te vinden over het digitaliseren van gedrukte teksten en de daarbij voorkomende problemen.

⁴ Belangrijk zijn allereerst alineas, witregels en andere structuurkenmerken. Daarnaast worden dialoogstreepjes, weglatingsstreepjes, gedachtestreepjes en koppeltekens van elkaar onderscheiden; ook apostrofs, aan- en afhalingen bij dialogen, vette, cursieve en gespatieerd gezette tekst worden expliciet gemarkeerd.

⁵ De Vlaamse onderzoekster Liesbeth van Melle bereidt momenteel een historisch-kritische editie voor van Maurice Gilliams' romans *Elias* (1936) en *Winter te Antwerpen* (1953). Het betreft hier een promotieonderzoek bij de Vakgroep Nederlandse literatuur en Algemene Literatuurwetenschap van de universiteit van Gent met als promotor Prof.dr. A.M. Musschoot; co-promotor is Dr. Y. T'Sjoen. De codering en collaties van het digitale materiaal worden in samenwerking met het Constantijn Huygens Instituut verricht op basis van de voor het Hermansproject ontwikkelde standaarden.

⁶ Zie voor een uitgebreide uiteenzetting over microverfilming, digitalisering, codering en de daarbij ondervonden complicaties ['Peter Kegel en Bert Van Elsacker, 'In progress – The Complete Works of Willem Frederik Hermans: a report on the integration of digital technologies in the publication of a multi-volume edition'](#) (Kegel en Van Elsacker 2003).

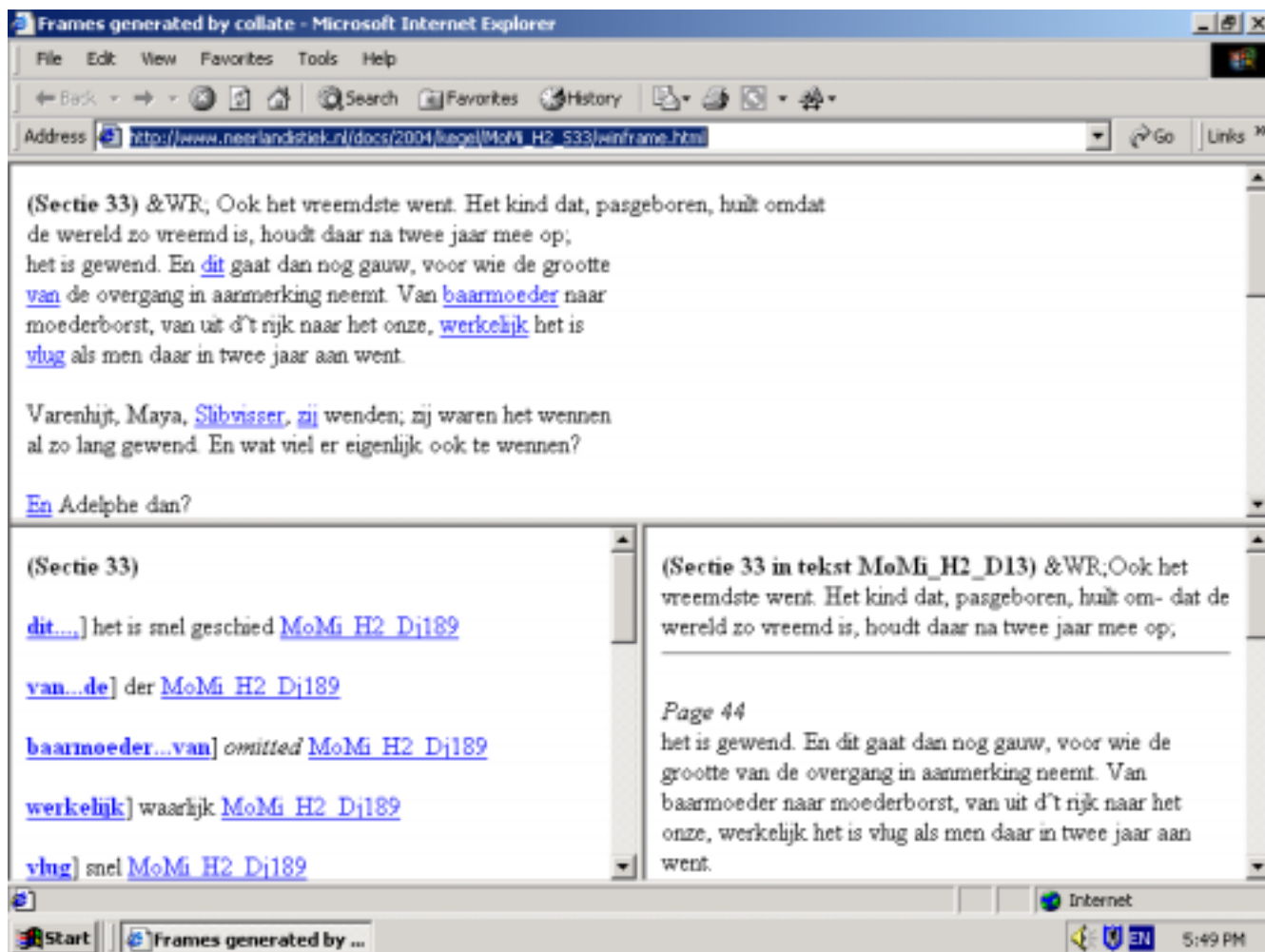
worden vergeleken. Bovendien kunnen gedetailleerde collatievoorschriften worden gegeven: een gebruiker van het programma kan bijvoorbeeld lijsten opgeven van woorden en tekens die als equivalent moeten worden beschouwd, of instellen in welke mate woorden op elkaar moeten lijken om de status van variant te krijgen. Het variantenoverzicht dat daarbij ontstaat kan op veel verschillende manieren worden gepresenteerd.

