

Tilburg University

Methoden van systematische rechtspraakanalyse

Verbruggen, Paul

Publication date:
2021

Document Version
Publisher's PDF, also known as Version of record

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):
Verbruggen, P. (editor) (2021). *Methoden van systematische rechtspraakanalyse: Tussen juridische dogmatiek en data science*. Boom juridisch.

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 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 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.

Methoden van systematische rechtspraakanalyse



**METHODEN VAN SYSTEMATISCHE
RECHTSPRAAKANALYSE**

TUSSEN JURIDISCHE DOGMATIEK EN DATA SCIENCE

REDACTIE:
P.W.J. VERBRUGGEN

BOOM JURIDISCH
Den Haag
2021

Omslagontwerp: Textcetera, Den Haag
Opmaak binnenwerk: Ambrac, Deventer

© 2021 De auteurs | Boom juridisch

Behoudens de in of krachtens de Auteurswet gestelde uitzonderingen mag niets uit deze uitgave worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen of enige andere manier, zonder voorafgaande schriftelijke toestemming van de uitgever.

Voor zover het maken van verveelvoudigingen uit deze uitgave is toegestaan op grond van artikel 16h Auteurswet of de reprorechtregeling van Stichting Reprorecht dient daarvoor een billijke vergoeding te worden voldaan aan Stichting Reprorecht (Postbus 3051, 2130 KB Hoofddorp, www.reprorecht.nl). Voor het verveelvoudigen en openbaar maken van (een) gedeelte(n) uit deze uitgave als toelichting bij het onderwijs, bijvoorbeeld in een (digitale) leeromgeving of een reader (art. 16 Auteurswet), dient een regeling te worden getroffen met Stichting Uitgeversorganisatie voor Onderwijslicenties (Postbus 3060, 2130 KB Hoofddorp, www.stichting-uvo.nl).

No part of this book may be reproduced in any form, by print, photoprint, microfilm or any other means without written permission from the publisher.

ISBN 978-94-6290-975-5
ISBN 978-90-5189-019-8 (e-book)
NUR 820

www.boomjuridisch.nl

INHOUDSOPGAVE

1	Methoden van systematische rechtspraakanalyse: Een inleiding tot het veld <i>Paul Verbruggen</i>	7
2	Het verrichten van een gestructureerde rechtspraakanalyse – De rol van excuses in de civiele rechtspraak en de medische tuchtrechtspraak <i>Lianne Wijntjens</i>	27
3	Mogelijkheden en beperkingen bij de kwantitatieve analyse van jurisprudentie voor de bestudering van geldend recht <i>Pieter Wolters</i>	47
4	Kwalitatieve inhoudsanalyse van rechtspraak met NVivo – Lessen uit een systematisch jurisprudentieonderzoek naar rechterlijke toetsing in het sportrecht <i>Paul Verbruggen</i>	71
5	Lessen uit een grootschalige rechtspraakanalyse met studenten – Een empirisch onderzoek naar samenwoning in het Belgisch relatievermogensrecht <i>Elise Goossens, Katalien Bollen en Alain-Laurent Verbeke</i>	99
6	Methodologische kansen en beperkingen van gebruik van statistiek en <i>machine learning</i> bij de bestudering van rechtspraak <i>Michel Vols</i>	125
7	Exploring and Searching Judicial Opinions with Top2Vec <i>Arthur Dyevre</i>	143
8	Methodologische aspecten van een data pipeline voor rechtspraakanalyse <i>Eric Tjong Tjin Tai</i>	163
9	Systematische jurisprudentieanalyse als uitdaging voor onderwijs en onderzoek <i>Frederik Peeraer en Rob van Gestel</i>	185

INHOUDSOPGAVE

Auteurslijst

213

1 **METHODEN VAN SYSTEMATISCHE RECHTSPRAAKANALYSE: EEN INLEIDING TOT HET VELD**

*Paul Verbruggen**

1 **INLEIDING**

Welke stappen moet een rechtswetenschapper zetten om rechtspraak systematisch te analyseren? De vraag stellen is eenvoudiger dan haar beantwoorden. Klassiek ‘juridisch-dogmatisch’ onderzoek betreft de beschrijving van het geldende recht, of hoe dat recht behoort te luiden.¹ Een groot gedeelte van de aandacht van de traditionele rechtswetenschapper gaat daarbij uit naar het in kaart brengen van patronen en ontwikkelingen in jurisprudentie, alsmede het doen van uitspraken over de betekenis en relevantie daarvan.² Met de vraag hoe hij of zij tot die inzichten komt en daarover verantwoording aflegt, begeeft de rechtswetenschapper zich echter op minder bekend terrein. Dat is des te meer zo naar mate het aantal te analyseren uitspraken groter wordt. Hoe verzamel ik de rechtspraak? Welke uitspraken selecteer ik wel en niet voor verdere analyse en hoe voorkom ik daarbij vooringenomenheid? Op welke wijze lees en analyseer ik de geselecteerde uitspraken om een patroon bloot te leggen? Hoe waarborg ik de betrouwbaarheid van de gevonden inzichten? Hoe presenteer ik die bevindingen zodat de relevantie voor juridische theorievorming of de rechtspraktijk duidelijk wordt? En hoe kunnen software, *artificial intelligence*

* Mr. dr. P.W.J. Verbruggen is universitair hoofddocent privaatrecht, Tilburg Law School, tevens houder van de TPR-Wisselleerstoel aan de KU Leuven (2019-2021). Deze bundel is tot stand gekomen met behulp van financiering toegekend onder het NWO Veni-project nr. 451-16-011.

1 J.M. Smits, *Omstreden rechtswetenschap: Over aard, methode en organisatie van de juridische discipline*, Den Haag: Boom Juridische uitgevers 2009, p. 28. Zie voor de diversiteit van onderzoek die schuil gaat achter het begrip juridisch-dogmatisch onderzoek J.B.M. Vranken, *Mr. C. Assers Handleiding tot de beoefening van het Nederlands Burgerlijk Recht. Algemeen Deel****. Een synthese*, Deventer: Kluwer 2014, nr. 7 e.v.

2 Deze opvatting delen zelfs degenen met geheel verschillende visies op de taak van rechtswetenschapper. Zie bijvoorbeeld: R.M. Wibier, ‘De taak van de rechtswetenschapper. Recht als volwaardige wetenschappelijke discipline’, *NJB* 2020/814, afl. 13, p. 883-887 (die op p. 883 stelt dat ‘de rechtswetenschapper zich vooral zou moeten bezighouden met het in kaart brengen, ordenen en systematiseren van het geldende recht.’) en A. Verheij, ‘Voor de rechtbank staat een been. Over de taak en aard van de rechtswetenschap’, *NJB* 2020/2005, afl. 30, p. 2228-2236 (die betoogt dat de scheidslijn tussen descriptief onderzoek zoals Wibier dat voorstaat en normatief of empirisch juridisch-onderzoek niet goed te trekken is, zodat een pluriforme beoefening van de rechtswetenschap moet worden aangemoedigd).

(AI) of andere vormen van *data science*-technieken de analyse en presentatie ondersteunen? Op deze methodologische vragen blijven handboeken over rechtswetenschappelijk onderzoek de rechtswetenschapper het antwoord grotendeels schuldig.

Dit boek heeft als doel inzicht te geven in de methodologische vraagstukken die spelen bij het doen van grootschalig systematisch jurisprudentieonderzoek. Daartoe brengt het auteurs uit Nederland en België samen die in verschillende innovatieve projecten ervaring hebben opgedaan met het opzetten en uitvoeren van systematische rechtspraakanalyses. De bijdragen van deze onderzoekers schetsen in heldere bewoordingen welke methoden zij hebben gebruikt voor hun jurisprudentieonderzoek, de resultaten die daaruit zijn voortgekomen en de diverse uitdagingen die zij in het onderzoeksproces hebben gekend. Op basis daarvan brengen zij de kansen, beperkingen en lessen in kaart die de door hen aangewende methoden bieden voor rechtswetenschappelijk onderzoek. Aldus vormt het boek niet alleen een kleurrijk palet van diverse onderzoeksthema's en -methoden. Ook ontsluit het een nieuw terrein binnen het lopende debat in de lage landen over methodologie in de rechtswetenschap.

Deze inleiding beoogt het kader te scheppen voor de hoofdstukken die volgen. Daartoe verkent zij het veld van methoden van systematische rechtspraakanalyse. Eerst volgt een omschrijving van wat systematische rechtspraakanalyse behelst (2). Die omschrijving krijgt nader vorm door te bespreken hoe deze analyse zich verhoudt tot juridisch-dogmatisch (3) en empirisch-juridisch onderzoek (4). Betoogd wordt dat systematische rechtspraakanalyse, vanwege haar nadruk op methodologische verantwoording, die twee vormen van rechtswetenschappelijk onderzoek verbindt (5). Na deze positionering van het kernthema van deze bundel volgt een introductie tot de individuele hoofdstukken en de daarin gebruikte methoden. Deze variëren van kwalitatieve tot kwantitatieve methoden, soms ondersteund door statistiek of technieken van *data science* (6). Afgesloten wordt met enkele centrale lessen die te trekken zijn uit de hoofdstukken (7).

2 SYSTEMATISCHE RECHTSPRAAKANALYSE

De digitale beschikbaarheid van rechtspraak, de overvloed aan informatie die daardoor voorhanden is en de mogelijkheden die software biedt in de ondersteuning van het verzamelen en analyseren van rechtspraak, maakt dat het doen van grootschalig systematisch

jurisprudentieonderzoek aantrekkelijk is geworden. Maar wat houdt systematische rechtspraak- of jurisprudentieanalyse in?³

Met het doen van een systematische rechtspraakanalyse beoogt de onderzoeker verbanden, patronen of ontwikkelingen in uitspraken van (tucht)rechters, arbiters en andere juridische geschilbeslechters op een gestructureerde en repliceerbare wijze in kaart te brengen. De onderzoeker baseert zich bij die analyse uitsluitend op de tekst van de juridische uitspraken, die hij of zij met behulp van labels ('codering') op consistente wijze interpreteert en van betekenis voorziet. Het doel daarbij is niet primair om zich uit te laten over de juistheid of wenselijkheid van juridische oordelen. De onderzoeker is geïnteresseerd in specifieke juridische of niet-juridische elementen of factoren die tot uiting komen in de tekst van de uitspraken. Voorbeelden betreffen de samenstelling van het rechtscollege, de aan- of afwezigheid van professionele juridische bijstand, de rechterlijke waardering van rechtsgronden of argumenten van partijen, de invulling van een juridisch begrip of leerstuk, of de navolging van een leidend arrest van het hoogste rechtscollege door feitenrechters in den lande. Afhankelijk van hoe en in welke mate deze aspecten naar voren komen in de analyse, kan de onderzoeker lijnen of tendens in de jurisprudentie blootleggen.

Systematische rechtspraakanalyse is derhalve in de eerste plaats beschrijvend van aard: zij geeft zicht op wat partijen, rechters en andere geschilbeslechters in juridische uitspraken doen. Het onderzoeksobject is de tekst waarin de juridische geschilbeslechting zijn beslag heeft gekregen. Dit statische, beschrijvende karakter dat is gericht op de inhoud van juridische teksten sluit echter niet uit dat de onderzoeker op basis van zijn of haar bevindingen zich verhoudt tot bestaande normatieve opvattingen of aannames in academische literatuur, beleid of wetgeving. Zoals blijkt uit de diverse hoofdstukken van dit boek, kan systematische rechtspraakanalyse de opmaat zijn voor juridische theorievorming en biedt zij mogelijkheden tot het toetsen van normatieve standpunten over hoe het recht zou moeten luiden. Ook kan zij beogen nieuwe inzichten te genereren die relevant zijn voor rechters, toezichthouders, advocaten, notarissen of andere praktijkjuristen en of leemtes te signaleren die interessant zijn voor verder onderzoek. Een wellicht (nog) ambitieuzer doel van systematisch jurisprudentieonderzoek is het op basis van de geanalyseerde data voorspellen van de rechterlijke toepassing van rechtsregels in een bepaald feitencomplex.⁴ Zo bezien kan sys-

3 De termen rechtspraakanalyse en jurisprudentieanalyse zijn synoniemen en de auteurs in deze bundel gebruiken ze door elkaar heen.

4 Vgl. de doelen van 'systematic content analysis' zoals uiteengezet in: M. Hall & R. Wright, 'Systematic Content Analysis of Judicial Opinions', *California Law Review* 2008, afl. 1, p. 85-100; M. Salehijam, 'The Value of Systematic Content Analysis in Legal Research', *Tilburg Law Review* 2018, afl. 1, p. 38-39 en O. Brook, 'Politics of Coding: On Systematic Content Analysis of Legal Text', in: M. Bartl, P. Cebulak &

tematische rechtspraakanalyse een belangrijke bijdrage leveren aan zowel de rechtspraktijk als aan rechtswetenschappelijk onderzoek.

Systematische rechtspraakanalyse bestaat in vele maten en soorten. Zij is geen onderzoeksmethode op zich, maar een verzamelnaam voor de toepassing van diverse sociaalwetenschappelijke methoden van data-analyse op juridische uitspraken.⁵ Die methoden kunnen, mede afhankelijk van het onderzoeksdoel en aantal te analyseren uitspraken, meer kwalitatief of kwantitatief van aard zijn. Ook de mate waarin de methoden worden ondersteund door software of technieken van *data science* varieert.⁶ Wat de verschillende methoden van analyse echter samenbindt, is dat ze alle gericht zijn op het systematisch verzamelen, selecteren, coderen en analyseren van tekst. Het doel is namelijk het op betrouwbare en reproduceerbare wijze in kaart brengen van betekenisvolle verbanden of patronen.

3 VERHOUDING MET JURIDISCH-DOGMATISCH ONDERZOEK

Het doen van systematisch rechtspraakanalyse sluit nauw aan bij wat de klassieke rechtswetenschapper gewend is te doen, indien hij of zij zich toelegt op het onderzoeken van een reeks van rechterlijke uitspraken betreffende een bepaald leerstuk of een bepaalde casus.⁷ Hall en Wright benoemen die gelijkenis in hun befaamde publicatie over ‘Systematic Content Analysis of Judicial Opinions’ als volgt:

‘On the surface, content analysis appears simple, even trivial, to some. Using this method, a scholar collects a set of documents, such as judicial opinions on a particular subject, and systematically reads them, recording consistent features of each and drawing inferences about their use and meaning. This method comes naturally to legal scholars because it resembles the classic scholarly exercise of reading a collection of cases, finding common threads that link the opinions, and commenting on their significance.’⁸

J. Lawrence (red.), *Behind the Method: The Politics of European Legal Research*, Cheltenham: Edward Elgar Publishing 2022, te verschijnen, par. 3, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3887536.

5 Zo wordt bijvoorbeeld in deze bundel onder meer gebruikgemaakt van ‘*structured content analysis*’ (Wijntjens), ‘*cross-case analysis*’ (Verbruggen), ‘*regressieanalyse*’ (Vols), ‘*natural language processing*’ (Dyevre), en ‘*network analysis*’ (Tjong Tjin Tai) als methoden van systematische rechtspraakanalyse. Iedere methode kent zijn eigen richtlijnen voor dataverzameling, -selectie, en/of -analyse.

6 Zie uitgebreid over de rol van computerondersteunde rechtspraakanalyse de bijdragen in M.A. Livermore & D.N. Rockmore (red.), *Law as Data: Computation, Text and the Future of Legal Analysis*, New Mexico: Santa Fe Institute Press 2019.

7 T.F.E. Tjong Tjin Tai & P.W.J. Verbruggen, ‘Nieuwe onderzoeksmethoden in de rechtswetenschap’, te verschijnen.

8 Hall & Wright 2008, p. 64.

Niet alleen is er een sterkte gelijkenis met klassiek juridisch-dogmatisch onderzoek. Dat type onderzoek is tevens onmisbaar voor een goede positionering en waardering van systematische rechtspraakanalyse. Zo benadrukt Vols in zijn hoofdstuk in deze bundel onder meer het fundamentele belang van het theoretisch-juridische kader (veelal een dogmatisch leerstuk) voor zowel de focus en vormgeving van de analyse, als voor het duiden van de relevantie van de gevonden inzichten voor theorie en praktijk. Peeraer en Van Gestel stellen in vergelijkbare zin in hun algemene beschouwing over systematische jurisprudentieanalyse als uitdaging voor juridisch onderwijs en onderzoek, dat een dergelijke analyse geenszins een vervanging is van, maar een aanvulling op traditionele interpretatieve vormen van jurisprudentieonderzoek. Naar hun idee leent systematische jurisprudentieanalyse zich vooral voor de studie in de 'breedte' van het onderzoeksobject en niet zozeer in zijn 'diepte'.⁹ De gevonden inzichten kunnen worden verklaard, geduid of gefalsificeerd met behulp van juridische interpretatiemethoden, zoals de grammaticale, wetsystematische en teleologische interpretatietechnieken. Dergelijk methoden en technieken kunnen natuurlijk ook worden aangewend om de betekenis van de blootgelegde patronen verder te onderzoeken. Deze complementariteit van onderzoeksmethoden is in de sociale wetenschappen bekend onder de term *triangulation*, dat wil zeggen de activiteit waarmee de onderzoeker beoogt de bevindingen met behulp van diverse methoden en technieken te ondersteunen of uit te diepen.¹⁰

Toch vergt het opzetten en uitvoeren van een systematisch jurisprudentieonderzoek bijzondere aandacht van de klassiek opgeleide rechtswetenschapper. Ten eerste omdat het doel van onderzoek anders is. Het primaire doel is, als gezegd, het beschrijven en analyseren van specifieke aspecten van een reeks van juridische uitspraken met als oogmerk patronen of ontwikkelingen in die reeks inzichtelijk te maken. De gegevens zoals die objectief blijken uit de tekst van de uitspraken staan centraal, niet de juistheid of wenselijkheid van de geanalyseerde oordelen, individueel of als geheel. De traditionele rechtswetenschapper is echter getraind om interpretatief en oordelend te zijn. Wat bedoelt de rechter met deze overweging? Is het oordeel van de rechter rechtens? Die natuur moet zeker niet terzijde worden geschoven bij het doen van systematisch jurisprudentieonderzoek. Zij is onmisbaar bij het verklaren, duiden of bekritisieren van gevonden patronen. Maar voordat die patronen systematisch zijn gedocumenteerd, zal de rechtswetenschapper zijn of haar mening over

9 Zie par. 3 van de bijdrage van Peeraer en Van Gestel in deze bundel (onder verwijzing naar M. Hall, *Coding Case Law for Public Health Law Evaluation: A Methods*, Robert Wood Johnson Foundation, Princeton NJ 2011, p. 6, <http://publichealthlawresearch.org/sites/default/files/downloads/resource/CodingCaseLaw-Monograph-Hall2011.pdf>).

10 Zie in het algemeen: J. Drisko, 'Triangulation' in: *Oxford Bibliographies* [online resource] (2017) doi: 10.1093/OBO/9780195389678-0045, www.oxfordbibliographies.com/view/document/obo-9780195389678/obo-9780195389678-0045.xml#firstMatch.

de geanalyseerde uitspraken in reserve moeten houden om de noodzakelijke afstand en objectiviteit te bewaren tot het onderzoeksobject.¹¹

De tweede reden van oplettendheid betreft de methoden van systematische rechtspraak-analyse. Die methoden vereisen van de onderzoeker dat hij of zij in expliciete bewoordingen inzicht geeft in de wijze van dataverzameling, -selectie en -analyse teneinde de betrouwbaarheid en reproduceerbaarheid van de analyse te borgen. De klassiek opgeleide rechtswetenschapper kan op dat vlak veel leren, zoals hierna zal worden belicht (zie par. 5). Voor juridisch-dogmatisch onderzoek zijn de conventies voor het opzetten en uitvoeren van onderzoek, ook jurisprudentieonderzoek, (nog) niet uitgekristalliseerd.¹² Rechtspraakanalyse krijgt in de in België en Nederland beschikbaar zijnde handboeken over methoden van juridisch onderzoek weinig aandacht.¹³ De aandacht die er is voor jurisprudentieanalyse, is vooral gericht op kunnen doorgronden van individuele rechterlijke uitspraken rekening houdende met hun structuur.¹⁴ Dat is uiteraard een essentiële vaardigheid voor iedere rechtswetenschapper, maar daarmee is er nog geen antwoord op de vraag hoe te werk te gaan bij het verzamelen, selecteren en analyseren van een kluwen aan rechtspraak om daar vervolgens op systematische wijze patronen in te ontdekken.

De beperkte aandacht voor rechtspraakanalyse als methode van juridisch-dogmatisch onderzoek is natuurlijk niet verwonderlijk. Lang niet al het juridisch onderzoek richt zich uitsluitend op rechtspraak, laat staan op de systematische analyse daarvan. De interpreta-

-
- 11 Illustratief zijn de methodologische opmerkingen van Vranken in zijn systematische analyse van de precedentwerking van vijftig effectenleearresten van de Hoge Raad. In de verantwoording van zijn analyse merkt hij op: 'Ter vermijding van misverstand benadruk ik dat mijn werkwijze meebrengt dat ik de uitspraken van de Hoge Raad over effectenlease niet inhoudelijk beoordeel. (...) Het is een ingrijpende beperking in mijn onderzoek naar precedentwerking die ik mijzelf moet opleggen, want het zou best kunnen zijn dat de vijftig uitspraken inhoudelijke vernieuwingen bevatten met een meerwaarde die (veel) verder reikt dan waarop ik mij richt: de afwikkeling, getalsmatig, aan de hand van de match van rechtskwesties in de Hoge Raad-uitspraken met die in andere effectenlease-zaken. Om dezelfde reden ga ik niet in op de vraag of de uitspraken van de Hoge Raad inhoudelijk anders en beter hadden gekund voor (een van de) betrokkenen in de effectenleaseprocedures.' J.B.M. Vranken, 'De afwikkeling van massaschade door de rechter lijkt onbegonnen werk. Een onderzoek aan de hand van de precedentwerking van vijftig uitspraken van de Hoge Raad over effectenlease', *NJB* 2020/2805, afl. 41, p. 3146.
- 12 J.B.M. Vranken, *Mr. C. Assers Handleiding tot de beoefening van het Nederlands Burgerlijk Recht. Algemeen Deel****. Een synthese*, Deventer: Kluwer 2014, nr. 87.
- 13 Een positieve uitzondering is het studieboek van Vols, die een apart hoofdstuk wijdt aan jurisprudentieanalyse. Zie M. Vols, *Juridisch onderzoek: Honderd vragen en antwoorden*, Den Haag: Boom juridisch 2020, p. 119-125.
- 14 Zie bijvoorbeeld. B. van Klink & A. Broekers-Knol, 'De vergeten kunst van het arrest-lezen', *AA* 2015, afl. 9, p. 723-728 en H.T.M. Kloosterhuis, *Juridische methoden. Casusoplossen, jurisprudentie-analyse en argumenteren*, Den Haag: Boom juridisch 2020, hfdst. 5 t/m 7. Voor België zie bijvoorbeeld F. Eggermont e.a., *Praktijkboek rechtsmethodologie*, Brugge: die Keure 2019, p. 233 e.v. en J. Baeck, *Juridische vaardigheden*, Brugge: die Keure 2021, p. 47 e.v.

tieve, normatieve studie van een enkele uitspraak of een beperkt aantal uitspraken, in het bijzonder die van hoogste rechters, is belangrijk voor zowel de rechtspraktijk, als voor het bestaan van een pluriforme rechtswetenschap.¹⁵ Benadrukt moet daarom worden dat systematische rechtspraakanalyse zoals hier omschreven deze normatieve juridische analyses van jurisprudentie niet vervangt. Zij bouwt daarop voort, vult aan en ontsluit nieuwe terreinen voor juridisch-dogmatisch onderzoek.¹⁶

4 VERHOUDING MET EMPIRISCH-JURIDISCH ONDERZOEK

Hall en Wright stellen in hun hierboven aangehaalde publicatie dat systematische rechtspraakanalyse verder gaat dan het beter lezen van jurisprudentie. Het brengt ook de voor de sociale wetenschappen kenmerkende consistentie of objectiviteit in de bestudering van jurisprudentie, waarmee een ‘*distinctively legal form of empiricism*’ wordt gecreëerd.¹⁷ Die kwalificatie roept de vraag op hoe systematische rechtspraakanalyse zich verhoudt tot empirisch-juridisch onderzoek, dat aan populariteit wint in de lage landen.

In empirisch-juridisch onderzoek, ook wel *empirical legal research* of *Empirical Legal Studies* (ELS) genoemd, wordt met behulp van gegevens verkregen uit observaties van de werkelijkheid (de empirie) onderzoek gedaan naar het recht.¹⁸ Daarbij is aandacht voor het systematisch verzamelen en analyseren van empirische gegevens overeenkomstig algemeen aanvaarde methoden en technieken, die zowel van kwalitatieve als kwantitatieve aard kunnen zijn.¹⁹ Wat betreft de toepassing van empirisch-juridisch onderzoek wordt onderscheid gemaakt tussen drie onderling verbonden pijlers: (1) onderzoek naar de

15 Zie daarover uitgebreid G. van Dijck, *Kwaliteit van de juridische annotatie. Een empirische studie naar kenmerken en kwaliteitsindicatoren*, Den Haag: Boom Juridische uitgevers 2011 en R. van Gestel & M.V.R. Snel, ‘Rehabilitatie van de vakpublicatie’, *NJB* 2019/2621, afl. 42, p. 3152-3160.

16 Vgl. G. van Dijck & D. van Kuppevelt, ‘Netwerkanalyse van werkgeversaansprakelijkheid: Deelonderwerpen en precedentes in het kader van werkgeversaansprakelijkheid ex artikel 7:658/7:661 BW’, *NJB* 2019/527, afl. 10, p. 661.

17 Hall & Wright 2008, p. 64.

18 Zie M. Smit e.a., ‘25 jaar empirisch-juridisch onderzoek in Nederland. Een inleiding tot de Encyclopedie ELS’, in: C. Bijleveld e.a. (red.), *Nederlandse Encyclopedie Empirical Legal Studies: Encyclopedie van 25 jaar empirisch juridisch onderzoek in Nederland*, Boom juridisch 2020, p. 11-12 (met verdere literatuurverwijzingen).

19 P. Cane & H.M. Kritzer, *The Oxford handbook of empirical legal research*, Oxford: Oxford University Press 2010, p. 4; L. Epstein & A.D. Martin, *An introduction to empirical legal research*, Oxford: Oxford University Press 2014, p. 3; F.L. Leeuw & H. Schmeets, *Empirical legal research: A guidance book for lawyers, legislators and regulators*, Cheltenham: Elgar Publishing 2016, p. 3-5; G. van Dijck, S. Sverdlöv & G. Buck, ‘Empirical Legal Research in Europe: Prevalence, Obstacles, and Interventions’, *Erasmus Law Review* 2018, p. 107-108; E. Hirsch Ballin, *Advanced Introduction to Legal Research Methods*, Cheltenham: Elgar Publishing 2020, p. 43-44; en K. van den Bos, *Inleiding empirische rechtswetenschap*, Den Haag: Boom juridisch 2021, p. 20.

assumpties die aan het recht ten grondslag liggen; (2) onderzoek naar de wijze van het *functioneren* van het recht in de praktijk; en (3) onderzoek naar de *effecten* van het recht.²⁰

Object van empirisch-juridisch onderzoek is dus de observeerbare wereld van het recht en de daarbij betrokken actoren, processen en omstandigheden. Rechtspraak en andere vormen van juridische geschilbeslechting maken deel uit van die wereld. Een voorbeeld van empirisch-juridisch onderzoek dat zich uitsluitend met rechtspraak bezighoudt, betreft het onderzoek dat voorheen onder de term *jurimetrics* bekendstond.²¹ Onderzoekers in die traditie proberen met behulp van geavanceerde vormen van statistiek de uitkomsten van juridische procedures te verklaren of te voorspellen aan de hand van veelal economische, sociale of politieke factoren, zoals de financiële middelen die partijen in juridische geschilbeslechting ter beschikking staan of de afkomst of politieke voorkeuren van rechters.²² Het recente en door algoritmen gedreven onderzoek ter voorspelling van het gedrag van individuele rechters (*predictive justice*) roept prangende ethische vragen op.²³

Of klassieke rechtspraakanalyse, gericht op de beschrijving van het geldende recht zoals zich dat vormt in jurisprudentie, empirisch onderzoek is, is onderwerp van discussie. De verklaring voor die discussie schuilt volgens Van Schaaijk in het feit dat jurisprudentie zowel een vindplaats is voor het geldende recht ('rechtsbron'), als voor gedrag van actoren of het verloop van processen ('kennisbron').²⁴ Traditioneel jurisprudentieonderzoek wordt voor velen buiten het veld geplaatst dat empirisch-juridisch onderzoek beslaat. Een reden daarvoor is dat onderzoekers die op basis van juridische bronnen analyses uitvoeren niet snel zullen zeggen dat ze empirisch-juridisch onderzoek doen, maar juridisch-dogmatisch onderzoek.²⁵ Dat is temeer zo indien de analyse kwalitatief van aard is en de tekst van de rechtspraak op argumentatieve hermeneutische waarde interpreteert en beoordeelt.

Toch zien verscheidene onderzoekers ook deze klassieke vorm van rechtspraakanalyse als empirisch onderzoek, mits zij plaatsvindt op basis van algemeen aanvaarde sociaalweten-

20 C.C.J.H. Bijleveld, *Research methods for empirical legal studies: An introduction*, Den Haag: Eleven International Publishing 2021, te verschijnen. Deze drieslag staat ook centraal in de 'civilologie', dat als toepassingsgebied het brede privaatrecht kent. Zie W.H. van Boom, I. Giesen, & A.J. Verheij (red.). *Capita Civilologie: Handboek empirie en privaatrecht*, Den Haag: Boom Juridische uitgevers 2013, p. 1.

21 Zie uitgebreid Hall & Wright 2008, p. 94-97 (met verdere literatuurverwijzingen).

22 Zie bijvoorbeeld: W. Wijtvliet & A. Dyevre, 'Judicial Ideology in Economic Cases: Evidence from the General Court of the European Union', *European Union Politics* 2020, afl. 1, p. 25-45.

23 Zie uitgebreid G.M. de Vries, 'Rechtswetenschappelijk onderzoek naar rechters door kwantitatieve analyse van jurisprudentie: Waar ligt de ethische grens bij het profileren van rechters?', *RMThemis* 2021, afl. 2, p. 50-68.

24 G.A.F.M. van Schaaijk, *Praktijkgericht juridisch onderzoek*, Den Haag: Boom juridisch 2019, p. 172-173.

25 Zie bijvoorbeeld Cane & Kritzer 2010, p. 5. De auteurs erkennen tegelijkertijd dat de grens tussen empirisch-juridisch en juridisch-dogmatisch onderzoek niet eenvoudig te trekken is.

schappelijke methoden en technieken. Zo bespreken de auteurs van de Nederlandse Encyclopedie *Empirical Legal Studies*, waarin op 34 deelterrein van het recht de *state-of-the-art* van Nederlands empirisch-juridisch onderzoek uiteengezet wordt, met regelmaat kwalitatieve inhoudsanalyses van rechtspraak.²⁶ Algemener nog schrijven Van Dijck, Snel en Van Golen in hun handboek over methoden van rechtswetenschappelijk onderzoek dat ‘jurisprudentieanalyse strikt genomen ook valt onder het observeren van de werkelijkheid’ en dus als empirisch onderzoek betiteld kan worden.²⁷

De discussie over de vraag in hoeverre systematische rechtspraakanalyse empirisch-juridisch onderzoek is, is echter een weinig vruchtbare. Terwijl de een steeds tekstanalyse van juridische geschilbeslechting betreft, kent de ander een veel grotere variëteit aan te analyseren (rechts)bronnen, actoren en procedures. De twee overlappen elkaar in het systematisch verzamelen, selecteren en analyseren van rechtspraak als observeerbare werkelijkheid van het recht. Het is in de methodologische aanpak en verantwoording daarvan dat systematische jurisprudentieanalyse overeenkomsten vertoont met het omvangrijkere werkterrein van de empirische rechtswetenschap.²⁸

5 **METHODOLOGISCHE VERANTWOORDING: EEN BRUGFUNCTIE**

Op het vlak van die methodologische verantwoording in het doen van systematische rechtspraakanalyse kunnen rechtswetenschappers leren van juridisch-empirisch onderzoek. Het gebrek aan conventies over methoden en technieken voor juridisch-dogmatisch onderzoek deed (doet?) onderzoekers uit andere disciplines twijfelen aan de wetenschap-

26 Antokolskaia en Schrama stellen het zo voor het familierecht: ‘Jurisprudentieanalyse en interviews zijn de meeste gebruikte empirische methoden door juristen die onderzoek doen over het familierecht’. M. Antokolskaia & W. Schrama, ‘Familierechtelijk onderzoek “goes empirical”’: Van pionieren naar een gangbare onderzoeksmethode’, in: Bijleveld e.a. 2020, p. 332.

27 G. van Dijck, M. Snel & T. van Golen, *Methoden van rechtswetenschappelijk onderzoek*, Den Haag: Boom juridisch 2018, p. 91. Zie voorts W.H. van Boom, ‘Empirisch privaatrecht: Enige beschouwingen over de rol van empirisch onderzoek in de hedendaagse privaatrechtswetenschap’, *Tijdschrift voor Privaatrecht* 2013, afl. 1, p. 7-84 (onder nr. 34) en N. Elbers e.a., ‘Empirisch-juridisch onderzoek in Nederland: Bespiegelingen over de stand van zaken in de rechtswetenschap, het juridisch onderwijs en de rechtspraktijk’, *Recht der werkelijkheid* 2018, afl. 1, p. 30. De bijdragen in Livermore & Rockmore 2019 kwalificeren rechtspraak als ‘*observational data*’ die met computerondersteunde methoden geanalyseerd kan worden.

28 Zie in die zin ook Van Boom, Desmet en Mascini, die stellen: ‘There is no sense in debating whether or why a quantitative tekst analysis of, say, court decisions is in fact ‘more empirical’ than the classic legal analysis of case law. We feel that any effort to systematically observe and understand social events, circumstance or processes relevant to the social operation of the law merits inclusion. Central to this approach is the *systematic* collection and use of observations, or data, to answer research questions.’ W.H. van Boom, P. Desmet & Peter Mascini, ‘Empirical legal research: Charting the terrain’, in: W.H. van Boom, P. Desmet & P. Mascini (red.), *Empirical Legal Research in Action. Reflections on Methods and their Applications*, Cheltenham: Edward Elgar 2018, p. 8.

pelijke status van juridisch onderzoek.²⁹ Voor klassieke jurisprudentieanalyse in het bijzonder lijkt de verantwoording vanaf de zijlijn beschouwd ondermaats te zijn. Precieze gegevens over de methodologische kwaliteit van jurisprudentieanalyses zijn er niet, maar het proefschrift van Tijssen geeft wel een inkijk. Tijssen onderzocht (op empirische wijze) hoe in negentig Nederlandse juridische proefschriften verantwoording werd afgelegd over onder meer de selectie van bronnen en de keuze en het gebruik van onderzoeksmethoden. In 73% van de dissertaties werd jurisprudentie als bron voor onderzoek gebruikt. Voor wetgeving betrof dat 86%. In minder dan de helft van de proefschriften waarin jurisprudentie en/of wetgeving als bronnen werden gebruikt, werd echter verantwoording afgelegd over het gebruik en de selectie van die bronnen en bestond de grootst mogelijke twijfel of die verantwoording in het proefschrift kon worden getraceerd.³⁰ Ook werd in slechts een derde van de dissertaties een verantwoording aangetroffen voor het gebruik van verschillende vormen van inhoudsanalyse (door Tijssen omschreven als ‘bureauonderzoek’) als methode van onderzoek.³¹

De bevindingen uit het onderzoek van Tijssen kunnen mede verklaard worden door de aard en doelstelling van juridisch-dogmatisch onderzoek. Enerzijds ligt het in die context grotendeels voor de hand welke rechtspraak op een bepaald terrein relevant is. Het zou dus vragen ‘naar de bekende weg’ zijn als de promovendus expliciet de selectie van uitspraken zou verantwoorden. Anderzijds bestaat er bij dit type onderzoek het impliciete streven naar volledigheid, waarmee verondersteld wordt dat alles op het onderzochte thema is bekeken, inclusief alle relevante rechtspraak.³² Als er echter wordt afgeweken van wat gangbaar is – in het onderzoek is aangeduid als ‘olifantenpaden’ – dient er volgens de door Tijssen bevestigde experts wel degelijk een expliciete verantwoording te zijn.³³

Een dergelijk afwijken van de norm kan bijvoorbeeld een kwantitatieve analyse van rechtspraak zijn. In dat geval voelen rechtswetenschappers op hun klompen aan dat ze moeten verantwoorden wat hun aanpak is geweest. Het onderzoek van Vols en Jacobs geeft verdere inzichten in de wijze waarop juristen dat in Nederland gedaan hebben. Zij

29 C.J.J.M. Stolker, “Ja, geléerd zijn jullie wel!” Over de status van de rechtswetenschap’, *NJB* 2003, afl. 15, p. 766-778.

30 H.E.B. Tijssen, De juridische dissertatie onder de loep: *De verantwoording van methodologische keuzes in juridische dissertaties*, Den Haag: Boom juridische uitgevers 2009, p. 125-126.

31 Tijssen 2009, p. 132-133.

32 Tijssen 2009, p. 146-147. Zie voorts J.B.M. Vranken, ‘Als een arrest niet bevalt, laten wij het gewoon weg’, in: H. Nieuwenhuis (red.), *Ex libris*, Deventer: Kluwer 2009, p. 77-94, waarin zes gezichtspunten worden geformuleerd aan de hand waarvan een antwoord kan worden verkregen op de vraag in hoeverre het in juridisch-dogmatisch onderzoek, in het bijzonder in proefschriften, vanuit wetenschappelijk perspectief geaccepteerd is om onwettelijke rechterlijke uitspraken achterwege te laten.

33 Tijssen 2009, p. 145-146 en 148. Verderop (p. 152-161) blijkt overigens dat het helemaal niet zo duidelijk is wat die olifantenpaden precies zijn.

brachten namelijk voor de periode 2006 tot midden 2016 het gebruik van kwantitatieve jurisprudentieanalyse in Nederland in kaart.³⁴ De onderzochte publicaties (n=22) laten een zeer divers beeld zien: ze bestrijken verschillen rechtsgebieden en rechtscolleges, en variëren nogal in het aantal geanalyseerde uitspraken (van n=24 tot n=8118, met een mediaan van n≈215).³⁵ Hoewel de onderzoekers in de publicaties op enigerlei wijze rekenschap geven van de wijze van dataverzameling en de gebruikte onderzoeksmethoden en -technieken, zijn Vols en Jacobs kritisch. Zo werd in nagenoeg alle publicaties niet verwezen naar methodologische literatuur, werd er slechts beperkt omschreven welke zoektermen en/of -criteria waren gebruikt ten behoeve van de dataverzameling, en was er vaak geen aandacht voor de wijze waarop de uitspraken zijn gecodeerd. Zonder gedetailleerdere informatie (o.a. over welke variabelen zijn onderscheiden bij het codeerproces, hoeveel mensen betrokken zijn geweest bij het coderen en hoe de onderzoekers codeerfouten hebben proberen te voorkomen) leverden de publicaties in op betrouwbaarheid en reproduceerbaarheid. Ook werd in de meerderheid van de publicaties geen geavanceerde statistische analysemethode gebruikt en kwam de analyse neer op het recht toe recht aan tellen van de data.³⁶

De toename van het aantal onlinedatabases van jurisprudentie in binnen- en buitenland en de opkomst van nieuwe computerondersteunde *data science*-technieken voor de verzameling en analyse van jurisprudentie, schept tot voor kort ongekende mogelijkheden tot het doen van systematische rechtspraakanalyse voor de moderne rechtswetenschapper.³⁷ Deze groei van mogelijkheden van analyse rechtvaardigt echter ook een verhoogde aandacht voor de methodologische verantwoording van die analyse. Naar mate systematisch jurisprudentieonderzoek in omvang en/of complexiteit groeit, zal ook de mate waarin de rechtswetenschapper verantwoording aflegt over zijn analyse moeten groeien.³⁸ De betrouwbaarheid en reproduceerbaarheid van het onderzoek komen anders onder druk te staan.

De juridisch onderzoeker die door middel van systematische jurisprudentieanalyse het geldende recht wil analyseren zal zich dus van het belang van methodologische verantwoording moeten vergewissen. Door zich een van de vele kwalitatieve of kwantitatieve methoden

34 M. Vols & J. Jacobs, 'Juristen als rekenmeesters: over de kwantitatieve analyse van jurisprudentie', in: P.A.J. van den Berg & G. Molier (red.), *In dienst van het recht*, Den Haag: Boom juridisch 2017.

35 Vols & Jacobs 2017, p. 98-99.

36 Vols & Jacobs 2017, p. 99-100. Of deze inzichten nog steeds representatief zijn voor de huidige stand van kwantitatieve rechtspraakanalyse, vergt nader systematisch onderzoek naar de literatuur verschenen na 2016.

37 Zie over die technieken, hun toepassing en randvoorwaarden Vols, Dyevre en Tjong Tjin Tai in deze bundel.

38 Zie in vergelijkbare zin over het belang van methodologische verantwoording in juridisch onderzoek in het algemeen Tijssen 2009, p. 148.

uit de sociale wetenschappen op een verantwoord niveau eigen te maken, leert de rechtswetenschapper op een ander wijze kijken naar jurisprudentie.³⁹ Zijn of haar interpretatieve, normatieve inborst zal daarbij niet verloren gaan.⁴⁰ Het zal in veel gevallen juist de combinatie met juridisch-dogmatisch onderzoek zijn die een systematische rechtspraakanalyse wetenschappelijke relevantie, originaliteit en (dus) kwaliteit geeft. Een analyse die losstaat van het doctrinaire discours en de toepassing van methode als – in de woorden van Peeraer en Van Gestel – *l'art pour l'art*, brengt de rechtswetenschap niet verder. Empiristen die het recht enkel en alleen bestuderen door het gebruik van sociaalwetenschappelijke methoden missen die noodzakelijke juridisch-theoretische inbedding. Het verwijt in dat geval is dat zij niet weten wat ze meten.⁴¹

Systematische rechtspraakanalyse, zo benadrukken diverse auteurs in deze bundel, heeft de potentie de verbinding te leggen tussen juridisch-dogmatisch onderzoek en empirisch-juridisch onderzoek. Enerzijds ontwikkelen klassieke rechtswetenschappers met dit type onderzoek hun kritische vermogen door op andere wijze jurisprudentie te bestuderen en daarover verantwoording af te leggen overeenkomstig bestaande conventies en spelregels in de sociale wetenschappen. Anderzijds noopt systematische rechtspraakanalyse empiristen ertoe het onderzoek dat zij doen over het recht in te bedden in de juridische doctrine.⁴² Zonder die theoretische inbedding loopt hun onderzoek immers het risico antwoorden te geven op vragen die niet juridisch relevant zijn.⁴³

39 Dat rechtswetenschappers dat in het algemeen zouden moeten doen is voor sommigen vanzelfsprekend. Zie voor een pleidooi in die zin: A. Dyevre, W. Wijtvliet & N. Lampach, 'The Future of European Legal Scholarship: Empirical Jurisprudence', *Maastricht Journal of Comparative and International Law* 2019, afl. 3, p. 348-371.