Het programma is oorspronkelijk ontworpen voor het onderzoek naar de vele manuscripten van Chaucers *Canterbury Tales*. Vooral het voorbereiden van de cd-rom uitgave van *The Wife of Bath's Prologue* (Cambridge UP, 1996) is bepalend geweest voor de technische verfijning van het programma. Collate staat aan de basis van een aantal digitale publicaties rondom *The Canterbury Tales*, die gezien kunnen worden als schoolvoorbeelden van een nieuwe editiewetenschappelijke praktijk: ze stimuleren nieuwe vormen van onderzoek en maken andere leeswijzen van de middeleeuwse teksten mogelijk.⁷

Bij het Hermansproject wordt Collate gebruikt voor de vergelijking van een groot aantal gedrukte teksten: de collaties moeten op een overzichtelijke manier gedetailleerde overzichten geven van alle bedoelde en onbedoelde wijzigingen die zich in de loop van de tijd in de teksten van Hermans hebben gemanifesteerd. Een full-text weergave in een HTML-formaat, een standaard uitvoermogelijkheid van Collate, vormt een goed uitgangspunt voor het onderzoek van de tekstgeschiedenis. We demonstreren dit aan de hand van een kort fragment uit het verhaal 'Atonale', dat (vanaf de tweede druk) deel uitmaakt van de verhalenbundel *Moedwil en misverstand*.⁸

⁷ Zie voor een overzicht van gerealiseerde en komende publicaties de website van het [Canterbury Tales Project](#). Gedetailleerde informatie over de mogelijkheden van Collate voor het Chauceronderzoek is onder andere te vinden bij [Peter M. W. Robinson, *New methods of editing, exploring, and reading The Canterbury Tales*](#) (Robinson 1998). Collate is ook bij andere grote editieprojecten van groot belang, bijvoorbeeld bij de '[Nestle-Aland](#)'-editie van het Griekse Nieuwe Testament (Institut für Neutestamentliche Textforschung (INTF), Universiteit van Münster) en bij het Zwitserse [Parzival-Projekt](#) (Universiteit van Basel); bij dit laatste project berust de elektronische editie bijna geheel op HTML-Collatebestanden.

⁸ Het getoonde gegevensoverzicht bevat ruw werkmateriaal: de elektronische bestanden hebben nog enkele voorlopige coderingen, zoals &WR voor witregel en &APO voor apostrof; ook kunnen in alle bronnen met uitzondering van de basistekst nog scan- en codeerfouten voorkomen; de onderverdeling van de tekst in secties is noodzakelijk om de teksten voor collatie geschikt te maken.



[afbeelding 1. HTML selectie van sectie 33 MoMi]

Een korte uitleg bij de drie frames. Het bovenste frame geeft de volledige tekst van 'Atonale' weer, naar de dertiende druk van *Moedwil en misverstand* (D13, JS 25)⁹, die als basistekst voor de editie dient; alle varianten zijn zichtbaar als hyperlinks. Door op een variant te klikken verschijnt deze in een gelemmatiseerd apparaat, in het frame linksonder, met een sigle aanduiding van de drukken die de variant hebben. Vanuit dit apparaat kan worden doorgedrukt naar de directe context van de variant, die in het frame rechtsonder verschijnt. Uit deze documentatie kunnen alle veranderingen die de novelle vanaf de eerste tijdschriftpublicatie in *Podium* tot en met de laatste druk in *Moedwil en misverstand* heeft ondergaan vrij eenvoudig worden afgelezen.¹⁰

Voor het primaire doel van de *Volledige Werken*, het bezorgen van een betrouwbare tekst, zijn lang niet alle varianten van belang. Het gaat ons specifiek om die tekstplaatsen waar, door de jarenlange overdracht van de tekst van de ene naar de andere informatiedrager, onbedoelde fouten of

⁹ Verwijzingen naar publicaties van Hermans worden steeds aangegeven via nummeringen uit de beschikbare bibliografieën: JS-nummers verwijzen naar de nummering in *Het bibliografische universum*. De drukaanduiding wordt weergegeven met een sigle en correspondeert, tenzij anders vermeld, met de gegevens van de bibliografie. DJ-nummers verwijzen naar *Schrijven is verbluffen; bibliografie van de verspreide publicaties van Willem Frederik Hermans*. (Delvigne en Janssen 1996).

¹⁰ Voor het onderzoek naar de tekstgeschiedenis van dit verhaal zijn op basis van *Het bibliografische universum* de voorpublicatie uit *Podium* (DJ 189) en de tweede, vierde, achtste, negende druk en dertiende druk van *Moedwil en misverstand* gecollationeerd (JS 13, 16, 20, 21 en 25).

onduidelijkheden zijn ontstaan. In dat opzicht is de zin "Dit alleen om de mate waarin Adelphe Varenhijt hinderde uit te drukken." relevant. Deze passage komt voor tot en met de zevende druk van *Moedwil en misverstand*, waarna de achtste en negende druk een variant bevatten: "Dit alleen om een denkbeeld te geven van de mate waarin Adelphe Varenhijt hinderde". Bestudering van de correctie-exemplaren uit het Hermansarchief leert dat die wijziging een direct gevolg is van een herziening die Hermans aanbracht in een correctie-exemplaar van de zevende druk.



[afbeelding 2. Correctie-exemplaar van *Moedwil en misverstand* D7, JS 19. Collectie Willem Frederik Hermans]

Merkwaardig genoeg blijkt dit fragment in de basistekst weer te zijn 'hersteld' naar de oude versie. De variant staat bovendien niet op zichzelf: ook op andere plaatsen in de bundel (niet zichtbaar in het getoonde fragment) komen varianten voor waarbij eerdere herzieningen zijn vervallen. Op basis van

een voorlopige bestudering van het varianten- en archiefmateriaal lijkt het erop dat het hier gaat om onbedoelde veranderingen in de tekst: ergens in het productieproces van een nieuwe druk van *Moedwil en misverstand* moet ten onrechte zijn teruggegrepen op een oude legger.¹¹ Uit aanvullend onderzoek blijkt dat vanaf de tiende druk (D10, JS 22) weer de oude versie van de tekst wordt gevolgd. De tekstgeschiedenis van *Moedwil en misverstand* is daarmee ingewikkelder dan *Het bibliografische universum* veronderstelt; bij de definitieve tekstconstitutie van de bundel worden daarom meer teksten gecollationeerd.¹²

Alhoewel het hier slechts een zeer beperkt fragment betreft, toont het voorbeeld direct de kracht van Collate: een complexe tekstgeschiedenis wordt, tot in de kleinste details, in kaart gebracht. Voor het Hermansproject is zo'n instrument voor analyse van het zeer omvangrijke en complexe bronnenmateriaal van cruciaal belang. De presentatie van de gegevens in HTML-formaat maakt dat de editor, met behulp van een willekeurige internetbrowser, alle veranderingen in een handzaam elektronisch formaat kan bestuderen. De productie van het variantenmateriaal in HTML maakt het bovendien mogelijk om de gegevens digitaal ter beschikking te stellen aan andere onderzoekers.

4 Collate en XML-TEI

Ondanks de duidelijke voordelen van een dergelijke weergave is een meer dynamisch gebruik van het digitale materiaal wenselijk, waarbij de editoren bij het tekstkritisch onderzoek informatie aan de gegevens kunnen toevoegen. Hierbij kan worden gedacht aan een directe verantwoording van ingrepen, aan regularisaties van inconsistenties in spelling, maar ook aan verwijzingen naar archiefinformatie, aan annotaties bij onduidelijke tekstplaatsen, aan toelichtingen bij en indexering van persoonsnamen of aan bibliografische verwijzingen. De digitale documentatie zou idealiter ook als uitgangspunt moeten kunnen dienen voor verder onderzoek door editiewetenschappers, letterkundigen en overige onderzoekers.

Een dynamische full-text onderzoeksdocumentatie op basis van de variantenoverzichten wordt mogelijk door gebruik te maken van XML-TEI-standaarden. Eerste voorwaarde is dan de conversie van de Collatebestanden tot een dergelijk formaat. Collate biedt daartoe in beginsel de mogelijkheden, maar een aantal technische aanpassingen is dan noodzakelijk. In deze paragraaf, die een overwegend technisch karakter zal hebben, geven we achtereenvolgens een beschrijving van de verschillende standaarden, de gevolgde werkwijze en de daarbij opgedane expertise.

De codeertaal XML (eXtensible Markup Language) berust op het principe om aan tekst of gegevens systematisch gespecificeerde codes toe te voegen, zogenaamde "markup". In het algemeen geldt dat XML-bestanden niet gebonden zijn aan specifieke programma's en daardoor bestand zijn tegen de veroudering en ontoegankelijkheid van het opslagformaat: de markup bestaat enkel uit gewone, 'platte' tekst, en blijft daardoor in principe tot in lengte van dagen leesbaar en bewerkbaar. De markup is meta-

¹¹ Met dank aan Johan Jansen (Rijks Universiteit Groningen, opleiding Nederlandse Taal en Cultuur), die in het kader van een stage bij het Hermansproject het digitale materiaal van *Moedwil en misverstand* voorbereidde en verkennend onderzoek deed naar de tekstgeschiedenis van de bundel.

¹² In het algemeen geldt dat het op grond van collatiegegevens of archiefinformatie nodig kan zijn om meer teksten te collationeren. De technische procedure voorziet daarin.

informatie: hij maakt duidelijk welke functie de eigenlijke data hebben. Een paragraaf kan bijvoorbeeld de code "para" krijgen. De keuze van codes is vrij. Inmiddels zijn er vrij veel codeerstandaarden gerealiseerd voor de meest uiteenlopende tekstsoorten. Dergelijke modellen geven dus een nadere invulling van de XML-standaard voor een bepaald toepassingsgebied. Het [Text Encoding Initiative](#) (TEI) heeft een zeer uitgebreide codeerstandaard uitgewerkt voor toepassingen binnen de humaniora, met specifieke beschrijvingsmodellen voor proza, toneel en poëzie, maar ook voor bijvoorbeeld transcripties van spraak, woordenboeken, en kritische edities.¹³

Het programma Collate is in staat om TEI-conforme documenten te produceren, maar enkel in de vorm van een basistekst en een extern variantenapparaat. Voor tekstkritisch onderzoek is een in de basistekst geïntegreerd variantenoverzicht echter efficiënter en overzichtelijker dan het gebruik van afzonderlijke documenten. De standaarduitvoer van Collate moest daarom worden aangepast: na een woord of een passage in de basistekst moesten onmiddellijk de varianten in de andere tekstversies volgen. Fundamentele beperkingen van Collate maken het onmogelijk om het programma direct de gewenste uitvoer te laten produceren. Collate kan wel de volledige tekst met een semi-gestructureerd variantenoverzicht uitvoeren, maar niet in XML-formaat. We hebben een speciaal configuratiebestand gemaakt om Collate deze tussenvorm te doen produceren. Vervolgens bewerkt een reeks bij het CHI ontwikkelde programma's de Collate-uitvoer en maakt er geheel automatisch correcte XML-TEI van. De programma's voor de XML-omzetting zijn geschreven in AWK, een gespecialiseerde programmeertaal voor tekst- en dataverwerking.¹⁴ De bruikbaarheid van deze ontwikkelde software blijft niet beperkt tot het Hermansproject. In de loop van 2004 zullen we de software en een handleiding publiceren op de projectwebsite.

¹³ De achtergronden van XML en TEI zijn uitgebreid aan de orde gekomen in '[Digitale edities en letterenonderzoek, een verkenning](#)', een *Neerlandistiek.nl*-artikel van Peter Boot en Els Stronks (Boot en Stronks 2002); voor meer informatie verwijzen we naar vooral de [derde paragraaf](#) van hun bijdrage. Relevante informatie met betrekking tot XML bevat ook '[What is XML and Why Should Humanists Care?](#)' van C.M. Sperberg-McQueen. Een volledig overzicht van alle beschrijvingsprotocollen van het Text Encoding Initiative is te vinden via de website van het [TEI Consortium](#).