40 Het gebruik van sociaalwetenschappelijke methoden voor het doen van jurisprudentieonderzoek leidt niet noodzakelijkerwijs tot empirisch onderzoek en laat ruimte voor een argumentatieve, hermeneutische benadering. Vgl. Salehijam 2018, p. 35 voor de methode van 'systematic content analysis'.

41 G. Davies, 'The Relationship between Empirical Legal Studies and Doctrinal Legal Research', *Erasmus Law Review* 2020, afl. 2, p. 9 (met verdere literatuurverwijzingen).

42 Galligan stelt over die noodzaak: 'The ability to recognize laws, to perceive the differences between different types of law, to grasp their significance as rules, to separate legal rules from the non-legal, to identify an institution as legal, to understand the relations between law and morality, and between state law and social norms – the very subject matter of legal theory – all seem a necessary foundation for empirical research. The same applies to the analysis of data and conclusions based on the analysis, where we would expect the ideas and concepts of legal theory to be of use.' D. Galligan, 'Legal theory and empirical research', in: Cane & Kitner 2010, p. 988.

43 De samenstellers van de Nederlandse Encyclopedie *Empirical Legal Studies* stellen dat empirisch onderzoek over het recht niet tot het domein van ELS behoort indien het niet tot juridisch relevante antwoorden leidt. Bijleveld e.a. 2020, p. 12-13.

6 **WAT VOLGT IN DEZE BUNDEL**

De hoofdstukken die volgen in deze bundel vormen een greep uit het grootschalig systematische jurisprudentieonderzoek dat recentelijk in Nederland en België is verricht. De onderzoeken die beschreven worden, zijn zeer divers wat betreft de bestudeerde rechtsgebieden (bestuursrecht, Europees recht, mensenrechten, privaatrecht), de omvang van de data (van n=76 tot het volledige aantal uitspraken in de verworven database) en de methodologie (kwalitatief en/of kwantitatief, eventueel ondersteund door speciale analysesoftware en *data science*-technieken). Voorts is het ene onderzoek in het kader van een proefschrift uitgevoerd en maakt het andere deel uit van een ERC-, FWO- of NWO-gefinancierd onderzoeksproject. Weer een ander is een geheel op zichzelf staand project. In acht genomen deze diversiteit schetsen de auteurs in hun bijdragen steeds de achtergrond van hun onderzoek, de aangewende methoden en technieken van dataverzameling, -selectie en -analyse, de verkregen resultaten en de bredere lessen die zij uit het onderzoeksproces trekken.

Lianne Wijntjes (Tilburg) trapt in het eerste hoofdstuk af met haar jurisprudentieonderzoek naar de rol van excuses in de civiele rechtspraak en de medische tuchtrechtspraak. Met dat kwalitatieve onderzoek gaf ze antwoord op de vraag in hoeverre en op welke wijze (tucht)rechters het aanbieden van excuses van de medisch hulpverlener meewegen in juridische aansprakelijkheidsprocedures. Haar onderzoek, dat ze in het kader van haar proefschrift (*cum laude*) verrichtte, omvatte bijna 4000 elektronisch beschikbare uitspraken. Met behulp van een gestructureerde inhoudsanalyse heeft Wijntjes deze uitspraken ontleed. In haar bijdrage beschrijft ze hoe zij de uitspraken heeft geselecteerd, gecodeerd en geanalyseerd en welke resultaten daarmee zijn verkregen. De analyse leidde onder meer tot een categorisering van manieren waarop excuses een rol spelen in de relevante procedures. Ook gaf zij inzicht in de argumenten die (tucht)rechters gebruiken om door de benadeelde gevorderde excuses in procedures toe of af te wijzen. Ten slotte konden verschillende argumentatiestructuren in kaart worden gebracht om te ontrafelen hoe excuses als bewijsmiddel meewegen in het (tucht)rechterlijke oordeel.

De fijnmazige inhoudsanalyse die Wijntjes presenteert, laat zien dat excuses slechts in zeer specifieke omstandigheden tot erkenning van civielrechtelijke aansprakelijkheid leiden. Ook in de gevallen waarin excuses als bewijs worden toegelaten, blijkt dat zij in geen enkel oordeel van doorslaggevend belang waren in de toe- of afwijzing van de vordering. Met deze resultaten kon Wijntjes de in de literatuur bestaande opvatting over de vrees van erkenning van aansprakelijkheidsrecht door het aanbieden van excuses ontcrachten.

Pieter Wolters (Nijmegen) bespreekt de kansen en kanttekeningen bij de kwantitatieve analyse van rechtspraak voor de bestudering van geldend recht. Hij vertrekt daarbij vanuit de ervaringen die hij opdeed in zijn studies naar de matiging van contractuele boetebedingen en de verjaring van vorderingen tot schadevergoeding. In beide studies stond de rechterlijke invulling van civielrechtelijke open normen aan de hand van de omstandigheden van het geval centraal. Het doel was daarbij om diepere inzichten te bieden in de rechterlijke toepassing van die open normen dan dat traditionele kwalitatieve analyses van jurisprudentie tot dan toe had geboden. Op basis van dergelijke analyses kon namelijk niet meer worden gezegd dan dat die toepassing afhangt 'van de omstandigheden van het geval'.

Wolters betoogt dat in dit kader de meerwaarde van het doen van kwantitatief jurisprudentieonderzoek is gelegen in het blootleggen van verbanden tussen relevant geachte omstandigheden en rechterlijke oordelen. Anders dan kwalitatief onderzoek kan op basis van statistiek perspectief worden aangebracht in dergelijke verbanden, in het bijzonder door aan te geven hoe vaak bepaalde combinaties van relevant geachte omstandigheden voorkomen en met welk effect. Wolters waarschuwt echter voor de nodige subjectiviteit die schuilgaat achter de ogenschijnlijk objectieve statistiek. De resultaten leunen op de gebezigde codering en de wijze waarop die codering rechterlijke doelredeneringen of onduidelijkheden in de tekst van de uitspraken weet af te vangen. Verder zijn de resultaten van de kwantitatieve analyse weinig verklarend gebleken in de context van het onderzoek van Wolters. Hij betoogt dan ook dat een combinatie van kwantitatieve en kwalitatieve analyse noodzakelijk is.

Paul Verbruggen (Tilburg) presenteert zijn kwalitatieve jurisprudentieonderzoek naar de rechterlijke toetsing van regulerend handelen door Nederlandse sportbonden. Deze bonden zitten in de top van de piramidale organisatiestructuur van de sport en bijgevolg kan hun handelen allesbepalend zijn voor sporters en andere betrokkenen in de uitoefening en organisatie van de sport. De stap naar de burgerlijke rechter om het handelen van de sportbond te laten toetsen op zijn 'redelijkheid en billijkheid' of 'zorgvuldigheid' is dan niet ver weg, zeker niet als de sportieve of financiële belangen groot zijn. De vraag is echter hoe de rechter die toetsing vormgeeft. Het onderzoek van Verbruggen legt patronen in die rechterlijke toetsing bloot voor wat betreft de aangewende rechtsgronden, de gebruikte toetsstenen en de rechtsgevolgen.

In zijn bijdrage bespreekt Verbruggen hoe hij op basis van de zogenoemde *cross-case analysis*-methode die patronen in kaart heeft gebracht. Ter ondersteuning van deze vorm van kwalitatieve inhoudsanalyse maakte Verbruggen gebruik van *NVivo*. Dit standaardsoftwareprogramma helpt onderzoekers in het beheer van kwalitatieve data, het organiseren van de ideeën en inzichten verkregen door analyse van die data, het doorzoeken van data

en bijbehorende inzichten, het visualiseren daarvan en het erover rapporteren.⁴⁴ Aan de hand van zijn onderzoek laat Verbruggen zien hoe NVivo met zijn diverse functionaliteiten de kwalitatieve analyse van jurisprudentie kan ondersteunen. De flexibiliteit die NVivo biedt voor dataverzameling, -selectie en -analyse geeft de onderzoeker het voordeel de codering en analyse stapsgewijs en in verschillende rondes te doen. Ook bespoedigt NVivo de analyse van rechtspraak in context door de onderzoeker via codes in staat te stellen terug te keren naar de data. De analyse werkt het beste als de uitspraken van min of meer gelijke rang zijn en aldus dezelfde autoriteit genieten, hetgeen uiteraard niet altijd zo is.

Wijntjens, Wolters en Verbruggen konden voor hun analyses putten uit de rechtspraakdatabase die de Nederlandse Raad voor de rechtspraak via rechtspraak.nl beschikbaar maakt. De afwezigheid van een sterk ontwikkelde digitale database in België stelt rechtswetenschappers aldaar voor een enorme uitdaging. Hoewel de Belgische wetgever reeds in 2019 aankondigde vonnissen en arresten van feitenrechters en het Hof van Cassatie op geanonimiseerde wijze en integraal te publiceren in een publiek toegankelijke elektronische database, laat die database vooralsnog op zich wachten.⁴⁵

Elise Goossens (VU Brussel, KU Leuven en UAntwerpen), **Katalien Bollen** (KU Leuven) en **Alain-Laurent Verbeke** (KU Leuven, Harvard, Tilburg) wachtten niet af en namen de uitdaging aan. Samen met een 150-tal Masterstudenten aan de Leuvense rechtenfaculteit verzamelden zij ter griffie een kleine achthonderd uitspraken van rechtbanken van eerste aanleg en hoven van beroep op het terrein van samenwoning in het relatievermogensrecht. Dit kwantitatieve onderzoeksproject beoogde in de eerste plaats met origineel bronnen – 96% van de data is ongepubliceerde rechtspraak – het lopende nationale beleidsdebat over feitelijke samenwoning van empirisch-wetenschappelijke inzichten te voorzien. Tevens diende het project een pedagogisch doel, namelijk door studenten te trainen in het doen van andersoortig juridisch onderzoek. De aanpak en ervaringen van Goossens, Bollen en Verbeke zijn uniek en nodigen anderen uit om actief studenten, promovendi en andere jonge onderzoekers te betrekken in het doen van systematische rechtspraakanalyse, zowel in de fasen van dataverzameling, als in codering en analyse.

Michel Vols (Groningen) bespreekt aan de hand van drie onderzoeksprojecten hoe methoden van kwantitatief jurisprudentieonderzoek in rechtswetenschappelijk onderzoek kunnen worden gebruikt. De drie projecten betreffen civielrechtelijke huisuitzettingen, de bestuursrechtelijke sluiting van gebouwen en het automatisch classificeren van uitspraken van het Europees Hof voor de Rechten van de Mens (EHRM) over schending of niet-

⁴⁴ Zie uitgebreid P. Bazeley & K. Jackson, *Qualitative Data Analysis with NVivo*, Londen: SAGE 2013.

⁴⁵ Zie uitgebreid Peeraer & Van Gestel in deze bundel.

schending van mensenrechten. In dat laatste project is een algoritme ontwikkeld met als doel te onderzoeken in hoeverre beslissingen in EHRM-uitspraken (schending of niet-schending) met behulp van *machine learning*-technieken correct kunnen worden geclassificeerd op basis van data verzameld in die uitspraken, zonder dat in die data de daadwerkelijke beslissing te vinden is. Voor ieder project schetst Vols de juridische theorieën ('leerstukken') en niet-juridische theorieën die een rol spelen in het kaderen van de onderzoeksvragen en de aangewende methoden voor dataselectie, -verzameling en -analyse.

Voor Vols leiden de drie projecten gezamenlijk tot een vijftal lessen, die tevens hun weerklank vinden in de bijdragen van andere auteurs (zie nader par. 7). Zo betoogt Vols dat theorieën, ook juridisch-dogmatische, in de opzet en verantwoording van het onderzoek en methoden expliciet de aandacht moeten krijgen om de dataverzameling, -selectie en -analyse goed te kaderen. Ook als sluitstuk van kwantitatieve jurisprudentieanalyse is juridisch-dogmatische kennis echter onmisbaar. De resultaten behoeven namelijk doorgaans normatieve duiding, waarbij moet worden teruggepakt op relevante leerstukken van het geldende recht. Verder onderstreept Vols de noodzaak om na te denken over de representativiteit van elektronische data en de reproduceerbaarheid van de uitgevoerde analyse, de wenselijkheid om zich als rechtswetenschapper verder te ontwikkelen door zich nieuwe methoden van onderzoek aan te leren, en de mogelijkheden die samenwerking met andere disciplines (econometrie en computationele linguïstiek in het geval van Vols) biedt.

Arthur Dyevre (KU Leuven) stelt in zijn Engelstalige bijdrage de vraag aan de orde in hoeverre algoritmen het zoekproces van de jurist naar relevante rechtspraak kunnen ondersteunen. Iedere praktijkjurist en rechtswetenschapper begrijpt dat de zoekformulieren die beheerders van elektronische databases van rechtspraak aanbieden hun beperkingen kennen. Het (slimme) gebruik van zoektermen, booleaanse operatoren en classificatiesystemen brengt de jurist op weg, maar ook dan staan er veel vals positieven tussen de zoekresultaten. Het wegfilteren daarvan kost tijd en energie die anders geïnvesteerd had kunnen worden in de inhoudelijke analyse van de rechtspraak.

Dyevre onderzoekt in zijn bijdrage in welke mate het door computationeel-linguïsten ontwikkelde Top2Vec-algoritme de zoektocht van juristen naar juridisch relevante jurisprudentie kan verlichten. Na een beschrijving van het algoritme en de technieken die daaraan ten grondslag liggen (o.a. *natural language processing*), test Dyevre Top2Vec op een digitaal corpus van Belgische jurisprudentie zoals voorheen beschikbaar gemaakt door Jure-Juridat.⁴⁶ Het door hem gebruikte prototype blijkt – na training met behulp van ruim

⁴⁶ Sinds 15 december 2020 is deze database niet meer online beschikbaar voor het publiek. Jure-Juridat is vervangen door de JUPORTAL (<https://juportal.be/home/welkom>).

10.000 Nederlandstalige uitspraken – in staat te zijn om onderwerpen die veelvuldig ter sprake komen in de uitspraken uit te lichten en te clusteren in topics (*topic modelling*). Ook kan het op basis van gerichte zoektermen topics in jurisprudentie aanwijzen. Het prototype ondervangt daarbij tevens spelfouten of woordonderbrekingen in de rechtspraak, iets dat met reguliere zoekfuncties een hardnekkig probleem is. Dyevre stelt dat algoritmen zoals Top2Vec de potentie hebben de spreekwoordelijke speld in de hooiberg te vinden. Ze kunnen voor juristen van grote waarde zijn bij het zoeken naar relevante jurisprudentie en het genereren van een complete dataset voor de verdere inhoudelijke analyse van rechtspraak.

Eric Tjong Tjin Tai (Tilburg) bespreekt meer in het algemeen het gebruik van *data science*-technieken bij rechtspraakanalyse. Die technieken lijken veelbelovend als het gaat om het systematisch analyseren van rechtspraakdata, omdat die data ‘big’ zijn en in toenemende mate digitaal beschikbaar worden voor het publiek. Toch is voorzichtigheid geboden, waarschuwt Tjong Tjin Tai. *Big data*-projecten gebruiken namelijk vaak een zogenoemde *data pipeline*, een computersysteem dat de te analyseren data op geautomatiseerde wijze bewerkt en zo geschikt maakt voor analyse. Aan de hand van een binnen Tilburg Law School gebouwd experimenteel systeem voor onder meer de analyse van rechtspraak uit de database van rechtspraak.nl, beschrijft Tjong Tjin Tai de technische werking van een *data pipeline* en laat hij enkele toepassingen zien, waaronder netwerkanalyse van jurisprudentie. Die technische werking – en de daarmee gepaard gaande beperkingen en problemen – wegen door in de betrouwbaarheid en reproduceerbaarheid van de uitgevoerde analyse. Tjong Tjin Tai betoogt dan ook dat onderzoekers die bij het doen van rechtspraakanalyse gebruikmaken van een *data pipeline*, meer dan tot nu toe in de literatuur gebruikelijk is, inzicht moeten geven in de problemen die zich hebben voorgedaan in het opzetten van de *pipeline*. Daarbij hoort ook verantwoording over de keuzes die gemaakt zijn om de problemen het hoofd te bieden. Het hoofdstuk verschaft zo diverse aanwijzingen om het ontluikende debat over de methodologie van *data science* bij rechtspraakanalyse verder te brengen.

Frederik Peeraer (UGent en UAntwerpen) en **Rob van Gestel** (Tilburg en KU Leuven) sluiten de bundel af met hun algemene beschouwing over grootschalige systematische jurisprudentieanalyse als een uitdaging voor het juridisch onderzoek en onderwijs. Deze vorm van juridisch onderzoek, zo betogen ook zij, slaat een brug tussen methoden van juridisch-dogmatisch en empirisch-juridisch onderzoek. Dat impliceert tevens dat systematische rechtspraakanalyse niet moet worden gezien als een vervanging van traditioneel jurisprudentieonderzoek. Dat onderzoek blijft nodig voor de normatieve duiding van de verkregen resultaten uit de systematische analyse. Zij is dus een aanvulling op juridisch-dogmatisch werk, één die tegemoetkomt aan de groeiende roep aan het adres van rechts-

wetenschappers om hun methoden en technieken van onderzoek explicieter te verantwoorden. Datatoegang en datadeling zijn echter belangrijke obstakels die het volledige potentieel van dit type juridisch onderzoek kan beperken.

Peeraer en Van Gestel stellen voorts dat systematische rechtspraakanalyse ook voor het onderwijs op de rechtenfaculteiten een brug kan vormen tussen twee ogenschijnlijk botsende opvattingen in de Europese academie. Zij kan namelijk het traditionele perspectief van de praktijkjurist (advocaat, rechter, notaris, enz.) verbinden met de opmars van het methodenonderwijs in met name onderzoeksmasters en promovendiprogramma's. Peeraer en Van Gestel ontwikkelen verschillende scenario's voor de toekomst van systematische jurisprudentieanalyse in onderzoek en onderwijs. Het prikkelendst is het scenario waarin het juridisch onderzoek en onderwijs onder invloed van een veranderende vraag in de rechtspraak en technologische ontwikkelingen drastisch herzien worden. Systematische rechtspraakanalyse zou in dit scenario deel gaan uitmaken van het brede palet van onderzoeksmethoden waarmee studenten en jonge onderzoekers bekend worden gemaakt.

7 CONCLUSIE

Dit boek heeft, als gezegd, tot doel inzicht te geven in de methodologische vraagstukken die spelen bij het doen van grootschalig systematisch jurisprudentieonderzoek. De hoop is lezers te enthousiasmeren voor het opzetten en uitvoeren van dit type juridisch onderzoek. Zij kunnen daarbij voortbouwen op de door de auteurs toegepaste methoden en technieken, en zich laten inspireren door de lessen die daaruit getrokken zijn.

Het voert te ver al die lessen hier te bespreken; daarvoor zij verwezen naar de individuele hoofdstukken. Wanneer men echter die hoofdstukken gezamenlijk bekijkt, staan drie lessen centraal. Bij wijze van conclusie verdienen die drie kort de aandacht hier. Een eerste les betreft de centrale rol die traditioneel, juridisch-dogmatisch onderzoek speelt in het kaderen van systematische rechtspraakanalyse en het duiden van haar resultaten. Het ene type onderzoek kan het andere niet vervangen, zoals in paragraaf 3 reeds is gesteld. Een gecombineerde toepassing kan wel tot nieuwe en verrijkende inzichten voor de rechtstheorie en -praktijk leiden.

Een tweede les is die van het afleggen van verantwoording over de gebezigde methoden van dataverzameling, -selectie en -analyse, en het ontsluiten van die data aan het publiek. Als onderzoekers zich voor hun jurisprudentieanalyse voegen naar aanvaarde sociaalwetenschappelijke methoden en technieken, kan belangrijke winst worden geboekt in het zich rekenschap geven van de analyse en daaruit voortvloeiende bevindingen. Zoals bleek

uit paragraaf 5, staat zonder dergelijke methodologische verantwoording de kwaliteit van het onderzoek ter discussie.

Een derde centrale les ten slotte is die van de meerwaarde en noodzaak van *team science*. Alle in deze bundel gepresenteerde systematisch rechtspraakanalyses zijn tot stand gekomen door een vorm van samenwerking tussen meerdere onderzoekers. Vaak betrof het juristen, zowel rechtenstudenten als ervaren rechtswetenschappers. Soms werd er echter samengewerkt met niet-juristen, waaronder IT-specialisten, econometristen of taalkundigen. Waardering van dergelijk teamwerk heeft de aandacht van universitaire beleidsmakers. Naast het erkennen en waarderen van individuele prestaties, bestaat de intentie *team science* meer te belonen.⁴⁷

Systematisch jurisprudentieanalyse stimuleert dat teamwerk. Ten eerste maakt samenwerking de tijdsinvestering die nodig is voor het doen van een systematisch jurisprudentieonderzoek lichter. Fundamenteler is wellicht de overweging dat een vorm van *team science* nodig is voor de betrouwbaarheid van de analyse. Zo is het staande praktijk dat bij het coderen van de rechtspraak meerdere codeurs betrokken zijn. Ook bieden de auteurs in deze bundel voorbeelden van analyses waarin de samenwerking met niet-juristen heeft geleid tot innovatief interdisciplinair onderzoek.

De kritische vraag die bij dergelijke samenwerking altijd opkomt, is in hoeverre de betrokken rechtswetenschapper moet begrijpen wat de niet-jurist doet. Tjong Tjin Tai vraagt zich bijvoorbeeld terecht af: ‘Controleert de jurist de data scientist daadwerkelijk inhoudelijk?’. Een zekere mate van controle zal wel degelijk moeten worden georganiseerd binnen het team. Dat vergt dat ook de rechtswetenschapper zich tot op zekere hoogte bekwaamt in de toegepaste methoden en technieken. Het zal niet verbazen dat de auteurs in deze bundel die investering toejuichen. Waartoe zij kan leiden, vormt de inhoud van dit boek.

47 VSNU, NFO, KNAW, NWO en ZonMw, ‘Ruimte voor ieders talent: Naar een nieuwe balans in het erkennen en waarderen van wetenschappers’ Position Paper, november 2019, www.scienceguide.nl/wp-content/uploads/2019/11/283.002-Erkennen-en-Waarderen-Position-Paper_NL_web.pdf en VSNU, ‘VSNU, NWO, NFO en ZonMw geven impuls aan verandering in het waarderen en belonen van wetenschappers’, november 2018, https://vsnu.nl/nl_NL/nieuwsbericht.html/nieuwsbericht/489-vsnu-nwo-nfo-en-zonmw-geven-impuls-aan-verandering-in-het-waarderen-en-belonen-van-wetenschappers.



2 HET VERRICHTEN VAN EEN GESTRUCTUREERDE RECHTSPRAAKANALYSE

De rol van excuses in de civiele rechtspraak en de medische tuchtrechtspraak

Lianne Wijntjens*

1 INLEIDING

Wanneer grote hoeveelheden uitspraken in kaart dienen te worden gebracht is het verstandig een gestructureerde aanpak te hanteren. Dit maakt de onderzoeksresultaten beter te controleren en meer valide. Een kwalitatieve methode die hiervoor gebruikt kan worden is de gestructureerde inhoudsanalyse.¹ Dit is een onderzoeksmethode die wordt gebruikt om op een systematische en replicerbare wijze teksten of andere onderzoeksobjecten (zoals afbeeldingen of geluidsopnamen) te analyseren en categoriseren.² Deze methode kan ook worden gebruikt voor het analyseren van jurisprudentie en is bij uitstek geschikt om grote lijnen te ontwaren in grote hoeveelheden uitspraken.³ In deze bijdrage zal ik aan de hand van mijn promotieonderzoek⁴ omschrijven hoe een gestructureerde jurisprudentie

* Mr. dr. L.A.B.M. Wijntjens is als universitair docent verbonden aan Tilburg University.

1 Zie hierover bijvoorbeeld M.W. Bauer, 'Classical Content Analysis: a Review', in: M.W. Bauer & G. Gaskell, *Qualitative Researching with Text, Image and Sound*, Londen: SAGE Publications 2000 en K. Krippendorff, *Content Analysis. An Introduction to Its Methodology*, Los Angeles: SAGE Publications 2013.

2 Zie Krippendorff 2013, p. 24 waar een 'content analysis' wordt gedefinieerd als 'a research technique for making replicable and valid inferences from texts (or other meaningful matter) to the contexts of their use' en A. Bryman, *Social research methods*, Oxford: Oxford University Press 2012, p. 289 waarin de volgende definitie wordt gehanteerd: 'content analysis is an approach to the analysis of documents and texts (which may be printed or visual) that seeks to quantify content in terms of predetermined categories and in a systematic and replicable manner'. Zie ook Bauer 2000 waarin een overzicht wordt gegeven van de in de literatuur gehanteerde definities.

3 In feite kan een inhoudsanalyse worden toegepast op alle typen tekst. Zie M.A. Hall & R.F. Wright, 'Systematic Content Analysis of Judicial Opinions', *California Law Review* (96) 2008, afl. 1, p. 63-122, waarin *best practices* worden gegeven voor het verrichten van een systematische inhoudsanalyse van rechterlijke uitspraken.

4 L.A.B.M. Wijntjens, *'Als ik nu sorry zeg, beken ik dan schuld?' Over het aanbieden van excuses in de civiele procedure en de medische tuchtprocedure*, Den Haag: Boom juridisch 2020.

analyse kan worden uitgevoerd.⁵ Ten behoeve van dat onderzoek heb ik bijna vierduizend uitspraken geanalyseerd om te achterhalen welke rol excuses spelen in de civiele rechtspraak en de medische tuchtrechtspraak. Ik zal laten zien hoe ik dit heb aangepakt en welke resultaten de gestructureerde aanpak mij heeft opgeleverd. In paragraaf 2 omschrijf ik eerst kort de achtergrond en het doel van de door mij verrichte jurisprudentieanalyse. Vervolgens sta ik in paragraaf 3 uitgebreid stil bij de gehanteerde methode. De belangrijkste onderzoeksresultaten bespreek ik heel beknopt in paragraaf 4. Ik sluit in paragraaf 5 af met enkele inzichten die ik gedurende het onderzoeksproces heb opgedaan over het verichten van een gestructureerde jurisprudentieanalyse.

2 ACHTERGROND EN ONDERZOEKSDOEL

Uit onderzoek blijkt dat het aanbieden van excuses het verloop van het schadeafwikkelingsproces kan beïnvloeden.⁶ Enerzijds blijkt dat excuses tegemoetkomen aan de immateriële behoeften van benadeelden en op deze manier het herstelproces bevorderen.⁷ Benadeelden blijken wanneer zij excuses ontvangen dan ook minder geneigd te zijn een juridische procedure te starten of andere juridische interventies te plegen en juist eerder geneigd te zijn om te schikken.⁸ Anderzijds wordt in de literatuur gesteld dat (de schaduw van) een aansprakelijkheidsprocedure schadeveroorzakers ervan kan weerhouden om excuses aan te bieden. De grote potentie die excuses hebben in het schadeafwikkelingsproces zou hierdoor niet ten volle kunnen worden benut.

In de literatuur is er veel aandacht voor de potentiële negatieve juridische consequenties van het aanbieden van excuses.⁹ Zo wordt wel gesteld dat het aanbieden van excuses een

5 Zie voor een ander voorbeeld van een promotieonderzoek waarin een gestructureerde inhoudsanalyse centraal stond: E. van Wijk-Verhagen, *Redelijkheid en billijkheid in het personen- en familierecht, meer in het bijzonder in het huwelijks- en echtscheidingsvermogensrecht. Wonderolie van flexibiliteit of gif van rechtsonzekerheid?*, Den Haag: Boom juridisch 2018. Van Wijk-Verhagen onderzocht hoe de redelijkheid en billijkheid worden toegepast in de Nederlandse jurisprudentie op het gebied van het personen- en familierecht.

6 Zie voor een uitgebreid literatuuronderzoek naar de werking en effectiviteit van excuses Wijntjens 2020, hfdst. 2.

7 Zie o.a. R.M.E. Huver e.a., *Slachtoffers en aansprakelijkheid. Een onderzoek naar behoeften, verwachtingen en ervaringen van slachtoffers en hun naasten met betrekking tot het civiele aansprakelijkheidsrecht. Deel 1: Terreinverkenning*, Den Haag: WODC 2007.

8 Zie o.a. J.K. Robbennolt, 'Apologies and legal settlement: an empirical examination', *Michigan Law Review* (102) 2003, afl. 3, p. 460-516 en J.K. Robbennolt, 'Apologies and settlement levers', *Journal of Empirical Legal Studies* (3) 2006, afl. 2, p. 333-373.

9 Zie o.a. K.P. Hoogenboezem & M.C. Hees, 'Het verbod tot erkenning van aansprakelijkheid', *WPNR* (154) 2014, afl. 7025, p. 638-642 ('Thans geldt echter nog altijd dat door de verzekerde verstrekte tekst en uitleg of excuses door de benadeelde redelijkerwijze kunnen worden opgevat als erkenning van aansprakelijkheid

erkenning van aansprakelijkheid zou kunnen impliceren. Ter onderbouwing van deze stelling wordt vaak verwezen naar enkele (en steeds dezelfde) uitspraken of naar anekdotisch bewijs (bijvoorbeeld een bij de auteur bekend voorbeeld uit de praktijk).¹⁰ Er is echter nog niet eerder gestructureerd onderzocht in hoeverre en op welke wijze excuses een rol spelen in de rechtspraak. Met mijn promotieonderzoek heb ik getracht deze kennislacune op te vullen. Ik heb er daarbij voor gekozen mij specifiek te richten op twee juridische procedures: de civiele procedure en de medische aansprakelijkheidsprocedure.¹¹ De onderzoeksvraag die ik door middel van het gestructureerde jurisprudentie onderzoek hebben willen beantwoorden is:

In hoeverre en op welke wijze spelen excuses een rol in de civiele rechtspraak en de medische tuchtrechtspraak, en werkt dit belemmerend of bevorderend voor het aanbieden van excuses?

Het doel van het jurisprudentieonderzoek was daarmee het verkrijgen van meer inzicht in de mate waarin, en de wijze waarop, excuses een rol spelen in het oordeel van de civiele rechter en de medische tuchtrechter. Binnen de civiele procedure en medische tuchtrecht-procedure dienen (tucht)rechters verschillende beslissingen te nemen en deze beslissingen kunnen elk op een verschillende wijze worden beïnvloed door excuses. Een gestructureerde uiteenzetting van de wijze waarop excuses de besluitvorming in deze procedures beïnvloeden ontbrak in de literatuur. Met mijn promotieonderzoek wilde ik meer inzicht op dit punt verkrijgen, aangezien de wijze waarop (tucht)rechters omgaan met excuses van invloed kan zijn op hoe zij dat in toekomstige gevallen doen (precedentwerking). Procespartijen

en dat deze opvatting van de benadeelde in beginsel doorslaggevend is. Om die reden doet de verzekerde er helaas goed aan beperkt, en zeer secuur, tekst en uitleg te geven en terughoudend te zijn met het maken van excuses'. Zie o.a. in de internationale literatuur: C. Farmer, 'Striking a Balance: A Proposed Amendment to the Federal Rules of Evidence Excluding Partial Apologies', *Belmont Law Review* (2) 2015, p. 249, waarin het aanbieden van excuses 'legally dangerous' wordt genoemd en J.R. Cohen, 'Advising Clients to Apologize', *Southern California Law Review* (72) 1999, afl. 5, p. 1011, waarin wordt opgemerkt 'although a physician may wish to tell a patient when he has made a mistake, lawyers often order doctors to say nothing'.

- 10 Zie o.a. Hoogenboezem & Hees 2014, p. 638-642; L.A.B.M. Wijntjens, 'Excuses en aansprakelijkheid: In hoeverre en op welke wijze beïnvloeden excuses het oordeel van de rechter?', *WPNR* (147) 2016, afl. 7092, p. 82-88; A.M. Zwart-Hink, 'Moet wie excuses aanbiedt ook schade vergoeden?: Mythe en werkelijkheid over het verband tussen excuses en aansprakelijkheid', *NJB* 2017, afl. 38, p. 2800-2808 en G. van Dijk, 'The Ordered Apology', *Oxford Journal of Legal Studies* 2017, afl. 3, p. 562-587. Zie voor een uitgebreid overzicht van het wetenschappelijke debat (onder verwijzing naar literatuur) op dit punt Wijntjens 2020, par. 3.2.
- 11 De keuze voor de civiele procedure lag voor de hand aangezien het Nederlandse wetenschappelijke debat over de rol van excuses in het recht zich voornamelijk hierop richt. Er is gekozen om een vergelijking te maken met de medische tuchtprocedure aangezien door verschillende auteurs wordt gesuggereerd dat de civiele rechtspraak en de medische tuchtrechtspraak elkaar wederzijds kunnen inspireren in de wijze waarop met excuses wordt omgegaan (zie Zwart-Hink 2017 en E.M. Deen, 'Centraal Tuchtcollege voor de Gezondheidszorg 08-03-2018, ECLI:NL:TGZCTG:2018:68, (X/Y). Het Centraal Tuchtcollege over de verantwoordelijkheid van zorgverleners bij schadeafwikkeling, het erkennen van fouten en excuses', *GZR Updates* 2018-0144).

en hun vertegenwoordigers zullen hun gedrag hierop aanpassen. Dit brengt mee dat de rol die excuses spelen in een juridische procedure, de bereidheid van procespartijen om excuses aan te bieden kan beïnvloeden. Indien blijkt dat het aanbieden van excuses juridische risico's met zich meebrengt, dan zullen zij hiermee logischerwijs terughoudendheid betrachten. Anderzijds zullen zij eerder geneigd zijn om excuses aan te bieden indien dit in de procedure positieve consequenties kan hebben.

3 METHODE

In deze paragraaf bespreek ik welke aanpak ik heb gehanteerd bij het uitvoeren van de systematische jurisprudentieanalyse. Deze aanpak kan worden onderverdeeld in drie fasen: de selectie van de uitspraken (par. 3.1), het codeerproces (par. 3.2) en de analyse van de data (par. 3.3).

3.1 *Selectie van de uitspraken*

De eerste stap van de systematische jurisprudentieanalyse betrof het selecteren van relevante uitspraken. De methodiek van het verrichten van een gestructureerde inhoudsanalyse vereist dat dit op een systematische en transparante wijze gebeurt, zodat de werkwijze door andere onderzoekers te herhalen is.¹² Voor de beide onderzochte procedures zijn alle elektronisch gepubliceerde uitspraken tot en met oktober 2018 meegenomen. Voor het verzamelen van de civiele uitspraken is gebruikgemaakt van de open data die publiekelijk beschikbaar worden gesteld via de website van de Rechtspraak.¹³ Het betreft een groot databestand waarin alle (geanonimiseerde) uitspraken zijn opgenomen die ook raadpleegbaar zijn via de uitsprakendatabank van Rechtspraak.nl.¹⁴ Gezocht is binnen alle uitspraken die in het databestand waren gecodeerd als civiel. Daarvan bevonden zich op dat moment 140.056 uitspraken in het databestand.

Voor het verzamelen van de uitspraken van de Tuchtcolleges voor de Gezondheidszorg is een andere aanpak gehanteerd. De reden hiervoor is dat er geen open databestand

12 Hall & Wright 2008, p. 79.

13 www.rechtspraak.nl.

14 Een belangrijke beperking hierbij is gelegen in het feit dat niet alle civiele rechtspraak is terug te vinden in het gebruikte open databestand van Rechtspraak.nl. Slechts een zeer klein gedeelte van de uitspraken wordt gepubliceerd. De afweging of een uitspraak gepubliceerd dient te worden, wordt gemaakt door de rechtbank zelf. Met andere woorden, de uitspraken worden niet louter toevallig gepubliceerd, maar dit is afhankelijk van het beleid van het desbetreffende gerecht of zelfs van het oordeel hieromtrent van de uitspraak doende rechter zelf. Dit brengt mee dat de geselecteerde civiele uitspraken mogelijk niet representatief zijn voor de gehele populatie met daarin alle andere niet-gepubliceerde uitspraken.

beschikbaar is waarin alle gepubliceerde medische tuchtrechtspraak is opgenomen. Er is daarom gezocht op de online uitsprakendatabank van Overheid.nl waar de tuchtrechtspraak wordt gepubliceerd.¹⁵ Binnen het domein ‘gezondheidszorg’ waren op dat moment 8.283 gepubliceerde tuchtrechtuitspraken terug te vinden.

Vervolgens zijn de uitspraken doorzocht met de volgende trefwoorden: excuus, excuses, spijt, verontschuldiging(en), sorry en berouw.¹⁶ Deze trefwoorden zijn gekozen na een inventarisatie van de eerder gevonden jurisprudentie waarin excuses een rol spelen. Daarnaast is gekeken naar eerder onderzoek waarin met behulp van zoektermen is gezocht naar jurisprudentie waarin excuses een rol spelen.¹⁷ Voor de civiele rechtspraak hebben de zoekacties 3.452 treffers opgeleverd en voor de medische tuchtrechtspraak 872 treffers. Nadat de dubbele treffers waren verwijderd, bleven er respectievelijk 3.117 en 821 uitspraken over.

3.2 Codeerproces

Het coderen van teksten betekent dat er aan een tekstfragment een samenvattend label (code) wordt gekoppeld.¹⁸ Op deze manier kunnen tekstfragmenten met dezelfde code op een later moment goed met elkaar worden vergeleken. Het coderen van uitspraken versterkt de objectiviteit en de repliceerbaarheid van het jurisprudentieonderzoek.¹⁹ Het coderen van de uitspraken heeft plaatsgevonden in twee fasen. In de eerste fase zijn de uitspraken globaal bekeken en heeft een eerste categorisatie plaatsgevonden, waarbij met name werd bekeken of de uitspraken wel of niet relevant waren voor het onderzoek. Om dit te bewerkstelligen is voor beide procedures een overzicht gecreëerd waarin steeds de paragraaf of paragrafen terug waren te vinden waarin de gebruikte zoektermen voorkwamen. Voor de civiele uitspraken is op basis van de open data een bestand gegenereerd waarin van alle gevonden uitspraken steeds het ECLI-nummer en het tekstfragment waarin de specifieke zoekterm voorkwam, terug waren te vinden. Voor de uitspraken van de Tuchtcolleges voor de Gezondheidszorg zijn voor de gevonden uitspraken steeds het ECLI-nummer en het tekstfragment waarin de specifieke zoekterm voorkwam, handmatig ingevoerd in een Excelbestand. Op basis van deze overzichten kon bij de meeste uitspraken in één oogopslag een categorisatie van de uitspraak plaatsvinden. Indien dat niet lukte, dan werd de betref-

15 Uitspraken van de Tuchtcolleges voor de Gezondheidszorg worden wel allemaal gepubliceerd. Hier geldt echter dat op de gebruikte onlinedatabase van Overheid.nl slechts uitspraken zijn gepubliceerd die zijn gedaan vanaf 2010. Uitspraken van voor die datum zijn dus niet meegenomen in dit onderzoek.

16 Hierbij zijn mede alle combinaties tussen deze trefwoorden meegenomen.

17 Zie Van Dijck 2017.

18 H. Boeije, *Analyseren in kwalitatief onderzoek. Denken en doen*, Amsterdam: Boom uitgevers 2014, p. 95.

19 Hall & Wright 2008, p. 81.

fende uitspraak alsnog integraal bekeken. In een Excelbestand zijn de gevonden uitspraken vervolgens gecodeerd. Bij de eerste globale categorisatie werd gelet op de volgende aspecten:

1. Allereerst werd bekeken of de gevonden uitspraak relevant was voor het onderzoek. Van relevantie voor het onderzoek was sprake wanneer uit de uitspraak bleek dat door een of meer procespartijen excuses waren aangeboden of dat uit de uitspraak bleek dat er juist geen excuses waren aangeboden.
2. Wanneer was vastgesteld dat de uitspraak relevant was voor het onderzoek werd in het Excelbestand aangegeven of uit de uitspraak bleek dat er wel of juist geen excuses waren aangeboden.
3. Vervolgens is bekeken of de wel/niet aangeboden excuses worden genoemd in het oordeel van de (tucht)rechter. Hiervan was sprake wanneer in het gedeelte van de uitspraak waarin de rechtbank/het tuchtcollege de vordering(en)/de klacht beoordeelt (meestal te vinden onder kopjes als 'de beoordeling van het geschil', 'het geschil en de beoordeling daarvan', 'de overwegingen van het college' of 'beoordeling') wordt benoemd dat minimaal een van de partijen wel/geen excuses heeft aangeboden. Doel van deze categorie was om een overzicht te creëren van alle uitspraken waarin excuses op enige wijze een rol spelen in het oordeel van de (tucht)rechter.
4. Er waren twee categorieën uitspraken die interessant zijn voor het onderzoek, maar die niet volledig werden gedekt door de vorige categorie. In de eerste plaats ging het daarbij om uitspraken waarin excuses waren gevorderd. Dit betreft een voor het onderzoek interessante categorie. In deze gevallen bleek echter dat in het gedeelte van de uitspraak waarin een oordeel wordt gegeven niet altijd terug werd gekomen op de gevorderde excuses (bijvoorbeeld omdat de vordering in zijn geheel werd afgewezen). Om die reden is deze categorie uitspraken afzonderlijk gecodeerd. Hetzelfde gold voor een categorie uitspraken die enkel was terug te vinden in de medische tuchtrechtspraak. Hierbij ging het om uitspraken waarin een van de formeel aan het tuchtcollege gerichte klachten het verwijt betrof dat door de beroepsbeoefenaar geen excuses waren aangeboden. Ook deze categorie uitspraken is afzonderlijk gecodeerd.

In de tweede fase van het codeerproces werden de onder 3 (wel/niet aanbieden van excuses wordt benoemd in het oordeel) en 4 (excuses werden gevorderd of er werd geklaagd over het uitblijven van excuses) gevonden uitspraken nader geanalyseerd en gecodeerd. Hierbij zijn verschillende hoofd- en subcategorieën gemaakt die corresponderen met de verschillende wijzen waarop excuses een rol spelen in de uitspraken.