¹⁴Zie voor meer informatie: http://www.cs.hmc.edu/tech_docs/qref/awk.html

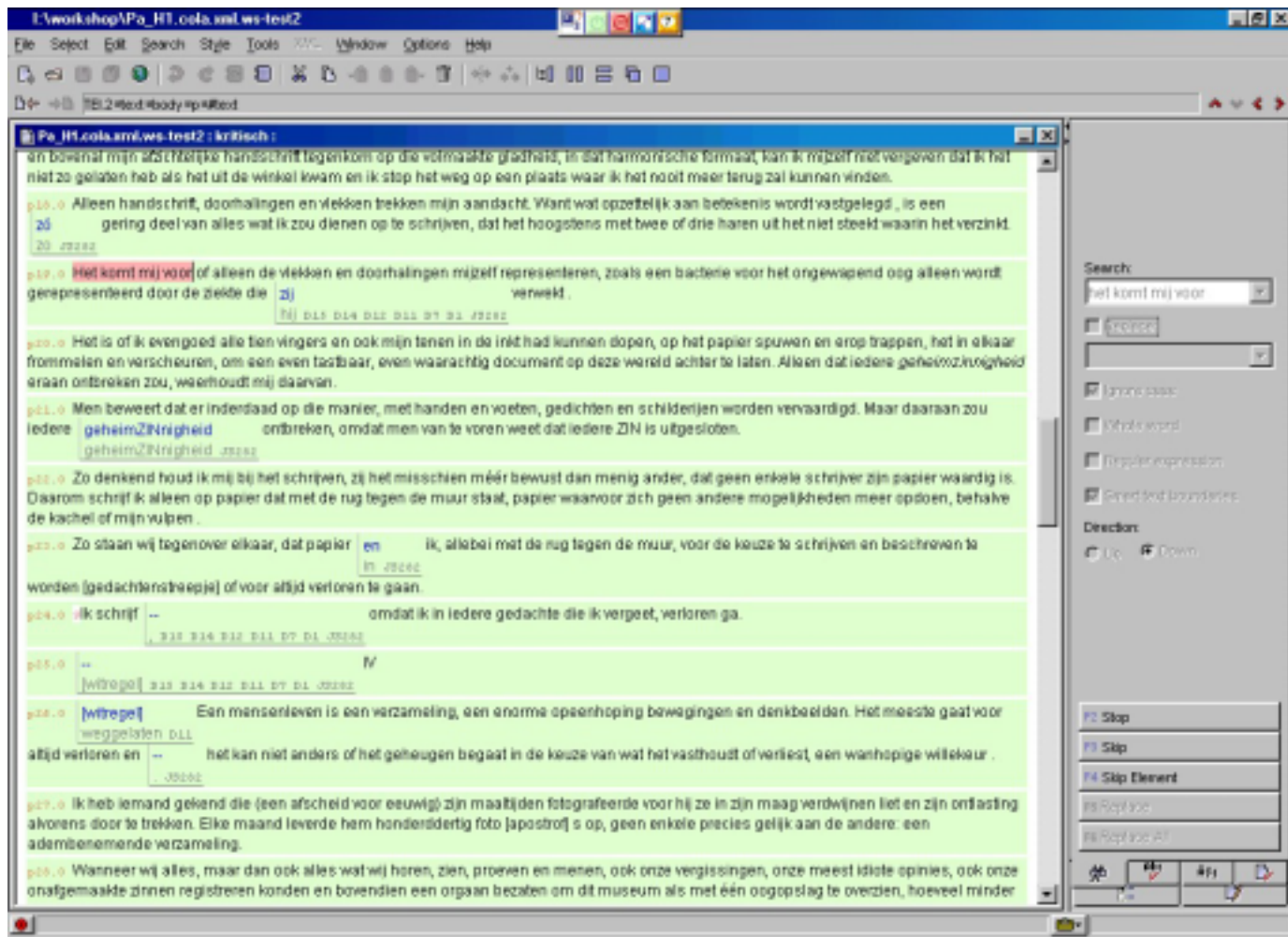
Het onderstaande voorbeeld, een fragment van *Paranoia* (D16, JS 114), toont hoe zo'n in XML-TEI gecodeerde basistekst inclusief kritisch apparaat eruitziet:

```
<p id="p18.0">Alleen handschrift, doorhalingen en vlekken trekken mijn
<lb n="8" />aandacht. Want wat opzettelijk aan betekenis wordt
vast<lb n="9" />gelegd , is een <anchor id="A19" /><app><lem>zó</lem><rdg
type="opm" wit="JS282">zo</rdg></app><anchor id="A20" />gering deel van
alles wat ik zou dienen op te <lb n="10" />schrijven, dat het hoogstens
met twee of drie haren uit het niet <lb n="11" />steekt waarin het
verzinkt.<lb n="12" /></p>
<p id="p19.0">Het komt mij voor of alleen de vlekken en doorhalingen <lb
n="13" />mijzelf representeren, zoals een bacterie voor het ongewapend
<lb n="14" />oog alleen wordt gerepresenteerd door de ziekte die <anchor
id="A21" /><app><lem>zij</lem><rdg type="gra"
wit="D15 D14 D12 D11 D7 D1 JS282">hij</rdg></app><anchor id="A22" />ver <lb n="15" />wekt
<lb n="16" /></p>
```

[afbeelding 3. fragment van *Paranoia* in XML-TEI]

Het aantal coderingen in dit kleine stukje tekst uit 'Preamble' is relatief gering. Desondanks is de XML-TEI-presentatie in broncode niet erg overzichtelijk: het is voor het menselijk oog ondoenlijk om in zo'n gecodeerd bestand snel en systematisch de tekstgeschiedenis te onderzoeken. Om het editeurswerk te vergemakkelijken hebben we daarom gezocht naar een softwareprogramma dat, met behoud van alle XML-TEI-codes, de tekstuele gegevens op een aantrekkelijke manier kan presenteren. De functionaliteit van een groot aantal XML-editors werd hiertoe uitgebreid getest. Uiteindelijk is gekozen voor het programma [XML-Mind XML Editor](#) (XXE). Dit product is 'documentgeoriënteerd': het lijkt op een reguliere tekstverwerker, werkt min of meer op vergelijkbare wijze en is daardoor gebruiksvriendelijk. XXE is een veelgebruikt, stabiel programma; het is bovendien gratis en de technische documentatie van het programma (handleidingen en ondersteuning via het internet) is adequaat.

We laten het hierboven in broncode getoonde fragment nogmaals zien, maar nu in XXE. De documentatie ziet er veel overzichtelijker uit: de basistekst is nu voorzien van een synoptische presentatie van de varianten. De belangrijkste XML-codes worden op een grafische manier weergegeven.



[afbeelding 4: XXE met Paranoia]

5 Digitale onderzoeksomgeving

Dankzij de beschikbaarheid van het primaire materiaal in XML-TEI-formaat kunnen de resultaten van het tekstkritische onderzoek en andere relevante informatie in een digitale onderzoeksomgeving worden gedocumenteerd. De documentatie is sober van aard en blijft beperkt tot duidelijk afgebakende categorieën. Het gaat daarbij eerst en vooral om informatie die betrekking heeft op de totstandkoming van de uiteindelijke tekst van de editie, de leestekst: ingrepen van editeurs worden op gestandaardiseerde wijze beschreven en verantwoord, indien nodig met toelichtingen in noten; voornamelijk onduidelijke tekstplaatsen, waarbij pas op grond van aanvullend onderzoek (bijvoorbeeld correctie-exemplaren of informatie uit secundaire bronnen) een beslissing kan worden genomen, worden systematisch gemarkeerd. Daarnaast worden, met het oog op de beschrijving van de overlevering van de tekst in de commentaar, varianten op elementair niveau getypeerd. We beschrijven via twee voorbeelden welke informatie aan de collatiebestanden wordt toegevoegd en hoe die verslaglegging eruitziet in XML-TEI-formaat. Om de coderingen goed te kunnen verduidelijken, geven we de voorbeelden steeds weer in de XML-broncode.

Voor het eerste voorbeeld keren we terug naar het fragment uit 'Atonale'; de documentatie bevat dezelfde variantengegevens als de eerder in dit artikel getoonde HTML-presentatie.¹⁵ Het 'Atonale'-fragment in XML-TEI ziet er als volgt uit:

```
<p id="p123.0">Ook aan Varenhijt had zij niet te wennen. Zij stond tegenover
  <lb n="10"/>hem in aanvallende houding <anchor id="A684"/><app><lem>--</lem>
<rdg type="inh" wit="DJ189">, zij was agens</rdg></app>. Onze parasieten wennen niet aan
<lb n="11"/>ons, wij hoogstens aan hen. Dit alleen om <anchor id="A685"/><app><lem ed="PK"
revDate="2003-10-07" edRat="var"><seg status="onzeker" id="inh016" type="inh">een
denkbeeld te geven van</seg></lem><rdg type="inh" wit="D13 D4 D2 JS189">de mate waarin
</rdg>
<rdg type="inh" wit="D9 D8">een denkbeeld te geven van </rdg></app>de mate waarin
  <lb n="12"/>Adelphe Varenhijt hinderde <anchor id="A686"/><app><lem ed="PK"
revDate="2003-10-07" edRat="var"><seg status="onzeker" id="inh017"
type="inh">{weggelaten}</seg></lem><rdg type="inh" wit="D13 D4 D2 JS189">uit te
drukken</rdg><rdg type="inh" wit="D9 D8">
{weggelaten}</rdg></app><note resp="PK" target="inh016 inh017" type="edrat"><date>2003-10-
07</date>De correctie-exemplaren van D9 D10 D11 D12 rechtvaardigen de tekstkeuze van D13
niet</note>. Want een parasiet <lb n="13"/>was zij niet; hijzelf was de laatste om haar als
zodanig te on<lb n="14"/>dervinden. </p><p id="p124.0">Hij had haar geroepen; hij had haar
een taak geschapen.<lb n="15"/></p>
```

[afbeelding 5: fragment van 'Atonale' (*Moedwil en misverstand*) in XML-TEI]

We gaan in op de belangrijkste coderingen. Bij het voorbeeld zijn we er voorlopig vanuit gegaan dat de versie met de passage "Dit alleen om de mate waarin Adelphe Varenhijt hinderde uit te drukken" onbedoeld in de basistekst terecht is gekomen. De tekst wordt daarom hersteld naar die van de drukken acht en negen. De wijzigingen ten opzichte van de basistekst zijn terug te vinden in de toegevoegde coderingen. Twee rood gemarkeerde elementen <lem>, met daarin de herstelde lezingen, zijn aan de documentatie toegevoegd; de attributen bij dit element geven onder andere aan welke editeur voor de ingreep verantwoordelijk is (ed="PK") en op grond waarvan de beslissing is genomen (edRat="var": de emendatie is verricht op grond van de gegevens uit het variantenoverzicht). De oorspronkelijke lezing van de basistekst is verplaatst naar een 'reading'-element (<rdg>), waarbij via het attribuut "wit" (voor 'witness') wordt aangegeven welke drukken de alternatieve lezing bevatten. Een toegevoegd <note> element geeft een meer gedetailleerde toelichting bij de ingreep.

Het tweede voorbeeld heeft betrekking op het beschouwend werk van Hermans en is vooral van belang met het oog op het vervaardigen van de editieregisters. In de *Volledige Werken* worden de door Hermans vervaardigde registers, die talrijke eigenaardigheden bevatten, gehandhaafd. Daarnaast

¹⁵ Wel zijn, om onnodige vervuiling van de bestanden te voorkomen, voorafgaand aan het tekstkritische onderzoek de resterende OCR-fouten uit de XML-TEI-bestanden verwijderd.

worden de teksten per band ontsloten door een naam- en titelregister. Tevens wordt een centraal digitaal register voorbereid, dat al het gebundelde en ongebundelde werk van Hermans zal omvatten.

We streven ernaar om de productie van de registers zoveel mogelijk te automatiseren. Eerste voorwaarde daarbij is dat alle namen van historische personen en titels van publicaties en kunstwerken systematisch worden gecodeerd. Bovendien moet de handmatige codering van al deze gegevens tot een minimum worden beperkt. Om dit te bereiken is programmatuur ontwikkeld die het mogelijk maakt om eenmaal gecodeerde namen steeds opnieuw te gebruiken. Dit hergebruik van gecodeerde naam- en titelgegevens betreft niet enkel een afzonderlijke titel, maar omvat al het te bewerken materiaal, in ons geval dus alle beschouwende teksten van Hermans.

Om de codering zoveel mogelijk te automatiseren worden eenmaal gecodeerde namen opgeslagen in een steeds groeiend centraal bestand, dat de basis vormt voor de codering van alle afzonderlijke titels. Van elke nog onbewerkte tekst worden de al in het centrale bestand aanwezige naam- en titelgegevens automatisch gecodeerd. De al gedeeltelijk verrijkte tekst wordt vervolgens door de editeur of een andere projectmedewerker gecontroleerd en aangevuld: de toegekende elementen en attributen worden gecontroleerd, incidentele onterechte coderingen, bijvoorbeeld voor metaforisch gebruik van namen van historische personen ('een Vlaamse Shakespeare') worden verwijderd en ten slotte worden nieuwe namen die ontbraken in het centrale bestand handmatig gecodeerd. Daarbij hoeft elke naamvorm slechts eenmaal gecodeerd te worden, waarna alle identieke vormen weer automatisch worden aangevuld. Nadat de codering van een nieuwe tekst is afgerond wordt een bestandslijst aangemaakt die alle informatie voor het register van de specifieke titel bevat; bovendien wordt het centrale bestand automatisch uitgebreid met de nieuw toegevoegde naam- en titelcoderingen. Onderstaand voorbeeld toont een kort gecodeerd fragment uit *Boze Brieven van Bijkaart* (D3, JS 348). Alle historische persoonsnamen in de tekst zijn voorzien van een element <name>, met daarbij een attribuut "reg" dat de gestandaardiseerde vorm van die naam bevat.

```
<lb n="9"></lb>van
de nobele <name reg="Allende, Salvador">Allende</name>. </p>
<p id="p313.0"><name reg="Allende, Salvador">Allende</name>, volkomen wettig
gekozen, net als <name reg="Hitler, Adolf">Hitler</name> indertijd. Of <lb
n="10"></lb>als <name reg="Nixon, Richard">Nixon</name> die zelfs met de grootste
meerderheid uit de Amerikaanse <lb n="11"></lb>geschiedenis gekozen is. </p>
<p id="p314.0">De laatste foto van <name reg="Allende, Salvador">Allende</name>
toont de vlezige socialist met een <lb n="12"></lb>vergiets op zijn kop en onder z
[apostrof] n arm z [apostrof] n eigen, hoogstpersoonlijke <lb
n="13"></lb>machinegeweer. <name reg="Brezjnev, Leonid">Brezjnev</name> had hem
dat cadeau gegeven, of <lb n="14"></lb>was het <name reg="Castro, Fidel">Fidel
Castro</name>? </p>
<p id="p315.0">Het is toch maar een goed ding dat <name
reg="Uyl, Joop den">Joop den Uyl</name> (door de <lb n="15"></lb>Franse
radio <name reg="Uyl, Joop den">Dzjoep de Nul</name> genoemd) niet ook
een hoogstpersoonlijk <lb n="16"></lb>machinegeweer bezit, en dat nog
wel terwijl hij toch <lb n="17"></lb>ook geheel wettig gekozen is. [apostrof] t is een ding
dat ik wel in hem <lb n="18"></lb>waardeer. Zal
wel door mijn reactionaire inborst komen. </p>
```