De codering van uitspraken heeft aldus plaatsgevonden op twee niveaus. Eerst heeft er een globale categorisatie plaatsgevonden en vervolgens zijn bepaalde categorieën uitspraken nader geanalyseerd. Om de codering op beide niveaus op een consistente wijze uit te voeren is gebruikgemaakt van codeerschema's. Deze codeerschema's zijn in eerste instantie afgeleid

uit de literatuur over de rol van excuses in rechterlijke uitspraken en eerder gevonden jurisprudentie waarin excuses een rol spelen. De analyse van deze literatuur en jurisprudentie heeft geleid tot verwachtingen over hoe excuses een rol kunnen spelen in het oordeel van de (tucht)rechter, waaruit verschillende vooraf opgestelde codes zijn afgeleid. Deze codes werden tijdens het codeerproces verder verfijnd (bijvoorbeeld door het opstellen van nieuwe en subcategorieën). Er is aldus gebruikgemaakt van een gecombineerde deductieve en inductieve aanpak.²⁰ Enerzijds is een aantal codes opgesteld voorafgaand aan het codeerproces, hetgeen kan worden gekwalificeerd als een deductieve aanpak. Anderzijds zijn de uitspraken die niet met de vooraf opgestelde codes gelabeld konden worden alsnog inductief geanalyseerd. Naarmate het codeerproces vorderde, werd duidelijk dat er tussen de twee onderzochte procedures verschillen bestaan in de wijze waarop excuses een rol spelen in de beoordeling. Om die reden zijn voor beide procedures afzonderlijke codeerschema's opgesteld.²¹

3.3 Analyse van de data

De volgende stap van jurisprudentieanalyse betreft het analyseren van de data. De analyse van de data heeft deels plaatsgevonden tijdens het codeerproces. In het onderzoek wordt in de kern namelijk een zeer vergaande categorisatie gemaakt van de rechtspraak waarin excuses een rol spelen. Veel uitspraken zijn om die reden meermaals bekeken. Eerst heel globaal waarbij enkel werd gekeken of de betreffende uitspraak relevant was voor onderzoek. Vervolgens steeds gedetailleerder waarbij steeds fijnmazigere categorieën werden aangebracht en ontdekt. Hierbij werd het codeertabel steeds aangescherpt en aangepast. De volgende stap was het nader analyseren van de uitspraken die waren gevonden binnen de verschillende categorieën. Hierbij zijn verschillende technieken toegepast. In het vervolg van deze paragraaf zal ik deze analysetechnieken toelichten aan de hand van concrete voorbeelden. Achtereenvolgend bespreek ik het in kaart brengen van argumenten (par. 3.3.1), het aantonen van verbanden (par. 3.3.2) en het opstellen van argumentatieschema's (par. 3.3.3).

20 Van oudsher wordt kwalitatieve data-analyse geassocieerd met de gefundeerde theoriebenadering, waarbij de onderzoeker bij de data-analyse louter inductief te werk gaat (zie bijvoorbeeld Boeije 2014). Tegenwoordig is het echter algemeen aanvaard dat ook een deductieve aanpak meerwaarde kan hebben bij de analyse van teksten. Zie bijvoorbeeld de door Ritchie en Lewis omschreven 'thematic framework'-methode. Zie voor een voorbeeld van een studie waarin een inductieve en deductieve aanpak worden gecombineerd: A. van Staa & K. de Vries, 'Directed content analysis: een meer deductieve dan inductieve aanpak bij kwalitatieve analyse', *Kwalon* (19) 2014, afl. 3, p. 46-54. Zie ook de bijdrage van Verbruggen in deze bundel.

21 Zie voor de uiteindelijke codeerschema's Wijntjens 2020, bijlage 1.

3.3.1 Het in kaart brengen van argumenten

Een eerste analysetechniek betreft het in het kaart brengen van de in uitspraken genoemde argumenten. Dit hebt ik bijvoorbeeld gedaan bij het analyseren van de categorie civiele uitspraken waarin excuses waren gevorderd (n=175). Bij de analyse van deze uitspraken heb ik erop gelet welke argumenten werden genoemd voor het toe- of afwijzen van een dergelijke vordering. Van de 175 gevonden uitspraken waarin excuses werden gevorderd, werd deze vordering in 27 gevallen toegewezen. Opvallend was dat de beslissing om een vordering tot het aanbieden van excuses toe te wijzen in geen enkel geval wordt toegelicht. In de 148 uitspraken waarin een vordering tot het aanbieden van excuses werd afgewezen, werd in veel gevallen wel een toelichting gegeven.²² In deze uitspraken droeg de rechter minimaal één argument aan voor het afwijzen van de vordering tot het aanbieden van excuses. Hierbij bleek dat er onderscheid kon worden gemaakt tussen principiële en contextgebonden argumenten. Er is sprake van een principieel argument wanneer een argument wordt aangedragen waarmee wordt bepleit dat een vordering tot het aanbieden van excuses *in zijn algemeenheid* niet toewijsbaar is (dus ongeacht welke omstandigheden er spelen). Van een contextgebonden argument is sprake wanneer een argument wordt aangedragen waarom de vordering tot het aanbieden van excuses *in dit specifieke geval* niet toewijsbaar is. De analyse heeft uiteindelijk geleid tot een overzicht van argumenten die civiele rechters hanteren bij het afwijzen van een vordering tot het aanbieden van excuses (zie tabel 1 voor een samenvattend overzicht). Vervolgens was het mogelijk om deze argumenten te toetsen aan de literatuur. Op basis hiervan kon ik concluderen dat de door de civiele rechters genoemde argumenten reeds in de literatuur waren weerlegd of genuanceerd.²³ Daarmee werd duidelijk dat de in de rechtspraak gebezigde argumentatie geen of weinig steun vindt in de wetenschap.

22 In 62 van deze uitspraken werd de gehele vordering afgewezen (bijvoorbeeld omdat er volgens de rechter geen sprake was van onrechtmatigheid) en dus ook de vordering tot het aanbieden van excuses. Daarnaast zijn er 23 uitspraken gevonden waarin de vordering tot veroordeling tot rectificatie werd afgewezen en daarmee dus ook de daarin opgenomen excuses. Er is één uitspraak gevonden waarin de beslissing werd aangehouden voor verdere bewijslevering. In deze gevallen (86) werd aldus niet toegekomen aan een inhoudelijk oordeel over de vorderingen tot het aanbieden van excuses. In de overige 62 gevallen werd wel een oordeel gegeven over de vordering tot het aanbieden van excuses. Deze uitspraken zijn verder geanalyseerd, waarbij is bekeken welke argumenten de rechter gebruikt voor het afwijzen van een dergelijke vordering. In 10 uitspraken werd niet uitgelegd waarom de vordering tot het aanbieden van excuses werd afgewezen. In de overige 52 uitspraken droeg de rechter wel minimaal één argument aan voor het afwijzen van de vordering tot het aanbieden van excuses.

23 Zie hierover uitgebreid Wijntjens 2020, par. 3.6.2.

Tabel 1 Samenvattend overzicht van de in de rechtspraak gehanteerde argumenten voor het afwijzen van een vordering tot het aanbieden van excuses

Principiële argumenten: er wordt bepleit dat een vordering tot het aanbieden van excuses <i>in zijn algemeenheid</i> niet toewijsbaar is (dus ongeacht welke omstandigheden er spelen).	<ul style="list-style-type: none"> – <i>Het morele argument: excuses zijn uitingen van persoonlijke gevoelens en kunnen daarom niet worden afgedwongen.</i> – <i>Het effectiviteitsargument: excuses zijn niet effectief (bijvoorbeeld omdat ze geen waarde hebben, niet oprecht zijn of niet zinvol zijn).</i> – <i>Strijd met de vrijheid van meningsuiting: het opleggen van excuses is in strijd is met de vrijheid van meningsuiting.</i>
Contextgebonden argumenten: er wordt bepleit dat de vordering tot het aanbieden van excuses <i>in dit specifieke geval</i> niet toewijsbaar is.	<ul style="list-style-type: none"> – <i>Onvoldoende belang: de eiser heeft onvoldoende belang bij toewijzing van de vordering tot het aanbieden van excuses.</i> – <i>Beslissing voldoende: de eiser moet in de beslissing (waarin bijvoorbeeld is vastgesteld dat er onrechtmatig is gehandeld) voldoende genoegdoening of rehabilitatie vinden.</i> – <i>Juridische basis onvoldoende vast komen te staan: de juridische basis voor een vordering tot het aanbieden van excuses is onvoldoende vast komen te staan.</i>

3.3.2 Verbanden aantonen

Bij een andere categorie uitspraken heb ik specifiek gelet op verbanden die in de betreffende uitspraken worden gelegd. Binnen de bestudeerde rechtspraak heb ik verschillende uitspraken gevonden waarin het feit dat een procespartij wel of juist geen excuses heeft aangeboden wordt benoemd als een relevante omstandigheid bij de invulling van een open norm. Hierbij gaat het om gevallen waarin de rechter bij het bepalen van de toepasselijkheid van een rechtsregel verschillende omstandigheden van het geval tegen elkaar afweegt, en het wel/niet aanbieden van excuses daarbij meeweegt. Hier heb ik erop gelet dat de rechter of het tuchtcollege een direct verband legt tussen de wel/niet aangeboden excuses en de toepasselijkheid van een rechtsregel. Om dit verband aan te tonen heb ik gelet op signaalwoorden. Indicatoren die op dergelijk verband wijzen zijn bijvoorbeeld: ‘onder deze omstandigheden’, ‘daarom’, ‘om die reden’, ‘gezien het voorgaande’, ‘dus’, ‘derhalve’.

Door deze aanpak heb ik inzichtelijk kunnen maken dat excuses door zowel de civiele rechter als de medische tuchtrechter worden meegewogen als een relevante omstandigheid bij de invulling van open normen. Dit meewegen blijkt in beide procedures (in hoofdlijnen) op twee niveaus plaats te vinden. Allereerst kunnen excuses een rol spelen wanneer de toepassing van de materiële norm vereist dat een inschatting moet worden gemaakt van hoe een procespartij zich in de toekomst gaat gedragen. Excuses vormen in dat kader een indicatie dat men heeft geleerd van eerder gemaakte fouten:

‘Verweerder heeft echter een en andermaal en ook nog eens uitdrukkelijk ter terechtzitting zijn *spijt betuigd* over de gedane uitspraken. *Daarmee* heeft verweerder in voldoende mate laten zien dat hij in dit opzicht fout is geweest. Een andere sanctie is dan ook niet nodig. Dat betekent dat ook dit onderdeel van de klacht ongegrond zal worden verklaard.’²⁴

‘*Ook acht de kantonrechter van belang* dat werknemer steeds open kaart heeft gespeeld met betrekking tot hetgeen is voorgevallen, dat hij ten overstaan van werkgeefster het boetkleed heeft aangetrokken en dat hij aan de schoonmaakster een *excuusbrief* heeft geschreven.’²⁵

Op de tweede plaats blijken excuses een rol te kunnen spelen bij het bepalen van de hoogte van de immateriële schadevergoeding (in de civiele procedure) of het bepalen van de op te leggen maatregelen (in de medische tuchtprocedure). Dit type beslissingen vertoont overeenkomsten met strafmaatbeslissingen in strafrechtelijke zaken:

‘Bij het bepalen van de daarvoor toe te kennen genoegdoening neemt de rechtbank *mede in aanmerking* dat de Stichting, zoals eerder is overwogen, *welgemeende excuses* heeft aangeboden en diverse pogingen heeft gedaan om de eerder gemaakte fouten te herstellen (...).’²⁶

‘Anderzijds heeft verweester haar fouten volledig erkend en hiervoor *excuses* aangeboden aan patiënte en haar naasten. *Daarom* ziet het College geen aanleiding een zwaardere maatregel dan een berisping op te leggen.’²⁷

‘Wat betreft de maatregel overweegt het college *het volgende*. (...) Voorts overweegt het college dat verweerder heeft *nagelaten excuses* aan klager aan te bieden.’²⁸

De wijze waarop excuses in de gevonden uitspraken als relevante omstandigheid worden meegewogen in het oordeel van de (tucht)rechter kan worden gezien als stimulerende factor voor het aanbieden van excuses. Hierbij is er sprake van een dubbele prikkel: het aanbieden van excuses wordt beloond terwijl het niet aanbieden van excuses wordt afgestraft. De sterke focus in de literatuur op potentiële negatieve consequenties gaat eraan

24 RTG Amsterdam 22 mei 2015, ECLI:NL:TGZRAMS:2015:50.

25 Rb. Zwolle-Lelystad 10 februari 2006, ECLI:NL:RBZLY:2006:AV1908.

26 Rb. Den Haag 12 december 2001, ECLI:NL:RBSGR:2001:AD7073.

27 RTG Den Haag 13 juni 2017, ECLI:NL:TGZRSGR:2017:97.

28 RTG Eindhoven 9 november 2015, ECLI:NL:TGZREIN:2015:92.

voorbij dat het aanbieden van excuses regelmatig in het voordeel van de aanbieder wordt meegewogen. Meer bekendheid hiervan bij schadeveroorzakers en hun procesvertegenwoordigers kan wellicht ertoe bijdragen dat er meer bereidheid ontstaat om excuses aan te bieden.²⁹

3.3.3 Argumentatieschema's

Een andere manier om groepen uitspraken meer diepgaand te analyseren is het in kaart brengen van argumentatiestructuren. In mijn onderzoek heb ik deze techniek gebruikt om inzichtelijk te maken op welke wijze en hoe zwaar excuses als bewijsmiddel meewegen in het rechterlijk oordeel.³⁰ Zowel binnen de civiele rechtspraak als binnen de medische tuchtrechtspraak heb ik verschillende uitspraken gevonden waarin aangeboden excuses worden gebruikt ter onderbouwing van bepaalde feiten die ten grondslag worden gelegd aan het oordeel. Excuses worden in deze zaken dus gebruikt als een bewijsmiddel. Na nadere analyse van de uitspraken bleek echter dat het niet zo was dat in deze gevallen de gehele beslissing wordt gedragen door het feit dat excuses zijn aangeboden. In de gevonden uitspraken werden steeds ook andere argumenten genoemd en de aangeboden excuses werden veelal gebruikt ter onderbouwing van een enkel onderdeel van het gehele feitencomplex. Om dit te onderbouwen heb ik de structuren van de argumentatie in de gevonden uitspraken inzichtelijk gemaakt door het opstellen van argumentatieschema's.³¹ Bij het opstellen van de argumentatieschema's is aansluiting gezocht bij de argumentatietheorie van Van Eemeren en Grootendorst.³² Hierin wordt onderscheid gemaakt tussen verschil-

29 Zie uitgebreid Wijntjens 2020, hfdst. 8.

30 Zie uitgebreid Wijntjens 2020, par. 5.2.1.1.

31 Hierbij dient te worden opgemerkt dat de argumentatieschema's een vereenvoudigde weergave zijn van de door de rechter gebruikte argumentatie ter onderbouwing van zijn oordeel. Bij het opstellen van de schema's is specifiek gekeken naar het gedeelte van de motivering dat betrekking had op het aanbieden van excuses. Over het algemeen wordt in een uitspraak een veelvoud van feiten, omstandigheden en argumenten genoemd die alle in de beoordeling worden betrokken. Het is ondoenlijk binnen een argumentatieschema volledig recht te doen aan deze complexiteit. De argumentatieschema's laten echter wel goed zien hoe de argumentatie van de rechter in grote lijnen verloopt en welk gewicht excuses hierbij in de schaal leggen.

32 F.H. van Eemeren & R. Grootendorst, 'Wat hebben we ook weer? Argumentatieleer!', in: W. van Paassen, G. Rijlaarsdam & F. Zwitserlood (red.), *Taalbeschouwing voor gevorderden. Een bundel artikelen over enkele aspecten van de taalbeschouwing in de bovenbouw*, 's-Hertogenbosch: Malmberg 1995, p. 112-130 en F.H. van Eemeren, R. Grootendorst & A.F. Snoeck Henkemans, *Argumentatie. Inleiding in het analyseren, beoordelen en presenteren van betogen*, Groningen: Wolters-Noordhoff 1995. Zie voor een toepassing van deze algemene argumentatietheorie op rechterlijke motiveringen H.J. Plug, 'De structuur van rechterlijke uitspraken: contra-argumentatie in de motivering', *Tijdschrift voor Taalbeheersing* (16) 1994, afl. 2, p. 85-93 en E.T. Feteris e.a., *Gewogen oordelen: essays over argumentatie en recht: bijdragen aan het zesde symposium juridische argumentatie 24 juni 2011*, Den Haag: Boom Juridische uitgevers 2012. Zie ook R.S. Summers & M. Taruffo, 'Interpretation and Comparative Analysis', in: D.N. McCormick & R.S. Summers (red.), *Interpreting Statutes*, Aldershot: Dartmouth Publishing Company 1991, waarin verschillende typeringen worden gegeven van rechterlijke motiveringen die in de omschrijving lijken op de door Van Eemeren e.a. getypeerde vormen van argumentatie.

lende vormen van argumentatie. In de gevonden uitspraken waren vormen van onderschik-kende, meervoudige en nevenschikkende argumentatie terug te vinden. Het kwam ook voor dat een combinatie werd gebruikt van deze argumentatievormen.³³

Een voorbeeld van nevenschikkende argumentatie is bijvoorbeeld terug te vinden in een aantal seksueelmisbruikzaken, waarin een excuusbrief van de gedaagde was ingebracht als bewijsmiddel.³⁴ Van nevenschikkende argumentatie is sprake wanneer de argumenten die samen ter ondersteuning van een standpunt worden aangevoerd onlosmakelijk bij elkaar horen. In een van de seksueelmisbruikzaken komt de rechtbank tot het oordeel dat de eiseres is geslaagd in het bewijs van de stelling dat zij door gedaagde seksueel was misbruikt.³⁵ De rechtbank wijst daarbij op verschillende bewijsmiddelen, waaronder een brief waarin gedaagde excuses heeft aangeboden:

‘Verder staat vast dat [gedaagde] aan [eiseres] na de confrontatie door de moeder van [eiseres] een briefje heeft geschreven, met de tekst: “De dingen die vanuit mij ten aanzien van jou zijn gebeurd hadden nooit mogen plaatsvinden. Ik ben degene die te ver is gegaan. Ik heb oprecht spijt van datgene wat er is gebeurd en vraag zo mogelijk daarvoor mijn excuses.”’³⁶

Vervolgens komt de rechter tot de conclusie dat de bewijsmiddelen, in onderlinge samenhang bezien, voldoende sterke aanwijzingen zijn dat sprake is geweest van seksueel misbruik:

‘De verklaringen van de getuigen [eiseres], [moeder van eiseres/getuige 1] en [betrokkene/getuige 2], de brief van 12 november 1992 van [eiseres] en de brief van [gedaagde] aan [eiseres], zijn, in onderlinge samenhang bezien, voldoende sterke aanwijzingen dat sprake is geweest van seksueel misbruik van [eiseres] door [gedaagde].’³⁷

Dat hier sprake is van nevenschikkende argumentatie kan worden afgeleid uit de door de rechter gebruikte zinsnede ‘in onderlinge samenhang bezien’. Het lijkt erop dat de rechter

33 Bij het lezen van de argumentatieschema’s dient de lezer het volgende in acht te nemen: (1) in de groene kaders staan de standpunten (het oordeel van de rechter) opgenomen, terwijl in de blauwe kaders het argument of de argumenten staan opgenomen die dit standpunt onderbouwen; (2) de pijlen in het schema laten het onderlinge verband van de argumenten zien.

34 Rb. Arnhem 14 maart 2012, ECLI:NL:RBARN:2012:BW0516 en Hof Arnhem 16 oktober 2001, ECLI:NL:GHARN:2001:AD4916.

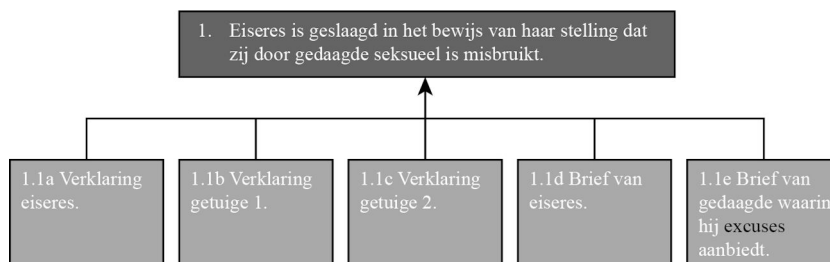
35 Rb. Arnhem 14 maart 2012, ECLI:NL:RBARN:2012:BW0516.

36 Rb. Arnhem 14 maart 2012, ECLI:NL:RBARN:2012:BW0516, r.o. 2.7.

37 Rb. Arnhem 14 maart 2012, ECLI:NL:RBARN:2012:BW0516, r.o. 2.9.

hier bedoelt dat alleen de bewijsmiddelen tezamen zijn oordeel kunnen dragen. De bewijsmiddelen op zichzelf kunnen dan geen zelfstandige ondersteuning zijn voor het oordeel.

Figuur 1 Argumentatieschema bij Rb. Arnhem 14 maart 2012, ECLI:NL:RBARN:2012:BW0516



Een ander voorbeeld van nevenschikkende argumentatie is de situatie waarin de rechter in zijn beoordeling een tegenargument noemt en dit tegenargument vervolgens verworpt.³⁸ Dat tegenargument was door de in het ongelijk gestelde partij opgeworpen en pleit dus tegen het uiteindelijke oordeel van de rechter. We spreken hier van verwijzing naar contra-argumentatie.³⁹ Wanneer een standpunt wordt onderbouwd door een argument en een verworpen contra-argument is de samenhang tussen de argumenten nevenschikkend omdat de verwerping van het contra-argument geen zelfstandige ondersteuning kan zijn voor het standpunt.⁴⁰ In de jurisprudentieanalyse is een aantal uitspraken gevonden waarin de rechter een door de in het ongelijk gestelde partij genoemd contra-argument verworpt door te verwijzen naar door die partij aangeboden excuses.⁴¹

³⁸ In een civiele procedure is het niet ongevoerd dat de rechter in de motivering van zijn uiteindelijke beslissing verwijst naar door de procespartijen naar voren gebrachte en bekritiseerde argumentatie. De motiveringsplicht zoals neergelegd in art. 30, 30p, 230 en 287 Wetboek van Burgerlijke Rechtsvordering (Rv) brengt mee dat de rechter dit in sommige gevallen zelfs moet doen. Het oordeel van de rechter moet ten minste zodanig gemotiveerd worden dat het voldoende inzicht geeft in de gedachtegang die daaraan ten grondslag ligt teneinde de beslissing voor partijen (met name de in het ongelijk gestelde partij) en derden (zoals de hogere rechter) controleerbaar en aanvaardbaar te maken.

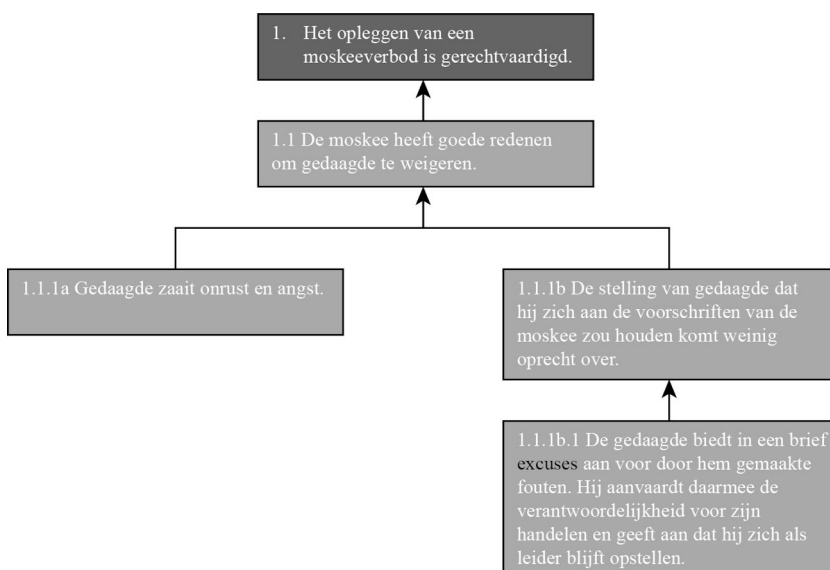
³⁹ Zie uitgebreid over contra-argumentatie in de rechterlijke motivering Plug 1994. Zij onderscheidt nog twee andere situaties waarin in een rechterlijk betoog kan worden verwezen naar contra-argumentatie: de rechter noemt een contra-argument dat tegen zijn *argumentatie* pleit en verworpt vervolgens dit contra-argument en de rechter erkent een contra-argument, maar houdt vol dat zijn standpunt verdedigbaar is. Op deze situaties wordt hier verder niet ingegaan aangezien hiervan geen voorbeelden zijn gevonden in de jurisprudentieanalyse.

⁴⁰ Plug 1994, p. 87.

⁴¹ Zie naast de hieronder uitgewerkte voorbeelden: Hof Den Haag 25 februari 2014, ECLI:NL:GHDHA:2014:1524; Rb. Noord-Nederland 24 februari 2016, ECLI:NL:RBNNE:2016:716; Rb. Rotterdam 26 september 2016, ECLI:NL:RBROT:2016:7329.

Een voorbeeld hiervan is te vinden in een zaak waarin – kort gezegd – een stichting vordert om aan een (ex-)imam een verbod op te leggen om nog in een door haar gehouden moskee te komen.⁴² Om te beoordelen of een dergelijk verbod op zijn plaats is, bekijkt de voorzieningenrechter of de stichting een goede reden heeft om de gedaagde te weigeren. De stichting heeft als een dergelijke reden aangevoerd dat de gedaagde onrust en angst zaait binnen de geloofsgemeenschap en dat de anderen onder leiding van gedaagde daaraan deelnemen. Gedaagde betwist dit door te stellen dat hij zich wel degelijk aan de voorwaarden voor de toegang tot de moskee houdt (contra-argument). De rechtbank oordeelt dat dit weinig oprecht overkomt en wijst daarbij – onder meer – op een brief waarin gedaagde excuses aanbiedt voor door hem gemaakte fouten bij zijn pogingen de eenheid in de geloofsgemeenschap te herstellen. Volgens de rechtbank aanvaardt hij daarmee de verantwoordelijkheid voor zijn handelen en geeft hij aan dat hij zich als leider blijft opstellen. Het tegenargument van gedaagde wordt aldus, onder verwijzing naar door hem aangeboden excuses, verworpen.

Figuur 2 Argumentatieschema bij Rb. 's-Hertogenbosch 25 februari 2009, ECLI:NL:RBSHE: 2009:BH4029



De analyse van de uitspraken laat zien dat voor beide onderzochte procedures geldt dat uitspraken zijn gevonden waarin naar excuses wordt verwezen ter onderbouwing van feiten

⁴² Rb. 's-Hertogenbosch 25 februari 2009, ECLI:NL:RBSHE:2009:BH4029.

waarop het oordeel is gebaseerd. De betekenis van excuses als bewijsmiddel is in deze gevallen echter niet doorslaggevend. De civiele rechter of het medisch tuchtcollege verwijst behalve naar de aangeboden excuses namelijk ook steeds naar andere argumenten ter onderbouwing van een bepaald feit waarop het oordeel wordt gebaseerd. Er zijn geen uitspraken gevonden waarin een uitspraak enkel wordt onderbouwd door te verwijzen naar aangeboden excuses. Het onderzoek laat dus zien dat excuses bewijskracht kunnen hebben. Hoewel dit op het eerste gezicht een stevige juridische belemmering lijkt, heb ik door middel van de gestructureerde jurisprudentieanalyse de belangrijke kanttekening kunnen maken dat rechters hier erg genuanceerd mee omgaan en in geen enkel geval doorslaggevende betekenis toekennen aan aangeboden excuses.

4 RESULTATEN

Een belangrijke onderzoeksbevinding is dat excuses op verschillende manieren een rol spelen in de geanalyseerde rechtspraak. Het Nederlandse wetenschappelijke debat was tot dusver met name gericht op twee vragen.⁴³ De eerste is of excuses kunnen leiden tot het erkennen van aansprakelijkheid.⁴⁴ Wat opvalt, is dat in deze discussie sprake is van zwart-witdenken: of auteurs betogen dat het aanbieden van excuses gelijk kan staan aan het erkennen van aansprakelijkheid en adviseren schadeveroorzakers hiermee terughoudend te zijn, of auteurs betogen dat excuses geen erkenning van aansprakelijkheid kunnen of zouden moeten kunnen inhouden. De tweede vraag die in de literatuur aan bod komt, is of een rechtsvordering tot het aanbieden van excuses kan worden ingesteld (door eisers) of toegewezen (door rechters).⁴⁵ Ook deze wetenschappelijke discussie kenmerkt zich door een scherpe tegenstelling: auteurs betogen ofwel dat excuses niet kunnen worden gevorderd/toegewezen ofwel dat deze mogelijkheid wel bestaat of zou moeten bestaan. Uit de uitgevoerde gestructureerde jurisprudentieanalyse blijkt dat beide binaire discussies geen recht doen aan de complexiteit van de werkelijkheid. Excuses blijken op veel meer manieren een rol te spelen in de geanalyseerde rechtspraak. Daarnaast blijken excuses lang niet altijd een duidelijke doorslaggevende betekenis te hebben. De jurisprudentieanalyse

43 Zie uitgebreid over de stand van het Nederlandse wetenschappelijke debat ten aanzien van deze vragen (onder verwijzing naar literatuur): Wijntjens 2020, hfdst. 3.

44 Zie o.a.: Zwart-Hink 2017; Hoogenboezem & Hees 2014; J.L. Smeehuijzen & A.J. Akkermans, 'Medische aansprakelijkheid: over grote problemen, haalbare verbeteringen en overschatte revoluties', in: *Ontwikkelingen rond medische aansprakelijkheid* (Preadviezen Vereniging voor Gezondheidsrecht), Den Haag: Sdu Uitgevers 2013, p. 42 en L.A.B.M. Wijntjens, 'Excuses en aansprakelijkheid: In hoeverre en op welke wijze beïnvloeden excuses het oordeel van de rechter?', *WPNR* (147) 2016, afl. 7092, p. 82-88.

45 Zie o.a.: J.E. Hulst e.a., 'Excuses in het privaatrecht', *WPNR* (139) 2008, afl. 6772, p. 778-784 en G. van Dijck, 'Hebben afgedwongen excuses zin?', *NTBR* 2017, afl. 10, p. 298-302.

laat zien dat er zes modaliteiten zijn te onderscheiden in de wijze waarop excuses een rol spelen in de civiele rechtspraak:⁴⁶

1. De rechter gebruikt aangeboden excuses om bepaalde feiten vast te stellen (excuses als bewijsmiddel).
2. De rechter beoordeelt of aangeboden excuses een erkenning van aansprakelijkheid (of een andere juridische kwalificatie) inhouden.
3. De rechter gebruikt het feit dat excuses wel/niet zijn aangeboden als een omstandigheid bij de invulling van een open norm.
4. De rechter beoordeelt of er sprake is van een geschonden verplichting tot het aanbieden van excuses (in welk geval de schending van deze verplichting zou leiden tot aansprakelijkheid).
5. De rechter verwijst naar excuses in een obiter dictum.
6. De rechter beoordeelt een vordering tot het aanbieden van excuses.

In tabel 2 is te zien hoe vaak deze categorieën voorkomen. Opvallend is dat in de civiele rechtspraak en in de medische tuchtrechtspraak dezelfde categorieën naar voren komen. Wel verschilt de mate waarin en de wijze waarop aan die categorieën invulling wordt gegeven. Deze verschillen kunnen grotendeels worden verklaard door te kijken naar de doelen van de procedures.⁴⁷

Tabel 2 Steekproefgegevens en aantal gevonden uitspraken per categorie⁴⁸

Categorie	Civiele rechtspraak		Medische tuchtrechtspraak	
	Aantal uitspraken	Percentage van het totaal aantal relevante treffers	Aantal uitspraken	Percentage van het totaal aantal relevante treffers
Relevante treffers	1254		644	
– Wel excuses aangeboden	915	72,97%	563	87,42%
– Geen excuses aangeboden	358	28,63%	96	14,91%
De wel/niet aangeboden excuses worden benoemd in het oordeel	258	20,57%	310	48,14%
1. Excuses en het vaststellen van feiten	45	3,61%	4	0,62 %
2. Aanbieden van excuses en het erkennen van aansprakelijkheid/tuchtrechtelijk verwijtbaar handelen	11	0,88%	13	2,02%

⁴⁶ Zie hierover uitgebreid Wijntjens 2020, hfdst. 5.

⁴⁷ Zie hierover Wijntjens 2020, hfdst. 8.

⁴⁸ Wijntjens 2020, p. 226.

Categorie	Civiele rechtspraak		Medische tuchtrechtspraak	
	Aantal uitspraken	Percentage van het totaal aantal relevante treffers	Aantal uitspraken	Percentage van het totaal aantal relevante treffers
3. Excuses als omstandigheid bij de invulling van een open norm	160	12,76%	166	25,78%
4. Verplichting tot het aanbieden van excuses	1	0,08%	60	9,32%
5. Obiter dictum over het aanbieden van excuses	1	0,08%	3	0,47%
6. Excuses gevorderd	175	13,96%	1	0,16%

De uitspraken die onder de zes overkoepelende categorieën werden geplaatst zijn nader geanalyseerd, waarbij ik gebruik heb gemaakt van de in paragraaf 3.3.3 besproken analyse-technieken. Uit deze analyse bleek dat de onderzochte procedures inderdaad, zoals in de literatuur wordt gesteld, juridische belemmeringen kunnen opwerpen voor het aanbieden van excuses (categorie 1 en 2). Na nadere analyse bleek echter wel dat deze belemmeringen sterk genuanceerd moeten worden. Slechts in zeer zeldzame gevallen en onder zeer specifieke omstandigheden kunnen excuses een erkenning van civielrechtelijke aansprakelijkheid impliceren.⁴⁹ Daarnaast heb ik kunnen constateren dat excuses zowel in civiele procedure als in de medische tuchtprocedure kunnen worden gebruikt als bewijsmiddel. Rechters blijken echter aan excuses als bewijsmiddel in geen enkel geval doorslaggevende betekenis toe te kennen. Op basis hiervan heb ik geconcludeerd dat de nadruk die wordt gelegd op de juridische risico's van het aanbieden van excuses onterecht is.⁵⁰ Veel vaker blijken excuses door de (tucht)rechter juist in het voordeel van aanbieder te worden meegewogen (categorie 3) of worden excuses in de onderzochte rechtspraak juist gestimuleerd of gefaciliteerd (categorie 4, 5 en 6). De uitkomsten van de gestructureerde rechtspraakanalyse brengen daarmee een belangrijke nuancering in het wetenschappelijke debat over de rol van excuses in de rechtspraak.

5 INZICHTEN

Wat heeft de gestructureerde aanpak mij opgeleverd? Het systematisch analyseren van de rechtspraak heeft ertoe geleid dat ik een fijnmazig overzicht heb kunnen creëren van de wijze waarop excuses een rol spelen in de civiele rechtspraak en de medische tuchtrechtspraak. Dat heeft ervoor gezorgd dat ik aannames die werden gedaan in de literatuur over

⁴⁹ Zie hierover uitgebreid Wijntjens 2020, par. 5.3.1.

⁵⁰ Wijntjens 2020, p. 305.

de rol van excuses in de (civiele) rechtspraak heb kunnen weerleggen of nuanceren (*myth busting*). Hieronder deel ik enkele inzichten die ik gedurende het onderzoeksproces heb opgedaan over het verrichten van een gestructureerde rechtspraakanalyse.

5.1 *Onderzoeksdesign*

Zoals elk onderzoek begint een gestructureerde rechtspraakanalyse met het formuleren van een goede onderzoeksvraag. Hierbij heb ik mij steeds afgevraagd: wat wil ik nu eigenlijk te weten komen en hoe verhoudt dit zich tot eerder gedaan onderzoek? De methode dient logischerwijs voort te vloeien uit de onderzoeksvraag. In mijn onderzoek ging het om een beschrijvende onderzoeksvraag. Ik was namelijk geïnteresseerd in de wijze waarop excuses een rol spelen in de civiele en medische tuchtrechtspraak. In de literatuur werden hierover al eerder allerlei standpunten ingenomen, maar een uitgebreid rechtspraakonderzoek was nog niet uitgevoerd. Ook bij het coderen van de uitspraken heb ik aansluiting gezocht bij de theorie. Eerst heb ik door middel van een literatuuronderzoek het wetenschappelijke debat omtrent de rol van excuses in het civiele recht in kaart gebracht. Bij het opstellen van de (initiële) codeerschema's is bij deze theorie aangesloten. Dit gaf richting en houvast bij het verzamelen van de data. Meer concreet werd op deze manier helder wat precies het doel was van de rechtspraakanalyse en welke informatie relevant was om uit de gevonden uitspraken te halen. Van belang is dat niet alle gegevens uit de gevonden uitspraken gecodeerd hoeven te worden. Wanneer je bijvoorbeeld wilt onderzoeken of vrouwelijke rechters anders omgaan met excuses vergeleken met mannelijke rechters dan dien je te coderen wat het geslacht is van de rechter die de uitspraak heeft gedaan. Ben je niet geïnteresseerd in genderverschillen dan is het ook niet nodig om dit te coderen. In paragraaf 3.3 van deze bijdrage heb ik daarnaast laten zien dat verschillende technieken kunnen worden gebruikt bij het gestructureerd analyseren van grote hoeveelheden uitspraken. Uitspraken kunnen niet alleen worden gecategoriseerd maar het is ook mogelijk om verbanden en argumenten in uitspraken bloot te leggen en met elkaar te vergelijken. Ook de keuze voor de analyse-techniek hangt af van het doel van het onderzoek.

5.2 *Is de informatie die je nodig hebt ook daadwerkelijk af te leiden uit de uitspraak?*

Bij het verrichten van een gestructureerde jurisprudentieanalyse dient rekening te worden gehouden met het feit dat mogelijk niet alle benodigde informatie is af te leiden uit een uitspraak. In eerste instantie was ik bijvoorbeeld geïnteresseerd in de wijze waarop rechters (onbewust) worden beïnvloed door het feit dat een procespartij excuses had aangeboden. Deze vraag bleek echter onmogelijk te kunnen worden beantwoord door het verrichten

van een jurisprudentieanalyse. Uit de uitspraken kon namelijk niet worden afgeleid of een procespartij excuses had aangeboden en of de rechter hiervan op de hoogte was. Het is immers mogelijk dat het feit dat er excuses waren aangeboden wel bij de rechter bekend was maar dat dit niet in de uitspraak is vastgelegd (bijvoorbeeld omdat de rechter/griffier⁵¹ het niet van belang achtte voor de beslissing). Voor het beantwoorden van de bovengenoemde vraag is het verrichten van een vignettenstudie met als respondenten (tucht)rechters dus veel geschikter.⁵²

5.3 Samenwerking

Het analyseren van grote hoeveelheden uitspraken is een tijdrovende klus. Het bekijken en coderen van honderden uitspraken kan gerust monnikenwerk worden genoemd. Binnen een promotietraject heeft de promovendus veelal de tijd om een dergelijke tijdrovende klus aan te gaan. Binnen andere onderzoeksprojecten kan het echter noodzakelijk zijn om het coderen van de uitspraken door meerdere personen te laten uitvoeren.⁵³ Afstemming en samenwerking is dan cruciaal. Maar ook in onderzoeksprojecten waarbij het coderen en analyseren van de data door één persoon wordt uitgevoerd verdient het aanbeveling om samenwerking met anderen op te zoeken. In mijn onderzoeksproject heb ik bijvoorbeeld een groepje masterstudenten bereid gevonden om mee te helpen bij het uittesten van de codeertabellen. Teneinde de betrouwbaarheid en de objectiviteit van het jurisprudentieonderzoek te waarborgen zijn 75 uitspraken gecodeerd door zowel mijzelf als deze masterstudenten. Nadat een aantal uitspraken was gecodeerd, hebben wij steeds een vergelijking gemaakt tussen de door de verschillende codeurs toegepaste codes. Eventuele verschillen werden besproken. Soms werd daarbij alsnog overeenstemming bereikt over de toepasselijke code. In andere gevallen was het nodig om de codeertabel aan te scherpen. De cyclus van het coderen van een aantal uitspraken en het bespreken daarvan werd herhaald totdat er een grote mate van overeenstemming bestond in de door de verschillende codeurs toegepaste codes. Op deze manier heb ik kunnen waarborgen dat de codeerschema's navolgbaar waren en dientengevolge ook voor andere onderzoekers bruikbaar zijn. Overigens was deze werkwijze niet alleen voor mij als onderzoeker nuttig, ook de deelnemende studenten gaven aan veel geleerd te hebben van het deelnemen aan het onderzoek. Deze manier van werken geeft aan studenten de mogelijkheid een kijkje in de keuken te nemen bij het ver-

51 N.L. Holvast, 'Wie neemt daadwerkelijk de beslissingen binnen de rechtbanken? De betrokkenheid van juridisch ondersteunend personeel in de rechterlijke besluitvorming', *NJB* 2017 (33), p. 2379-2385.

52 Zie bijvoorbeeld: J.K. Robbenolt & R.M. Lawless, 'Bankrupt apologies', *Journal of Empirical Legal Studies* (10) 2013, afl. 4, p. 771-796 en J.J. Rachlinski, C. Guthrie & A.J. Wistrich, 'Contrition in the courtroom: Do apologies affect adjudication?', *Cornell Law Review* (98) 2013, afl. 5, p. 1189-1243.

53 Zie bijvoorbeeld de bijdrage van Goossens, Bollen & Verbeke elders in deze bundel.

LIANNE WIJNTJENS

richten van wetenschappelijk onderzoek. Daarnaast geldt dat niet iedere rechtswetenschapper kennis heeft van kwalitatieve of kwantitatieve methoden die kunnen worden gebruikt voor het analyseren van grote hoeveelheden uitspraken. In dat kader kan het nuttig zijn om advies in te winnen van experts binnen andere wetenschapsdisciplines. Bij mijn promotieonderzoek gold bijvoorbeeld dat een van de promotoren een methodespecialist was verbonden aan Tilburg School of Social and Behavioral Sciences. De verbinding zoeken met wetenschappers van andere wetenschapsdisciplines kan de kwaliteit van de rechtspraak-analyse mijns inziens verhogen.

3 MOGELIJKHEDEN EN BEPERKINGEN BIJ DE KWANTITATIEVE ANALYSE VAN JURISPRUDENTIE VOOR DE BESTUDERING VAN GELDEND RECHT

*Pieter Wolters**

1 INLEIDING

‘Traditioneel’ juridisch of ‘juridisch-dogmatisch’ onderzoek bestudeert in de eerste plaats¹ het geldende recht. De methode is voornamelijk kwalitatief en gaat uit van een zekere rangorde van bronnen. Het geldende recht blijkt in de eerste plaats uit de wet of een gezaghebbende² uitspraak van de hoogste rechter. Als zo’n bron niet beschikbaar is, kan de onderzoeker onder andere aansluiten bij uitspraken en wetsbepalingen die zien op vergelijkbare situaties of zich baseren op (de heersende mening in) de literatuur. Ten slotte kan ook ‘lagere’ of ‘feitenrechtspraak’ een rol spelen bij het duiden van het geldende recht. Hoewel deze uitspraken hun gezag verliezen als de hoogste rechter een nieuwe, andersoortige uitspraak doet, bieden zij inzicht in de wijze waarop een bepaalde kwestie, ten minste voorlopig, wordt beslist.

Dit traditionele kwalitatieve juridische onderzoek leidt echter niet altijd tot het gewenste inzicht. De geanalyseerde uitspraken zien immers slechts op een bepaalde casus. Het is niet duidelijk of en in hoeverre hetzelfde resultaat ook geldt in andere gevallen. Dit speelt

* Mr. P.T.J. (Pieter) Wolters is universitair hoofddocent burgerlijk recht en onderzoeker bij het Onderzoekcentrum Onderneming & Recht van de Radboud Universiteit.