[afbeelding 6: fragment van *Boze Brieven van Bijkaart* in XML-TEI]

Titels van boeken en artikelen worden op een vergelijkbare manier gecodeerd. Ze worden voorzien van een attribuut waarin de auteur wordt vermeld, waardoor titels automatisch onder een auteursingang in het register opgenomen kunnen worden. Het beschouwend werk zal ook in bescheiden mate voorzien worden van toelichtingen en annotaties. Ook die worden uiteindelijk opgenomen in de XML-TEI-bestanden.

6 Meerwaarde van de digitale documentatie

De uitgave van de *Volledige Werken* wordt gedurende een periode van twaalf tot vijftien jaar door verschillende medewerkers voorbereid. Heldere procedures met betrekking tot de te volgen werkwijze, de verslaglegging van het onderzoek en de opslagstructuur van de documentatie zijn daarom noodzakelijk. De coderingen in XML-TEI dragen in belangrijke mate bij aan een zo transparant mogelijke werkomgeving: in de XML-TEI-documentatie wordt vermeld wie voor de uitvoering van specifieke werkzaamheden verantwoordelijk is; bijdragen van de verschillende projectmedewerkers en onderzoekers, zoals de verantwoording van emendaties en toelichtingen bij de tekst via annotaties, worden geïndividualiseerd, waarbij keuzes en argumentaties expliciet worden gemaakt. Eenmaal genomen beslissingen zijn altijd te achterhalen, waardoor de werkzaamheden van de editor beter

controleerbaar zijn dan in een traditionele omgeving. De digitale duurzaamheid van het XML-TEI-opslagformaat garandeert bovendien een blijvende toegang tot het onderzoeksmateriaal tijdens en na afloop van het project.

Maar de verslaglegging met behulp van XML-TEI biedt meer grote voordelen. Zo dient de XML-TEI-documentatie direct als kopij voor de uitgave van de *Volledige Werken* door De Bezige Bij. Door de expliciete coderingen zijn alle inhoudelijke tekstelementen duidelijk gescheiden van 'accidentele' vormtechnische en typografische aspecten (paginaopmaak, letterkeuze en corpgrootte etc.). De vormgever van de editie kan bovendien gebruik maken van de in de elektronische documentatie gedefinieerde betekenis van meerduidige leestekens. Verschillende vormen van toelichtingen bij de teksten, zoals de registerinformatie, annotaties en het overzicht van tekstemendaties, kunnen eveneens worden afgeleid uit de XML-TEI-documentatie en met behulp van stylesheets worden omgezet naar de voor de boekuitgave gewenste presentatievorm. De teksten zelf blijven in deze fase vrij van vormtechnische informatie.¹⁶

Een dergelijke werkwijze maakt dat de XML-TEI-documentatie in een later stadium eenvoudig kan worden hergebruikt. Te denken valt hierbij aan heruitgaven van afzonderlijke titels, die terugvoeren op de teksten uit de *Volledige Werken*. Ook voor meer specifieke doeleinden, zoals schooluitgaven met toelichtingen of (in de toekomst) herspelde uitgaven, kan de elektronische documentatie steeds worden gebruikt.

De *Volledige Werken* zijn meer dan een boekuitgave alleen. Vanaf het begin van het project richten we ons erop om onderzoeksmateriaal elektronisch te publiceren. Op basis van de elektronische data kunnen, voor verschillende groepen gebruikers, deelselecties uit het onderzoeksmateriaal op het internet worden aangeboden. Afhankelijk van het soort materiaal kan dit gepubliceerd worden voorafgaand aan, parallel aan, of na verschijning van de gedrukte delen van de *Volledige Werken*. De systematische documentatie van alle persoonsnamen en titelgegevens in XML-TEI maakt het bijvoorbeeld mogelijk om op het internet een digitaal register te presenteren, dat alle teksten ontsluit die op dat moment in druk zijn gepubliceerd. Dat digitale register vormt tevens de basis voor een eventuele afsluitende registerband bij de gedrukte editie.

Ook kunnen, bij wijze van 'casestudy', beredeneerde selecties uit het variantenmateriaal elektronisch gepubliceerd worden: het kan dan gaan om materiaal dat van belang is geweest voor de totstandkoming van de editie, of om materiaal dat om literair-historische of tekstgenetische redenen interessant is.¹⁷

Een laatste voorbeeld betreft de publicatie van primaire bibliografische gegevens bij afzonderlijke titels. Ook die komen in een elektronische omgeving goed tot hun recht. Stemma's en gedetailleerde beschrijvingen van primaire bronnen kunnen op een aantrekkelijke manier worden getoond: documentatie van dit materiaal in XML-TEI maakt immers dat dergelijke gegevens flexibel kunnen worden bevraagd en gepresenteerd. Bijkomend voordeel van elektronische publicatie van bibliografisch materiaal is natuurlijk het gegeven dat in een elektronische omgeving de gegevens eenvoudiger geactualiseerd kunnen worden. De exacte mogelijkheden voor elektronische publicatie van primair

¹⁶ Om het complexe productieproces in goede banen te leiden, vindt op dit moment al overleg plaats met de zetterij die betrokken zal zijn bij de productie van de *Volledige Werken*.