1 Deze beschrijving is uiteraard sterk vereenvoudigd. Na een beschrijving van het geldende recht volgt in veel situaties een oordeel over het wenselijke recht. Deze stap blijft in deze bijdrage buiten beschouwing. Zie voor een uitgebreidere omschrijving van juridisch-dogmatisch onderzoek bijvoorbeeld G. van Dijck, M. Snel & T. van Goolen, *Methoden van rechtswetenschappelijk onderzoek*, Den Haag: Boom juridisch 2018, p. 84-85; R.A.J. van Gestel, ‘Revitalisering van juridisch-dogmatisch onderzoek’, *NJB* 2021, afl. 4, p. 288-290.

2 Een uitspraak is bijvoorbeeld minder gezaghebbend als zij relatief oud is, ziet op andere omstandigheden of als er veel kritiek bestaat vanuit de literatuur. Hetzelfde geldt overigens voor wetsbepalingen.

in het bijzonder bij zogenoemde ‘open normen’ zoals de redelijkheid en billijkheid.³ De precieze rechtsgevolgen van deze open normen hangen af van ‘de’ of zelfs ‘alle’ omstandigheden van het geval. Omdat de omstandigheden per geval verschillen, zijn de rechtsgevolgen van een open norm moeilijk te voorspellen. Hoewel het in sommige gevallen mogelijk is om een ‘harde subregel’ te formuleren of uit de jurisprudentie te distilleren, worden deze normen in de regel door de Hoge Raad verworpen.⁴ De feitenrechtspraak is bovendien, zeker op het eerste gezicht, lang niet altijd duidelijk of consequent. Het is niet altijd duidelijk welke omstandigheden de doorslag geven. Vergelijkbare feitencomplexen kunnen bovendien gepaard gaan met uiteenlopende conclusies.

In deze gevallen is het met traditioneel kwalitatief juridisch onderzoek niet goed mogelijk om het geldende recht te duiden, anders dan door te constateren dat het afhangt van de omstandigheden van het concrete geval. Hoewel de werking van open normen in duidelijke gevallen kan worden voorspeld, blijft er een grote hoeveelheid moeilijke gevallen over waarin de genoemde bronnen niet leiden tot een helder antwoord. Kwantitatieve analyse van jurisprudentie kan in deze gevallen een uitkomst bieden. Zij kan leiden tot een cijfermatig begrip van de rol die de verschillende relevante omstandigheden in de rechtspraak hebben en daarmee, bij gebrek aan richtinggevende bronnen met meer gezag, in de werking van het geldende recht. Tegelijkertijd kent deze methode ook verschillende beperkingen en valkuilen die haar succes in de weg staan.

Deze bijdrage is bedoeld om inzicht te verschaffen in de mogelijkheden van kwantitatieve analyse van jurisprudentie voor de bestudering van geldend recht. Zij beoogt in het bijzonder duidelijker te maken in welke situaties de methode wel of niet geschikt is. Ik ga daarom met name in op de beperkingen en valkuilen van deze methode.⁵

Ik put in deze bijdrage in het bijzonder uit mijn ervaringen bij mijn onderzoek naar de matiging van boetebedingen op grond van artikel 6:94 lid 1 Burgerlijk Wetboek (BW)⁶ en

3 Zie over dit begrip onder andere P.T.J. Wolters, *Alle omstandigheden van het geval. Een onderzoek naar de omstandigheden die de werking van de redelijkheid en billijkheid beïnvloeden* (diss. Nijmegen) (Serie Onderneming en Recht deel 77), Deventer: Kluwer 2013[a], p. 2-3, met verwijzing naar verdere literatuur.

4 Zie hierover uitgebreid P.T.J. Wolters, ‘Regels over de werking van de redelijkheid en billijkheid? Een analyse van de rechtspraak van de Hoge Raad’, *MvV* 2013, afl. 9, p. 249-255.

5 Zie over deze methode tevens de bijdragen in deze bundel van Goossens, Bollen & Verbeke; Vols.

6 Dit onderzoek ziet naast de matiging van boetebedingen ook op de vernietiging van boetebedingen in algemene voorwaarden op grond van art. 6:233 onder a BW. Het is gepubliceerd in P.T.J. Wolters, ‘De invloed van de verhouding tussen schade en boete op de matiging van boetebedingen. Een empirische analyse’, *AV&S* 2016, afl. 3, p. 109-117; P.T.J. Wolters, ‘De matiging van boetebedingen bij koopovereenkomsten ten aanzien van onroerend goed. Een empirische analyse van de invloed van het gebruikelijke karakter’, *WPNR* 2017[a], afl. 7136, p. 122-128; P.T.J. Wolters, ‘De toetsing van boetebedingen in de algemene voorwaarden. Consistentie in de lagere rechtspraak?’, *RMThemis* 2017, afl. 3, p. 116-128; P.T.J. Wolters,

3 *MOGELIJKHEDEN EN BEPERKINGEN BIJ DE KWANTITATIEVE ANALYSE VAN
JURISPRUDENTIE VOOR DE BESTUDERING VAN GELDEND RECHT*

(in mindere mate) naar de beperking van de verjaring van vorderingen tot schadevergoeding als gevolg van blootstelling aan asbest.⁷ Ik begin met een nadere beschrijving van de rol van kwantitatieve analyse voor de bestudering van het geldende recht (par. 2). Hierna bespreek ik (de beperkingen bij) de selectie van de uitspraken (par. 3), de codering (par. 4) en de analyse van de resultaten (par. 5). Ik concludeer met een reflectie over deze beperkingen. Ik maak hierbij ook een vergelijking met traditioneel juridisch onderzoek zonder kwantitatieve analyse van jurisprudentie (par. 6).

2 **BESCHRIJVING METHODE**

Kwantitatieve analyse van jurisprudentie voor de bestudering van het geldende recht gaat in de kern om tellen. Zij telt hoe vaak een rechter tot een bepaalde conclusie komt en welke omstandigheden tot deze conclusies leiden. Deze getallen worden vervolgens gebruikt om statistisch significante conclusies te trekken over ‘het geldende recht’.

De methode gaat er hiermee van uit dat het recht kan worden afgeleid uit de feitenrechtspraak. Deze beslissingen zijn voor rechtszoekenden immers van groot belang. Een rechtszoekende wil weten welke rechtsgevolgen er voor hem, uitgaande van zijn omstandigheden, gelden. De feitenrechtspraak is voor rechtszoekenden misschien zelfs wel van groter belang dan het geldende recht zoals uitgelegd in een abstracte uitspraak van de hoogste rechter.⁸ Hiermee is echter niet gezegd dat de methode slechts past in een traditie van *legal realism*. Het nut van de methode is niet beperkt tot het bestuderen van de invloed van niet-juridische factoren zoals het ontbijt van de rechter.⁹ Integendeel, de methode kan

‘De vernietiging en matiging van boetebedingen in huurovereenkomsten ten aanzien van onroerende zaken in de feitenrechtspraak’, *WR* 2017, afl. 10, p. 571-577; P.T.J. Wolters, ‘De boetebedingen van de “hebberige” schuldeiser in het licht van de vernietiging en matiging van contractuele boetes’, *ORP* 2018, afl. 1, p. 39-49.

7 Dit onderzoek is gepubliceerd in P.T.J. Wolters, ‘Het vaste gewicht van de gezichtspunten van Van Hese/De Schelde’, *AV&S* 2015[a], afl. 1, p. 15-25; P.T.J. Wolters, ‘Naschrift. Feiten, omstandigheden, waarderingen en totaalafwegingen’, *AV&S* 2015[b], afl. 4, p. 143-144.

8 Een gebruiker van algemene voorwaarden zal bijvoorbeeld eerder geïnteresseerd zijn in de wijze waarop feitenrechtters in Nederland omgaan met de ambtshalve toetsing van algemene voorwaarden en de gevolgen van de vernietiging van een beding dan in de uitspraken van het Hof van Justitie die hiervoor uiteindelijk leidend zijn. Vergelijk bijvoorbeeld Redactieraad van het LOVCK&T, *Ambtshalve toetsing III*, 2018. Het is overigens ook mogelijk om ‘het recht’ nog meer te betrekken op hetgeen er in de werkelijkheid gebeurt. Rechtssociologie bestudeert het recht bijvoorbeeld (onder andere) als set van regels die het handelen van partijen bepalen, zelfs als deze afwijken van het geldende recht volgens de juridisch-dogmatische methode. Vergelijk bijvoorbeeld R. Cotterrell, ‘Sociology of law’, in: D.S. Clark (red.), *Encyclopedia of Law and Society: American and Global Perspectives*, Thousand Oaks, CA: Sage 2007, p. 1413-1418.

9 Geparafraseerd van: ‘Law is what the judge had for breakfast’. Zie hierover D. Priel, ‘Law Is What the Judge Had for Breakfast: A Brief History of an Unpalatable Idea’, *Buffalo Law Review* (68) 2020, afl. 3, p. 899-930. Zie over de mogelijke invloed van de maaltijd van een rechter op voorwaardelijke vrijlatingen (‘parole’)

ook worden gebruikt om de invloed van factoren te bestuderen die *wel* van belang zijn in een juridisch-dogmatische bestudering van het recht. Een voorwaarde blijft dan echter wel dat de feitenrechtspraak ook volgens deze traditionele methode een belangrijke rechtsbron is. De kwantitatieve analyse van jurisprudentie voor de bestudering van het geldende recht is hierdoor met name van belang voor open normen en andere vraagstukken waarover de wet en hoogste rechter geen duidelijk antwoord geven (zie ook par. 1).

Mijn onderzoek naar de matiging van boetebedingen biedt een voorbeeld. Een boetebeding is op grond van artikel 6:91 BW (kortgezegd) een contractuele clause die bepaalt dat de schuldenaar een 'boete' moet betalen indien hij in de nakoming van de hoofdverbintenis tekortschiet. De rechter kan deze contractuele boete op grond van artikel 6:94 lid 1 BW matigen als de billijkheid dit klaarblijkelijk eist. Het ligt hierbij voor de hand om veel gewicht toe te kennen aan de verhouding tussen de werkelijke schade en de boete.¹⁰ Uit HR 27 april 2007, NJ 2007/262 (*Intrahof/Bart Smit*) blijkt echter dat een rechter niet alleen hierop moet letten, maar ook op verschillende andere omstandigheden. De precieze invloed van de verschillende omstandigheden, waaronder de verhouding tussen schade en boete, blijft echter onduidelijk. Hij blijkt evenmin uit de verschillende individuele uitspraken van feitenrechters. Hoewel de uitspraken over het algemeen aangeven welke omstandigheden er in het concrete geval van belang zijn, maken zij lang niet altijd duidelijk welke factoren precies de doorslag geven. Het blijft in de woorden van Barendrecht in veel gevallen bij 'een machtswoord gekoppeld aan de vaststelling van de feiten'.¹¹

Een kwantitatieve analyse van jurisprudentie maakt het mogelijk om meer te weten te komen over de invloed van de verschillende factoren. Door bij iedere uitspraak te noteren (coderen) of de boete is gematigd en van welke omstandigheden er sprake was, is het mogelijk om te ontdekken welke omstandigheden significant vaak leiden tot, of ten minste gepaard gaan met (zie ook par. 5.2), de matiging van de boete.

Ik heb mij hierbij beperkt tot enkele omstandigheden die juridisch relevant zijn voor de werking van artikel 6:94 lid 1 BW en daarmee onderdeel uitmaken van het geldende recht. Het gaat hierbij met name om de omvang van de boete, de omvang van de werkelijke schade, de hoedanigheid van de partijen (professioneel of particulier) en de vraag of het

S. Danziger, J. Levav & L. Avnaim-Pesso, 'Extraneous factors in judicial decisions', *PNAS* 2011, afl. 17, p. 6889-6892.

10 Zie hierover uitgebreider Wolters 2016, p. 109-112.

11 J.M. Barendrecht, *Recht als model van rechtvaardigheid. Beschouwingen over vage en scherpe normen, over binding aan het recht en over rechtsvorming* (diss. Tilburg), Deventer: Kluwer 1992, p. 12. Vergelijk ook Barendrecht 1992, p. 75-76; Asser/Vranken *Algemeen deel*^{***} 2005/29; V. van den Brink, 'Alle omstandigheden van alle gevallen', *NTBR* 2012, afl. 1, p. 2; Wolters 2013a, p. 254.

3 *MOGELIJKHEDEN EN BEPERKINGEN BIJ DE KWANTITATIEVE ANALYSE VAN
JURISPRUDENTIE VOOR DE BESTUDERING VAN GELDEND RECHT*

gaat om een ‘gebruikelijke’ boete zoals de boete bij koopovereenkomsten ten aanzien van onroerende zaken. Mijn belangrijkste hypothese was hierbij dat de verhouding tussen schade en boete, ondanks het nadrukkelijke oordeel van de Hoge Raad om ook op andere omstandigheden te letten, veruit de belangrijkste omstandigheid is. Een ‘wanverhouding’ tussen een lage werkelijke schade en een hoge boete wijst sterk in de richting van matiging.

Ik heb geen aandacht besteed aan niet-juridische omstandigheden die eigenlijk geen rol zouden mogen spelen, maar waarvan het voorstelbaar is dat zij toch invloed hebben op de uitkomst. Ik ga bijvoorbeeld niet in op de identiteit van de rechter,¹² de locatie van de rechterlijke instantie of de woon- of vestigingsplaats van de partijen. Hoewel deze kwesties ongetwijfeld interessant kunnen zijn vanuit een sociologische of politicologische benadering, vallen zij buiten de bestudering van het geldende recht in de traditionele zin.

3 **SELECTIE**

Een kwantitatieve analyse van jurisprudentie leidt slechts tot significante resultaten als er voldoende uitspraken worden onderzocht. Deze uitspraken dienen bovendien op een systematische, reproduceerbare en representatieve wijze te worden geselecteerd. Het is niet mogelijk om algemene eisen te formuleren aan de omvang van de steekproef.¹³ Wel is het van belang om te beseffen dat de statistische ‘kracht’ van de verschillende testen afhankelijk is van deze omvang. Een grotere steekproef leidt tot meer zekerheid over de significantie van de resultaten. Zij maakt het tevens mogelijk om ‘kleinere’ effecten te detecteren en om de samenhang tussen meer dan twee factoren te bestuderen (zie ook par. 5). Zij kost echter ook meer tijd. Het bepalen van de omvang van de steekproef is hierdoor dus altijd een afweging tussen kwaliteit en praktische haalbaarheid.

Het is bovendien niet altijd mogelijk om voldoende uitspraken te vinden. Dit geldt in het bijzonder voor analyses over omstandigheden die in relatief weinig uitspraken een rol spelen. Het is hierdoor niet mogelijk om statistisch significante conclusies te trekken over de rol van deze omstandigheden (zie ook par. 5). Artikel 6:94 lid 1 BW en de beperking van de verjaring van vorderingen tot schadevergoeding als gevolg van blootstelling aan asbest zijn in dit licht bijzonder. Bij artikel 6:94 lid 1 BW ligt de rol van de verhouding

12 Zie hierover bijvoorbeeld G.M. de Vries, ‘Rechtswetenschappelijk onderzoek naar rechters door kwantitatieve analyse van jurisprudentie: waar ligt de ethische grens bij het profileren van rechters?’, *RMThemis* 2021, afl. 2, p. 50-68; zie Vols in deze bundel.

13 Zie over deze kwestie bijvoorbeeld A. Field, *Discovering statistics using SPSS*, Thousand Oaks, CA: Sage 2009, p. 58; A. Acheson, ‘Sample Size’, in: N.J. Salkind (red.), *Encyclopedia of Research Design*, Thousand Oaks, CA: Sage 2010, p. 1300.

tussen schade en boete zo voor de hand (par. 2) dat de rechter er in een groot deel van de zaken aandacht aan besteedt (par. 5). Bij de beperking van de verjaring van vorderingen tot schadevergoeding als gevolg van blootstelling aan asbest heeft de Hoge Raad benadrukt dat de rechter aandacht *moet* besteden aan verschillende omstandigheden.¹⁴

Het vinden van jurisprudentie is relatief gemakkelijk. De overheid publiceert een grote hoeveelheid uitspraken in een openbare, doorzoekbare database (<https://uitspraken.rechtspraak.nl/>). De zoekopdracht '6:94 BW matiging' leidde op 2 december 2015 bijvoorbeeld tot 118 uitspraken uit de periode 2013-2015 waarin de rechter een uitspraak heeft gedaan over de toepassing van artikel 6:94 lid 1 BW. Deze uitspraken zijn allemaal geselecteerd. De selectie is hierdoor systematisch en reproduceerbaar.

Deze selectie is echter geen representatieve weergave van alle relevante conflicten. De overheid publiceert allereerst niet alle uitspraken. Het ligt voor de hand dat relatief 'interessante' zaken vaker worden gepubliceerd.¹⁵ Dit zou in het geval van boetebedingen bijvoorbeeld kunnen betekenen dat er relatief veel uitspraken worden gepubliceerd waarin een boetebeding wordt gematigd. In de door mij onderzochte uitspraken matigde de rechter de bedongen boete in 45% van de gevallen.¹⁶ Hieruit kan echter niet in algemene zin worden geconcludeerd dat een beroep op artikel 6:94 lid 1 BW in 45% van de gevallen succesvol is.¹⁷ Het is zeer waarschijnlijk dat er een grote hoeveelheid kansloze, ongemotiveerde beroepen op matiging niet zijn gepubliceerd die hierdoor buiten de selectie zijn gebleven.¹⁸ Dit probleem kan worden verholpen door ook ongepubliceerde uitspraken te bestuderen. Het moge echter duidelijk zijn dat dit aanzienlijke extra inspanningen oplevert.¹⁹

14 HR 28 april 2000, NJ 2000/430 (*Van Hese/Koninklijke Schelde Groep*); J.L. Smeehuijzen, *De bevrijdende verjaring* (serie Recht en Praktijk deel 162), Deventer: Kluwer 2008, p. 257; J.P. Quist, *Gezichtspunten in het privaatrecht, in het bijzonder het arbeidsrecht* (diss. Rotterdam), Den Haag: Boom Juridische uitgevers 2014, p. 330; Wolters 2015a, p. 17.

15 Zie ook art. 5 onder e en 6 lid 1 Besluit selectiecriteria uitsprakendatabank Rechtspraak.nl. Uitspraken worden onder andere gepubliceerd als zij een 'jurisprudentievormend karakter' hebben en er 'niet uitsluitend standaardformuleringen zijn gebruikt'; www.rechtspraak.nl/Uitspraken/paginas/selectiecriteria.aspx. Zie over het probleem van representativiteit in deze bundel ook Vols en Peeraer & Van Gestel.

16 Wolters 2016, p. 113.

17 Zie over andere problemen met dergelijke conclusies ook P.T.J. Wolters, 'Plas/Valburg en de klassieke tegenstelling tussen rechtszekerheid en billijkheid. Reactie op "Plas/Valburg: veel rechtsonzekerheid en ondermaats resultaat in de feitenrechtspraak"', *AA* 2015, afl. 6, p. 475-478.

18 Hierbij is van belang dat het beroep op art. 6:94 lid 1 BW niet noodzakelijk de reden is geweest voor de publicatie van de uitspraak of de gang naar de rechter. Het kan ook slechts een van de twistpunten zijn. Het is bijvoorbeeld mogelijk dat de schuldenaar in de eerste plaats meent dat de boete niet is verschuldigd en subsidiair een beroep doet op matiging. In dergelijke situaties kunnen kansloze, ongemotiveerde beroepen op matiging ook worden gepubliceerd.

19 Zie ter illustratie de bijdrage in deze bundel van Goossens, Bollen & Verbeke.

3 *MOGELIJKHEDEN EN BEPERKINGEN BIJ DE KWANTITATIEVE ANALYSE VAN JURISPRUDENTIE VOOR DE BESTUDERING VAN GELDEND RECHT*

Het is bovendien maar de vraag in hoeverre dit werkelijk zou leiden tot een beter inzicht in het geldende recht. Ook ongepubliceerde uitspraken vormen immers geen representatieve afspiegeling van alle juridische conflicten. De keuze om naar de rechter te gaan wordt onder andere bepaald door de verwachte opbrengst daarvan.²⁰ Als rechters een beroep op artikel 6:94 lid 1 BW eerder honoreren, zullen ook schuldenaren met relatief zwakke argumenten hier vaker een beroep op doen. Dit zal er dan weer toe leiden dat het percentage succesvolle beroepen daalt.

Kwantitatieve analyse van jurisprudentie voor de bestudering van geldend recht is door deze problemen rondom representativiteit niet geschikt om te onderzoeken hoe vaak een bepaalde individuele variabele in de rechtspraak van belang is. Zij is slechts zinvol voor het bestuderen van de invloed van ten minste twee variabelen op elkaar. In mijn onderzoek naar boetebedingen onderzoek ik onder andere of de verhouding tussen schade en boete (de onafhankelijke variabele) invloed heeft op de matiging op grond van artikel 6:94 lid 1 BW (de afhankelijke variabele). Bij deze kwestie lijkt het niet waarschijnlijk dat er systematische verschillen bestaan tussen gepubliceerde en ongepubliceerde uitspraken. Met andere woorden: zelfs als succesvolle beroepen op artikel 6:94 lid 1 BW vaker worden gepubliceerd, is het niet waarschijnlijk dat de verhouding tussen schade en boete hierbij een andere rol heeft gespeeld dan in ongepubliceerde uitspraken. Het is in zo'n geval voor de bestudering van het geldende recht niet nodig om ook ongepubliceerde uitspraken te bestuderen.

Feit blijft echter dat het zonder bestudering van de ongepubliceerde uitspraken niet mogelijk is om dit zeker te weten. Het is hierdoor opnieuw een afweging tussen kwaliteit en praktische haalbaarheid. Een onderzoeker die ervoor kiest om slechts gepubliceerde uitspraken in de analyse te betrekken, dient in ieder geval goed na te denken over mogelijke systematische verschillen tussen gepubliceerde en ongepubliceerde uitspraken.

4 *CODERING*

Na de selectie van de uitspraken volgt de codering. In deze fase noteren een of meer codeurs bij iedere uitspraak verschillende variabelen. Zij noteren bij een onderzoek naar het geldende recht bijvoorbeeld tot welke conclusie een rechter komt (wel of geen matiging) en van welke omstandigheden er sprake was (verhouding tussen boete en werkelijke schade). Dit proces is lang niet altijd eenvoudig. Dit leidt tot belangrijke beperkingen van de mogelijkheden van kwantitatieve analyse van jurisprudentie voor de bestudering van geldend recht.

²⁰ Zie ook A. Mein & F. de Meere, *Motieven van burgers om (niet) naar de rechter te gaan* (rapport voor de Raad voor de rechtspraak), 2018.

Deze beperkingen liggen voor een groot gedeelte in elkaars verlengde. Ik ga daarom pas in paragraaf 4.5 in op de mogelijke oplossingen. De beperkingen zijn bovendien niet beperkt tot de codering. Uit paragraaf 5 blijkt dat zij ook invloed hebben op de duidelijkheid en betrouwbaarheid van de resultaten.

4.1 *Onduidelijke omstandigheden; afhankelijkheid van de rechter*

De verschillende problemen bij de codering vloeien voort uit het feit dat de relevante omstandigheden in veel gevallen onduidelijk zijn. Het is bijvoorbeeld moeilijk te bepalen of en in welke mate er sprake is van een ‘verwijtbare gedraging van een van de partijen’.²¹ Deze omstandigheid speelt in het kader van artikel 6:94 lid 1 BW bijvoorbeeld een rol met betrekking tot de omvang van de werkelijke schade. Dit niet vreemd: een boetebeding is immers onder andere bestemd om lastig vast te stellen schade te fixeren. De partijen twisten dan ook regelmatig over de vraag of de schuldeiser van een contractuele boete werkelijk nadeel heeft ondervonden door de tekortkoming van de schuldenaar. Dit verklaart waarom de procedure bij de rechter niet alleen bedoeld is om een oordeel te vellen over de toepassing van artikel 6:94 lid 1 BW bij een bepaald feitencomplex. Zij is ook bedoeld om de relevante omstandigheden zelf vast te stellen. De onderzoeker kan daarom bij de kwantitatieve analyse uitgaan van de omstandigheden *zoals zij door de rechter zijn vastgesteld*. Dit heeft echter het nadeel dat de onderzoeker *afhankelijk* wordt van deze vaststelling door de rechter.

Deze afhankelijkheid is min of meer onvermijdelijk. De onderzoeker is immers niet in staat om de relevante omstandigheden zelf vast te stellen. Hij heeft in de regel slechts beperkte toegang tot het partijdebat, is niet aanwezig geweest bij de zitting en is, anders dan de rechter, geen expert op dit gebied. Het vaststellen van de omstandigheden is bovendien zeer tijdrovend. Het komt in min of meerdere mate neer op het overdoen van het werk van de rechter. Dit neemt niet weg dat een opzet waarin de onderzoeker de relevante omstandigheden zelf vaststelt, van groot belang kan zijn om te ontdekken in hoeverre de beperkingen met betrekking tot de codering van belang zijn. Een dergelijke opzet staat echter verder af van de kwantitatieve analyse *van jurisprudentie*. Ik laat de mogelijkheden hiertoe daarom verder onbesproken.

21 Zie over deze factor in het algemeen Wolters 2013a, p. 111-114, en specifiek in de context van de matiging van boetebedingen Wolters 2016, p. 117.

4.2 Onduidelijke vaststellingen

De rechter stelt de relevante omstandigheden niet altijd duidelijk vast. De onderzoeker is bij de codering echter afhankelijk van deze vaststellingen (par. 4.1). Het is daarom niet altijd duidelijk hoe hij de uitspraak precies moet coderen. De potentiële onduidelijkheid van de vaststellingen wisselt per factor. Zij hangt sterk samen met de (on)duidelijkheid van de omstandigheid zelf (par. 4.1). Relatief duidelijke omstandigheden blijken in de regel ook helder uit de uitspraak. De inhoud van de contractuele boete, en daarmee de vraag of er sprake was van een gebruikelijke boete, is in de meeste gevallen bijvoorbeeld goed uit de uitspraak te halen.

Dit geldt echter niet voor de in mijn onderzoek cruciale omvang van de werkelijke schade. De rechter komt hierbij lang niet altijd tot een helder oordeel. Hoewel de rechter in een deel van de onderzochte uitspraken (51 van de 118) een concreet oordeel geeft over de omvang van de werkelijke schade, beperkt hij zich in 39 andere zaken tot een opmerking over de omvang van de schade. Deze opmerkingen zijn niet altijd helder of expliciet. Het is in deze gevallen niet altijd duidelijk wat de rechter precies vaststelt over de omvang van de werkelijke schade.²²

4.3 Noodzaak tot categorisering

Kwantitatieve analyse is niet mogelijk aan de hand van 39 losstaande opmerkingen. Het is noodzakelijk om ze te ordenen en in te delen in categorieën. Dit versterkt de in paragraaf 4.2 besproken beperking. Een onduidelijke vaststelling van de omstandigheden maakt het moeilijk om te beoordelen in welke categorie een opmerking thuishoort. De kwantitatieve analyse is hierdoor uiteindelijk gebaseerd op een twijfelachtige categorisering van onduidelijke uitspraken.

Het is bijvoorbeeld niet altijd duidelijk of de rechter oordeelt dat de schade *duidelijk* lager is dan de boete of dat hij deze discrepantie slechts als *mogelijkheid* signaleert. Het verschil tussen de categorieën ‘de schade is duidelijk lager dan de boete’ en ‘de schade en boete lopen mogelijk uiteen’ is daarom niet altijd duidelijk.²³

Het is niet altijd mogelijk om deze problemen bij de categorisering op te lossen. De vaststellingen door de rechter zijn nu eenmaal niet altijd duidelijk. Wel is het van groot belang om kritisch te kijken naar de verschillende categorieën. De indeling dient recht te doen

²² Zie ook par. 4.3; 4.4; 5.2.

²³ Wolters 2016, p. 114. Zie ook par. 4.4; 5.2.

aan de vaststellingen door de rechter, maar ook de onderzoeker in staat stellen om de uitspraken consistent te coderen.

4.4 Vooringenomenheid van de onderzoeker

Een onderzoeker begint niet zomaar aan een kwantitatieve analyse. Hij doet dit omdat hij een bepaalde hypothese heeft over de invloed van een relevante omstandigheid. Ik begon mijn onderzoek naar de invloed van de verhouding tussen de schade en boete bijvoorbeeld met de veronderstelling dat deze omstandigheid ondanks *Intrahof/Bart Smit* een dominante rol speelt bij de toepassing van artikel 6:94 lid 1 BW.

De onderzoeker codeert de uitspraken met de hypothese in zijn achterhoofd. Dit brengt in combinatie met de afhankelijkheid van de rechter (par. 4.1), de onduidelijke vaststellingen (par. 4.2) en de noodzaak tot categorisering (par. 4.3) een risico met zich. Een onderzoeker kan bewust of onbewust geneigd zijn om de omstandigheden te coderen op een wijze die overeenstemt met zijn hypothese.

Stel bijvoorbeeld dat de opmerking over de werkelijke schade niet duidelijk is. De onderzoeker twijfelt over de categorieën ‘de schade is duidelijk lager dan de boete’ en ‘de schade en boete lopen mogelijk uiteen’ (par. 4.3). Hij ziet dat de rechter de boete uiteindelijk *niet* heeft gematigd. Dit gebeurt volgens de hypothese met name als er geen wanverhouding tussen schade en boete bestaat. De onderzoeker zou in dit geval geneigd kunnen zijn om de opmerking in te delen in de categorie ‘de schade en boete lopen mogelijk uiteen’. Een oordeel waarin de rechter de boete ondanks de duidelijk lagere schade niet matigt, stemt immers niet overeen met de hypothese.

Dit effect kan worden versterkt als deze oordelen slechts zelden voorkomen. Het zijn in die gevallen uitschieters of ‘*outliers*’. Een onderzoeker zal bijzondere aandacht aan deze *outliers* wijden.²⁴ Stel bijvoorbeeld dat de rechters een boete bij een wanverhouding tussen schade en boete *bijna* altijd matigen. De uitspraken waarin dit ondanks de wanverhouding niet gebeurt, zijn natuurlijk extra interessant. Het is daarnaast van belang om te bekijken of deze *outliers* niet het gevolg zijn van een verkeerde kwalificatie. Een eventuele fout kan dan worden gecorrigeerd. Dit leidt echter wel tot een situatie waarin afwijkende uitspraken twee keer worden bekeken en er dus ook twee keer een kans bestaat dat de onderzoeker de omstandigheid ten onrechte codeert op een wijze die overeenstemt met de hypothese.

24 Zie ook D.J. Sheskin, ‘Outlier’, in: N.J. Salkind (red.), *Encyclopedia of Research Design*, Thousand Oaks, CA: Sage 2010, p. 980.

Uitspraken die bij de hypothese en de overige resultaten passen, worden echter maar een keer onderzocht.

4.5 *Mogelijke oplossingen*

De in de vorige subparagrafen besproken problemen kunnen op verschillende wijzen worden geadresseerd. Het is in de eerste plaats van groot belang om in het codeboek goed bij te houden hoe de uitspraken zijn gecodeerd. De onderzoeker dient bijvoorbeeld bij te houden hoe hij met onduidelijke formuleringen (par. 4.2) en grensgevallen (par. 4.3) omgaat. Ik heb in de codeboeken bijvoorbeeld nader omschreven hoe ik ben omgegaan met de verschillende onduidelijke opmerkingen over de werkelijke schade. Dit maakt het coderwerk beter controleerbaar en kan inconsistenties voorkomen (zie par. 4.4).

De problemen bij de codering kunnen daarnaast worden geadresseerd door te werken met meerdere codeurs. De codeurs kunnen in dit geval elkaars werk controleren en overleggen over twijfelgevallen. Door meerdere codeurs naar dezelfde uitspraken te laten kijken, is het bovendien mogelijk om de resultaten met elkaar te vergelijken en de interbeoordelaarsbetrouwbaarheid te berekenen.²⁵ Als zij dezelfde uitspraken op een andere manier coderen, komt beter aan het licht waar de onduidelijkheden precies zitten. Het verdient daarom de aanbeveling om serieus te onderzoeken of het mogelijk is om bij het coderen samen te werken. Dit kan bijvoorbeeld met een collega-onderzoeker, maar ook met studenten.

Het nadeel van deze oplossing is evident. Het kost dubbel zoveel tijd om twee codeurs naar een uitspraak te laten kijken. De training van de verschillende codeurs kost bovendien ook tijd, zeker nu het van belang is dat zij op dezelfde manier coderen. Het zal daarom niet voor iedereen en bij ieder onderzoek mogelijk zijn om een tweede codeur te betrekken. Het moet immers gaan om iemand die, het liefst op korte termijn, veel tijd kan en wil vrijmaken voor het relatief saaie coderwerk. Het zal voor zo iemand vaak de vraag zijn of hij of zij zijn tijd niet beter kan besteden aan het schrijven van een ander artikel. Er bestaat dus opnieuw (par. 3) een afweging tussen kwaliteit en praktische haalbaarheid. Ik heb in mijn onderzoeken daarom geen gebruikgemaakt van een volwaardige²⁶ tweede codeur.

25 Dit kan bijvoorbeeld met 'Cohen's kappa'. Zie over deze test bijvoorbeeld M.J. Warrens, 'Cohen's kappa is a weighted average', *Statistical Methodology* 2011, afl. 6, p. 473-475. Zie ook de bijdragen van Verbruggen en Vols (beiden met verwijzing naar nadere literatuur).

26 Ik had wel de beschikking over een onderzoeksassistent die mij werk uit handen heeft genomen bij de selectie en codering van de uitspraken. De uiteindelijke codering was echter steeds van mijn hand.

De in de vorige subparagrafen besproken problemen kunnen daarnaast deels worden ondervangen door transparant en controleerbaar te werken. Het verdient daarom de aanbeveling om het codeboek en de dataset inclusief de interbeoordelaarsbetrouwbaarheid openbaar te maken. Dit gaat eenvoudig via een daartoe bestemde *repository*, zoals de door mij gebruikte <https://easy.dans.knaw.nl/>.²⁷ Een voordeel bij kwantitatieve analyse van *gepubliceerde* jurisprudentie ten opzichte van andere empirische onderzoeken is bovendien dat de uitspraken openbaar zijn. De dataset is hierdoor te controleren door anderen.

De hiervoor besproken oplossingen lossen echter niet alle in de vorige subparagrafen besproken problemen op. Een tweede codeur kan de vooringenomenheid van de onderzoeker adresseren (par. 4.4) en samen met een codeboek bijdragen aan een consistente categorisering (par. 4.3). Deze oplossingen veranderen echter niets aan de afhankelijkheid (par. 4.1) van de onduidelijke oordelen van de rechter (par. 4.2).

Een laatste mogelijkheid is om de codering te automatiseren. Een uitgebreide bespreking van de mogelijkheden hiervan valt buiten het bereik van deze bijdrage. Ik beperkt mij hier tot een korte bespreking van de wijze waarop deze mogelijkheid zich verhoudt tot de besproken beperkingen.

Het belangrijkste voordeel van geautomatiseerde codering is snelheid. Handmatige codering kost veel tijd (zie ook par. 3), zeker als het de voorkeur verdient om dit door meerdere codeurs te laten doen. Een eenmaal ontwikkeld algoritme kan echter zeer snel een grote hoeveelheid jurisprudentie analyseren. Het helpt hierbij bovendien dat juridische uitspraken bestaan uit machineleesbare (digitale) tekst en zijn opgebouwd volgens een vast stramien.

Het is echter maar de vraag in hoeverre een algoritme in staat is om onduidelijke opmerkingen over de omvang van de schade te herkennen en categoriseren, laat staan of zij dit beter kunnen dan menselijke codeurs.²⁸ Een algoritme is daarnaast ook niet noodzakelijk objectiever. Het dient immers getraind te worden met behulp van door mensen uitgevoerde coderingen. De vooringenomenheid in de trainingsdata kan hierdoor terugkeren in het algoritme.²⁹ Automatische codering is daarom vooral nuttig met betrekking tot relatief duidelijke omstandigheden. Zij is minder geschikt voor de kwantitatieve analyse die steunt op de relatief onduidelijke omstandigheden die bij open normen vaak van belang zijn.

27 De vindplaatsen zijn opgenomen in de verschillende artikelen. Zie in het bijzonder de gegevens bij Wolters 2016, <https://doi.org/10.17026/dans-xxz-q9ke>.

28 Zie voor een toepassing van deze methode ook Vols in deze bundel. Zie voor kanttekeningen bij gestelde successen De Vries 2021, p. 51.

29 Zie ook T. Baer, *Understand, Manage, and Prevent Algorithmic Bias*, New York, NY: Apress 2019.

5 RESULTATEN

Na de codering volgt de analyse van de gegevens. In deze fase gebruikt de onderzoeker de verzamelde gegevens om statistisch significante conclusies te trekken over 'het geldende recht'. Hij analyseert of en in hoeverre de onafhankelijke variabelen invloed hebben op de afhankelijke variabelen (zie ook par. 3). Hij test bijvoorbeeld of de verhouding tussen schade en boete van belang is voor de toepassing van artikel 6:94 lid 1 BW.

Kwantitatieve analyse heeft hierbij een meerwaarde ten opzichte van kwalitatieve analyse. Stel bijvoorbeeld dat een onderzoeker enkele uitspraken vindt waarin een wanverhouding tussen schade en boete gepaard gaat met matiging. Hij vindt echter ook een uitspraak waarin een rechter ondanks de wanverhouding niet matigt. Hij vindt op basis van kwalitatieve analyse van de uitspraken geen relevante verschillen die deze wisselende oordelen kunnen verklaren. Het is bovendien niet duidelijk hoe vaak de rechters tot de verschillende resultaten komen. Het is immers ook mogelijk dat een van de uitspraken een uitzondering is en dat de feitenrechtspraak voor het overige wel consistent is. Kwalitatieve analyse leidt in dit geval daarom niet tot duidelijke inzichten, anders dan dat de feitenrechtspraak niet (volledig) consistent is.

Kwantitatieve analyse kan meer duidelijkheid geven. De onderzoeker kan preciezer laten zien in hoeveel gevallen de wanverhouding wel of niet gepaard gaat met matiging. Hij kan bovendien berekenen of deze uitkomst statistisch significant is en hoe sterk dit effect is. In mijn onderzoek laat ik bijvoorbeeld zien dat er een sterk verband bestaat tussen de verhouding van de schade en de boete en de matiging. De rechter matigt de boete significant vaker bij een wanverhouding.

Dit geldt met betrekking tot de uitspraken waarin de werkelijke schade is vastgesteld,³⁰ de uitspraken waarin de rechter een opmerking maakt over de schade³¹ en alle uitspraken

30 $r_b = -0,577$, $p < .001$.

31 $X^2(1) = 25,558$, $p < .001$, Cramer's $V = .856$. De vijf soorten opmerkingen zijn hiervoor teruggebracht tot twee categorieën. De eerste twee opmerkingen vormen de eerste categorie waarin de rechter benadrukt dat de werkelijke schade laag is. De andere drie opmerkingen vormen de tweede categorie.

gezamenlijk³² (tabel 1, paneel A).³³ Het belang van deze omstandigheid blijkt bovendien uit het feit dat in uitspraken waarin niets is bepaald over de werkelijke schade relatief weinig wordt gematigd³⁴ en dat de rechter zich in 90 van de 118 uitspraken concreet (51) of door middel van een opmerking (39) uitlaat over de werkelijke schade (tabel 1, paneel B). De verhouding tussen schade en boete blijkt dus inderdaad, ondanks het nadrukkelijke oordeel van de Hoge Raad om ook op andere omstandigheden te letten, een zeer belangrijke omstandigheid te zijn.

Tabel 1 De opmerkingen over of vaststelling van de werkelijke schade

	Aantal	% van alle uitspraken	Aantal gematigd	% gematigd
A.				
Verhouding < 0.5 of een opmerking die benadrukt dat de schade laag is	55	47	45	82
Geen informatie over werkelijke schade	28	24	7	25
Verhouding > 0.5 of andere opmerking	35	30	1	3
Totaal	118	100	53	45
B.				
Opmerking over of vaststelling van de schade	90	76	46	51
Geen informatie over werkelijke schade	28	24	7	25
Totaal	118	100	53	45

Mijn onderzoek laat daarnaast zien dat de rechter de gebruikelijke boetes bij koopovereenkomsten ten aanzien van onroerende zaken significant minder snel matigt. De rol van deze omstandigheid is echter veel beperkter dan die van de verhouding tussen schade en boete (tabel 2).³⁵ Bij andere relevante omstandigheden waren er onvoldoende gegevens

32 $X^2(2) = 59.785, p < .001$, Cramer's $V = .712$. De uitspraken zijn ingedeeld in drie categorieën. In de eerste groep zitten de uitspraken waarin de rechter een opmerking maakt die de omvang van de werkelijke schade relateert ('de schade is duidelijk lager dan de boete' en 'de rechter erkent het bestaan van schade, maar relateert de omvang') en de uitspraken waarin de werkelijke schade minder is dan de helft van het totale toewijsbare bedrag (verhouding < 0.5). In de tweede categorie zitten de uitspraken waarin niets is gesteld over de omvang van de werkelijke schade. In de derde groep zitten uitspraken met een andere opmerking over de omvang van de schade en de uitspraken waarin de werkelijke schade groter is dan de helft van het totale toewijsbare bedrag (verhouding > 0.5).

33 Zie voor een uitgebreidere beschrijving van de resultaten Wolters 2016, p. 114-115. Hier zijn ook additionele tabellen opgenomen.

34 $X^2(1) = 5.885, p < .05$, Cramer's $V = .223$.

35 $X^2(1) = 5.888, p < .05$, Cramer's $V = .189$. De significantie bestaat ook als wordt gecontroleerd voor de verhouding tussen schade en boete. *Loglinear analyse*, $X^2(2) = 6.609, p < .05$. Zie uitgebreider Wolters 2017a, p. 126.

3 *MOGELIJKHEDEN EN BEPERKINGEN BIJ DE KWANTITATIEVE ANALYSE VAN
JURISPRUDENTIE VOOR DE BESTUDERING VAN GELDEND RECHT*

beschikbaar om kwantitatieve conclusies te trekken. Er waren bijvoorbeeld onvoldoende gepubliceerde uitspraken over de matiging van andere gebruikelijke boetebedingen.³⁶

Tabel 2 De opmerkingen over of vaststelling van de werkelijke schade

	Aantal	% van uitspraken binnen de groep	Aantal gematigd	% gematigd
Groep A. Geen gewoonte				
Verhouding < 0.5 of een opmerking die benadrukt dat de schade laag is ('wanverhouding')	39	51	34	87
Verhouding > 0.5 of andere opmerking	19	25	0	0
Geen informatie over werkelijke schade	18	24	4	22
Totaal	76	100	38	50
Groep B. Gebruikelijke boete koop onroerende zaak				
Verhouding < 0.5 of een opmerking die benadrukt dat de schade laag is	33	37	22	67
Verhouding > 0.5 of andere opmerking	37	42	4 ³⁷	11
Geen informatie over werkelijke schade	19	21	2	11
Totaal	89	100	28	21

De beperkingen bij de codering doen afbreuk aan de betrouwbaarheid van de resultaten. Deze resultaten zijn hierdoor immers gebaseerd op een twijfelachtige categorisering van onduidelijke uitspraken (par. 4.3) en zouden kunnen zijn beïnvloed door de hypothese (par. 4.4). Een uitgewerkt codeboek en een hoge interbeoordelaarsbetrouwbaarheid kunnen deze beperkingen slechts deels wegnemen (par. 4.5). Een analyse die is gebaseerd op een onduidelijke opmerking (par. 4.2) over de onduidelijke omvang van de werkelijke schade (par. 4.1) blijft immers gepaard gaan met onzekerheden.

De beperkingen kunnen daarnaast worden ondervangen door te analyseren of de resultaten ook kunnen worden gebaseerd op relatief betrouwbare gegevens. Het gaat hierbij om vaststellingen van relevante omstandigheden die relatief duidelijk en objectief zijn in te

³⁶ Wolters 2017a, p. 127. Zie ook Wolters 2016, p. 115-117.

³⁷ De rechter matigde ondanks een beperkt verschil tussen schade en boete in Rb. Utrecht 21 oktober 2009, ECLI:NL:RBUTR:2009:BK1300 (verhouding 0.53); Rb. Rotterdam 9 december 2009, ECLI:NL:RBROT:2010:BM2406 (het bestaan van schade is aannemelijk); Rb. Rotterdam 10 maart 2010, ECLI:NL:RBROT:2010:BM2406 (verhouding 0.71); Hof Amsterdam 1 februari 2011, ECLI:NL:GHAMS:2011:BP7570 (verhouding 0.58).

delen. In mijn onderzoek geldt bijvoorbeeld dat de uitspraken met een concreet oordeel over de omvang van de werkelijke schade over het algemeen duidelijker zijn dan de opmerkingen over deze omvang (par. 4.2). Het is daarom van belang dat de significantie van de resultaten ook bestaat als de analyses slechts worden gebaseerd op uitspraken met een concreet oordeel over de omvang van de werkelijke schade.

Tegelijkertijd is het geen optie om alle onbetrouwbare gegevens weg te laten. Ook onduidelijke vaststellingen maken immers deel uit van de relevante jurisprudentie. Het verdient daarom de voorkeur om de analyses zowel met als zonder de onbetrouwbare gegevens uit te voeren. Het is een gunstig teken als, zoals in mijn onderzoek, de resultaten in beide gevallen significant en ongeveer hetzelfde zijn. Dit suggereert dat de problemen bij de codering geen grote invloed hebben gehad op de resultaten.

5.1 *Het verschil tussen de werkelijke en vastgestelde omstandigheden*

De vaststelling van de relevante omstandigheden door de rechter is niet alleen onduidelijk (par. 4.2), zij kan ook subjectief zijn. Het is mogelijk dat andere rechters in dezelfde gevallen tot andere vaststellingen komen. Dit speelt in het bijzonder bij relatief onduidelijke omstandigheden, maar kan ook voorkomen bij duidelijkere feitencomplexen.

Dit blijkt bijvoorbeeld uit mijn onderzoek naar de beperking van de verjaring van vorderingen tot schadevergoeding als gevolg van blootstelling aan asbest. Rechters oordelen consequent dat de verjaring niet wordt beperkt in situaties waarin de schuldenaar niet of slechts in beperkte mate rekening behoort te houden met de aanspraak. Zij zijn echter inconsequent over de vraag of er sprake is van een dergelijke situatie. Dit hangt mede af van de ontwikkeling van de kennis over de gevaren van asbest. De rechters noemen echter sterk uiteenlopende momenten vanaf wanneer de schuldenaar rekening had kunnen houden met de aanspraak. Zij wijzen op het moment dat bekend was dat blootstelling aan asbest tot andere ziektes (zoals asbestose) kan leiden, de asbestconferentie van 1964, het proefschrift van Stumphius in 1969, de maatschappelijke commotie rondom asbest halverwege de jaren negentig, een eerdere aansprakelijkheidstelling van het aangesproken bedrijf door een andere werknemer en een eerdere toewijzing van een vordering van een niet-werknemer.³⁸ Een zeer duidelijke en objectieve omstandigheid als het verschijnen van een proefschrift leidt hierdoor tot uiteenlopende vaststellingen over de mogelijkheden om rekening te houden met een aanspraak.

³⁸ Wolters 2015a, p. 20 en 23-24; Wolters 2015b, p. 142. Zie ook par. 5.2.

3 MOGELIJKHEDEN EN BEPERKINGEN BIJ DE KWANTITATIEVE ANALYSE VAN JURISPRUDENTIE VOOR DE BESTUDERING VAN GELDEND RECHT

De afhankelijkheid van de subjectieve vaststelling van de omstandigheden door de rechters vertroebelt hierdoor de door kwantitatieve analyse verkregen inzichten. Het geldende recht ziet op de rechtsgevolgen die onder bepaalde omstandigheden optreden. Een rechtszoekende wil weten welke rechtsgevolgen er voor hem, uitgaande van zijn omstandigheden, gelden (par. 2). Er bestaat echter een verschil tussen de *werkelijke* omstandigheden en de door de rechter *vastgestelde* omstandigheden. De kwantitatieve analyse ziet uiteindelijk op de verhouding tussen de vastgestelde omstandigheden en de werking van de rechtsnorm. Zij geeft slechts een beperkt en indirect inzicht in de invloed van de werkelijke omstandigheden op het oordeel. Zij maakt met andere woorden wel duidelijk dat de verjaring niet wordt beperkt in situaties waarin de aangesproken partij geen rekening hoeft te houden met aansprakelijkheid, maar niet welke omstandigheden tot dit oordeel leiden. Figuur 1 geeft een overzicht van dit probleem.

Figuur 1 De werkelijke en vastgestelde omstandigheden



Hierdoor geldt opnieuw (par. 5) dat kwantitatieve analyse met name tot betrouwbare resultaten over het geldende recht leidt met betrekking tot omstandigheden die relatief objectief kunnen worden vastgesteld. De discrepantie tussen de werkelijke omstandigheden en de vaststelling daarvan is bij deze omstandigheden (hoogstwaarschijnlijk) beperkter.

5.2 *Causatie en correlatie*

De subjectiviteit van de vaststelling van de omstandigheden leidt bovendien tot onzekerheid over het oorzakelijk verband tussen de vastgestelde omstandigheden en het oordeel van de rechter. Het geldende recht vereist dat de rechtsgevolgen in overeenstemming met figuur 1 afhankelijk zijn van de vastgestelde omstandigheden. De toepassing van artikel 6:94 lid 1 BW dient bijvoorbeeld afhankelijk te zijn van de relevante omstandigheden van het geval, waaronder de verhouding tussen de werkelijke schade en de boete.

Het tegenovergestelde is echter ook mogelijk. De vaststelling van de omstandigheden staat immers niet los van het oordeel over de matiging. Zij is mede bedoeld om dit oordeel te rechtvaardigen.³⁹ Een rechter heeft hierdoor de mogelijkheid om de vaststelling van de omstandigheden aan te passen aan zijn oordeel. Een rechter die de boete wil matigen, kan bijvoorbeeld geneigd zijn om een gedraging van de schuldeiser als verwijtbaar te beoordelen.

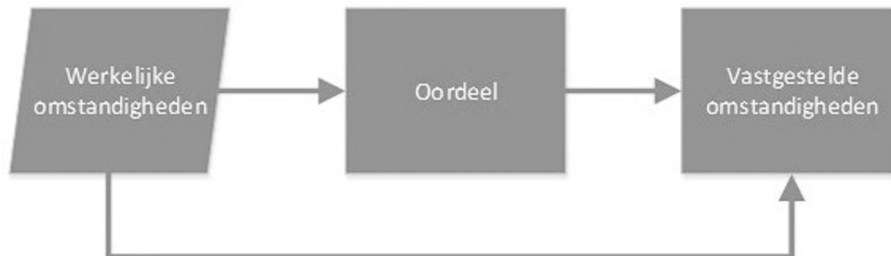
Ook de opmerkingen over de werkelijke schade (par. 4.2) maken dergelijke doelredeneringen mogelijk. Stel bijvoorbeeld dat de werkelijke schade niet precies is vast te stellen, maar dat de feiten dwingen tot de conclusie dat de schuldeiser ten minste enig relevant nadeel heeft geleden. De schade lijkt echter lager te zijn dan de contractuele boete. Een rechter die een boete niet wil matigen, kan de nadruk leggen op het bestaan van relevante schade en overwegen dat de boete niet 'buitensporig' is. Een rechter die wel wil matigen zal benadrukken dat de schade 'duidelijk' lager is dan de boete zonder zich expliciet uit te laten over de eventuele buitensporigheid. Eenzelfde feitencomplex kan hierdoor leiden tot uiteenlopende opmerkingen die gepaard gaan met afwijkende oordelen over de toepassing van artikel 6:94 lid 1 BW.⁴⁰

In deze gevallen is er sprake van *omgekeerde causaliteit*. De door de rechter genomen beslissing over de matiging beïnvloedt de vaststelling van de relevante omstandigheden. Het risico op deze omgekeerde causaliteit verschilt per omstandigheid. Het is kleiner bij omstandigheden die relatief objectief kunnen worden vastgesteld. De rechter heeft in dat geval immers minder ruimte om de vastgestelde omstandigheden aan zijn oordeel aan te passen. De werkelijke omstandigheden zullen in dat geval altijd ten minste enige invloed uitoefenen op de vastgestelde omstandigheden. Uit paragraaf 5.1 blijkt echter dat subjectiviteit en daarmee doelredeneringen ook mogelijk zijn bij deze relatief duidelijke omstandigheden. Figuur 2 geeft een overzicht van deze beperking.

39 W. Mendelson, 'The Neo-Behavioral Approach to the Judicial Process: A Critique', *Am. Pol. Sci. Rev.* (57) 1963, afl. 3, p. 601-602; W.M. Sage, 'Judicial Opinions Involving Health Insurance Coverage: Trompe L'Oeil or Window on the World?', *Ind. L. Rev.* (31) 1998, afl. 1, p. 66-67; M.A. Hall & R.F. Wright, 'Systematic Content Analysis of Judicial Opinions', *California Law Review* (96) 2008, afl. 1, p. 95-99; Wolters 2015a, p. 23-24; zie ook noot 40 *infra*.

40 Wolters 2016, p. 114. Zie ook Wolters 2015a, p. 23-24; par. 5.1.

Figuur 2 Omgekeerde causaliteit



Het voorgaande dient niet te worden geïnterpreteerd als een beschuldiging in de richting van de rechters. Een wisselende waardering van de achterliggende omstandigheden betekent niet dat de rechters deze vaststelling bewust aanpassen aan hun oordeel. Het is bovendien maar de vraag in hoeverre dit een probleem zou zijn. De rechter heeft enige vrijheid bij de motivering van zijn uitspraak. De vraag of en in hoeverre hij deze vrijheid op de in deze paragraaf besproken wijze mag gebruiken, valt buiten het bestek van deze bijdrage.⁴¹

Het is voor kwantitatieve analyse van jurisprudentie voor de bestudering van geldend recht echter wel van belang om rekening te houden met dit risico op omgekeerde causaliteit. Het is niet met zekerheid vast te stellen of en in hoeverre de vastgestelde omstandigheden werkelijk hebben geleid tot het oordeel van de rechter.

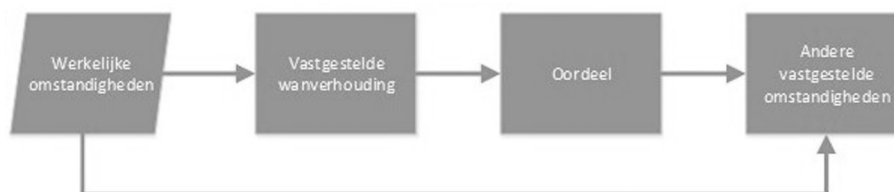
In mijn onderzoek naar de matiging van boetebedingen trek ik daarom alleen conclusies over *correlatie*. Ik concludeer dat een wanverhouding tussen schade en boete significant vaak gepaard gaat met matiging. Dit suggereert dat een wanverhouding ook tot deze matiging *leidt*, maar dat is niet met zekerheid vast te stellen. Het is immers ook mogelijk dat dit effect ten minste gedeeltelijk wordt verklaard doordat rechters een wanverhouding vaststellen *omdat* zij van oordeel zijn dat de boete moet worden gematigd.

Ik denk overigens dat de kans op dergelijke omgekeerde causaliteit met betrekking tot de verhouding tussen schade en boete relatief klein is. De Hoge Raad benadrukt dat aan deze specifieke omstandigheid niet al te veel gewicht dient te worden toegekend (par. 2). Het ligt daarom niet voor de hand om juist met betrekking tot deze omstandigheid een doelre-

⁴¹ Zie ook J.M. Polak, *Theorie en praktijk der rechtsvinding* (diss. Leiden), Zwolle: W.E.J. Tjeenk Willink 1953, p. 15-16; H.F. Crombag, J.L. de Wijkerslooth & M.J. Cohen, 'Over het legitimeren van rechterlijke beslissingen', in: H.F. Crombag e.a., *Het rechterlijk oordeel*, Zwolle: W.E.J. Tjeenk Willink 1973, p. 7-8; Asser/Scholten *Algemeen deel** 1974, p. 132-133.

denering toe te passen. Het is waarschijnlijker bij andere omstandigheden: een rechter die op basis van een wanverhouding wil matigen, kan er extra omstandigheden 'bijhalen' die dit oordeel ondersteunen (zie figuur 3). Dit blijft echter speculatie. Juist de in deze subparagraaf besproken beperking leidt ertoe dat dit niet kan worden achterhaald door middel van kwantitatieve analyse.

Figuur 3 Causaliteit en omgekeerde causaliteit



5.3 Tijdelijk karakter

De door kwantitatieve analyse verkregen inzichten over het geldende recht zijn naar hun aard tijdelijk en historisch. Zij zijn gebaseerd op de toepassing van de norm in een bepaalde periode. Deze toepassing kan echter veranderen. Dit kan bijvoorbeeld komen door een wetwijziging of een nieuwe uitspraak van de Hoge Raad. Het is daarnaast mogelijk dat de achtergrond verandert. Het is bijvoorbeeld voorstelbaar dat de rol van modelovereenkomsten en de daarin opgenomen boetebedingen bij koopovereenkomsten ten aanzien van onroerende zaken een minder grote rol gaan spelen in de praktijk. Dit zou tot gevolg kunnen hebben dat de rechters deze boetes niet langer minder snel matigen (zie par. 5).

Kwantitatieve analyse maakt op zichzelf niet duidelijk waarom een bepaalde omstandigheid van belang is. Zij biedt daarom geen inzicht in de achtergronden die ertoe kunnen leiden dat deze invloed verandert. Het is daarom noodzakelijk om kwantitatieve analyse te combineren met kwalitatieve analyse. Alleen als de onderzoeker begrijpt waarom een bepaalde omstandigheid van belang is, kan hij beoordelen of deze achtergrond op een later moment nog steeds bestaat. Ten aanzien van artikel 6:94 lid 1 BW zijn er sinds het verschijnen van mijn onderzoek geen noemenswaardige ontwikkelingen meer geweest.⁴² Er is dus geen reden om aan te nemen dat de conclusies niet meer gelden. Het blijft echter mogelijk dat er ongezien en zonder een duidelijk aanwijsbare reden een verschuiving is opgetreden en

⁴² HR 16 februari 2018, NJ 2018, 100 (*Turan/Easystaff*) ziet slechts op de vraag hoeveel er moet worden gematigd, maar niet op de in mijn onderzoek besproken vraag of er moet worden gematigd. Zie voor empirische gegevens over de vraag hoeveel er wordt gematigd ook Wolters 2018.

3 *MOGELIJKHEDEN EN BEPERKINGEN BIJ DE KWANTITATIEVE ANALYSE VAN
JURISPRUDENTIE VOOR DE BESTUDERING VAN GELDEND RECHT*

rechters artikel 6:94 lid 1 BW anders zijn gaan toepassen. Een dergelijk risico is minder groot bij kwalitatieve analyse: een uitspraak van de Hoge Raad of een wetswijziging is immers opvallender.

6 **CONCLUSIE EN REFLECTIE**

Deze bijdrage geeft een overzicht van de mogelijkheden en beperkingen van kwantitatieve analyse van jurisprudentie voor de bestudering van geldend recht. Ik ben in het bijzonder ingegaan op de beperkingen. De belangrijkste beperkingen zijn als volgt:

- a. Kwantitatieve analyse is alleen mogelijk ten aanzien van omstandigheden waarover voldoende uitspraken bestaan (par. 3 en 5).
- b. De gepubliceerde uitspraken zijn niet representatief (par. 3).
- c. De relevante omstandigheden zijn niet altijd duidelijk (par. 4.1) en blijken ook niet duidelijk uit de uitspraak van de rechter (par. 4.2).
- d. Het is niet altijd duidelijk hoe de omstandigheden moeten worden gecategoriseerd (par. 4.3).
- e. De vooringenomenheid van de onderzoeker kan de codering beïnvloeden (par. 4.4).
- f. De selectie en codering is tijdrovend, zeker als de hiervoor gesignaleerde beperkingen worden geadresseerd (par. 3, 4.1 en 4.5).
- g. Kwantitatieve analyse ziet slechts op het verband tussen de vastgestelde omstandigheden en het oordeel van de rechter. Zij ziet niet op de rol van de werkelijke omstandigheden (par. 5.1).
- h. Kwantitatieve analyse leidt niet tot zekerheid over het oorzakelijk verband tussen de vastgestelde omstandigheden en het oordeel van de rechter (par. 5.2).
- i. De resultaten zijn naar hun aard tijdelijk (par. 5.3).

Deze beperkingen zijn van groot belang. Zij zien zowel op de betrouwbaarheid van de verzamelde gegevens als op de resultaten van de analyse. Het lijkt hierdoor dat de mogelijkheden van kwantitatieve analyse beperkt zijn. Deze conclusie is echter voorbarig. Kwantitatieve analyse van jurisprudentie bestaat niet in een vacuüm. Zij is een alternatieve methode voor de bestudering van geldend recht. Haar mogelijkheden en beperkingen moeten worden afgezet tegen de mogelijkheden en beperkingen van traditioneel, kwalitatief, juridisch-dogmatisch onderzoek.

Een groot deel van de besproken beperkingen is niet beperkt tot kwantitatieve analyse. Ook bij een kwalitatieve analyse van feitenrechtspraak zijn de besproken uitspraken niet representatief (b), zijn de relevante omstandigheden niet duidelijk (c) en kan de vooringenomenheid van de onderzoeker(s) (enkelvoud of meervoud, zie ook par. 4.5) de interpre-

tatie beïnvloeden (e). Daarnaast verandert de kwalitatieve aard van de analyse niets aan het feit dat de vastgestelde omstandigheden kunnen afwijken van de werkelijke omstandigheden (g) en kunnen zijn beïnvloed door het oordeel van de rechter (h). Ook een kwalitatieve analyse dient bovendien te worden bijgewerkt na het verschijnen van nieuwe uitspraken of als de achtergrond verandert (i). Deze beperkingen bestaan bij alle vormen van jurisprudentieanalyse voor de bestudering van geldend recht. Hoewel kwantitatieve analyse een andere omgang met deze beperkingen kan vereisen, leiden zij niet tot een fundamentele reden om deze methode niet toe te passen.

De volgende beperkingen van kwantitatieve analyse blijven over:

- a. Kwantitatieve analyse is alleen mogelijk ten aanzien van omstandigheden waarover voldoende uitspraken bestaan.
- d. Het is niet altijd duidelijk hoe de omstandigheden moeten worden gecategoriseerd.
- f. De selectie en codering is tijdrovend, zeker als de hiervoor gesignaleerde beperkingen worden geadresseerd.

Kwantitatieve analyse is niet geschikt voor iedere rechtsvraag (a). Zij kan alleen tot significante resultaten leiden als er voldoende uitspraken zijn met betrekking tot de omstandigheden waarop de hypotheses zien. Dit is niet altijd het geval bij iedere norm die afhankelijk is van 'alle omstandigheden van het geval'. Kwantitatieve analyse vereist daarnaast meer tijd dan juridisch-dogmatisch onderzoek (f). Met name het coderen kost veel tijd. Voordat hij met kwantitatieve analyse begint, dient de onderzoeker zich daarom af te vragen of de meerwaarde opweegt tegen de extra tijdsinvestering. Het is bijvoorbeeld niet de moeite waard om deze methode te gebruiken als het geldende recht duidelijk blijkt uit een gezaghebbende uitspraak van de Hoge Raad en er geen aanwijzingen zijn dat de feitenrechtspraak hiervan afwijkt. Een vergelijkbare afweging moet worden gemaakt ten aanzien van de vraag of het noodzakelijk is om een tweede codeur te betrekken. De voordelen met betrekking tot betrouwbaarheid moeten worden afgewogen tegen de extra tijdsinvestering.

De moeilijkheden bij de categorisering zijn van fundamentele aard. Een kwalitatieve analyse van onduidelijke uitspraken leidt niet tot duidelijke inzichten over het geldende recht. Deze onduidelijke uitspraken worden bij kwantitatieve analyse echter ingedeeld in verschillende onderscheidbare categorieën. Deze gegevens leiden vervolgens, in het ideale geval, tot heldere inzichten over het geldende recht.

Kwantitatieve analyse kan hierdoor leiden tot duidelijke conclusies op basis van onduidelijke uitspraken. Dit is een grote meerwaarde van deze methode. Het wordt echter problematisch als de gegevens zelf onbetrouwbaar zijn. De door kwantitatieve analyse geschapen helderheid is in die gevallen slechts illusoir. Dit blijkt echter niet duidelijk uit de gegevens. Zij geven

3 *MOGELIJKHEDEN EN BEPERKINGEN BIJ DE KWANTITATIEVE ANALYSE VAN
JURISPRUDENTIE VOOR DE BESTUDERING VAN GELDEND RECHT*

weer hoeveel uitspraken er in een bepaalde categorie vallen, maar in het geval van gegevens die slechts door één onderzoeker zijn gecodeerd (par. 4.5) niet of en in hoeverre de onderzoeker bij de indeling in categorieën heeft getobd en getwijfeld. Ook een uitgebreid codeboek en een tweede codeur kunnen niet alle bezwaren wegnemen. Zelfs als het coderen volgens alle regels van de kunst geschiedt, blijven de onderzoekers afhankelijk van onduidelijke vaststellingen van onduidelijke feiten (par. 4.5).

Kwantitatieve analyses zijn bovendien moeilijker te controleren dan juridisch-dogmatisch onderzoek. Een kwalitatieve analyse die is gebaseerd op een of twee uitspraken is eenvoudig te toetsen. Hoewel kwantitatieve analyse controleerbaar kan worden gemaakt door de dataset en het codeboek openbaar te maken (par. 4.5), zal een lezer niet snel inspecteren of de uitspraken juist zijn gecodeerd. Hoewel onderzoekers in een ideale wereld de tijd nemen om elkaars (empirische) onderzoek te herhalen en te controleren, besteden zij hun tijd in de praktijk liever aan nieuw, opvallender onderzoek. Het is daarom van groot belang om zo duidelijk mogelijk te zijn over de beperkingen van het onderzoek. De onderzoeker moet onder andere duidelijk maken of en in hoeverre de conclusies zijn gebaseerd op onbetrouwbare gegevens en of zij ook kunnen worden getrokken op basis van betrouwbare gegevens (par. 5).

Het voorgaande maakt duidelijk dat kwantitatieve analyse is gebonden aan verschillende belangrijke beperkingen en voorwaarden. Hieruit volgt echter niet dat de methode ongeschikt is voor de bestudering van het geldende recht. Ook een kwalitatieve analyse van jurisprudentie is aan veel van dezelfde beperkingen gebonden. Een kwantitatieve analyse vereist slechts een andere omgang met deze beperkingen. Zij biedt bovendien een grote meerwaarde. Kwantitatieve analyse voor de bestudering van geldend recht kan leiden tot duidelijke en significante nieuwe conclusies die niet mogelijk zijn op basis van een kwalitatieve analyse. Bovendien heeft ook de conclusie dat de feitenrechtspraak onduidelijk en tegenstrijdig is een grote waarde. Juist het feit dat zij is gebaseerd op een uitgebreide, systematische en kwantitatieve analyse zorgt voor een duidelijkheid en betrouwbaarheid die niet kan worden bewerkstelligd met behulp van kwalitatieve analyse (par. 5). Kwantitatieve analyse van jurisprudentie leidt hierdoor tot een groter inzicht in het geldende recht.



4 KWALITATIEVE INHOUDSANALYSE VAN RECHTSPRAAK MET NVIVO

Lessen uit een systematisch jurisprudentieonderzoek naar rechterlijke toetsing in het sportrecht

*Paul Verbruggen**

1 INTRODUCTIE

In dit hoofdstuk doe ik verslag van een kwalitatief jurisprudentieonderzoek naar de rechterlijke toetsing van besluitvorming en anderszins regulerend handelen door sportbonden.¹ Sportbonden zijn machtige spelers binnen de sportwereld. Niet alleen bepalen zij de regels van het spel. Ook zijn zij verantwoordelijk voor de organisatie, het bestuur en het beleid van de sport op nationaal niveau, doorgaans in afstemming met een internationale bond. Een dergelijke ‘monopolistische’ structuur van organisatie is kenmerkend voor de sport en haar private regulering.² Het gevolg is dat regulerend handelen door sportbonden de sportieve en financiële belangen van sporters, begeleiders en andere betrokkenen (zoals scheidsrechters, clubs, organisatoren en sponsors) hard kan raken. Zeker als deze partijen zich beroeps- of bedrijfsmatig bezighouden met de sport, kunnen die belangen groot zijn.

Met een zekere regelmaat wordt dan ook de stap gezet naar de burgerlijke rechter om het handelen van de sportbond rechtens te laten toetsen. Nu sportbonden doorgaans zijn opgericht als vereniging, krijgt die toetsing hoofdzakelijk vorm in het verenigingsrecht. De sportbond en degenen die krachtens de wet en statuten betrokken zijn bij zijn organisatie moeten zich op voet van artikel 2:8 Burgerlijk Wetboek (BW) jegens elkander gedragen overeenkomstig hetgeen door redelijkheid en billijkheid wordt verlangd. Of de sportbond dat in het concrete geval heeft gedaan, beoordeelt de rechter marginaal.

* Mr. dr. P.W.J. Verbruggen is universitair hoofddocent privaatrecht, Tilburg University en houder van de TPR-Wisselleerstoel aan de KU Leuven (2019-2021). Met dank aan Dennis van Gulik, Steff Huijbers, Huguette Knols en Marijne Mevius voor hun onderzoeksondersteuning als leden van het studentenpanel in het hier gepresenteerde onderzoek.

1 P.W.J. Verbruggen, ‘Sportbonden voor de rechter: Een systematische jurisprudentieanalyse naar de rechterlijke toetsing van regulerend handelen door sportbonden’, *AA* 2021, afl. 9, p. 827-840.

2 S. Jellinghaus & G. Hahn, ‘Sportrecht en de organisatiestructuur van sport’, in: S. Jellinghaus & G. Hahn (red.), *Capita sportrecht*, Oud-Turnhout/s-Hertogenbosch: Gompel & Svacina 2018, p. 21. Zie ook S. Weatherill, *Principles and Practice in EU Sports Law*, Oxford: Oxford University Press 2017, p. 1.

Onduidelijk is echter waar de burgerlijke rechter bij die marginale toetsing precies op let. De rechterlijke toetsing van regulerend handelen door verenigingen, in het bijzonder sportbonden, is daarmee een ‘blinde vlek’ die nog nauwelijks is verkend.³ Een praktisch gevolg daarvan is rechtszoekenden in onzekerheid verkeren over hun rechtspositie tegenover machtige sportbonden.

Om dit terrein beter in kaart te brengen is een systematisch jurisprudentieonderzoek opgezet. In deze bijdrage zet ik ten eerste kort uiteen wat de achtergrond, het doel en de centrale vraagstelling zijn van dit onderzoek (par. 2). Vervolgens ga ik uitgebreid in op de methoden die zijn aangewend om de rechtspraak te verzamelen, te selecteren en te analyseren (par. 3). Gedurende deze laatste fase van het onderzoek heb ik gebruikgemaakt van *NVivo*. Dat is een softwareprogramma dat in de sociale wetenschappen veelvuldig aangewend wordt voor de kwalitatieve analyse van data zoals tekst, afbeeldingen, geluid en video.⁴ Het gebruik daarvan heeft mijn onderzoek niet alleen ondersteund in het organiseren van en het behouden van het overzicht op de grote hoeveelheid te analyseren tekst. Ook heeft het tot nieuwe, verrijkte resultaten geleid met een hoger niveau van betrouwbaarheid en validiteit dan het geval waarin de software niet zou zijn gebruikt. Deze onderzoeksresultaten bespreek ik kort en ik zet daarbij uiteen hoe die zich verhouden tot bestaande aannames en standpunten in de literatuur (par. 4). Bij wijze van conclusie breng ik aan de hand van mijn ervaringen in kaart welke methodologische aandachtspunten er bestaan in de uitvoering van het geschetste jurisprudentieonderzoek (par. 5).

2 ACHTERGROND, DOEL EN VRAAGSTELLING VAN HET ONDERZOEK

Sportbonden zijn verenigingen van sportverenigingen. Zij vervullen een centrale rol in de regulering van sport.⁵ Sportbonden stellen de sport- en spelregels vast, alsmede de reglementen voor competitie (toernooien en wedstrijden), niveauselectie (promotie en degradatie) en de handhaving daarvan. Het grondwettelijk beschermde recht op vrijheid van vereniging geeft sportbonden een grote mate van autonomie in deze private regulering van de sport. Deze regulering volgt in de eerste plaats uit de statuten, reglementen en besluiten van sportbonden. Deze zijn bindend voor sporters, begeleiders en andere betrokkenen vanwege hun lidmaatschap bij die bond. Gebondenheid kan ook volgen uit hoofde van een overeenkomst, bijvoorbeeld die betreffende arbeid, licentie of sponsoring.

3 Vgl. K.A.M. van Vught, ‘Verenigingstuchtrecht voor de rechter: Soeverein straffen in eigen kring?’, *AA* 2019, afl. 12, p. 960.

4 QSR International, ‘NVivo qualitative data analysis’, www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home.

5 Zie Verbruggen 2021, p. 829-830.

In alle gevallen wordt de binding echter geclausuleerd door de wet en door de redelijkheid en billijkheid. Strijdigheid met de wet of de redelijkheid kan leiden tot de nietigheid, vernietigbaarheid of buiten toepassing verklaring van de betrokken regels.⁶ Een schending kan tevens een onrechtmatige daad van de bond jegens een lid opleveren, als gevolg waarvan de sportbond een gebod, verbod of plicht tot schadevergoeding kan worden opgelegd.

Of de vereniging in de gegeven omstandigheden in strijd met de redelijkheid en billijkheid handelde, beoordeelt de burgerlijke rechter op basis van de concrete omstandigheden van het geval. Daarbij toetst de rechter niet inhoudelijk, maar slechts marginaal.⁷ Het is dus niet aan de rechter te bepalen of, en zo ja welke maatregelen nodig waren in het concrete geval. Hij kan alleen achteraf beoordelen of de sportbond in de gegeven omstandigheden in redelijkheid tot de genomen maatregel heeft kunnen komen. De vraag is echter wat deze marginale toets behelst. Welke randvoorwaarden bezigt de rechter in die beoordeling? Hoe verhouden deze criteria zich tot elkaar en met welk gevolg worden ze toegepast?

Deze vragen hebben slechts tot op zekere hoogte aandacht gekregen van sportrechtjuristen.⁸ Auteurs in dat domein wijzen onder meer op het belang van algemene rechtsbeginselen in de rechterlijke toetsing. Daarbij is met name de aandacht voor het handelen op basis van bestaande bevoegdheden en conform de voorgeschreven procedures, hoor en wederhoor, motivering en het verbod op willekeur. Een discussie op basis van een systematische jurisprudentieanalyse ontbreekt daarin echter. Ook in de algemene literatuur over het verenigingsrecht treft men verwijzingen aan naar rechtsbeginselen die de marginale toetsing door de rechter gestalte geven. Zo wijst een gezaghebbende auteur op toetsingscriteria als misbruik van bevoegdheid, willekeur, precedentwerking en voorts 'elk ander criterium dat de rechter in het desbetreffende geval gepast acht'.⁹ De vraag rijst derhalve welke randvoorwaarden of beginselen de rechter precies betreft in zijn toetsing. Ook komt de vraag op naar het relatieve belang van die voorwaarden of beginselen. Op welke voorwaarden of beginselen beoordeelt de rechter het regulerend handelen van sportbonden vaker dan andere en leidt die beoordeling dan ook tot een toewijzing van de vordering?

In het licht van het voorgaande heeft het hier gepresenteerde jurisprudentieonderzoek tot doel gehad patronen bloot te leggen in de wijze waarop de burgerlijke rechter handelingen van sportbonden ter regulering van de sport toetst op hun juridische houdbaarheid. De

6 Vgl. art. 2:8 BW, art. 2:14 BW, art. 2:15 BW, art. 3:40 BW en art. 6:248 BW.

7 Zie in het algemeen C.H.C. Overes, T.J. van der Ploeg & W.J.M. van Veen, *Van vereniging en stichting, coöperatie en onderlinge waarborgmaatschappij*, Deventer: Kluwer 2013, p. 111-113.

8 Zie Verbruggen 2021, p. 830-831.

9 F. Kollen, *De vereniging in de praktijk* (Serie Recht en Praktijk nr. 89), Deventer: Kluwer 2008, p. 13.

centrale vraag die ten grondslag lag aan deze rechtspraakanalyse was welke randvoorwaarden de rechter aanlegt in de beoordeling van dat regulerend handelen, hoe deze voorwaarden zich tot elkaar verhouden en met welk resultaat ze worden toegepast. Aldus beoogde het onderzoek ook bij te dragen aan de theorievorming over de wijze waarop de burgerlijke rechter (en algemener het privaatrecht) handelingen van private regelgevers normeert.¹⁰

3 METHODE VAN RECHTSPRAAKANALYSE

Op welke wijze is het jurisprudentieonderzoek opgezet? Ik onderscheid hier de onderzoeksfasen van dataverzameling, -selectie en -analyse. In die laatste fase is gebruikgemaakt van *NVivo*, dat vanaf paragraaf 3.2.2 specifiek de aandacht krijgt.

3.1 *Dataverzameling en -selectie*

3.1.1 **Onderzoeksubject**

Het hierboven geschetste onderzoeksthema omvat geen vastomlijnde categorie van rechtspraak. Het thema wordt evenmin beheerst door één specifieke rechtsgrond, procedure of rechtbank. Hoe moet worden gezocht naar relevante rechtspraak is daardoor niet evident. Voor een goede dataverzameling en -selectie was het dus zaak om het object van onderzoek zo nauwkeurig mogelijk te omschrijven.

Er is derhalve gekozen om te werken met een *standaardcasuspositie* waaraan de rechterlijke uitspraak moest voldoen om te worden opgenomen in het onderzoek. Die positie betreft steeds een geschil tussen de sportbond en een betrokkene, waarbij een regulerende handeling van de bond de uitoefening van een economische activiteit door de betrokkene beïnvloedt. Als *betrokkenen* kunnen onder meer worden aangemerkt sporters, hun begeleiders (trainers, coaches, ouders of medische staf), verenigingen (clubs, afdelingen), officials (scheidsrechters, arbiters, juryleden) en organisatoren of sponsors. Zij zijn op grond van het verenigingsrecht, eventueel aangevuld door het contractenrecht, gebonden aan het *regulerend handelen* van de sportbond. Dat handelen kan bestaan uit normstelling (het maken, wijzigen of intrekken van collectieve regels of beleid), het houden van toezicht (de controle op naleving of enig andere oordeelsvorming over gestelde regels of beleid), of het handhaven van regels (het opleggen van sancties of andere (tucht)maatregelen bij niet-naleving).

10 Zie uitgebreid over die normerende functie van het privaatrecht P. Verbruggen (red.), *Regulating Private Regulators: Understanding the Role of Private Law (European Review of Private Law 2019, afl. 2, themanummer)*, Alphen a/d Rijn: Kluwer Law International 2019.

De litigieuze regulerende handeling van de sportbond moet de uitoefening van een *economische activiteit* van de betrokkene beïnvloeden. Regulering betreft immers het sturen en regelen van economische activiteiten. Er is sprake van een economische activiteit indien de betrokkene beroepsmatig of tegen betaling sport bedrijft, begeleidt of organiseert.¹¹ Verondersteld is dat bij wedstrijden, competities of toernooien op topsportniveau, zoals die plaatsvinden op het hoogste nationale niveau, op een nationaal, Europees of wereldkampioenschap of op de Olympische Spelen, sprake is van een economische activiteit. Uitgesloten van de analyse is derhalve rechtspraak over amateursport, tenzij er ondubbelzinnig sprake is van betaling aan de betrokkene voor diens activiteit.¹²

3.1.2 Dataverzameling

Deze paradigmatisch casuspositie vormde de basis voor het verzamelen van de rechtspraak. Het zoekformulier van de elektronische database van www.rechtspraak.nl is gebruikt om de gehele populatie van gepubliceerde rechterlijke uitspraken die beantwoorden aan de casuspositie te vinden.¹³ Deze wijze van dataverzameling kent uiteraard zijn beperkingen die tevens de resultaten van het onderzoek raken. Ten eerste worden niet alle rechterlijke uitspraken in de database opgenomen.¹⁴ Ten tweede zijn de zoekfuncties van het zoekformulier behorende bij de database (<https://uitspraken.rechtspraak.nl/>) beperkt. Zo kan er niet op partijnaam gezocht worden. Hoewel het mogelijk is het complete databestand van [rechtspraak.nl](http://www.rechtspraak.nl) te downloaden¹⁵ en vervolgens met behulp van data science-technieken te analyseren,¹⁶ heb ik dat vanwege het gebrek aan de vereiste technische kennis en ondersteuning niet gedaan. Ook zou die oplossing, zonder de zoekfunctie open te stellen voor derden, de transparantie en reproduceerbaarheid van de rechtspraakanalyse beperken. Er is dus gewoon gebruikgemaakt van [rechtspraak.nl](http://www.rechtspraak.nl) om de volledige populatie van publiceerde rechterlijke uitspraken te verzamelen.

11 Onder betaling worden tevens begrepen sponsoring en prijzengeld. Alleen een onkostenvergoeding is ontoereikend. Zie eveneens HvJ EG 12 december 1974, zaak 36/74, ECLI:EU:C:1974:140, r.o. 4, (*Walrave/Union Cycliste Internationale*) en HvJ EG 14 juli 1976, zaak 13/76, ECLI:EU:C:1976:115, r.o. 12 (*Donà/Manter*).

12 Zie Verbruggen 2021, p. 832.

13 Er bestaat ook veel arbitrale rechtspraak, bijvoorbeeld van het Instituut Sportrechtspraak. Zij is echter vanwege het onderzoeksdoel, dat ziet op de toetsing door de burgerlijke rechter, buiten beschouwing gelaten.

14 Zie Besluit selectiecriteria uitsprakendatabank www.rechtspraak.nl/Uitspraken/paginas/selectiecriteria.aspx. Zie over de *selection bias* die dit besluit tot gevolg heeft A.T. Marseille & M. Wever, 'Het verstrekken van niet-gepubliceerde uitspraken door de Rechtspraak: Gunst of recht?', *NJB* 2019/96, p. 123-124. Zie ook de bijdragen van Vols en Peeraer & Van Gestel in deze bundel.

15 Beschikbaar via www.rechtspraak.nl/Uitspraken/Paginas/Open-Data.aspx.

16 Zie over de mogelijkheden en beperkingen van dergelijke technieken, in het bijzonder *natural language processing* en *topic modeling*, de bijdragen van Dyevre en Tjong Tjin Tai in deze bundel.

Het gevolg van deze aanpak was dat het proces van dataverzameling behoorlijk tijdrovend is geweest. In een periode van twee maanden (oktober-november 2019) zijn voor iedere sportbond die lid is van de sportkoepel NOC*NSF en NOC*NSF zelf (n=78) verschillende zoekacties geformuleerd gebaseerd op de naam van de sportbond, zijn afkorting of alternatieve aanduidingen. Voor iedere zoekactie zijn de gebruikte zoektermen en eventuele Booleaanse operatoren in een Excel-bestand opgeslagen en er is per zoekopdracht een PDF-bestand gemaakt van de gevonden treffers zoals gepresenteerd op rechtspraak.nl.¹⁷ Gezamenlijk leverde die zoekacties ruim tweeduizend zoekresultaten op, betreffende meer dan zeshonderd unieke civielrechtelijke uitspraken.

3.1.3 Dataselectie

Iedere uitspraak is vervolgens geopend en op basis van een eerste lezing is beoordeeld of de uitspraak in grove mate voldeed aan de standaardcasuspositie. In het merendeel van de uitspraken was dat niet het geval. Na deze eerste lezing zijn ongeveer 120 unieke uitspraken geselecteerd voor nadere analyse. In al deze gevallen was de betrokken sportbond partij in een civielrechtelijk geschil, meestal als gedaagde in eerste aanleg en slechts zelden als eiser.

Bij een tweede en derde lezing van de gevonden uitspraken is nauwkeurig beoordeeld of de uitspraak voldeed aan de standaardcasuspositie zoals hierboven geschetst. Deze opvolgende leesrondes waren noodzakelijk voor een beter begrip van de uitspraken en hun onderlinge samenhang en voor de ontwikkeling van het codeerschema (zie par. 3.2.3). Als gevolg van deze nadere bestudering is ongeveer een derde van de uitspraken uitgesloten van analyse. Deze zaken zagen niet op de regulering van de sport of bleken geen inhoudelijke overwegingen van de rechter te bevatten vanwege procedurele en/of bewijstechnische redenen.¹⁸

Na deze opvolgende leesrondes is het aantal uitspraken teruggebracht naar 75 unieke civielrechtelijke uitspraken. De uitspraken betreffen eind- of tussenuitspraken in eerste aanleg, in hoger beroep en in cassatie, alsmede uitspraken in kort geding. Besloten is om een uitspraak van een lagere rechter die in hoger beroep of cassatie wordt bekrachtigd of (deels) opzij wordt gezet, tevens op te nemen in de selectie. Hetzelfde geldt voor een bodemzaak die volgt op een kort geding. Iedere uitspraak, ongeacht of die in hogere

17 De gebruikte zoekacties en het corresponderend aantal zoekresultaten zijn te raadplegen in een CSV-bestand via Tilburg University Research Portal, 'Sportbonden voor de rechter: Een systematische jurisprudentieanalyse naar de rechterlijke toetsing van regulerend handelen door sportbonden', <https://research.tilburguniversity.edu/en/publications/sportbonden-voor-de-rechter-eeen-systematische-jurisprudentieanaly>. Deze bewaarplaats betreft het institutionele repository van Tilburg University.