¹⁷ Volledige elektronische documentatie van de varianten is in principe ook mogelijk, maar is voor een meer algemene doelgroep minder interessant. Wel zal, in ieder geval op termijn, het complete digitale materiaal in enigerlei vorm beschikbaar moeten zijn voor wetenschappelijk onderzoek.

bibliografisch materiaal worden de komende tijd nader onderzocht. Daarbij gaan we ook na in hoeverre het mogelijk is om de primaire bibliografische informatie uit het onderzoek te koppelen aan de gegevens uit *Het bibliografische universum*.

7 Enkele vooruitzichten

De variantenoverzichten in XML-TEI die aan de basis staan van het tekstkritisch onderzoek voor de *Volledige Werken* van Willem Frederik Hermans zijn op te vatten als een complete, zij het rudimentaire historisch-kritische gegevensverzameling. Deze bevat veel informatie, die slechts ten dele voor de leeseditie wordt gebruikt, maar zich zeer goed leent voor aanvullend onderzoek. Alhoewel er nog geen plannen in die richting bestaan, vormt een XML-TEI-documentatie zoals hierboven beschreven een goede basis voor een (elektronische of gedrukte) historisch-kritische editie of studie-uitgave; voor gedetailleerd tekstgenetisch onderzoek kunnen transcripties en/of afbeeldingen van handschriftelijke bronnen en typoscripten aan de XML-TEI-bronbestanden worden toegevoegd.

Maar de variantendocumentatie vormt tevens een interessante gegevensverzameling voor poëticaal, literair-historisch of andersoortig letterkundig onderzoek. De gegevens vormen een 'mer à boire' voor de editeur of andere letterkundige die bijvoorbeeld onderzoek wil doen naar specifieke aspecten van de tekstgeschiedenis van de novellen uit *Moedwil en misverstand*.¹⁸ Allerlei verhelderingen en veronderstellingen met betrekking tot de tekst en tekstgenese kunnen door de onderzoeker in de XML-TEI-bestanden worden gedocumenteerd.

Bij wijze van verkenning gaan we kort in op enkele meer algemene onderzoeksmogelijkheden bij omvangrijke elektronische tekstcorpora. We gebruiken daarbij eerder verricht onderzoek als uitgangspunt. Om te beginnen een voorbeeld uit de onderzoekspraktijk rondom het oeuvre van Hermans. Frans Janssen inventariseerde in zijn opstellenbundel *Bedriegers en bedrogenen* de varianten in de tiende druk van *De donkere kamer van Damokles* en onderscheidde die in vijf categorieën: drukfouten, spellingwijzigingen, stilistische wijzigingen, 'wijziging van gegevens uit de werkelijkheid buiten de roman' en 'wijziging van inconsistenties binnen de romanwerkelijkheid', de belangrijkste categorie.¹⁹ Hij gaf van alle categorieën een aantal voorbeelden, waarbij hij en passant ook wees op een aantal niet of pas in latere drukken gecorrigeerde inconsistenties, die soms juist weer een gevolg waren van tekstuele wijzigingen elders in de tekst.

Dergelijk onderzoek is bij uitstek geschikt voor elektronische documentatie, waarbij alle varianten systematisch gecodeerd kunnen worden. De materiaalverzameling kan bovendien steeds worden doorzocht en bezien worden op de onderlinge samenhang; werkhypotheses kunnen zodoende eenvoudig worden getoetst en bijgesteld. De bevindingen van Frans Janssen maken bovendien nieuwsgierig naar meer. In de eerste plaats kan met behulp van de XML-TEI-coderingen worden onderzocht hoe de latere wijzigingen in *De donkere kamer van Damokles* zich tot zijn model verhouden: blijkens *Het bibliografische universum* verschenen er immers na de 'tiende' druk nog minimaal zes

¹⁸ Uit het kleine fragment van 'Atonale' komen al flinke stilistische verschillen naar voren tussen de tijdschriftpublicatie uit *Podium* en de eerste boekversie van het verhaal. Ook de spelling is bij de bundeling van de tekst gemoderniseerd.

¹⁹ Janssen 1980. Volgens *Het bibliografische universum* gaat het hier om D11 (JS 156). Behalve deze 'tiende' druk hebben volgens de bibliografie alleen de drukken D2 en D3 varianten.

drukken met varianten. Ook is het interessant om na te gaan hoe specifiek dit variantenpatroon is voor juist deze roman 'als ordening die de chaos vertoont'.²⁰ Hangt de aard van de veranderingen nauw samen met het thema van de roman en staan de wijzigingen in *De donkere kamer* min of meer op zichzelf, of zijn ze representatief voor de soort veranderingen die Hermans ook in andere teksten aanbracht? Zulk detailonderzoek dat de individuele tekst overstijgt is in een traditionele omgeving moeilijk te verwezenlijken. In een XML-TEI-documentatie echter kan een groot aantal teksten systematisch worden bestudeerd. Doordat de elektronische bestanden kunnen worden verrijkt met gespecificeerde coderingen en vervolgens op verschillende manieren doorzoekbaar en bevragebaar zijn, kunnen nieuwe hypotheses over de werkwijze van Hermans geformuleerd worden.

Ook voor andere onderzoeksdisciplines kan een XML-TEI-documentatie veel opleveren. Een voorbeeld uit eerder onderzoek met betrekking tot Multatuli's *Max Havelaar* maakt dat duidelijk. De taalkundige Theo Janssen vroeg zich naar aanleiding van de historisch-kritische editie van *Max Havelaar* af of het gedrukte variantenapparaat van die editie een 'Fundgrube [zou] kunnen zijn voor de taalkunde'.²¹ Voor onderzoek naar sociolinguïstische ontwikkelingen, voor tekstwetenschappelijk onderzoek (in het geval van Multatuli het gebruik van demonstrativa) en voor de bestudering van syntactische en lexicale ontwikkelingen beantwoordde hij die vraag alvast bevestigend en bestempelde hij de documentatie als een potentiële 'goudmijn'. Daarbij wees hij erop dat de bevindingen wel gerelateerd moesten worden aan ander onderzoek. Juist het leggen van relaties tussen verschillende onderzoeksobjecten wordt vergemakkelijkt wanneer het materiaal in een XML-TEI-omgeving wordt bestudeerd en de verslaglegging van het onderzoek plaatsvindt via XML-TEI-conforme coderingen: grote tekstcorpora kunnen dan automatisch of semi-automatisch worden vergeleken met andere onderzoeksgegevens.