18 Zie Verbruggen 2021, p. 832-833.

instantie overeind blijft, bevat immers voorwaarden waarop het besluit of andersoortige handeling van de sportbond wordt geëvalueerd. Die kunnen dezelfde zijn als in de uitspraak in hogere instantie, zelfs als deze wordt vernietigd. De voorwaarden worden dan bevestigd als toetsingskader, maar hun invulling of onderlinge weging wijkt af. Als het hoger beroep wordt verworpen, kan evenzeer bevestiging van de voorwaarden plaatsvinden. Het aantal uitspraken dat zo is meegenomen is beperkt (8%), waardoor er geen onevenredig groot aantal uitspraken in het *sample* zitten die onderwerp zijn geweest van een beroeps- of bodemprocedure.

Ten slotte wordt opgemerkt dat een jaar na de eerste ronde van dataverzameling een tweede zoekactie is gehouden (oktober 2020). Die actie is uitgevoerd om eventuele nieuw gepubliceerde uitspraken mee te nemen in de analyse. Het resultaat van deze tweede ronde was dat één nieuwe uitspraak in de pool van uitspraken is opgenomen. Het totale aantal te analyseren uitspraken kwam daarmee op 76.¹⁹

3.2 Data-analyse

Voor de analyse van de rechtspraak is aansluiting gevonden bij een in de sociale wetenschappen veel gebruikte kwalitatieve methode voor de inhoudsanalyse van teksten, namelijk de *cross-case analysis* zoals omschreven door Miles en Huberman.²⁰ De stappen die deze methode voorschrijft voor het uitvoeren van een vergelijkende analyse, dat wil zeggen het ontwikkelen van het codeerschema, coderen, blootleggen van patronen in en tussen de uitspraken, zijn met behulp van NVivo uitgevoerd.²¹

3.2.1 Cross-case analysis

Voor de kwalitatieve analyse van data zoals tekst, afbeeldingen, geluid en video zijn in de sociale wetenschappen vele methoden ontwikkeld.²² De keuze voor deze methoden door de onderzoeker wordt mede bepaald door de voorhanden zijnde data (soort, hoeveelheid,

19 De geselecteerde uitspraken zijn in een CSV-bestand te raadplegen via Tilburg University Research Portal (voetnoot 17 *supra*).

20 M.B. Miles & A.M. Huberman, *Qualitative Data Analysis: An Expanded Sourcebook*, Thousand Oaks: SAGE Publications 1994, p. 172 e.v. zoals herzien door M.B. Miles, A.M. Huberman & J. Saldaña, *Qualitative Data Analysis: A Methods Sourcebook*, Thousand Oaks: SAGE Publications 2014, p. 100 e.v.

21 De gehele dataset, inclusief codeerschema, codering en vergelijkende analyses zijn als NVivo Project (NVP-bestand) te downloaden via Tilburg University Research Portal (voetnoot 17 *supra*).

22 Ik noem hier: *qualitative content analysis* van M. Schreier, *Qualitative Content Analysis in Practice*, Londen: SAGE Publications 2012; *thematic content analysis* van R.E. Boyatzis, *Transforming Qualitative Information: Thematic Analysis and Code Development*, Thousand Oaks: SAGE Publications 1998; en *grounded theory* van A.L. Strauss & J. Corbin, *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, Thousand Oaks: SAGE Publications 1998.

mate van gestructureerdheid, enz.) en het onderzoeksdoel. *Cross-case analysis* zoals voorgesteld door Miles en Huberman houdt in algemene termen in dat een specifiek en bepaald geval (een *case*) van een breder maatschappelijk fenomeen wordt bestudeerd en vergeleken met andere gevallen. Die vergelijkende analyse is erop gericht de overeenkomsten en verschillen tussen de gevallen in kaart te brengen en te verklaren, zodat de bevindingen het individuele geval overstijgen en zo te generaliseren zijn.²³ Met de gesignaleerde ontwikkelingen en patronen kunnen bestaande inzichten in de literatuur worden verrijkt en nieuwe hypothesen over de uitkomsten in soortgelijke zaken worden ontwikkeld.

Het hier gepresenteerde jurisprudentieonderzoek sluit goed aan bij dezelfde uitgangspunten. Als gezegd, richt het onderzoek zich op het blootleggen van patronen in rechterlijke uitspraken (de *cases*) die een beoordeling inhouden van contentieuze reguleringshandelingen door sportbonden (het maatschappelijke fenomeen). Zo beoogt het onderzoek bij te dragen aan theorievorming over de wijze waarop de burgerlijke rechter handelingen van private regelgevers normeert.

In meer algemene zin sluit de methode van Miles en Huberman voor vergelijkende inhoudsanalyse nauw aan bij wat een rechtswetenschapper gewend is te doen indien hij zich toelegt op het analyseren van een reeks rechterlijke uitspraken betreffende een bepaald leerstuk of een bepaalde casus. Hall en Wright benoemen die gelijkenis in hun beroemde artikel over 'Systematic Content Analysis of Judicial Opinions' als volgt:

'On the surface, content analysis appears simple, even trivial, to some. Using this method, a scholar collects a set of documents, such as judicial opinions on a particular subject, and systematically reads them, recording consistent features of each and drawing inferences about their use and meaning. This method comes naturally to legal scholars because it resembles the classic scholarly exercise of reading a collection of cases, finding common threads that link the opinions, and commenting on their significance. But content analysis is more than a better way to read cases. It brings the rigor of social science to our understanding of case law, creating a distinctively legal form of empiricism.'²⁴

Voor het doen van een rechtspraakanalyse levert *cross-case analysis* zoals omschreven door Miles en Huberman in de eerste plaats een systematische en gestructureerde wijze van

23 Miles & Huberman 1994, p. 173; Miles, Huberman & Saldaña 2014, p. 101.

24 M. Hall & R. Wright, 'Systematic Content Analysis of Judicial Opinions', *California Law Review* 2008, afl. 1, p. 64.

inhoudsanalyse op. Deze methode kent vier stappen die toegepast op rechtspraakanalyse als volgt kunnen worden omschreven:

1. verkennen van de uitspraak en het beschrijven van zijn kenmerken;
2. verklaren van de uitkomst van de uitspraak;
3. verkennen van patronen tussen verschillende uitspraken en het beschrijven van hun overeenkomsten en verschillen;
4. ordenen van uitspraken in gevalstypes en het verklaren van patronen.²⁵

De analyse van een individuele uitspraak (stap 1 en 2: *within-case analysis*) gaat logischerwijs vooraf aan de vergelijkende analyse van uitspraken (stap 3 en 4: *cross-case analysis*). Toch zijn de stappen niet lineair, maar iteratief. De methode moedigt het tussentijds analyseren en vergelijken van cases aan, ook voordat dataverzameling en -selectie afgerond is. Anders dan bijvoorbeeld Hall en Wright suggereren,²⁶ trekken Miles en Huberman dus geen scherpe grens tussen de fasen van dataverzameling en -selectie enerzijds en analyse anderzijds.²⁷ In de hier gepresenteerde rechtspraakanalyse zijn ook in verschillende rondes uitspraken verzameld, geselecteerd en geanalyseerd (zie par. 3.1.2 en 3.1.3).

Miles en Huberman schrijven het gebruik van *codes* en *codering* voor om een goed begrip te krijgen van de kenmerken van de data die belang zijn voor het doel van het onderzoek. Naar hun idee is coderen analyseren.²⁸ Codes zijn labels die een bepaalde betekenis geven aan de originele data of daarvan afgeleide informatie behorende bij een onderzoek.²⁹ Ze worden gekoppeld aan stukjes data (in het geval van tekst: een woord, frase, zin, alinea, paragraaf of geheel document) om gelijksoortige data aan te duiden of te categoriseren. De onderzoeker kan door te coderen data sneller terugvinden, onderscheiden of clusteren

25 Zie voor een vergelijkbaar stappenplan: L. Spencer e.a., 'Analysis in Practice', in: J. Ritchie e.a. (red.), *Qualitative Research Practice: A Guide for Social Science Students and Researchers*, Thousand Oaks: SAGE Publications 2014, p. 295-345.

26 Hall & Wright 2008, p. 107 ('once cases are selected, experts in content analysis outline four basic steps researchers should follow in coding any material').

27 Miles & Huberman 1994, p. 50.

28 Miles & Huberman 1994, p. 56 ('Coding is analysis. To review a set of [data] and to dissect them meaningfully, while keeping the relations between the parts intact, is the stuff of analysis. This part of analysis involves how you differentiate and combine the data you have retrieved and the reflections you make about this information.'). Zie ook: Miles, Huberman & Saldaña 2014, p. 72 ('[C]oding is analysis. Some research methodologists believe that coding is merely technical, preparatory work for higher level thinking about the study. But we believe that coding is deep reflection about and, thus, deep analysis and interpretation of the data's meanings.'). Zie voor een ander perspectief de bijdrage van Vols in deze bundel.

29 Zie ook J. Saldaña, *The Coding Manual for Qualitative Researchers*, Thousand Oaks: SAGE Publications 2013, p. 3-4.

zodat ze verder kunnen worden geanalyseerd.³⁰ Coderen is dus een heuristiek – een methode – om te vinden.³¹

3.2.2 CAQDAS: NVivo

Het coderen van rechterlijke uitspraken draait om het systematisch categoriseren en rangschikken van tekst en het eenvoudig kunnen terugvinden daarvan voor analyse. Een veel gebruikte methode is het werken met ordners, mappen, indexkaarten, post-its, tabjes, kleurmarkeringen, mindmaps, of samenvattingen op A3-formaat of een whiteboard. Deze aanpak is handzaam en effectief voor jurisprudentieonderzoek van beperkte omvang. Voor veel kwalitatieve projecten – zo ook het mijne – is deze methode vanwege het aantal te analyseren uitspraken en/of de te onderzoeken aspecten (variabelen) niet langer praktisch.³² Voor dergelijke projecten is het gebruik van *computer-assisted qualitative data analysis software* (CAQDAS) een uitkomst.³³

CAQDAS is niet bedoeld ter vervanging van het doen van kwalitatieve analyse. Het beoogt de kwaliteit en efficiëntie van de kwalitatieve analyse die de onderzoeker uitvoert juist te verbeteren.³⁴ In de eerste plaats ondersteunt CAQDAS onderzoekers in het slim organiseren en beheren van kwalitatieve data. De data van verschillende bronnen kunnen onder één project overzichtelijk bij elkaar worden gebracht en gestructureerd. Daarnaast maakt CAQDAS het voor onderzoekers mogelijk om aan de data nieuwe informatie te koppelen zoals codes, memo's en andere analytische gegevens, die het denkproces van de onderzoeker weergeven. Die informatie kan vervolgens worden geordend, samengevoegd of herzien, zodat in opeenvolgende rondes een steeds gedetailleerder en geïntegreerd begrip van de data wordt verkregen. CAQDAS verschaft in dit analytisch proces van samenvatten, interpreteren en vergelijken de nodige flexibiliteit. Voor gezamenlijke onderzoeksprojecten kan CAQDAS ook efficiënt zijn. Zo krijgen verschillende onderzoekers binnen een project toegang tot de data en kunnen zij die gezamenlijk beheren, analyseren en hun bevindingen controleren op betrouwbaarheid. Alles samengenomen wordt breed aangenomen dat CAQDAS de efficiency en nauwkeurigheid van kwalitatieve analyses bevordert en het eenvoudiger wordt om kwaliteit aan te tonen.³⁵

30 Miles & Huberman 1994, p. 56-57.

31 Miles, Huberman & Saldaña 2014, p. 73.

32 De 76 geselecteerde uitspraken beslaan een totaal van 662 pagina's A4-formaat die gecodeerd zijn aan de hand van 70 unieke codes, aangevuld met een veelvoud van analytische memo's en vergelijkende noties.

33 Miles & Huberman & Saldaña 2014, p. 93.

34 Zie in het algemeen: Miles, Huberman & Saldaña 2014, p. 46-52 en L. Spencer e.a., 'Analysis: Principles and Processes', in: Ritchie e.a. 2014, p. 287-291 (ieder met verdere verwijzingen naar standaardliteratuur over CAQDAS). Zie voor een discussie in het domein van de rechtswetenschap: M. Salehijam, 'The Value of Systematic Content Analysis in Legal Research', *Tilburg Law Review* 2018, afl. 1, p. 34-42.

35 Spencer e.a. 2014, p. 289.

NVivo behoort tot de meest geavanceerde CAQDAS-pakketten.³⁶ In algemene zin ondersteunt NVivo onderzoekers in het beheren van data, het organiseren van de ideeën en inzichten verkregen door analyse van de data, het doorzoeken van data en bijbehorende inzichten, het visualiseren daarvan en het erover rapporteren. NVivo is eerder door rechtswetenschappers gebruikt voor de analyse van rechtspraak, maar wijdverbreid is dat gebruik zeker niet.³⁷

Voor het hier gepresenteerde project zijn lang niet alle NVivo-functionaliteiten gebruikt die voor deze doeleinden beschikbaar zijn. In het analytisch proces is wel gebruikgemaakt van de functionaliteiten die NVivo biedt voor:

- het maken, opsplitsen, aggregeren en verbinden van codes (via *nodes*);
- het optekenen van gedachten over bestaande concepten, voorwaarden en patronen in aantekeningen (via *annotations* en *memos*);
- het maken van vergelijkende tabellen met de gecodeerde kerneigenschappen van de uitspraken (via *framework matrixes* en *case classification sheets*);
- het ontdekken en visualiseren van verbanden tussen uitspraken en codes (via *matrix coding queries*, *concept maps* en *comparison diagrams*);
- en het uitvoeren van betrouwbaarheidsanalyses (via *coding comparison queries*).

Hoe deze functionaliteiten gebruikt zijn voor het hier gepresenteerde jurisprudentieonderzoek wordt hierna duidelijk. In het gebruik van deze functionaliteiten maakt NVivo het ook steeds mogelijk om via hyperlinks terug te grijpen op de originele, primaire data en die (weer) in context te bestuderen. Zo wordt de onderzoeker in het ontwikkelen van meer algemene of abstracte inzichten – de kern van het analytisch proces – geholpen in het bewaren van de verbinding met de onderliggende data. Ook kan worden voorkomen dat de concepten en patronen te veel los komen te staan van de context. Zeker voor juridisch onderzoek naar de invulling van open normen, waarin de concrete feiten en omstandigheden doorslaggevend zijn, is dat van groot belang.³⁸

Voor zich te wenden tot NVivo doet een klassiek getrainde rechtswetenschapper er goed aan zich basisprincipes van kwalitatief onderzoek in de sociale wetenschappen eigen te

36 Zie uitgebreid P. Bazeley & K. Jackson, *Qualitative Data Analysis with NVivo*, Londen: SAGE Publications 2013 en D. Mortelmans, *Kwalitatieve analyse met Nvivo* (Leuven: Acco, 2017).

37 Een voorbeeld is het Gentse proefschrift van Myriam Salehijam, *Mediation and Commercial Contract Law Towards a Comprehensive Legal Framework*, Abington/New York: Routledge 2021. Ik ben niet bekend met systematisch jurisprudentieonderzoek dat in Nederland met behulp van NVivo is uitgevoerd.

38 Vgl. de zienswijzen van Wolters in deze bundel betreffende kwantitatieve rechtspraakanalyse.

maken.³⁹ Het werken met NVivo dwingt onderzoekers namelijk bij het opzetten en uitvoeren van hun onderzoek te denken in de typische stappen van een kwalitatieve analyse. NVivo is evenwel neutraal ten opzichte van de gebruikte methode van onderzoek.⁴⁰ Het softwarepakket leent zich dus ook voor het gebruik van andere kwalitatieve methoden dan *cross-case analysis* van Miles en Huberman.

NVivo is een gebruiksvriendelijk CAQDAS-pakket. Iedere enigszins geïnteresseerde rechtswetenschapper kan de basis van het softwareprogramma zelfstandig onder de knie krijgen met behulp van online tutorials,⁴¹ NVivo Help⁴² en het handboek van Bazeley en Jackson.⁴³ Het effectief gebruik van NVivo zal niettemin sterk verbeterd worden door het volgen van een geavanceerde cursus. QSR International (de ontwikkelaar van NVivo) biedt dergelijke cursussen aan, maar er zijn ook onafhankelijke experts die deze cursussen in een academische setting aanbieden.⁴⁴ De ervaringen uit de cursus voeden de inzichten die uit andere bronnen verkregen kunnen worden.

3.2.3 Ontwikkeling codeerschema

Als gezegd, wordt coderen in de methode van Miles en Huberman als analyse gezien. Omdat coderen leidt tot een beter begrip van de data en het onderzoeksonderwerp, leidt het tot verdere dataverzameling. Dit gegeven impliceert dat de fases van dataverzameling en -selectie nog niet afgerond hoeven te zijn voordat begonnen kan worden met het codeerproces.⁴⁵ Een eerste stap in dat proces is het ontwikkelen van een voorlopig codeerschema (of codeboek).⁴⁶ Die ontwikkeling kan deductief of inductief gedreven zijn.⁴⁷

39 Dit advies klinkt wat tegenstrijdig omdat rechtswetenschap soms wordt geschaard onder het brede domein van sociale wetenschappen en zijn methoden in algemene zin nauw aansluiten bij die kwalitatieve methoden uit dat domein. De nadruk die wordt gelegd in de methodologische literatuur van de sociale wetenschappen (zie onder meer voetnoot 22 *supra*) op het proces van dataverzameling, -selectie en -analyse, zal de rechtswetenschapper als overweldigd ervaren. Niettemin zal de rechtswetenschapper vanuit zijn of haar ervaring en perspectief eenvoudig aansluiting weten te vinden bij de voorgeschreven kwalitatieve methoden. Vgl. Salehijam 2018, p. 40.

40 Bazeley & Jackson 2013, p. 4.

41 QSR International, 'NVivo by QSR', www.youtube.com/user/QSRInternational/playlists.

42 QSR International, 'NVivo Help', <https://help-nv.qsrinternational.com/20/win/Content/table-of-contents.htm>.

43 Bazeley & Jackson 2013. Recent verscheen de derde editie (2019).

44 De ECPR Summer en Winter Schools in Methods and Techniques georganiseerd door de *European Consortium for Political Research* (www.ECPR.eu) bieden standaard een cursus 'Introduction to NVivo for Qualitative Data Analysis' of 'Advanced Qualitative Data Analysis' aan. Die laatste cursus volgde ik in 2018.

45 Of zoals Miles en Huberman stellen: 'Coding is not just something you do to "get the data ready" for analysis but (...) something that drives ongoing data collection. It is a form of early (and continuing) analysis. It typically leads to a reshaping of your perspective and of your instrumentation for the next round.' Miles & Huberman 1994, p. 65.

46 Zie algemeen voor verschillende strategieën voor de ontwikkeling van een codeerschema: Saldaña 2013.

47 Miles & Huberman 1994, p. 58-61 en Miles, Huberman & Saldaña 2014, p. 81.

Deductieve codes krijgen onder meer vorm door het conceptuele kader, de onderzoeksvraag en wat daarover al bekend is in de literatuur. Inductieve codes zijn daarentegen empirisch gedreven en komen voort uit de primaire data. Zij zijn daardoor contextueel van aard.

Voor de ontwikkeling van het codeerschema voor het hier beschreven jurisprudentieonderzoek is een combinatie van deductief en inductief coderen gebruikt. Die aanpak sluit goed aan bij het doel van het onderzoek, namelijk het blootleggen van patronen in de wijze waarop de burgerlijke rechter het regulerende handelen van sportbonden beoordeelt op hun juridisch houdbaarheid. Op basis van het juridische kader voor dergelijke rechterlijke toetsing en de bestaande inzichten in de literatuur daarover is een voorlopige lijst van codes opgesteld. Deze lijst is inductief aangevuld door opeenvolgende codeerrondes. In de eerste ronde zijn 30 uitspraken van de voorlopige selectie van ongeveer 120 unieke civielrechtelijke uitspraken (zie par. 3.1.3) in detail gelezen en gecodeerd. Dat aantal was voldoende om een goede indruk te krijgen van de vorderingen, verweren, feitelijke en juridische grondslagen, de rechterlijke overwegingen en uitkomsten in de uitspraken.

Op basis van deze eerste codeerronde zijn codes ontwikkeld voor de kenmerken van een uitspraak. Deze codes kunnen in NVivo worden weergegeven als *case nodes* of *cases*. Iedere uitspraak werd als een *case* gecodeerd, met daaraan gekoppeld verschillende categorieën (*case classifications*) en bijbehorende kenmerken (*case attributes*).⁴⁸ Voor de inhoudelijke kenmerken van de uitspraken zijn ook codes ontwikkeld, weergegeven in NVivo als *nodes*. Deze nodes betreffen de door eiser ingestelde vorderingen, de aangevoerde grondslag(en) daarvoor, de toewijzing of afwijzing van de vorderingen. Ook de concepten, gezichtspunten en omstandigheden betrokken in de rechterlijke beoordeling van de vorderingen zijn onder nodes gecodeerd.

Aandachtspunt in deze fase bij het maken van de nodes was zo dicht mogelijk bij de tekst blijven en dus zo beschrijvend mogelijk zijn. Dat was voor de formele cases met geijkte juridische categorieën, makkelijker gedaan dan voor de meer inhoudelijke nodes inzake de rechterlijke beoordeling. In de literatuur bestond immers geen duidelijk beeld van hoe die rechterlijke beoordeling plaatsvindt. Door bewust beschrijvende codes te hanteren in deze eerste codeerronde (en dus inductief te werken) werd het voorlopige codeerschema verrijkt met codes die gegrond zijn in de dataset.⁴⁹ Zo is voorkomen dat te snel geïnterpre-

48 Voorbeelden van gebruikte categorieën met hun bijbehorende kenmerken zijn 'rechtsprekende instantie' (rechtbank, hof, Hoge Raad), 'soort procedure' (eerste aanleg, hoger beroep, cassatie, kort geding, enz.), soort beslissing (tussenvonnis of -arrest, eindoordeel) en 'uitkomst' (afwijzing van alle vorderingen, toewijzing van minimaal een vordering).

49 Miles & Huberman 1994, p. 58. Zie ook de *grounded theory*-methode als verkondigd door Straus & Corbin 1998.

teerd werd en voorbij is gegaan aan de veelzijdigheid van de rechterlijke overwegingen en formuleringen.⁵⁰

Een tweede aandachtspunt bij het maken van de nodes was consistentie ten aanzien van het segment van tekst dat gecodeerd wordt. De eenheid van codering (*unit of coding*) kan een woord, frase, zin, alinea, paragraaf of zelfs een geheel document zijn.⁵¹ Het onderzoeksdoel bepaalt logischerwijs de keuze voor een eenheid. Voor een beschrijvende en vergelijkende analyse als het onderhavige onderzoek moet de eenheid voldoende context geven om snel (weer) te begrijpen waarom de code aan die tekst gekoppeld is, zonder daarbij excessief lang te zijn. Ook was het voor de beoogde analyse van belang dat de gecodeerde eenheid het samen voorkomen van nodes kan identificeren. Een tekstpassage kan namelijk iets zeggen dat tegelijkertijd past bij twee of meer verschillende codes. Door een eenheid van codering te kiezen die niet te eng, maar ook niet te breed is, kan in NVivo met behulp van *matrix coding queries* het naast elkaar bestaan of samenvallen van codes duidelijk worden gemaakt. De queries helpen de onderzoeker in het ontdekken en weergeven van patronen in de data (zie par. 3.2.5).

Na deze eerste ronde van coderen (en analyseren) zijn de ontwikkelde nodes kritisch naast elkaar gelegd om te bezien welke nodes eventueel konden worden samengevoegd of konden worden ondergebracht als (sub)categorie (*child nodes* in Nvivo) van een node van een abstracter niveau (een *parent node*). Om het codeerschema zo simpel mogelijk te houden (en dus dicht bij de tekst) is deze tweelagige structuur aangehouden.⁵² In dit proces van revisie is gebruikgemaakt van termen en concepten die zowel in de tekst als in de literatuur tot uitdrukking zijn gekomen. Voorbeelden hiervan zijn ‘terughoudendheid in de toets door de rechter’, ‘motivering’ en ‘hoor en wederhoor’. De naamgeving van de nodes is daarmee deels herzien en op een hoger interpretatief niveau gebracht. Ook zijn de nodes logisch geordend, zodat zij min of meer overeenkomen met de orde waarin zij worden behandeld in de uitspraak.⁵³

Een belangrijk aspect van de revisie betrof tevens het herzien en preciseren van definitiebeschrijvingen van de nodes. Om analytische waarde te hebben moeten codes helder

50 Zie ook Bazeley & Jackson 2013, p. 69 (‘Early work with text and concepts is about laying the foundation for identification of key themes in the data. Beware of jumping to conclusions too early, however. Constantly challenge your first ideas by drawing comparisons, by purposively sampling diverse cases, or by reviewing what the literature says on the topic.’).

51 Bazeley & Jackson 2013, p. 72 en 89.

52 Vgl. Miles & Huberman 1994, p. 61.

53 Dat wil zeggen: vorderingen, grondslagen, inhoudelijke overwegingen voor de beoordeling, toegewezen vorderingen, afgewezen vorderingen.

gedefinieerd zijn. Het is dan ook waarschijnlijker dat gedurende de gehele looptijd van een onderzoeksproject hetzelfde onderwerp onder dezelfde code wordt gecodeerd.⁵⁴ Ten behoeve van deze interne consistentie heeft iedere node als geheugensteuntje een korte omschrijving gekregen en, indien nodig, een beslisregel in de trant van ‘tekstpassage over A niet onder node X, maar onder node Y’.⁵⁵ Met een heldere definitiebeschrijving zijn nodes tevens onderling onderscheidend en wordt voorkomen dat een tekstpassage wordt gecodeerd onder twee nodes, terwijl die passage slechts een enkel onderwerp omvat.

Om de betrouwbaarheid van het codeerschema en de validiteit en reproduceerbaarheid van de analyse te vergroten is het voorlopige codeerschema voorgelegd aan een *studenten-panel*.⁵⁶ In vier verschillende sessies zijn de vier leden van dit panel bekend gemaakt met het doel en de achtergrond van het onderzoek, en getraind in het omgaan met het voorlopige codeerschema. Aldus is het schema getest op een totaal van tien verschillende uitspraken. De inzichten verkregen tijdens deze sessies hebben geleid tot het toevoegen en verwijderen van codes, het wijzigen van definitieomschrijvingen en het aanpassen in de structuur van het codeerschema. In de laatste twee sessies, waarvoor de panelleden vier uitspraken codeerden van in totaal 42 pagina’s, was sprake van een grote gelijkensis in de codering door de leden en mijzelf. De mate van overeenstemming in de codering tussen de panelleden en mijzelf – de zogenoemde *intercoder reliability rate* – varieerde tussen de 85 en 94%.⁵⁷ Dit is ook het niveau van betrouwbaarheid dat nodig wordt geacht voordat wordt overgegaan tot het coderen van de gehele data.⁵⁸

Ten slotte moet aangetekend worden dat al deze revisierondes eenvoudig waren uit te voeren met NVivo. Met het softwarepakket kan de naamgeving en definitieomschrijving van de codes te allen tijde worden gewijzigd en kunnen de codes doorlopend worden geordend, samengevoegd, gesplitst of anderszins herzien. De verbinding met de reeds

54 Dit geldt zowel voor individuele als gezamenlijke projecten. Zie Miles & Huberman 1994, p. 63 en Miles, Huberman & Saldaña 2014, p. 84.

55 Vgl. Boyatzis 1998, p. 30 e.v. over ‘good’ codes.

56 Zie in het algemeen over de praktijk van *code checking* en de waarde daarvan voor de betrouwbaarheid van het codeerschema en de validiteit van de onderzoeksresultaten: Miles & Huberman 1994, p. 64 en Miles, Huberman & Saldaña 2014, p. 84. Zie over de inzet en training van studenten voor codering: Hall & Wright 2008, p. 109-112 en de bijdrage van Elise Goossens, Katalien Bollen en Alain Laurent Verbeke aan deze bundel.

57 De test is uitgevoerd via MS Excel aan de hand van de formule die Miles & Huberman (1994, p. 64) voorstellen. Zie ook Miles, Huberman & Saldaña 2014, p. 85. De uitgevoerde test is als MS Excel XLSX-bestand te raadplegen via Tilburg University Research Portal (voetnoot 17 *supra*). De test is ook via NVivo uit te voeren op basis van verschillende gebruikersprofielen (zie verder par. 3.2.4).

58 Miles & Huberman 1994, p. 64 en Miles, Huberman & Saldaña 2014, p. 85.

gecodeerde tekstpassages blijft daarbij in stand.⁵⁹ Zonder de software zou veel tijd voor analyse verloren gaan.

3.2.4 Coderen

De diverse revisierondes hebben geleid tot een codeerschema met in totaal 71 unieke nodes.⁶⁰ De nodes zijn geordend in de categorieën: vorderingen, grondslagen, inhoudelijke overwegingen (waaronder bijvoorbeeld formele gebreken, hoor en wederhoor, motivering, transparantie en evenredigheid), toegewezen vorderingen en afgewezen vorderingen. De nodes zijn inductief en deductief geleid en zowel beschrijvend als interpretatief van aard.

Na de vaststelling van het codeerschema is het totale aantal te analyseren uitspraken (n=76) in anderhalve maand gecodeerd. Er is voor gekozen deze codeerronde zelfstandig uit te voeren. Praktische bezwaren stonden eraan in de weg deze codering gezamenlijk te doen, bijvoorbeeld met de leden van het studentenpanel. Het werken met meerdere codeurs vereiste namelijk toegang tot NVivo en de nodige training voor het juiste gebruik van de software, waarvoor geen financiële ruimte en tijd binnen het project bestond.⁶¹ In deze omstandigheden lag het voor de hand om de betrouwbaarheid van de codering en de validiteit en reproduceerbaarheid van de daaruit verkregen resultaten te waarborgen via een *intracoder reliability* test.⁶² Die test – ook wel een *test-retest reliability* toets genoemd – meet de mate van consistentie van codering gedaan door een enkele codeur op verschillende intervallen (t1 en t2).⁶³ Een tussenpose van zeven tot veertien dagen wordt als accuraat beschouwd voor kwalitatief onderzoek. Die periode is namelijk voldoende lang om het onwaarschijnlijk te maken dat de codeur op t2 wordt beïnvloed door de codering op t1. Tevens is zij kort genoeg om de werkelijke verschillen in het gebruik van het zich eigengemaakte codeerschema naar voren te laten komen.

Om de *intracoder reliability* te meten zijn tien uitspraken in het begin van de reeks en tien uitspraken in het midden van de reeks nogmaals gecodeerd. Voor alle twintig uitspraken

59 Zie voor de verschillende functionaliteiten die NVivo biedt voor het maken, herzien en structureren van een codeschema (inclusief de niet gebruikte tool *auto coding*): Bazeley & Jackson 2013, p. 95-117.

60 Het volledige codeerschema, inclusief *nodes*-omschrijvingen, is als CSV-bestand te raadplegen via Tilburg University Research Portal (voetnoot 17 *supra*).

61 In tegenstelling tot enkele andere universiteiten zit NVivo niet in het standaardsoftwarepakket beschikbaar voor studenten of wetenschappelijk medewerkers van Tilburg University. Een gebruikerslicentie voor het pakket moest dus gekocht worden.

62 Schreier schrijft over de relatieve sterkte van *inter*- en *intracoder* reliability: 'Conceptually speaking, inter-coder reliability is always 'stronger' and therefore preferable, as long as a second suitable coder is available. But if no suitable second coder can be found, intra-coder reliability based on one expert coder is preferable to inter-coder reliability based on the coding of the researcher and a second unsuitable coder.' Schreier 2012, p. 192.

63 Miles & Huberman 1994, p. 64. Zie ook Schreier 2012, p. 167.

is een interval van veertien dagen aangehouden voor hercodering. Via een *coding comparison query* is het via NVivo mogelijk om zowel het overeenkomstpercentage, als de Cohens Kappa coëfficiënt te berekenen.⁶⁴ Deze coëfficiënt is een statistische maat van overeenkomst tussen coderingen gecorrigeerd voor de kans dat een overeenkomst per toeval bestaat. Met een overeenkomst percentage van 99% en een Kappa van 0.98 was de interne consistentie van mijn codering zeer hoog.⁶⁵

3.2.5 Patronen blootleggen

Met het voltooiën van de codering stopt de kwalitatieve analyse uiteraard niet. Codering is immers een wijze waarop informatie systematisch kan worden teruggevonden, onderscheiden en geordend, zodat daaraan betekenis kan worden gegeven.⁶⁶ Voor het hier gepresenteerde jurisprudentieonderzoek volgt die betekenis uit het inzichtelijk maken van patronen in de wijze waarop de burgerlijke rechter handelingen van sportbonden ter regulering van de sport toetst op hun geldigheid. Om die patronen bloot te leggen, suggereren Miles en Huberman voor hun *within-case* en *cross-case analysis* onder meer het gebruik van *matrix displays*.⁶⁷ Dat zijn tabellen van rijen en kolommen waarin de kernelementen van een uitspraak zijn samengevat en georganiseerd, zodat de onderzoeker deze in één oogopslag kan bezien. Door de verschillende uitspraken (*cases*) in de rijen te plaatsen en de kernelementen (*nodes*) in de kolommen, kon ik per uitspraak de relaties tussen de kernelementen overzien en de uitkomst in het concrete geval verklaren. Tegelijkertijd kon ik ook de verschillende uitspraken met elkaar vergelijken, zodat de regelmatigheden of herhalingen tussen de verschillende uitspraken snel opgemerkt werden.

Hoewel deze vergelijkende analyse handmatig of via een programma als MS Excel kan worden gedaan, biedt NVivo wederom verschillende functionaliteiten om dit proces te ondersteunen. Matrix displays kunnen via de functie *Framework matrices* worden gemaakt.⁶⁸ Door de cases als rijen te gebruiken en de nodes als kolommen kan op basis van de gebruikte coderingen de matrix worden ingevuld met aantekeningen en citaten. Met de functie *Auto summary* vult NVivo de cellen automatisch aan met de tekst die in de uitspraak onder de

64 Bazeley & Jackson 2013, p. 293-296. Zie ook: QSR International, 'Run a coding comparison query', http://help-nv11.qsrinternational.com/desktop/procedures/run_a_coding_comparison_query.htm.

65 De uitkomst moet boven 0.75 Kappa zijn om stabiliteit aan te tonen. De test is als CSV-bestand te raadplegen via Tilburg University Research Portal (voetnoot 17 *supra*).

66 Vgl. voetnoot 29-31 *supra*. Zie ook Bazeley & Jackson 2013, p. 243 ('I've coded all my data. Now what do I do?').

67 Miles & Huberman 1994, p. 93-102 en p. 178-181.

68 Bazeley & Jackson 2013, p. 257-260. Zie ook: QSR International, 'About framework matrices', http://help-nv11.qsrinternational.com/desktop/concepts/about_framework_matrices.htm?rhsearch=framework%20matrixes&rhsyns=%20.

desbetreffende node is gecodeerd. De celtekst kan weer handmatig worden aangepast. Figuur 1 geeft een voorbeeld van een framework matrix voor drie geselecteerde uitspraken.

Figuur 1 Vergelijkende NVivo framework matrix

	A : 2. Grondslag(en) vordering	B : 4. Formele gebreken	C : 5. Strijdigheid met wet of verdrag	D : 21. MOOIE QUOTES
2 : Fortuna Sittard v KNVB	Art. 2:15 b/c + fundamentele rechtsbeginselen	transparantie- en motiveringsplicht geschonden; hoor en wederhoor, onafhankelijkheid en onpartijdigheid in besluitvorming niet gewaarborgd; niet de vereiste deskundigheid of expertise		Onafhankelijkheid: Licentiereglement is het voorkomen van ieder tegenstrijdig belang en het bewerkstelligen van een onafhankelijke en onpartijdige beslissing. Het gaat hier om
3 : Oudbondscoach v KNGU	Art. 6:74, art. 6:162		Laster: het doen van onrechtmatige uitlatingen die eiser in eer en goede naam schaden 4.2. Bij de beoordeling van de vraag of de KNGU onrechtmatig	Motivering, hoor en wederhoor: Dit alles klemt te meer omdat het onderzoek waar de KNGU naar verwijst niet naar behoren is uitgevoerd. [gedaagde sub 2] heeft, zo blijkt uit zijn uitlatingen, de grieven van de
4 : Visser v NTB en Ironman	Art. 2:15 c + fundamentele rechtsbeginselen + redelijkheid en billijkheid + play-of-field beginsel	Schending van in reglementen neergelegde bevoegdheden. In dat kader worden besproken: willekeur, vertrouwen en verwachtingen. Overwegingen van hoor en wederhoor en motivering staan daar los van.	Art. 6 EVRM (hoor en wederhoor) 4.34. Indien Ironman al een grondslag had gehad om [eiseres] te disqualificeren, dan had zij in ieder geval daarbij zorgvuldiger dienen te handelen met in het bijzonder	Hoor en wederhoor: Evenmin is gehandeld in overeenstemming met de algemene rechtsbeginselen zoals hiervoor genoemd, waaronder met name het beginsel dat iemand ten aanzien van wie een besluit wordt genomen, daarover dient

diverse media feitelijk onjuist en negatief over [...] [eiser] uitgelaten naar aanleiding van twee artikelen in het tijdschrift 'Tijdschrift'. Ten onrechte suggereerde de KNGU met deze uitlatingen en beschuldigingen dat [eiser] zich schuldig zou hebben gemaakt aan hetgeen hem in voornoemde artikelen in 'Tijdschrift' was verweten. De KNGU heeft deze suggestie niet met deugdelijk feitenmateriaal kunnen staven. De voorzieningenrechter heeft de KNGU veroordeeld tot het plaatsen van deze rectificatie.

Koninklijke Nederlandse Gymnastiek Unie;
5.5. veroordeelt de KNGU om aan [eiser] een divangsom te betalen van € 5.000,00 voor iedere dag, of gedeelte daarvan, dat zij niet aan de in 5.4. uitgesproken hoofdveroordeling voldoet, tot een maximum van € 50.000,00 is bereikt;
5.6. veroordeelt de KNGU in de proceskosten, aan de zijde van [eiser] tot op heden begroot op € 1.182,82;

Een aanvullende manier van zoeken van patronen met behulp van NVivo is via *queries*. NVivo biedt verschillende manieren aan waarop gecodeerde data kunnen worden doorzocht met een query.⁶⁹ Met een *matrix coding queries* kunnen verschillende cases, nodes en hun onderliggende categorieën tegenover elkaar worden afgezet in een tabel om associaties, zoals het naast elkaar bestaan of samenvallen van codes, duidelijk te maken.⁷⁰ Die queries kunnen zowel voor een individuele uitspraak, als voor verschillende uitspraken worden opgezet. In het eerste geval kan aan de hand van de (frequentie van) voorkomende nodes worden afgelezen hoe de uitspraak tot stand is gekomen (*within-case analysis*). In het tweede geval kan een vergelijkende analyse tussen meerdere cases en/of nodes worden gedaan (*cross-case analysis*).⁷¹ Figuur 2 illustreert de matrix coding query voor de drie uitspraken opgenomen in de hiervoor afgebeelde vergelijkende framework matrix.

69 Zie uitgebreid Bazeley & Jackson 2013, p. 244-255.

70 Bazeley & Jackson 2013, p. 89-90 en 250-253. Zie ook: QSR International, 'Run a matrix coding query', http://help-nv11.qsrinternational.com/desktop/procedures/run_a_matrix_coding_query.htm?rhsearch=matrix%20coding%20query&rhsyns=%20.

71 Bazeley & Jackson 2013, p. 257-258.

Figuur 2 Matrix coding query individuele cases

	A : Fortuna Sittard v KNVB ▾	B : Oudbondcoach v KNGU ▾	C : Visser v NTB en Ironman ▾
4b. Strijdigheid met reglement of andere interne regeling ▾	6	0	4
5. Strijdigheid met wet of verdrag ▾	0	1	1
7. Onafhankelijkheid, onpartijdigheid ▾	3	0	0
8. Deskundigheid, expertise ▾	2	0	0
9. Motivering ▾	1	2	2
10. Transparantie ▾	2	0	0
11. Vertrouwen en verwachtingen ▾	0	0	2
12. Hoor en wederhoor ▾	1	1	2
14. Willekeur ▾	0	0	1
15c. Belangen van eiser ▾	0	2	0
15e. Andermans belangen ▾	0	2	0
17. Mogelijkheid tot herziening of beroep ▾	1	0	0

Figuur 2 laat zien welke nodes minimaal één keer in de inhoudelijke overwegingen van de rechter in de desbetreffende uitspraak voorkomen. De tabel geeft zo een snel overzicht van hoe de rechterlijke beoordeling per uitspraak is opgebouwd. De vergelijking tussen de uitspraken leert onder meer dat in deze drie uitspraken, waarin steeds een vordering van de eiser werd toegewezen, de rechter aandacht besteedt aan de elementen van deugdelijke motivering van de litigieuze handeling door de sportbond en de vraag of de bond in die handeling hoor en wederhoor heeft toegepast. Niet de frequentie van de nodes is belangrijk in deze kwalitatieve analyse, maar hun betekenis in de context.⁷² Door op de cel te klikken haalt NVivo de gecodeerde stuk(ken) tekst naar voren en kan bijvoorbeeld worden bepaald of de overwegingen betreffende het achterwege blijven van een deugdelijke motivering en het toepassen van hoor en wederhoor van doorslaggevend belang waren voor de toewijzing van de vorderingen. In het geval van de zaak *Oudbondcoach/KNGU* was dat het geval, terwijl in de andere twee zaken deze overwegingen ondersteunende argumenten vormden voor de uitkomst voor partijen.

Deze strategie, van het doorzien van patronen in individuele uitspraken en die vergelijken met andere uitspraken, is voor verschillende categorieën van uitspraken uitgevoerd. Zo is gekeken naar associaties tussen (1) de nodes inzake de juridische grondslag van de vordering en die inzake de door de rechter gebezigde overwegingen; (2) de uitkomst van de uitspraak (toewijzing van minimaal één vordering of afwijzing van de vorderingen) en de rechterlijke overwegingen; en (3) de rechterlijke overwegingen onderling. Die associaties kunnen aan de hand van matrix coding queries worden weergegeven. NVivo maakt het mogelijk om deze queries vorm te geven door middel van Booleaanse operatoren (AND, OR, NOT) en zogenoemde *proximity operators* (NEAR, PRECEDING, SURROUNDING).⁷³ Het samenvallen van verschillende nodes in een tekstpassage wordt uitgelicht met een query

⁷² Het heeft derhalve ook weinig zin statistische analyses uit te voeren op deze frequenties.

⁷³ Bazeley & Jackson 2013, p. 262-264.

met de operator AND en het naast elkaar voorkomen van nodes (ongeacht hun volgorde in de gecodeerde tekst) met de operatoren NEAR en SURROUNDING.