8 Ten slotte

Veel verworvenheden op het gebied van digitaal publiceren zijn van recente datum. Onderzoek op basis van digitale documentatie vindt tot nu toe slechts beperkt plaats. Toch zal het werken met elektronische bestanden in toenemende mate gaan behoren tot de reguliere onderzoekspraktijk.²² Bij het Constantijn Huygens Instituut worden momenteel verschillende initiatieven ontplooid om de expertise op het gebied van digitaal onderzoek en elektronisch publiceren verder te ontwikkelen en uit te dragen; in toenemende mate willen we niet alleen de eigen onderzoekers, maar ook andere alfawetenschappers wijzen op de mogelijkheden van digitale informatieverwerking.²³

Binnen dat kader is recentelijk het nieuwe project EDITOR van start gegaan, waarbij EDITOR staat voor 'the Edition as a Digital Instrument for Text-based Open Research'. Doelstelling van het project is het ontwikkelen van programmatuur en standaards voor het toevoegen van onderzoeksinformatie aan

²⁰ Janssen 1980, p. 77.

²¹ Janssen 1994; met name de laatste paragraaf 'Max Havelaar in varianten'.

²² Zie voor een kernachtige beschrijving van de stand van zaken en de toepassingsmogelijkheden voor letterkundig en taalkundig onderzoek de internetpagina's van het Leeds Electronic Text Centre, met name het onderdeel '[Implications of the use of Electronic Text in Academic Work](#)'

²³ Recente voorbeelden van de steeds toenemende aandacht voor digitalisering zijn de oprichting van het CHI-weblog elektronisch publiceren en het voorstel aan de KNAW tot inrichting van een digitaal tekstarchief, TANGO (Tekst Archief voor Nederlands Geesteswetenschappelijk Onderzoek). Zie voor beide initiatieven de [CHI-website](#).

tekstbestanden en voor de raadpleging, manipulatie, en het exporteren van dat onderzoeksmateriaal. Het project wil editeurs, letterkundigen en (cultuur)historici stimuleren in het gebruik van digitaal onderzoeksmateriaal en digitale edities. De gerealiseerde hulpmiddelen moeten door onderzoekers zelfstandig kunnen worden gebruikt, maar tevens toepasbaar zijn in een elektronische omgeving waar verschillende onderzoekers met dezelfde data werken. Bij de ontwikkeling en implementatie van de benodigde hulpmiddelen wordt samengewerkt met enkele nationale en internationale instellingen.

Ook de verrichte inspanningen bij het Hermansproject zijn een uitgesproken voorbeeld van dit streven om alfawetenschappers meer te betrekken bij nieuwe vormen van digitaal onderzoek. Om anderen te interesseren voor die nieuwe mogelijkheden en om tegelijkertijd op de hoogte te blijven van de wensen vanuit het wetenschappelijke veld zal een gebruikersgroep met wetenschappers uit verschillende disciplines worden samengesteld. Uiteraard zullen de ontwikkelingen bij het EDITOR-project van zeer dichtbij worden gevolgd.

De digitale documentatie die ten behoeve van de *Volledige Werken* van Willem Frederik Hermans wordt gerealiseerd, zal, met het poëziedebuut *Kussen door een rag van woorden* (1944) tot en met de laatste roman *Ruisend gruis* (1995), meer dan een halve eeuw literair materiaal omvatten. Daarmee ontstaat uiteindelijk een laboratorium voor vernieuwend onderzoek naar het werk van een van de grootste Nederlandstalige schrijvers van de vorige eeuw.

Peter.Kegel@chi.knaw.nl

Bert.Van.Elsacker@chi.knaw.nl

Constantijn Huygens Instituut- KNAW

Prins Willem-Alexanderhof 5

Postbus 90754

2509 LT 's Gravenhage

Literatuuropgave

- Boot en Stronks 2002: Peter Boot en Els Stronks, 'Digitale edities en letterenonderzoek, een verkenning', in: *Neerlandistiek.nl* 02.08 (Gepubliceerd: 23 oktober 2002)
- Delvigne en Janssen 1996: Rob Delvigne en Frans Janssen, *Schrijven is verbluffen; bibliografie van de verspreide publicaties van Willem Frederik Hermans*. Amsterdam, [1996].
- Janssen en Van Stek 2001: Frans A. Janssen en Sonja van Stek, *Het bibliografische universum van Willem Frederik Hermans; bibliografie van de afzonderlijk verschenen werken*. Amsterdam, Den Haag, 2000]
- Janssen 1980: Frans A. Janssen, 'Varianten in orde en chaos'; Over de varianten in de tiende druk van *De donkere kamer van Damokles*. In: *Bedriegers bedrogenen; Opstellen over het werk van Willem Frederik Hermans*. Amsterdam, 1980, pp. 55-78.
- Janssen 1994: Theo A.J.M. Janssen, 'De verhouding tussen Saïdjah en Adinda in twee varianten'. In: *Over Multatuli* 33 (1994), afl. 33, pag. 30-39.
- Kegel en Van Elsacker 2003: Peter Kegel and Bert Van Elsacker, "In progress – The Complete Works of Willem Frederik Hermans: a report on the integration of digital technologies in the publication of a multi-volume edition", 2003: http://www.chi.knaw.nl/nieuwetijd/hermans/pdf/in_progress.pdf
- Robinson 1998: Peter M. W. Robinson, "New methods of editing, exploring, and reading *The Canterbury Tales*", 1998: http://www.cta.dmu.ac.uk/projects/ctp/ctp_desc2.html

Websites met meer informatie over de beschreven onderwerpen (in artikelvolgorde)

- Volledige Werken* Willem Frederik Hermans: <http://www.chi.knaw.nl/nieuwetijd/hermans/project.html>
- Metamorfoze: <http://www.kb.nl/coop/metamorfoze/>
- Virginia Electronic Text Center: <http://etext.lib.virginia.edu/helpsheets/scantext.html>
- Leeds Electronic Text Centre: <http://etext.leeds.ac.uk/>
- NIWI: <http://www.niwi.knaw.nl/>
- Canterbury Tales-project en Collate via: <http://www.cta.dmu.ac.uk/projects/ctp/>
- Nestle-Aland-editie: <http://nestlealand.uni-muenster.de/>
- Parzival-project: <http://www.parzival.unibas.ch/probed.html>.
- Text Encoding Initiative (TEI): <http://www.tei-c.org/>
- AWK: http://www.cs.hmc.edu/tech_docs/qref/awk.html
- XML-Mind XML Editor (XXE): <http://www.xmlmind.com/xmleditor/>