De inzichten verkregen uit de diverse matrix coding queries kunnen op verschillende manieren worden verwerkt in NVivo. Ten eerste kunnen ze handmatig worden toegevoegd als aantekeningen in een framework matrix om de verbanden in die matrix te scherpen. Ook kunnen ze worden gevisualiseerd. NVivo biedt met de tools *concept maps* en *comparison diagrams* diverse mogelijkheden daartoe.⁷⁴

Wellicht belangrijker nog is om de inzichten op te slaan in een analytische *memo*.⁷⁵ Met het optekenen van die inzichten krijgt de analyse vorm en wordt duidelijk welke regelmatigigheden of herhalingen er zijn en hoe die kunnen worden verklaard. Deze memo's geven de gevonden patronen betekenis en de vormen aldus de basis voor de verkregen onderzoeksresultaten.⁷⁶ De memo's vormen ook de basis voor wat Miles en Huberman *patrooncodes* noemen: analytische memo's waarin een thema, categorie, of verklaring wordt vastgelegd. Daar waar beschrijvende en interpretatieve codes de data samenvatten, geven patrooncodes die codes in combinatie betekenis.⁷⁷ In NVivo is een memo de omschrijving en achtergrond van de patrooncode, terwijl de label van de patrooncode als een *result node* kan worden opgeslagen.⁷⁸

4 RESULTATEN

Tot welke resultaten heeft het jurisprudentieonderzoek geleid? Ik licht een drietal resultaten uit en duid hun verhouding tot de bestaande aannames en standpunten in de literatuur.⁷⁹

Ten eerste is duidelijk geworden dat tussen de rechterlijke toetsing van het regulerend handelen van een sportbond op basis van verenigingsrechtelijke en verbintenisrechtelijke grondslagen geen grote principiële verschillen bestaan. De rechter toetst op basis van het verenigingsrecht, in het bijzonder artikel 2:14, 2:15 en 2:8 BW, of de sportbond de wet of enige interne regelingen (statuten, reglementen) heeft geschonden en of de litigieuze handeling in de gegeven omstandigheden naar maatstaven van redelijkheid en billijkheid onaanvaardbaar was. In dat laatste geval dient de rechter terughoudendheid te betrachten:

74 Bazeley & Jackson 2013, p. 217 e.v.

75 Bazeley & Jackson 2013, p. 245-246.

76 Zie ook Miles & Huberman 1994, p. 72 en Miles, Huberman & Saldaña 2014, p. 95-96.

77 Miles & Huberman 1994, p. 69-72 en Miles, Huberman & Saldaña 2014, p. 86-90.

78 Bazeley & Jackson 2013, p. 248.

79 Voor een volledige bespreking van de onderzoeksresultaten verwijs ik naar Verbruggen 2021, p. 834-840.

er is slechts ruimte voor rechterlijk ingrijpen als een verenigingsbesluit in redelijkheid niet genomen had kunnen worden.⁸⁰

Dezelfde verenigingsrechtelijke uitgangspunten gelden echter ook in grote lijnen voor de toetsing van vorderingen wanneer deze vorderingen hun grondslag vinden in het verbintenisrecht, zoals wanprestatie (art. 6:74 BW), onrechtmatige daad (art. 6:162 BW) of de vernietiging van een bindend advies (art. 7:904 lid 1 BW). Ook in die gevallen draait de rechtelijke toetsing om de vraag of de toepasselijke regels zijn gevolgd en de litigieuze handeling redelijk of zorgvuldig was. Hoewel anderen reeds op deze analogie wezen voor artikel 7:904 lid 1 BW,⁸¹ verschaft de hier gepresenteerde analyse diverse voorbeelden uit de rechtspraak inzake wanprestatie of onrechtmatige daad.⁸² Een dergelijke analoge toepassing van verenigingsrechtelijke uitgangspunten in verbintenisrechtelijke relaties moet worden begrepen tegen de achtergrond van de collectieve dimensie van de toepasselijke regels ter regulering van de sport. Die regels hebben immers tot doel de sport en de organisatie en uitoefening daarvan te uniformeren en gelijke voorwaarden voor alle betrokkenen te scheppen. Ook als de regels anders dan via verenigingsstructuren toepassing krijgen, moet bij die toepassing ruimte worden gelaten voor beleid en keuzes. Daarbij past een terughoudende rechterlijke toets naar de redelijkheid of zorgvuldigheid van handelen door de sportbond.

Ten tweede volgt uit het onderzoek dat de rechterlijke beoordeling een kern van elementaire voorwaarden bevat waaraan het regulerend handelen van de sportbond vaak getoetst wordt. Deze voorwaarden of beginselen kunnen neergelegd zijn in statuten en reglementen van de vereniging, en daarin in groter detail zijn uitgewerkt.⁸³ Maar ook als dat niet zo is, toetst de rechter het handelen regelmatig aan deze beginselen.⁸⁴ Zij geven derhalve aan waarop de marginale toetsing van de rechter zich frequent toespitst om te beoordelen of de sportbond redelijk en billijk, althans zorgvuldig heeft gehandeld.

⁸⁰ Vgl. Jellinghaus & Hahn 2018, p. 30-32.

⁸¹ Zie bijvoorbeeld Overes, Van der Ploeg & Van Veen 2013, p. 152 en Van Vught 2019, p. 961-962.

⁸² In dertien uitspraken (= 17%) verloopt de beoordeling van vorderingen gestoeld op wanprestatie of onrechtmatige daad langs de genoemde verenigingsrechtelijke uitgangspunten.

⁸³ Zo is bijv. in het Licentiereglement van de KNVB voor de organisatie van het Betaald Voetbal neergelegd dat ieder besluit van de Licentiecommissie moet voldoen aan onder meer de beginselen van motivering, transparantie en onafhankelijke en onpartijdige besluitvorming. Zie over de ernstige tekortkomingen in de naleving daarvan Rb. Utrecht (vzr.) 17 juni 2009, ECLI:NL:RBUTR:2009:BI8398 (*Fortuna Sittard/KNVB*).

⁸⁴ Zie bijv. Rb. Arnhem (vzr.) 11 januari 2008, ECLI:NL:RBARN:2008:BC2763, r.o. 4.6 (*Eiseres/Taekwondo Bond Nederland*), Rb. Oost-Nederland (vzr.) 26 maart 2013, ECLI:NL:RBONE:2013:BZ5500, r.o. 4.3.3 (*Oud-bondscoach/KNGU*), Rb. Gelderland (vzr.) 21 oktober 2015, ECLI:NL:RBGEL:2015:6447, r.o. 5.2 (*Jurylid/KNHS*) en Rb. Midden-Nederland (vzr.) 29 mei 2019, ECLI:NL:RBMNE:2019:2376, r.o. 4.34 (*Visser/NTB en Ironman*).

Om welke elementaire voorwaarden gaat het hier? Figuur 3 laat de matrix coding query zien waarbij de voorwaarden van de beoordeling als nodes afgezet worden tegen de uitkomst van de case.

Figuur 3 Matrix coding query frequentie codering afgezet tegen de uitkomst

	A : Case:Uitkomst = Toewijzing vordering ▾	B : Case:Uitkomst = Afwijzing vordering ▾
4a. Strijdigheid met statuut	3	10
4b. Strijdigheid met reglement of andere interne regeling	29	52
5. Strijdigheid met wet of verdrag	10	12
5a. Vrijheid van vereniging	1	0
6. Oneigenlijk gebruik van een bevoegdheid	1	0
7. Onafhankelijkheid, onpartijdigheid	5	7
8. Deskundigheid, expertise	3	3
9. Motivering	21	20
10. Transparantie	10	17
11. Vertrouwen en verwachtingen	5	16
12. Hoor en wederhoor	8	24
13. Gelijkheid	5	11
14. Willekeur	4	4
15a. Ernst van de schending door eiser	1	10
15b. Schuld of verwijtbaarheid van eiser	2	6
15c. Belangen van eiser	15	32
15d. Weging van bijzondere omstandigheden aan de zijde van eiser	1	9
15e. Andermans belangen	14	22
16. Subsidiariteit	4	8
17. Mogelijkheid tot herziening of beroep	3	7

Uit de query volgt duidelijk dat de rechterlijke toetsing ten eerste is gericht op de *naleving van interne procedurele regels* (nodes 4a en 4b). De frequentie van codering is het hoogst onder die nodes voor uitspraken zowel in kolom A (n=27), als in kolom B (n=49). Ook volgt dat de beoordeling vaak een afweging van de belangen van de eisers (node 15c) tegenover die van de sportbond of andere betrokkenen (node 15e) inhoudt. Deze afweging behelst, samen met de nodes van 15a, 15b en 15d, een toets van *evenredigheid of proportionaliteit*: staan het handelen van de sportbond en de daarmee gemoeide organisatorische en collectieve belangen in redelijke verhouding tot de ernst en verwijtbaarheid van het gedrag van de betrokkene en de mogelijke negatieve gevolgen voor hem of haar?

Als we verder zien, blijkt ook dat het element van *motivering* (node 9) frequent wordt geëvalueerd door rechters. Deze voorwaarde blijkt 21 keer betrokken te worden in de rechterlijke beoordeling die ten grondslag lag aan 17 van de 27 unieke uitspraken waarin minimaal één vordering is toegewezen. Het beginsel van motivering speelt dus een rol in 63% van de uitspraken waarin een vordering werd toegewezen. Verdere analyse leert dat in bijna de helft (8 uit 17 uitspraken) de gebrekkige motivering door de sportbond van doorslaggevende betekenis was voor de rechterlijke beoordeling. Ook het beginsel van *hoor en wederhoor* krijgt vaak de aandacht, meer nog in uitspraken waarin alle vorderingen zijn afgewezen (24 keer in 18 uitspraken), dan waarin een toewijzing volgt (8 keer in 7

uitspraken). Het horen van de betrokkenen en hen in staat stellen hun visie en belangen kenbaar te maken, gaat op basis van de gecodeerde uitspraken sportbonden dus beter af dan het deugdelijk motiveren van beslissingen.

Minder frequent, maar toch met enige regelmaat toetst de rechter het handelen van de sportbond aan andere rechtsbeginselen. Die toetsing ziet, zoals Figuur 3 laat zien, onder meer op de vraag of de sportbond vooringenomen is geweest en of er tegenstrijdige belangen waren die de besluitvorming hebben beïnvloed (7. Onafhankelijkheid/onpartijdigheid); of de sportbond voldoende openbaarheid heeft betracht en de relevante informatie aan betrokkenen heeft verleend (10. Transparantie); of de sportbond voorafgaand aan de besluitvorming of anderszins handelen verwachtingen heeft gewekt waarop door betrokkene redelijkerwijs mocht worden vertrouwd (11. Vertrouwen en verwachtingen); of de sportbond gelijke gevallen gelijk behandelt (13. Gelijkheid); of de sportbond arbitrair of willekeurig handelt (14. Willekeur); en of er een redelijk alternatief voorhanden was, zodat de betrokkenen door het handelen van de sportbond minder getroffen zouden worden in hun gerechtvaardigde belangen (16. Subsidiariteit).

De genoemde randvoorwaarden of beginselen komen ook in bepaalde combinaties voor in uitspraken. Dit derde resultaat wordt duidelijk aan de hand van Figuur 4.

Figuur 4 Matrix coding query aantal unieke uitspraken waarin nodes gezamenlijk voorkomen

	9. Motivering ▾	10. Transparantie ▾	11. Vertrouwen en verw... ▾	12. Hoor en wederh... ▾	13. Gelijkheid ▾
7. Onafhankelijkheid, on... ▾	6	4	6	6	2
9. Motivering ▾	32	7	9	15	8
10. Transparantie ▾	7	18	6	8	5
11. Vertrouwen en verw... ▾	9	6	15	10	6
12. Hoor en wederhoor ▾	15	8	10	25	6
13. Gelijkheid ▾	8	5	6	6	15
14. Willekeur ▾	4	3	4	3	3
16. Subsidiariteit ▾	4	3	4	4	4
17. Mogelijkheid tot herz... ▾	5	3	0	3	2

Deze matrix coding query geeft weer in hoeveel unieke uitspraken geselecteerde nodes gezamenlijk voorkomen in de gecodeerde dataset. Uit de tabel volgt bijvoorbeeld dat 57 uitspraken (75% van het totaal) een evaluatie van ofwel motivering (n=32) of wel hoor en wederhoor (n=25) omvat. Van de 25 uitspraken met een toets van het beginsel van hoor en wederhoor, betreffen er 15 (60%) gelijktijdig een beoordeling van de motivering. Dit bevestigt de sterke aanwezigheid van deze twee beginselen als de in de praktijk meest aangewende toetsstenen voor de rechterlijke beoordeling van het regulerend handelen van sportbonden.

Verder laat de query bijvoorbeeld een associatie zien tussen enerzijds motivering en vertrouwen en verwachtingen (n=9) en anderzijds motivering en gelijkheid (n=8). Het samenvallen van deze nodes kan logischerwijs verklaard worden doordat de rechter meer belang hecht aan een deugdelijke motivering daar waar de sportbond voorafgaand aan zijn besluitvorming gerechtvaardigde verwachtingen heeft gewekt bij de betrokkenen, maar daarvan alsnog is afgeweken. Ook als een geval vergelijkbaar is met dat van de betrokkene, zal de rechter van de sportbond verlangen dat hij deugdelijk motiveert waarom hij anders heeft beslist. De associatie (n=10) tussen enerzijds hoor en wederhoor en anderzijds vertrouwen en verwachtingen laat zich echter niet eenvoudig verklaren. Deze vaststelling benadrukt dat niet de aangeduide aantallen doorslaggevend zijn, maar de kwalitatieve betekenis van de achterliggende rechterlijke overwegingen in hun context. De NVivo queries zijn hulpmiddelen om tot dwarsverbanden en patronen te komen. Zij geven die patronen niet noodzakelijk weer.

In de literatuur is in algemene bewoordingen gewezen op het belang van rechtsbeginselen voor de rechterlijke beoordeling van verenigingshandelingen.⁸⁵ Van Vught wordt het meest concreet door te stellen dat als minimale ondergrens ('elementaire rechten') voor het verenigingstuchtrecht de rechtsbeginselen van hoor en wederhoor en motivering gelden.⁸⁶ Het hier gepresenteerde onderzoek bevestigt deze beginselen als kernvoorwaarden. Het vestigt echter ook de aandacht op de rol van andere beginselen en hun onderlinge samenhang in de rechterlijke toetsing van verenigingshandelingen.

In meer algemene zin onderstreept het onderzoek de gewichtige rol die rechtsbeginselen toekomen in de rechterlijke controle op privaatrechtelijke relaties waarin een gezagsverhouding centraal staat.⁸⁷ Sportbonden genieten als private verenigingen de grondwettelijke vrijheid om hun organisatie en regulering zelfstandig en autonoom vorm te geven. Vertoont hun handelen echter ernstige gebreken die indruisen tegen erkende rechtsbeginselen, dan

85 Zie bijvoorbeeld Overes, Van der Ploeg & Van Veen 2013, p. 153 ('Bij andere dan dit soort conflicten binnen de vereniging zullen beginselen van burgerlijke rechtspleging een rol kunnen spelen bij het bepalen van mogelijke strijd met de redelijkheid en billijkheid.'). Zie specifiek voor sportrecht: Jellinghaus & Hahn 2018, p. 28-32.

86 Van Vught 2019, p. 968.

87 Zie over de gemeenschappelijke rol en betekenis van rechtsbeginselen in de normering van zowel publiek-, als privaatrechtelijke handelingen in asymmetrische gezagsverhoudingen: D. Oliver, 'Common Values in Public and Private Law and the Public/Private Divide', *Public Law* 1997, p. 630-646 en W. van Gerven & S. Lierman, *Algemeen Deel. Veertig jaar later: privaat- en publiekrecht in een meergelaagd kader van regelgeving, rechtsvorming en regeltoepassing*, Mechelen: Kluwer 2010, p. 484-485 en 524-534 (wijzende op de gemeenschappelijke onderbouw van privaat- en publiekrecht en hun gedeelde 'beginselen van behoorlijk handelen' voor gezagsverhoudingen).

kan de burgerlijke rechter ingrijpen na daartoe te zijn aangezocht door de betrokkene. Daaraan staat de verenigingsautonomie niet in de weg.⁸⁸

Deze vaststelling rijmt ook met recente jurisprudentie die private geschilbeslechting in verenigingsverhoudingen onderwerpt aan fundamentele beginselen van goede procesorde volgens uit artikel 6 van het Europees verdrag voor de rechten van de mens (EVRM).⁸⁹ Ook in de rechtspraak geanalyseerd in het kader van het hier gepresenteerde onderzoek hanteren rechters deze maatstaf ter beoordeling van de juridische houdbaarheid van tuchtrechtelijke maatregelen, terwijl in de literatuur geen overeenstemming is over de mate waarin artikel 6 EVRM moet gelden voor verenigingstuchtrecht.⁹⁰ Door op deze wijze de rechterlijke toetsing vorm te geven biedt de burgerlijke rechter een belangrijke mate van rechtsbescherming aan sporters en andere betrokkenen die door het regulerend handelen van machtige sportbonden in hun belangen worden geraakt.

5 CONCLUSIE – METHODOLOGISCHE LESSEN

Het opzetten en uitvoeren van het hier omschreven kwalitatieve jurisprudentieonderzoek in NVivo nodigt uit tot een uiteenzetting van zijn methodologische lessen. Veel is al opgetekend, maar ik benadruk nog een aantal aandachtspunten bij wijze van conclusie van dit hoofdstuk.

Vooropgesteld moet worden dat het hier gepresenteerde jurisprudentieonderzoek naar zijn aard niet ver afstaat van de kwalitatieve analyse die een klassiek-juridische aanpak zou inhouden. Ook met die aanpak zou de beantwoording van de centrale onderzoeksvraag vermoedelijk tot relevante inzichten hebben geleid. Het systematisch verzamelen van uitspraken, het lezen en herlezen daarvan en het maken van notities en samenvattingen van terugkerende elementen om zo tot vergelijkende inzichten en patronen in diverse uitspraken te komen, is evenzeer zonder een expliciet sociaalwetenschappelijke kwalitatieve methode en de software van NVivo mogelijk.

⁸⁸ Vgl. Van Vught 2019, p. 965.

⁸⁹ Zie bijvoorbeeld: EHRM 2 oktober 2018, appl.nrs. 40575/10 en 67474/10, ECLI:CE:ECHR:2018:1002 JUD004057510 (*Mutu & Pechstein/Zwitserland*), waarover T. Barkhuysen, 'Straatsburg en de rechtsbescherming van topsporters', *NJB* 2018/2049, afl. 39, p. 2049.

⁹⁰ Overes, Van der Ploeg & Van Veen 2013 adviseren om art. 6 EVRM te volgen voor verenigingstuchtrechtspraak (p. 153). Van Vught 2019 is terughoudender (p. 968) en verwijst naar F.P. Ölçer, 'Eerlijk proces en tuchtrechtelijke procedures', *AA* 2016, afl. 7, p. 511 e.v., waarin wordt betoogd dat art. 6 EVRM niet een-op-een kan worden toegepast op (verenigings)tuchtrecht.

Naarmate een jurisprudentieanalyse grootschaliger is qua aantal te bestuderen uitspraken of de daarin gelegen variabelen, doet de onderzoeker er mijns inziens echter verstandig aan meer systematiek in zijn of haar analyse te brengen.⁹¹ Dat is precies wat een methode als *cross-case analysis* van Miles en Huberman biedt. Het doorwerken van deze methode leidt tot een gestructureerde manier van hoe de volledige dataset wordt geanalyseerd.⁹² Daarbij past bijvoorbeeld het organiseren van meerdere, opvolgende leesrondes, het maken van een overzicht van relevante onderwerpen en thema's, het vormen en herschikken van de codering en het komen tot patrooncodes. De hele exercitie staat in het teken van het verkrijgen van een diepere betekenis van de individuele uitspraken, hun onderlinge verwantschap bezien in het licht van de gehele dataset, en het niet te snel trekken van conclusies.

Cross-case analysis geeft een methodologisch raamwerk dat dichtbij de doelstellingen en intuïtie van de rechtswetenschapper staat. Mits systematisch doorgevoerd, zorgt de methode niet alleen voor een beter begrip van de dataset, maar leidt zij tot uitkomsten die een grotere mate van objectiviteit, betrouwbaarheid en validiteit genieten.⁹³ De methode werkt daarbij niet verstikkend: zij biedt de nodige flexibiliteit in de vormgeving en uitvoering van de analyse. Niet alle stappen en aanwijzingen die Miles en Huberman geven, behoeven opvolging: het gaat – ook volgens hen – om het aanwenden van de technieken voor zover ze bruikbaar zijn voor de beoogde analyse.⁹⁴ Hierdoor is de methode in mijn ervaring een geschikte methode om tot een systematische kwalitatieve jurisprudentieanalyse te komen.

Het gebruik van NVivo ondersteunt de uitvoering van die analyse op verschillende wijzen. De software stelt de onderzoeker in staat de te analyseren uitspraken slim te organiseren en te beheren, zodat het overzicht bewaard blijft. Ook biedt NVivo de gewenste flexibiliteit voor dataverzameling, -selectie en -analyse, zodat de codering en analyse stapsgewijs en in verschillende rondes kan plaatsvinden. Cruciaal voor de kwalitatieve analyse is de mogelijkheid om via de gebruikte codes terug te keren naar de relevante tekstpassages. Zo is de primaire bron steeds voorhanden, zodat de passage in de originele context opnieuw kan worden bestudeerd. Framework matrices en matrix coding queries zijn vervolgens slimme tools van NVivo voor het kunnen visualiseren van de data en het opsporen van patronen.

91 Zie tevens de bijdrage van Wijnjens in deze bundel over 'gestructureerd' jurisprudentieonderzoek.

92 Vgl. Salehijam 2018, p. 39 voor andere methoden van kwalitatieve inhoudsanalyse.

93 Vgl. Hall & Wright 2008, p. 64.

94 Miles & Huberman 1994, p. 5. Vgl. Hall & Wright 2008, p. 100 (waarin ze in plaats van methodologische 'regels', spreken over '*guidelines or standards*' omdat niet ieder project zich leent voor hun volledige toepassing).

Cross-case analysis en het gebruik van NVivo in het kader van rechtspraakanalyse vergt op sommige punten echter bijzondere aandacht. Een eerste punt betreft dat van status en autoriteit van de te analyseren rechterlijke uitspraken. Het sociaalwetenschappelijke domein waarbinnen deze methode en software oorspronkelijk zijn ontwikkeld, gaat uit van een *gelijkwaardigheid van de te bestuderen data*, zoals interviews, observaties of vignetten. Ook voor een vergelijkende rechtspraakanalyse zoals hier gepresenteerd geldt dat deze analyse het beste werkt als de uitspraken van min of meer gelijke rang zijn en aldus dezelfde autoriteit genieten.⁹⁵ Als bijvoorbeeld een arrest van de Hoge Raad uitgesproken bepalend is voor het te onderzoeken terrein, zal het gelijkwaardig behandelen van dat arrest in een vergelijkende kwalitatieve analyse met voornamelijk feitenrechtspraak tot onzuivere resultaten leiden. De (on)gelijkwaardigheid van de te analyseren uitspraken beïnvloedt derhalve de opzet en het doel van de rechtspraakanalyse. Door dit gegeven in acht te nemen hebben onderzoekers, zoals ook die in deze bundel, zich in hun werk voornamelijk tot doel gesteld blinde vlekken in wetenschappelijke literatuur, jurisprudentie of wetgevingsbeleid te belichten, dan wel de daarin bestaande aannames en conventies te ontkrachten, op terreinen waarbij de gelijkheid van uitspraken goeddeels onomstreden is.

Een verwant aandachtspunt betreft de *juridische correctheid of overtuigingskracht van de te analyseren rechtspraak*. Of een uitspraak, of elementen daarvan, goed of slecht is, staat vaak centraal in het werk van een rechtswetenschapper. Zo staat de annotator doorgaans een kritisch bespreking van de uitspraak in het licht van opvattingen in de literatuur of eerdere jurisprudentie voor.⁹⁶ Een onderzoeker die zich richt op het ontwaren van patronen in de rechterlijke beoordeling van een grotere verzameling van uitspraken zal echter niet zozeer geïnteresseerd zijn in deze normatieve dimensie van een individuele uitspraak. De onderzoeksvragen die voor dit type analyse leidend zijn, hebben als focus de tekst van de rechterlijke beoordeling, zoals de aangevoerde rechtsgronden, de feiten die de rechter betreft in zijn overwegingen, of hoe hij zijn beoordeling construeert.⁹⁷ Indien eenmaal de analyse is afgerond, kan de onderzoeker de normatieve implicaties van zijn of haar bevindingen bespreken. Zich daarbij verhouden tot opvattingen in bestaande literatuur, jurisprudentie en/of wetgevingsbeleid sterkt de wetenschappelijke inbedding en maatschappelijke relevantie van de rechtspraakanalyse.

95 Vgl. Hall & Wright 2008, p. 63-64 ('Content analysis also works best when the judicial opinions in a collection hold essentially equal value, such as where patterns across cases matter more than a deeply reflective understanding of a single pivotal case.') en p. 83-84 ('Coding and counting cases usually assumes that the information from one opinion is potentially as relevant as that from any other opinion').

96 Zie in het algemeen: G. van Dijk, *De kwaliteit van de juridische annotatie. Een empirische studie naar kenmerken en kwaliteitsindicatoren*, Den Haag: Boom Juridische uitgevers 2011, p. 96-97.

97 Niettemin is het bij een kwalitatieve analyse zoals hier gepresenteerd, mogelijk om een andere weging te geven aan uitspraken waarin de juridische houdbaarheid van de uitspraak discutabel is. Sterker nog, een correcte weergave van de onderzoeksresultaten zou een dergelijke weging duidelijk moeten maken.

PAUL VERBRUGGEN

Een laatste punt van aandacht betreft de *consistentie in de eenheid van codering*. Door consistent te zijn in deze eenheid kan met behulp van NVivo matrix coding queries worden ontwaard in welke mate nodes naast elkaar bestaan of samenvallen. De queries vormden zo een belangrijk instrument met behulp waarvan de beschreven patronen in kaart konden worden gebracht. Zonder consistent gebruik van de gekozen codeereenheid zou de potentie van de NVivo queries voor een kwalitatieve rechtspraakanalyse als hier gepresenteerd onbenut blijven of tot onzuivere bevindingen kunnen leiden.

5 **LESSEN UIT EEN GROOTSCHALIGE RECHTSPRAAKANALYSE MET STUDENTEN**

*Een empirisch onderzoek naar samenwoning in het
Belgisch relatievermogensrecht*

*Elise Goossens, Katalien Bollen en Alain-Laurent Verbeke**

1 **INLEIDING**

In deze bijdrage lichten wij ons recent empirisch onderzoek naar de Belgische rechtspraak inzake ongehuwd samenwonen toe. De gepresenteerde rechtspraakanalyse is om meerdere redenen atypisch te noemen. Ten eerste omwille van de betrokken onderzoekers. Wij voerden dit onderzoek uit samen met een 150-tal masterstudenten in de rechten.¹ Ten tweede omwille van de methodologie. In plaats van het bespreken van individuele vonnissen en arresten kozen we ervoor om een groot aantal vonnissen en arresten te coderen met als doel het identificeren van patronen, tendensen en relaties tussen de data aangaande de conflicten die feitelijke samenwoners ervaren. Het werken met studenten in de rechten in dit type onderzoek bracht een aantal uitdagingen met zich mee. Zo dienden we de methodologie hier en daar aan te passen om de beperkte statistische voorkennis, inclusief het coderen, in rekenschap te brengen. De reden om het project op deze manier aan te pakken, verwijst naar het dubbele doel van dit project: wetenschappelijk maar vooral ook pedagogisch. Door studenten actief te betrekken bij het verzamelen en analyseren van

* Prof. dr. E. Goossens is docent privaatrecht aan de Vrije Universiteit Brussel, postdoctoraal onderzoeker van het FWO Vlaanderen aan het Rector Roger Dillemans Instituut Familiaal Vermogensrecht aan de KU Leuven en gastprofessor aan de Universiteit Antwerpen. Dr. K. Bollen is Senior Expert Family & Business Dynamics bij Greenille Private Client Team (Deloitte Legal Lawyers) en bemiddelaar. Ze is als onderzoeker en docent verbonden aan de KU Leuven (Faculteit Psychologie en Faculteit Rechten) waar ze onderzoek doet naar bemiddeling, leiderschap en coaching. Prof. Dr. A.-L. Verbeke is gewoon hoogleraar en directeur Rector Roger Dillemans Instituut Familiaal Vermogensrecht KU Leuven, Professor of Law & Negotiation Harvard, Tilburg, UCP Lisbon en advocaat. Het onderzoek kaderde in het postdoctoraal onderzoeksproject van E. Goossens over de regulering van ongehuwd samenwonen (Fonds Wetenschappelijk Onderzoek – Vlaanderen; beurs nr. 40461).

1 Naast de studenten speelden ook Ingvild Debbaut, Marie Delbroek en Michelle Aerts, allen verbonden aan het Rector Roger Dillemans Instituut voor Familiaal Vermogensrecht van de KU Leuven, een cruciale rol bij het welslagen van dit project. De auteurs wensen hen allen te bedanken voor hun inspanningen.

rechtspraak, raken ze bekend met het voeren van empirisch onderzoek en worden ze uitgenodigd actief te reflecteren. We zetten in deze bijdrage uiteen hoe we dit onderzoek samen met de studenten hebben aangepakt en tot welke conclusies we zijn kunnen komen. Wij onthouden vooral dat een samenwerking met studenten als volwaardige onderzoekspartners voor alle betrokkenen een verrijkende ervaring is.

2 ONDERZOEKSOPZET

Koppels die beslissen om met elkaar door het leven te gaan, kunnen dit op verschillende manieren doen. Zij kunnen ervoor kiezen om te huwen, om ongehuwd samen te wonen en hun samenwoning formeel te registreren (in Nederland: ‘geregistreerd partnerschap’; in België: ‘wettelijke samenwoning’) of om gewoonweg feitelijk samen te wonen zonder registratie. Deze ‘feitelijke samenwoning’ (in Nederland doorgaans ‘ongehuwd samenwonen’ genoemd) is een blinde vlek in het Belgisch relatievermogensrecht.

Het Napoleontische uitgangspunt was dat samenwoners zich bewust buiten de wet plaatsten door niet te huwen, en de wet dus ook niet in een regeling moest voorzien: ‘*Les concubins se passent de la loi, la loi se désintéresse d’eux*’. Die premisse bepaalt vandaag nog steeds het familiaal vermogensrecht van feitelijke samenwoners. Het Belgisch Burgerlijk Wetboek² bevat geen bijzondere regeling voor feitelijke samenwoning. Dat is een bewuste keuze van de Belgische wetgever,³ bevestigd bij de hervorming van het erfrecht en huwelijksvermogensrecht, in werking sinds 1 september 2018.⁴

Dat uitgangspunt belet natuurlijk niet dat de feitelijke gedragingen van feitelijke samenwoners juridische gevolgen kunnen hebben, en dat er tussen hen regelmatig conflicten in de juridische sfeer rijzen. Deze worden doorgaans beslecht op grond van het gemene verbin-tenissen- en vermogensrecht.⁵ Anders dan in de Napoleontische tijden is de feitelijke samenwoning vandaag evenwel geen randfenomeen meer, integendeel, de laatste decennia

2 Sinds 1 november 2020 draagt het Belgisch Burgerlijk Wetboek van 21 maart 1804 officieel het opschrift ‘oud Burgerlijk Wetboek’. Het wordt gradueel vervangen door een nieuw Burgerlijk Wetboek. De regeling van de samenwoning is tot nu toe echter niet gewijzigd. Zie art. 2 lid 2 wet van 13 april 2019 tot invoering van een Burgerlijk Wetboek en tot invoering van boek 8 ‘Bewijs’ in dat Wetboek, *Belgisch Staatsblad (B.S.)* 14 mei 2019.

3 Minstens is er voorlopig geen politieke consensus om een wettelijke regeling voor feitelijke samenwoners uit te werken.

4 Wet van 31 juli 2017, *B.S.* 1 september 2017 en Wet van 22 juli 2018, *B.S.* 27 juli 2018.

5 Wel bevat deelwetgeving soms specifieke voorschriften voor feitelijke samenwoners, vaak om een punctuele gelijkstelling met gehuwden te bekomen. Zie hiervoor E. Goossens, ‘Kompas voor een nieuw samenwoningsrecht: neutraal, coherent en compensatie- en behoeftegericht’, *TBBR* 2019/9, p. 541 en C. Forder & A. Verbeke, *Gehuwd of niet: maakt het iets uit?*, Antwerpen: Intersentia 2005, en de verwijzingen daar.

is het aantal feitelijk samenwonende koppels sterk toegenomen. Hierdoor komt de vraag naar een juridisch kader voor deze vorm van samenleven meer en meer op de voorgrond.

Het is echter zeer moeilijk gebleken om een eenduidige visie te ontwikkelen over een gepast samenwoningsvermogensrecht. Naast ideologische en politieke verschillen in visie, is het onderwerp op zich gevarieerd en complex. De groep van feitelijke samenwoners is immers zeer divers en de betrokkenen hebben verschillende belangen en redenen om voor deze manier van samenwonen en leven te kiezen. Een veelgehoorde verzuchting is bovendien dat er te weinig informatie beschikbaar is over feitelijke samenwoners en de problemen die zij (juridisch) ervaren. Bij gebrek aan empirische data wordt het debat over de regulering van de feitelijke samenwoning dan ook vaak gevoerd op basis van anekdotiek, eigen ervaringen of buikgevoel.

Het opzet van het hier gepresenteerde onderzoek was om het debat aangaande de regulering van de feitelijke samenwoning te objectiveren, door een grootschalig empirisch onderzoek te voeren over de recente rechtspraak omtrent feitelijke samenwoning. We wilden in kaart brengen welke conflicten feitelijke samenwoners voor de rechtbank brengen, welke juridische figuren rechters gebruiken om deze conflicten te beslechten en in welke mate dergelijke gedingen kans van slagen hebben. Wij realiseerden dit onderzoek samen met een 150-tal masterstudenten in de rechten aan de KU Leuven binnen het kader van het mastervak Familiaal vermogensrecht (academiejaar 2018-2019). Onder ons toezicht stonden de studenten in voor de verzameling van de rechtspraak, de codering en de eerste analyse van de resultaten.

In tegenstelling tot de klassieke overzichten van de rechtspraak waar traditioneel een casusgewijze bespreking van individuele vonnissen en arresten wordt gehanteerd, lag in dit onderzoek de focus op het identificeren van patronen, tendensen en relaties tussen data. De rechtspraak werd geschematiseerd en gecodeerd: er werd onderzocht hoe vaak bepaalde zaken voorkomen in de vonnissen en arresten en hoe die zich verhouden tot elkaar. Dit liet ons toe om conclusies te trekken over de soorten van conflicten tussen feitelijke samenwoners en de betrokken rechtsgronden.

Het is duidelijk dat we in eerste instantie wilden komen tot een kwantitatief overzicht, op basis van codering van rechterlijke uitspraken. Dit impliceert dan ook dat wij niet met de partijen konden spreken of interviews afnemen, zodat meer kwalitatieve vragen waarom samenwoners niet trouwen, welke verwachtingen zij hebben over hun relatie, en in welke mate zij op de hoogte zijn van de juridische implicaties van de gekozen relatievorm, buiten het bestek van dit onderzoek vallen.

De doelstelling van het onderzoek was dus driedig:

1. wetenschappelijk-inhoudelijk: bijdragen aan het debat over de regulering van de feitelijke samenwoning door het verzamelen van gegevens over de feitelijke samenwoners en de problemen die zij ervaren;
2. wetenschappelijk-methodologisch: binnen de rechtswetenschap is de afgelopen jaren gepleit voor meer empirisch wetenschappelijke kennis en methoden in het recht. Meer empirisch onderzoek is nodig om te toetsen of 'law in the books' overeenkomt met de werkelijkheid, of 'law in action',⁶ maar ook om te komen tot *evidence-based* wetgeving en rechtspraak.⁷ Ons onderzoek trachtte hier een antwoord op te bieden door het voeren van empirisch-juridisch onderzoek;⁸
3. pedagogisch: studenten actief betrekken bij het verzamelen en analyseren van rechtspraak. Op deze manier raken studenten niet alleen bekend met het voeren van empirisch onderzoek, maar worden zij ook uitgenodigd om actief te reflecteren.

De focus van deze bijdrage ligt op de wetenschappelijk-methodologische en de pedagogische doelstelling van het onderzoek. We zetten uiteen welke uitdagingen en beperkingen wij ervaren bij een grootschalige rechtspraakanalyse die hoofdzakelijk werd gevoerd door studenten. In het bijzonder lichten we toe welke invloed die pedagogische doelstelling had op het wetenschappelijk-methodologische design van het empirisch onderzoek. De eigenlijke onderzoeksresultaten (d.i. de wetenschappelijk-inhoudelijke doelstelling) komen in deze bijdrage slechts summier aan bod.

3 METHODE

3.1 *Sample en dataverzameling*

Het onderzoek handelt over de vermogensrechtelijke gevolgen van de feitelijke samenwoning tussen koppels, zoals opgetekend in de Belgische rechtspraak van september 2014 tot en met september 2018. Enkel partnerrelaties werden betrokken; feitelijke samenwoning tussen familieleden, vrienden en kennissen viel buiten het bestek van het onderzoek. Het

6 H.G. Genn, M. Partington & S. Wheeler, *Law in the Real World: Improving Our Understanding of How Law Works: Final Report and Recommendations*, Londen: Nuffield Foundation 2006.

7 P. van der Laan, M. Smit & B.C. Welsch, 'Rechters, ga experimenteren! Over het belang van evalueren en experimenteren voor een effectieve (straf)rechtspleging', *Rechtstreeks* 2013/4, p. 15-36 en R.A.J. van Gestel, I. Giesen & W. van Boom, 'Een landelijk Centrum voor Methodologie en Empirische Rechtsbeoefening', *NJB* 2012, afl. 29, p. 2032-2035.

8 C. Stolker, *Rethinking the law school. Education, Research, Outreach and Governance*, Cambridge: Cambridge University Press 2015.

onderzoek omvat vonnissen en arresten uitgesproken tussen 1 september 2014 – de startdatum van de familierechtbanken⁹ – en 30 september 2018 – dit is net vóór de verzameling van de rechtspraak vanaf oktober 2018.

Het kwantitatief opzet van ons onderzoek impliceert dat we nood hadden aan een grote hoeveelheid aan vonnissen en arresten. Omdat we in kaart wilden brengen welke conflicten feitelijke samenwoners bij de rechtbank aanhangig maken, was het bovendien essentieel om ook lagere rechtspraak te betrekken, dat wil zeggen de vonnissen in eerste aanleg. De hogere rechtspraak – hoven van beroep, Hof van Cassatie en Grondwettelijk Hof – vertegenwoordigt immers niet noodzakelijk de grondstroom in de conflicten tussen samenwoners. Een groot probleem en oud zeer is dat in België maar een fractie van alle rechtspraak wordt gepubliceerd.¹⁰ Dat is in het bijzonder het geval voor de lagere rechtspraak. Inzage in de rechtspraak is dan ook alleen maar mogelijk door veldwerk op de griffies van de rechtbanken. Dat is een huzarenwerk waarvoor vele handen en hoofden nodig zijn.

De dataverzameling gebeurde door de masterstudenten Familiaal vermogensrecht aan de KU Leuven. We betrokken in dit onderzoek zowel de gepubliceerde als de ongepubliceerde rechtspraak van de Belgische rechtbanken van eerste aanleg (burgerlijke rechtbanken en familierechtbanken), hoven van beroep, het Hof van Cassatie en het Grondwettelijk Hof. Alle rechterlijke uitspraken werden als onafhankelijk beschouwd. Dit impliceert dat we geen rekening hielden met een eventuele samenhang tussen uitspraken. Indien een vonnis in eerste aanleg in de relevante onderzoeksperiode werd verbroken door een hof van beroep, werden beide uitspraken apart opgenomen en gecodeerd. Het was immers (onder meer) onze bedoeling om de rechtsgronden en elementen in de redenering van de rechtbank in kaart te brengen, en die elementen kunnen verschillen in eerste aanleg en in beroep. Toekomstig onderzoek zou rekening kunnen houden met die samenhang tussen de uitspraken aangaande eenzelfde dossier. Hieronder schetsen we de verschillende chronologische stappen die we gezet hebben in het onderzoek.

In september 2018 werden alle Belgische rechtbanken van eerste aanleg en hoven van beroep aangeschreven vanuit het Rector Roger Dillemans Instituut voor Familiaal Vermogensrecht van de KU Leuven (hierna: RRD) met de vraag of zij hun medewerking wilden

⁹ Wet van 30 juli 2013 betreffende de invoering van een familie- en jeugdrechtbank, *B.S.* 27 september 2013.

¹⁰ De wet van 5 mei 2019 belooft beterschap, door te voorzien in de verplichte publicatie van alle vonnissen en arresten in een publiek toegankelijke, elektronische databank (art. 7 wet van 5 mei 2019 tot wijziging van het Wetboek van strafvordering en het Gerechtelijk Wetboek wat de bekendmaking van de vonnissen en arresten betreft, *B.S.* 16 mei 2019.). De praktische realisatie van deze databank kan echter nog enige tijd op zich laten wachten. De inwerkingtreding van de wet van 5 mei 2019 werd al een eerste keer uitgesteld van 1 september 2020 naar 1 september 2021 (uiterste datum). Zie ook Peeraer en van Gestel in deze bundel.

verlenen aan het onderzoek en de studenten toegang wilden verlenen tot de betrokken rechtspraak. Van de 33 aangeschreven hoven en rechtbanken, zegden er uiteindelijk twintig toe (61%): alle vijf hoven van beroep en vijftien afdelingen van de rechtbanken van eerste aanleg (burgerlijke rechtbanken en/of familierechtbanken).¹¹

De eigenlijke dataverzameling vond plaats in oktober en november 2018.¹² Bij het merendeel van de betrokken hoven en rechtbanken gebeurde de dataverzameling door de studenten ter griffie. Als algemene richtlijn gold dat de studenten alle rechtspraak uit de betrokken periode moesten doornemen. Aan de meewerkende hoven en rechtbanken werd dan ook gevraagd om hun toegang te verlenen tot alle relevante uitspraken. Bij sommige rechtscolleges bleek de werklast voor de studenten evenwel buitensporig, waardoor we een deel van het opzoekingswerk lieten wegvallen (zie *infra*). Drie hoven en rechtbanken kozen ervoor om zelf geselecteerde uitspraken op te sturen.

Om te bepalen welke studenten naar welke hoven en rechtbanken zouden gaan, verdeelden de medewerkers van het RRDI de studenten in twintig groepen van zes tot twaalf studenten, om ze vervolgens toe te wijzen aan bepaalde griffies. Er werd bij de groepsverdeling rekening gehouden met de moedertaal van de studenten (Franstalige of tweetalige studenten werden bij voorkeur toegewezen aan Franstalige rechtscolleges), met de woonplaats van de studenten en met de verwachte werklast per griffie. Op verzoek van de student kon men aan een andere groep worden toegewezen indien dit om praktische redenen beter te regelen was. Ter griffie werd de rechtspraak gekopieerd en (met stift) geanonimiseerd om de vertrouwelijkheid en de privacy van de partijen te garanderen. De geanonimiseerde rechtspraak werd nadien gedigitaliseerd (ingescand). De data voor dit onderzoek bestaan dus uitsluitend uit geanonimiseerde rechtspraak.

De tijd die het kostte om de arresten en vonnissen op te zoeken, verschilde erg van griffie tot griffie. Dit had vooral te maken met (a) het totaal aantal aanwezige arresten of vonnissen en (b) of er binnen de griffie sprake was van een classificatiesysteem. In de mate van het mogelijke werd hier bij de groepsverdeling reeds op geanticipeerd (zie hoger). Waar gaandeweg de werklast tussen de groepen te veel bleek te verschillen, werd bijgestuurd. Een aantal groepen kregen extra opzoekwerk, bij andere groepen werd opzoekwerk weg-

11 Hoven van beroep: Antwerpen, Bergen, Brussel (enkel met betrekking tot Nederlandstalige arresten), Gent en Luik. Rechtbanken van eerste aanleg: Antwerpen (afdelingen Antwerpen, Mechelen en Turnhout), Brussel (Franstalige rechtbank), Limburg (afdelingen Hasselt en Tongeren), Luxemburg (afdelingen Aarlen, Marche-en-Famenne en Neufchâteau), Oost-Vlaanderen (Dendermonde en Gent), Waals-Brabant en West-Vlaanderen (afdelingen Brugge, Kortrijk en Veurne).

12 Met uitzondering van het hof van beroep te Brussel, waar de dataverzameling gebeurde in januari 2019 door onderzoekers en assistenten van het RRDI.

genomen. Sommige griffies ontvingen daarom een extra groep studenten. Als richtlijn werd evenwel één dag dataverzameling ter griffie per groep voorop gesteld. Indien dit quotum was overschreden, viel het overige onderzoekwerk weg.

Bijkomend werd ook de gepubliceerde rechtspraak uit de betrokken periode toegevoegd aan de dataset. Het betrof zowel de uitspraken van het Hof van Cassatie en het Grondwettelijk Hof, die grotendeels online beschikbaar zijn, als de gepubliceerde rechtspraak van de overige rechtscolleges. De studenten raadpleegden hiertoe de websites www.juridat.be en www.const-court.be, evenals een lijst van door ons geselecteerde vaktijdschriften.¹³ Uit deze zoekopdracht bleek dat de gepubliceerde rechtspraak uit de betrokken periode bijna uitsluitend bestaat uit uitspraken van het Hof van Cassatie, het Grondwettelijk Hof en van hoven en rechtbanken die reeds meewerkten aan ons onderzoek. Slechts één gepubliceerde uitspraak was van een rechtscollege dat nog niet meewerkte, namelijk de rechtbank van eerste aanleg te Luik (afdeling Luik).¹⁴

De finale dataset omvat 778 vonnissen en arresten. De initiële dataset omvatte 781 vonnissen en arresten. Drie arresten of vonnissen werden verwijderd aangezien deze foutief en/of onvolledig werden gecodeerd. Maar liefst 96% van de dataset betreft ongepubliceerde rechtspraak.¹⁵ Een dergelijke omvangrijke dataset is in de Belgische context zeer uitzonderlijk voor deze materie.

3.2 *Codering*

Om de inhoud van de vonnissen en arresten te analyseren, werd er gebruikgemaakt van een codeerschema. Het codeerschema biedt een bril om naar de rechtspraak te kijken en om deze te classificeren aan de hand van verschillende variabelen.

Een eerste versie van het codeerschema werd opgesteld door de onderzoekers die beslagen zijn in de thematiek alsook vertrouwd met de inhoud van dergelijke vonnissen en arresten. Bij het opstellen van het codeerschema werd er rekening gehouden met de beperkte statistische voorkennis en beperkte ervaring met coderen van de studenten aan de rechtsfaculteit. Aangezien het inhoudelijk coderen van de arresten en vonnissen niet voor de hand zou

13 We baseerden deze lijst op de tijdschriften die periodiek worden geraadpleegd door de assistenten van het RREDI in het licht van de jaarlijkse publicatie van de *Rechtspraakfiches familiaal vermogensrecht* in het *Tijdschrift voor Estate Planning*. Het betreft de volgende tijdschriften: NJW, RABG, TBBR, NNN, Notamus, RPP, T.Fam., Rev.dr.ULg, TBH, TEP, RTDF, Act.dr.fam., AFT, TGR-TWVR, FJF, JDF, JLMB, P&B, RNB, Successierechten, RW, TPR, VIP, Notariaat, T.Not., T.Verz., JT, NFM, Rec.gén.enr.not.

14 Rb. Luik 17 december 2014, RTDF 2015/2, p. 291.

15 751 vonnissen en arresten zijn ongepubliceerd.

liggen, werd ervoor gekozen om het codeerschema zo eenvoudig mogelijk te houden. Om die reden werd er geopteerd voor het programma Excel in plaats van het statistische softwareprogramma SPSS of het programma 'R', en werd het codeerschema zo opgesteld dat er zo veel mogelijk binair zou moeten worden gecodeerd (namelijk iets is al dan niet van toepassing, dus ja versus nee of '1' versus '0'). Toch werd er bij de verschillende parameters of variabelen ook ruimte gelaten om extra toelichting te geven (tekstueel) waardoor het onderzoek naast kwantitatieve elementen ook een aantal kwalitatieve elementen bevat. Het doel hiervan was om antwoorden te preciseren en/of verder toe te lichten. Zo werkten we bijvoorbeeld voor het bepalen van de leeftijd met meerkeuzecohorten (< 18 jaar; 18-25 jaar; 25-35 jaar enz.), alsook voor bijvoorbeeld de duurtijd van de feitelijke samenwoning (0-3 jaar; 3-6 jaar; 6-10 jaar). Daarnaast werden de studenten gevraagd om – indien mogelijk – respectievelijk de exacte leeftijd en/of duurtijd aan te geven. Verder werd ook stevast de mogelijkheid gegeven om een antwoord nader toe te lichten en/of belangrijke informatie toe te voegen die niet door het codeerschema gevat kon worden (bijvoorbeeld de mogelijkheid tot het toevoegen van een andere, extra rechtsgrond dan die welke reeds werden opgenomen in het codeerschema). Dit stelde ons in de mogelijkheid om specifieke informatie te verzamelen alsook relevante informatie te bekomen die we zouden verliezen indien we enkel zouden werken met het codeerschema.

Het codeerschema werd tweemaal aangepast op basis van de input van de studenten. Begin oktober 2018 kregen de studenten een inleidend methodologisch college over empirisch onderzoek. Tijdens dit college werden de principes van kwantitatief en kwalitatief onderzoek toegelicht, alsook een aantal beginselen van 'goed onderzoek' en werd ook het codeerschema voorgesteld waarmee de studenten zouden gaan werken. De feedback en suggesties van de studenten tijdens dat college resulteerden in een eerste, (zeer) beperkte aanpassing van het codeerschema.¹⁶

Een tweede, ingrijpendere wijziging van het codeerschema werd doorgevoerd op basis van de tussentijdse verslagen van de studenten en het bijbehorende plenair college midden november 2018. Alvorens aan het coderen van alle verzamelde rechtspraak te beginnen, hadden de studenten de opdracht gekregen om per groep alvast een *close reading* te organiseren van één geselecteerd vonnis of arrest (zie *infra*). In het tussentijds verslag moesten de studenten hun bevindingen inzake die *close reading* noteren en in voorkomend geval een voorstel doen tot aanvulling of wijziging van de variabelen. In overleg met de studenten werd op grond van de tussentijdse verslagen het codeerschema aangepast. Hoewel we initieel vooral opteerden voor een binair codeerschema om het coderen voor de studenten – met een beperkte codeerervaring – zo eenvoudig mogelijk te houden, bleek na de *close*

16 Bijv. toevoegen van de variabelen 'genderneutraal of transgender', naast 'mannelijk' en 'vrouwelijk'.

reading en de verslagen dat deze binaire codering onvoldoende nuance bood om de juridische werkelijkheid (en de nuances daarin) te vatten. Om die reden werd voor een aantal variabelen, zoals de rechtsgrond en de redenering van de rechter, afgestapt van een zuiver binaire benadering ('1' versus '0').

In plaats daarvan werd enerzijds geopteerd voor open tekstuele antwoorden (kwalitatief) en anderzijds voor een Likertschaal.¹⁷ Met behulp van de Likertschaal probeert men zaken meetbaar te maken die normaal gesproken lastig zijn om te meten in kwantitatief onderzoek. Op die manier kan men moeilijk te kwantificeren gegevens toch ondervragen en behandelen op intervalmeetniveau. Concreet maakt men iets meetbaar met behulp van een Likertschaal in combinatie met een bepaalde uitspraak of stelling. Voor ieder aangewende rechtsgrond werd bijvoorbeeld gebruikgemaakt van de volgende schaal: niet van toepassing (0), onontvankelijk (1), ontvankelijk maar ongegrond (vordering niet toegekend) (2), ontvankelijk en gedeeltelijk gegrond (vordering partieel toegekend) (3), ontvankelijk en volledig gegrond (vordering geheel toegekend) (4), uitstel (5), akkoord partijen (6). Om te bepalen hoe belangrijk een bepaalde factor doorwoog bij de besluitvorming van het rechtscollege, werd ook gebruikgemaakt van de volgende Likertschaal (gaande van 0-5), waarbij 0 stond voor niet van toepassing en 5 voor doorslaggevend. Concreet kregen de studenten de volgende antwoordalternatieven aangeboden: niet van toepassing (0), aanwezig (1), tussen aanwezig en belangrijk (2), belangrijk (3), tussen belangrijk en doorslaggevend (4) en doorslaggevend (5).

Die werkwijze leidde tot een finaal codeerschema waarvoor de krachten werden gebundeld en zowel gebruik werd gemaakt van de expertise van de onderzoekers als van de ervaringen van de studenten.

De volgende aspecten werden geanalyseerd en gecodeerd: (a) het rechtscollege, (b) de rechterlijke beslissing, (c) de feitelijke situatie van de partners, (d) de concrete eis, (e) de ingeroepen rechtsgrond en uitkomst, (f) elementen in de redenering van de rechtbank en (g) extra. De eerste onderdelen (a-d) werden hoofdzakelijk binair gecodeerd. Zoals hierboven werd aangehaald, werd er voor secties 'e' en 'g' gebruikgemaakt van zowel Likertschalen als open vragen waar men tekstueel (kwalitatief) een antwoord kon geven.

- a. *Het rechtscollege* verwijst naar het betrokken rechtscollege zoals bijvoorbeeld het Hof van Cassatie of het hof van beroep te Bergen, en, in geval van een rechtbank van eerste aanleg, naar de sectie (burgerlijke rechtbank of familierechtbank).

17 A. Joshi, S. Kale, S. Chandel & D.K. Pal, 'Likert scale: Explored and explained', *Current Journal of Applied Science and Technology* 2015, p. 396-403; G.M. Sullivan & A.R. Artino Jr, 'Analyzing and interpreting data from Likert-type scales', *Journal of graduate medical education* 2013/5, afl. 4, p. 541.

- b. *De rechterlijke beslissing* omvat de datum van de uitspraak, het rolnummer, het taalgebied, gender van de voorzitter, of de beslissing werd gepubliceerd of niet, of er vermelding werd gemaakt van het begrip ‘feitelijke samenwoning’ of ‘cohabitation de fait’, en over welk type conflict het gaat (bijvoorbeeld conflict tijdens de relatie, naar aanleiding van een relatiebreuk of naar aanleiding van een overlijden).
- c. *De feitelijke situatie van de partners*: gender van de partners, duurtijd van de relatie, duurtijd van de feitelijke samenwoning, leeftijd van de partners, inkomen van de partners, of (en hoeveel) kinderen de desbetreffende partners al dan niet samen hebben.
- d. *De concrete eis*: vermogensverdeling na relatiebreuk; vermogensverdeling na overlijden; toekenning onderhoudsgeld; vergoeding voor investering in goederen; vergoeding als compensatie voor huishoudelijke arbeid; afgifte eigen goederen; procedurele eis (bijvoorbeeld aanstelling van de notaris-vereffenaar); andere.
- e. *Rechtsgrond en uitkomst*: voor de 21 rechtsgronden dienden de studenten het volgende aan te geven: (1) van wie gaat de vordering uit? (bijvoorbeeld eiser, verweerder, appellant, enz.); (2) soort vordering (bijvoorbeeld hoofdvordering, tegenvordering, enz.); en (3) type van beslissing (bijvoorbeeld ontvankelijk maar ongegrond, onontvankelijk, enz.). De volgende 21 rechtsgronden werden opgenomen in het codeerschema; onverdeeldheid; natrekking en kostenleer; zaakwaarneming; samenlevingsovereenkomst; lening; bruikleen; borgtocht; lastgeving; onverschuldigde betaling; bijdrage in lasten van samenleving; natuurlijke verbintenis; verrijking zonder oorzaak; onrechtmatige daad; beding van aanwas of tontinebeding; schenking; testament; kostenleer; eigendom of revindicatie; subrogatie of indeplaatsstelling; billijkheid; procedurele aspecten; andere.
Bij de *concrete eis* (d) en de *ingeroepen rechtsgrond en uitkomst* (e) konden de studenten indien nodig meerdere antwoorden aanduiden. Daarnaast was er ook steevast in een antwoordoptie ‘andere’ voorzien indien de studenten de indruk hadden dat hun antwoord niet (voldoende) aanwezig was tussen de gegeven antwoordalternatieven.
- f. Voor de *elementen in de redenering van de rechtbank* geven de studenten aan in welke mate een bepaalde factor/overweging (juridische overweging;¹⁸ extra-juridische

18 Dit zijn overwegingen die louter verband houden met de toepassing van het recht. We onderscheiden hier de volgende variabelen: toepassingsvoorwaarden rechtsgrond; bewijsproblematiek; vergelijking met de regels inzake het huwelijk of de wettelijke samenwoning; of andere.

overwegingen;¹⁹ visie rechtbank over de stand van het recht²⁰) meespeelde bij de besluitvorming van de rechtbank. Zoals hierboven aangegeven doen de studenten dit aan de hand van een Likertschaal, nl. niet van toepassing (0), aanwezig (1), tussen aanwezig en belangrijk (2), belangrijk (3), tussen belangrijk en doorslaggevend (4) en doorslaggevend (5). Daarnaast werd aan de studenten gevraagd om een kruisverwijzing toe te voegen naar een citaat uit een bepaalde passage in de rechtspraak.

- g. *Extra*: Hier werden de studenten uitgenodigd om iets toe te voegen of op te merken dat nog niet eerder aan bod kwam maar wel relevant is gegeven de insteek van het onderzoek.

Met dit finale codeerschema gingen de studenten aan de slag om de arresten en vonnissen te gaan analyseren. Elke groep codeerde de toegewezen rechtspraak in een Excelfile.²¹ De verschillende Excelfiles van de groepen werden later bij elkaar gevoegd tot één moederfile. Het was deze moederfile die de basis vormde voor de verdere analyses. De codering door de studenten ging bovendien gepaard met een groepsmemo, waarin de studenten per groep de onderzoeksresultaten samenvatten en analyseerden. Daar waar het in de initiële *close reading* eerder ging om een *within case analysis* (namelijk bekijken hoe verschillende studenten naar één bepaalde casus kijken en deze interpreteren), ging het hier om een *cross case analysis*: wat zijn de opmerkelijkste resultaten over de verschillende geanalyseerde cases heen? De opmerkelijkste resultaten werden toegelicht, met bijzondere aandacht voor de samenhang tussen de variabelen en voor bepaalde patronen. Het opzet van de groepsmemo was dus om reliëf en duiding te brengen bij de Excelfiles. De groepsmemo was beschrijvend en duidend, maar niet normatief: de studenten analyseerden de onderzoeksresultaten *de lege lata*, maar deden geen voorstellen *de lege ferenda*. De onderzoeksresultaten werden voorgesteld en besproken in een plenair reflectiecollege midden december 2018. Dit reflectiecollege vormde de afsluiting en apotheose van de grootschalige onderzoekssamenwerking met de masterstudenten Familiaal vermogensrecht.

In februari 2019 startten vier masterstudenten met een verdiepend masterproefonderzoek op basis van de dataset. Zij werkten op vier deelthema's die aansloten bij de opmerkelijkste preliminaire resultaten van het empirisch onderzoek, zoals die bleken uit de moederfile,

19 Hier zochten we naar overwegingen aangaande de aard van de relatie en de situatie van de partners. Het codeerschema bevatte de volgende variabelen: verwachtingen/vertrouwen van de partners, verplichtingen gecreëerd door de samenwoning; inhoudelijke/conceptuele vergelijking van de feitelijke samenwoning met huwelijk of wettelijke samenwoning; bewuste keuze van de partijen om geen huwelijk of een wettelijke samenwoning aan te gaan (autonomiegedachte); bescherming van de economisch zwakkere partij; algemene billijkheidsoverwegingen, redelijkheid en billijkheid, rechtvaardigheid, *fairness*; of andere.

20 Blijkt uit de redenering van het rechtscollege een (kritisch-evaluerende of normatieve) visie over de huidige stand van het recht?

21 Extensie.xlsx.

de groepsmemo's en het plenair reflectiecollege. De studenten kregen hiervoor bijkomende training in gegevensverwerking via Excel van de ICTS-dienst van de KU Leuven. In hun masterproeven combineerden deze studenten een thematische analyse van de resultaten van het empirisch onderzoek met een klassiek juridisch-rechtsvergelijkend onderzoek. Dat wil zeggen dat zij over een deelaspect van de feitelijke samenwoning het geldende Belgische recht toetsten aan de dataset en vergeleken met een buitenlands rechtstelsel. In tegenstelling tot de groepsmemo's bevatten de masterproeven ook een normatief luik. De vier studenten rondden hun masterproefonderzoek succesvol af aan het einde van het academiejaar 2019-2020.²² De masterproeven zorgden voor bijkomende verdieping en maakten de sprong van beschrijvend naar voorzichtig evaluerend en normatief. Zij hadden evenwel geen invloed op de dataset. De moederfile bleef ongewijzigd: de studenten deden niet aan bijkomende gegevensverzameling (geen codering van nieuwe rechtspraak, noch interviews).

3.3 Betrouwbaarheid

Aangezien we werkten met een grote groep van studenten, was het van cruciaal belang om de betrouwbaarheid van ons onderzoek te garanderen. Betrouwbaarheid van een meting impliceert dat de ene onderzoeker/student tot vrijwel dezelfde resultaten zou komen als de andere. Pas dan is er sprake van betrouwbaar onderzoek. De volgende principes zijn hierbij van belang: (1) zorgen voor een representatieve steekproef, (2) objectiviteit bewaren, (3) kritisch data analyseren en (4) meerdere specialisten laten kijken naar de data (en/of de resultaten).²³

Hierboven gaven we al aan dat het onderzoek gebaseerd is op 778 vonnissen en arresten over de vermogensrechtelijke gevolgen van feitelijke samenwoning. De dataset omvat niet alle rechterlijke uitspraken over de feitelijke samenwoning in de betrokken periode (sep-

22 R. De Backer, *Verrijking zonder oorzaak als rechtsgrond voor vermogensverschuiving bij feitelijke samenwoners* (masterproef KU Leuven), 2020; S. Debie, *Minimale solidariteit voor feitelijke samenwoners: vanuit een macro-Europees perspectief naar een Belgisch juridisch-empirische werkelijkheid* (masterproef KU Leuven), 2020; C. Despierre, *Kan feitelijke samenwoning aanleiding geven tot rechtsgevolgen? Onder welke vereisten kan een feitelijke samenwoning vandaag, en moet een feitelijke samenwoning idealiter leiden tot rechtsgevolgen?* (masterproef KU Leuven), 2020; B. Minten, *Solidariteit met betrekking tot de gezinswoning: een ambitieus doel of een onhaalbare utopie? Een studie van de solidariteit tussen feitelijke samenwoners* (masterproef KU Leuven), 2020 – allemaal beschikbaar op KU Leuven e-theses via <https://bib.kuleuven.be/sbib/collectie/eindwerken>.

23 A. Fink & M.S. Litwin, *How to measure survey reliability and validity*, vol. 7, Thousand Oaks: Sage 1995; R. Heale & A. Twycross, 'Validity and reliability in quantitative studies', *Evidence-based nursing* 2015/18, afl. 3, p. 66-67; P.A. Moss, 'Can there be validity without reliability?', *Educational researcher* 1994/23, afl. 2, p. 5-12.

tember 2014-september 2018). Het gaat echter wel over een zeer aanzienlijke hoeveelheid uitspraken waarbij we het merendeel van de bevoegde Belgische rechtscollages konden betrekken en waarbij de uitspraken een vrij lange periode (vier jaar) beslaan. Om die redenen menen we dat de dataset eerder representatief is.

Om de betrouwbaarheid van het onderzoek te waarborgen werden de volgende stappen gezet. We vermeldden hierboven reeds dat de studenten een inleidend methodologisch college kregen over het doen van empirisch onderzoek. In dat college werd het codeerschema samen doorlopen en toegelicht. Er werd ruimte gelaten voor feedback en suggesties van de studenten. Daarnaast kregen de studenten de opdracht om individueel een vonnis of arrest te lezen (*close reading*) en (individueel) te coderen (*Assess*). Als eenmaal alle studenten van het groepje dit hadden gedaan, werden de coderingen van de verschillende studenten naast elkaar gelegd en werd er gekeken naar die zaken waar studenten verschilden in hun codering. Studenten kregen de opdracht om elkaar goed te bevragen en te luisteren naar de aangedragen argumenten om al dan niet te opteren voor een bepaalde codering, om vervolgens hierover in discussie te gaan en samen te komen tot een gedragen codering (*Discuss*). Het feit dat studenten elkaar uitleggen waarom ze al dan niet opteren voor een bepaalde codering, bevordert het inzicht in de materie (diepgang) en helpt om op een eenduidige manier zaken te interpreteren (*shared mental model*²⁴) en te coderen (aan de hand van criteria). De studenten rapporteerden hun bevindingen inzake de *close reading*-oefening in een tussentijds verslag (*Review*). Het principe van *Assess-Discuss-Review*²⁵ bevordert inzicht maar komt ook de betrouwbaarheid van de coderingen ten goede. Na de analyse van dit eerste arrest werd er een plenair college georganiseerd met de 150 studenten om enerzijds toelichting te geven bij moeilijkheden of vragen van studenten bij het coderen, anderzijds om het codeerschema te finetunen en te finaliseren.

Om na te gaan hoe betrouwbaar de metingen zijn, verwijst men ook wel naar de interbeoordelaarsbetrouwbaarheid (*intercoder reliability*).²⁶ Als je in een onderzoek met meerdere onderzoekers (codeerders) werkt, dan zouden idealiter de verschillende codeerders zo veel mogelijk dezelfde resultaten moeten (kunnen) bekomen. De interbeoordelaarsbetrouw-

24 G.C. Feng, 'Mistakes and how to avoid mistakes in using intercoder reliability indices', *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences* 2015/11, afl. 1, p. 13; J.E. Mathieu, T.S. Heffner, G.F. Goodwin, E. Salas & J.A. Cannon-Bowers, 'The influence of shared mental models on team process and performance', *Journal of applied psychology* 2000/85, afl. 2, p. 273.

25 Teaching by the Case Method – Christensen Center for Teaching and Learning – Harvard Business School (www.hbs.edu/teaching/case-method/Pages/default.aspx).

26 Bijv. A. Agresti, *Categorical Data Analysis*, New York: John Wiley and Sons 1990; W.H. Beyer, *CRC Standard Mathematical Tables*, 31ste ed., Boca Raton: CRC Press 2002, p. 536 en 571; S. Kotz, N.L. Johnson & C.B. Read (red.), *Encyclopedia of Statistical Sciences*, Wiley 2006; W.P. Vogt, *Dictionary of Statistics & Methodology: A Nontechnical Guide for the Social Sciences*, Londen: Sage 2005.

baarheid kan worden uitgedrukt in Cohen's kappa of in een percentage dat de mate van overeenkomst tussen de verschillende beoordelaars weergeeft, waarbij 1 (of 100%) een totale overeenstemming weergeeft en 0 (0%) verwijst naar geen enkele overeenstemming. Aangezien we niet uitsluitend werkten met categoriale variabelen, de antwoorden niet wederzijds exclusief waren en de beoordelaars niet op voorhand bewust werden gekozen en/of volledig onafhankelijk waren van elkaar, werd niet voldaan aan de voorwaarden om te werken met de Cohen's kappa. Bijgevolg opteerden we voor het werken met een percentage om de mate van interbeoordelaarsbetrouwbaarheid weer te geven. Daarnaast mag het duidelijk zijn dat de vragen waar men vrij iets kon beantwoorden of aanvullen (stukken tekst) buiten beschouwing werden gelaten voor de berekening van dit percentage.

Om de interbeoordelaarsbetrouwbaarheid na te kunnen gaan werd een deel van de rechtspraak die reeds door een groepje studenten werd gecodeerd, nadien nog eens opnieuw gecodeerd door twee studenten uit een andere groep om vervolgens te kijken naar de mate van overeenstemming. Meer specifiek kregen studenten per twee de opdracht om nog één ander arrest of vonnis te coderen. In totaal werden er zo 63 vonnissen en arresten dubbel gecodeerd (8% van alle arresten en vonnissen). Idealiter wordt een zo groot mogelijk gedeelte van de data dubbel gecodeerd. Veel hangt natuurlijk ook af van de totale omvang van de steekproef alsook het type van onderzoek. Wanneer men bijvoorbeeld interviews zal gaan coderen (waarvan de omvang doorgaans bijvoorbeeld 15 à 20 is) dan zal men al snel 25% van de data dubbel gaan coderen. Wanneer de steekproef veel omvangrijker is, dan is het vooral erg belangrijk om heel nauw het dubbel coderen op te volgen door bijvoorbeeld tussentijds reeds de interbeoordelaarsbetrouwbaarheid te berekenen en het codeerschema verder aan te passen mocht deze betrouwbaarheid nog niet voldoende hoog zijn. Daarnaast is ook het overleg tussen de codeerders belangrijk om zo de neuzen in dezelfde richting te krijgen. Of in andere woorden, dat men beschikt over een *shared mental model* om de data te interpreteren. Hoewel dit altijd nuttig en aangewezen is, is het extra belangrijk in een setting waarin er gewerkt wordt met studenten die zich nog volop de inhoud aan het eigen maken zijn. We hebben dit in ons onderzoek trachten te onderwerpen door de *close reading*, de *Assess-Discuss-Review*-methode zoals eerder ook aangehaald en het feit dat onderzoekers met expertise in het vakgebied arresten en vonnissen dubbel hebben gecodeerd. Bij nadere analyse was er slechts één vonnis of arrest waarbij er sprake was van een zeer lage overeenstemming tussen de beoordelaars.

Om het percentage van overeenstemming tussen de beoordelaars te berekenen werden alle antwoorden van de beide beoordelaars op de verschillende variabelen manueel gecodeerd. In geval van overeenstemming, kregen ze een '1', wanneer er geen sprake was van overeenstemming voor die desbetreffende variabele een '0'. Op basis van de mate van overeenstemming over alle variabelen heen, komt men tot het percentage van overeenstem-

ming. Er werd per variabele (en dus niet per antwoordcategorie) gekeken naar de mate van overeenstemming. Hiertoe werden voor sommige variabelen de verschillende binaire antwoorden samen genomen. De percentages van overeenstemming variëren hier van 55% tot en met 100%. De gemiddelde overeenstemming over alle 63 vonnissen en arresten heen bedraagt 78%. Wat een 'goed' percentage is voor de interbeoordelaarsbetrouwbaarheid wordt mede bepaald door het onderzoeksveld.²⁷ Voor dokters moet dit bijvoorbeeld 90% zijn (men moet het namelijk beiden zo goed als helemaal eens zijn of een patiënt kanker heeft of een specifieke behandeling nodig heeft). Doorgaans spreekt men vanaf 61% over een substantiële overeenkomst, wordt 75% reeds als zeer goed bevonden en spreekt men vanaf 81% over een 'near perfect agreement'. Het percentage van 78% weerspiegelt een (zeer) goede overeenstemming tussen de beoordelaars.

Verder is het interessant te zien dat de verschillen tussen de beoordelaars niet zozeer te situeren zijn bij de codering van de rechtsgrond en uitkomst, maar vooral bij de elementen in de redenering van de rechtbank en vervolgens bij de feitelijke situatie van de partners.

4 RESULTATEN

Het onderzoek leidde tot enkele opmerkelijke eerste vaststellingen.

De overgrote meerderheid van de vonnissen en arresten handelt over conflicten na een relatiebreuk. Slechts een klein deel van de uitspraken gaat over conflicten naar aanleiding van een overlijden. In eerste aanleg gaat het merendeel van de vonnissen over onverdeeldheden, met name de vereffening-verdeling van een huis, vaak met aanstelling van een notaris-vereffenaar. Het betreft hier doorgaans rechttoe rechtaan uitspraken, die relatief eenvoudig zijn. De zaken die tot voor de hoven van beroep komen, tonen een veel grotere verscheidenheid aan rechtsgronden.

Opmerkelijk is dat de meeste zaken door vrouwen worden ingeleid: in eerste aanleg worden 361 zaken ingeleid door vrouwen tegenover 244 door mannen. Het ligt niet voor de hand om hieruit conclusies te trekken op het vlak van gender, hoewel vaststaat dat er genderdynamieken spelen bij echtscheiding.²⁸

27 J. Cohen, 'A coefficient of agreement for nominal scales', *Educational and Psychological Measurement* 1960/20, p. 37-46.

28 Vgl. bijv. K. Bollen, A. Verbeke & M.C. Euwema, 'Money or children? Power sources in divorce mediation', *Journal of Family Studies* 2013/19, afl. 2, p. 159-173.

Een opvallende bevinding is verder dat in de vonnissen en arresten de feitelijke situatie van de partijen (bijvoorbeeld hun leeftijd, inkomen, of er kinderen zijn of niet) amper aan bod komt. Nochtans zou men op het eerste gezicht kunnen veronderstellen dat de feitenconstellatie hier een belangrijke rol speelt. Bij gebrek aan een specifieke wettelijke regeling voor feitelijke samenwoners moet de rechter immers het gemeen recht toepassen. Men zou kunnen denken dat de context van de feitelijke samenwoning hier tot een bijzondere interpretatie van het gemeen recht zou nopen, maar dit blijkt niet uit de dataset. Ofwel houden rechters hier minder dan verwacht rekening met de feitelijke situatie van de partijen, ofwel schrijven rechters eventuele casusspecifieke redelijkheids- en billijkheidsargumenten niet op bij het redigeren van een vonnis of een arrest. Om dit aspect verder uit te diepen, is er kwalitatief onderzoek nodig waarbij rechters en raadsleden worden geïnterviewd om te peilen naar de achterliggende motieven van rechterlijke beslissingen.

Daarnaast stelden de studenten vast dat *'cohabitation de fait'* minder een geïjkt begrip lijkt in de lagere Franstalige rechtspraak dan het geval is voor de feitelijke samenwoning in de lagere Nederlandstalige rechtspraak. De Franstalige rechtbanken van eerste aanleg gebruikten een brede waaier aan termen om de feitelijke samenwoning aan te duiden, wat het niet altijd eenvoudig maakte om de relevante vonnissen te identificeren. De Nederlandstalige rechtbanken van eerste aanleg spraken veel consistentere van feitelijke samenwoning.

Tot slot blijkt dat in slechts 2% van de geanalyseerde rechtspraak een samenlevingsovereenkomst tussen feitelijke samenwoners voorkwam. Dit lijkt eerdere vaststellingen over het gebrek aan populariteit van samenlevingsovereenkomsten te bevestigen. Men zou echter ook kunnen speculeren over de vraag of er misschien in geval van een samenlevingsovereenkomst minder conflicten door de rechter moeten worden beslecht.

5 DISCUSSIE

5.1 *Bijdrage van het onderzoek aan de rechtswetenschap*

Ons onderzoek is op verschillende manieren vernieuwend. Het is inhoudelijk vernieuwend aangezien een dermate grootschalige rechtspraakanalyse aangaande de feitelijke samenwoning nog niet eerder werd uitgevoerd. Anders dan gebruikelijk het geval is,²⁹ baseren we

29 Zie bijv. L. De Schrijver, 'Wettelijke en feitelijke samenwoning', in: G. Verschelden, I. Boone, S. Brouwers, L. De Schrijver & A. Van Thienen, 'Overzicht van rechtspraak. Familierecht (2012-2016)', *TPR* 2017, p. 655 en J. Lambrechts & A. Marissens, 'Overzicht van rechtspraak (2008-2016) – De feitelijke samenwoning', *T.Fam.* 2017, p. 4.

ons bovendien niet enkel op de gepubliceerde rechtspraak maar ook op ongepubliceerde rechtspraak. We krijgen zo meer inzicht in de problemen die feitelijke samenwoners ervaren³⁰ en op welke manier de hoven en rechtbanken het bestaande (gemeenrechtelijke) rechtskader toepassen op conflicten tussen feitelijke samenwoners. Deze inzichten helpen om de geschiktheid van het gemeen recht voor conflicten tussen samenwoners te evalueren. Zodoende kunnen zij ook bijdragen aan een theorievorming over de feitelijke samenwoning.

Verder is ons onderzoek vernieuwend omdat we niet de traditionele, casusgewijze methode van rechtspraakanalyse toepassen, maar onze vraagstelling empirisch-juridisch benaderen. Meer nog, we hebben ervoor geopteerd om voornamelijk kwantitatief te werk te gaan. Dit is in tegenstelling tot het voornamelijk kwalitatief onderzoek dat de meeste juristen uitvoeren omdat deze methode het dichtst staat bij de beschrijvende methoden waar juristen vertrouwd mee zijn.³¹ Hoewel een empirisch-juridische analyse van rechtspraak ook in België een opmars maakt,³² blijft een dergelijke benadering toch eerder uitzonderlijk.

Tot slot is ons onderzoek vernieuwend omdat er ook een pedagogische component aan ons onderzoek is verbonden. Het onderzoek is niet alleen relevant voor de academische kennis rondom deze thematiek maar voegt ook educationeel waarde toe: data verzamelen, reflecteren, analyseren, interpreteren en dus leren of inzicht verwerven gaan hand in hand. De masterstudenten pakten de rechtspraak letterlijk vast, benaderden deze vanuit een totaal andere invalshoek dan gewoonlijk, met een codeerschema, gingen daarover met elkaar in discussie, eerst in kleine groepen en dan in de plenaire colleges. Ze schreven hun collectieve bevindingen en dialoog ook uit in diverse memoranda (zie verder *infra*). In zekere zin sluit dit aan bij het concept van *action research* waarbij niet alleen het onderzoek centraal staat, maar waarbij men ook actief deelneemt aan het proces waardoor er een beter inzicht ontstaat in de problematiek en/of het onderzochte thema. Tegelijkertijd vormt het onderzoek een toepassing van *team science*: het wordt gedreven door een samenwerking tussen onderzoekers met verschillende achtergronden en verschillende niveaus van expertise (studenten, juristen, een psycholoog en ondersteunend personeel). Het gaat om

30 Zie bijvoorbeeld de vaststelling dat het merendeel van de conflicten in eerste aanleg handelt over de verdeling van een huis (*supra*).

31 A.J. Akkermans, 'Enkele praktijkervaringen met empirisch juridisch onderzoek op het snijvlak van gezondheid en recht', in: W.H. Van Boom, I. Giesen & M. Smit (red.), *Bundel van het 1e Nederlandse Civilologiecongres*, Den Haag: Boom Juridische uitgevers 2012, p. 83-108 en M.V. Antokolskaia, 'Opkomst van Empirical Legal Studies: een vloek, een zegen, of allebei?', *Tijdschrift voor Privaatrecht* 2016, afl. 2/3, p. 423-432.

32 Zie bijv. J. De Jaegere, *Judicial review and strategic behaviour: an empirical case law analysis of the Belgian Constitutional Court*, Antwerpen: Intersentia 2019 en het werk van A. Dyevre (vgl. zijn bijdrage in deze bundel).

multidisciplinair en collaboratief onderzoek over de grenzen van specifieke domeinen of disciplines heen, dat ook expliciet op de agenda van de NWO en de VSNU staat.³³

5.2 *Uitdagingen*

Geen enkel onderzoek verloopt zonder slag of stoot. We noteren verschillende uitdagingen waaruit we lessen kunnen trekken voor vervolgonderzoek.

5.2.1 *Medewerking van de rechtscolleges*

Een eerste uitdaging was de medewerking van de rechtscolleges. Hoger vermeldden we reeds dat twintig van de 33 aangeschreven hoven en rechtbanken hun medewerking toezegden (61%). Verschillende van deze rechtscolleges vroegen bijkomende informatie of een telefonisch onderhoud vooraleer zij toezegden. Andere wensten eerst onderling te overleggen alvorens hun medewerking te bevestigen. Bij de rechtscolleges die niet meewerkten, ontvingen we van de meeste geen antwoord. Een enkel rechtscollege weigerde met uitgebreide motivering. Uit deze contacten met de hoven en rechtbanken bleek een aantal gedeelde bezorgdheden. We geven hieronder een overzicht van die bezorgdheden en van ons antwoord erop.

De voornaamste bezorgdheid van de hoven en rechtbanken betrof de privacy van de betrokken partijen. Dat is een zeer legitieme zorg en bovendien ook voor ons onderzoek een belangrijke uitdaging, die een aparte toelichting verdient (*zie infra*).

Daarnaast wenste men erover te waken dat de dataverzameling door de studenten de normale werking van de griffie niet zou storen. Om die reden opteerden drie hoven en rechtbanken ervoor om zelf geselecteerde uitspraken op te sturen in plaats van de studenten te ontvangen. Er is hier een bepaalde onduidelijkheid over de criteria voor selectie van de uitspraken. We vertrouwen erop dat de hoven en rechtbanken zich deze taak ter harte hebben genomen. We kunnen echter niet uitsluiten dat het toegezonden materiaal slechts een deel van de relevante uitspraken omvatte. Op de griffie zelf liet de werklast van de griffiemedewerkers het niet toe om de studenten bij te staan bij de selectie van de rechtspraak. De studenten moesten bovendien de door de griffie gegeven classificatie van de uitspraken respecteren. Tijdens de colleges werden de studenten erop gewezen dat ze bij de selectie van de uitspraken screen te werk dienden te gaan, en gaven we daarbij ook een

³³ www.vsnunl/en_GB/news-items.html/nieuwsbericht/489-vsnunwo-nfu-en-zonmw-geven-impuls-aan-verandering-in-het-waarderen-en-belonen-van-wetenschappers.

aantal instructies. Daarnaast boden wij aan alle rechtscolleges aan om de kopiekosten ter griffie te betalen om eventuele financiële bezwaren weg te nemen.³⁴

Tot slot deelden sommige rechtscolleges ook hun frustratie mee over het uitblijven van een centrale databank waarop geanonimiseerde vonnissen toegankelijk zouden zijn voor iedereen, niet in het minst voor de magistraten zelf.

Een belangrijke factor om de hoven en rechtbanken tot medewerking te bewegen, bleek de goede reputatie van het RRDI en van zijn directeur, prof. Alain Laurent Verbeke. De communicatie met de hoven en rechtbanken gebeurde door het RRDI, niet door de studenten. Vaak verklaarde een rechtscollege zich pas bereid om mee te werken na een persoonlijk contact met een medewerker van het RRDI of met zijn directeur. Mogelijk verklaart dit ook waarom meer Nederlandstalige dan Franstalige rechtbanken van eerste aanleg meewerkten aan het onderzoek;³⁵ we kunnen veronderstellen dat het RRDI aan de andere kant van de taalgrens minder bekendheid geniet. Dit argument lijkt evenwel niet te spelen op niveau van de hoven van beroep: alle hoven van beroep werkten mee.

De studenten werden herhaaldelijk gewezen op hun rol en voorbeeldfunctie als vertegenwoordiger van de KU Leuven, de faculteit Rechtsgeleerdheid en het RRDI in de contacten met de hoven en rechtbanken. We onderstreepten dat de goede relatie tussen de faculteit en de hoven en rechtbanken ook door de studenten mee werd uitgedragen. Deze benadering lijkt te hebben gewerkt. De studenten toonden in onze contactmomenten duidelijk een respectvolle houding tegenover de rechtscolleges en koppelden terug wanneer zij twijfelden over de meest gepaste benadering. We ontvingen ook enkel positieve geluiden over het gedrag van de studenten bij het opzoekwerk ter griffie. Die volwassen houding van onze studenten ligt in de lijn van de ervaringen die wij al vele jaren hebben bij het uitrollen van allerlei taken en opdrachten in de master rechten en de master notariaat en die in de vreselijke coronatijden van de academiejaren 2019-2020 en 2020-2021 nog extra werd bevestigd.

5.2.2 *Privacy van de partijen*

Dan is er dus de grote uitdaging van de privacy van partijen. De ongepubliceerde rechtspraak die de studenten ter griffie raadpleegden, bevatte vertrouwelijke informatie over de identiteit van de partijen. Het is uiteraard cruciaal om het vertrouwelijk karakter van

34 Betaald door het Fonds Wetenschappelijk Onderzoek – Vlaanderen; beurs nr. 40461.

35 Van de vijftien afdelingen van rechtbanken van eerste aanleg die meewerkten aan het onderzoek, zijn er tien Nederlandstalig en vijf Franstalig.

persoonlijke informatie te waarborgen, zeker bij een delicate materie als familiale conflicten. Om die reden stelden wij de volgende werkwijze voorop.

Alle griffies ontvingen op voorhand de namen van de betrokken studenten. De studenten moesten de rechtspraak ter griffie kopiëren en vervolgens anonimiseren met een (zelf mee te nemen) alcoholstift. Alle persoonsgegevens van de partijen moesten worden geanonimiseerd (in het bijzonder hun identiteit en woonplaats). Dit gebeurde verplicht op de griffie. Eventuele instructies van de rechtscollages en griffies in dit verband moesten strikt worden gevolgd. Onder geen beding mochten niet-geanonimiseerde kopieën van rechtspraak de griffie verlaten. De geanonimiseerde kopieën werden na het verlaten van de griffie gedigitaliseerd (ingescand) en aan ons bezorgd.

Daarnaast werden de studenten uitvoerig onderwezen over het belang van een vertrouwelijke omgang met de gegevens. Zij ontvingen de instructie om het project en de rechtspraak-analyse met de hoogste kwaliteitstandaarden qua wetenschappelijke integriteit, privacy, vertrouwelijkheid en zorgvuldigheid te benaderen. Vertrouwelijke informatie mocht niet worden gedeeld, ook niet met ons.

De studenten beschikten voor hun rechtspraakanalyse enkel over de (geanonimiseerde) vonnissen en arresten die waren toegewezen aan hun groep. Enkel de onderzoekers van het RRDI hadden toegang tot de databank met de totaliteit van alle verzamelde rechtspraak. De volledige dataset, inclusief alle geanonimiseerde rechtspraak, werd nadien ook bezorgd aan de vier masterstudenten die hun masterproef schreven op basis van de data. Zij moesten bijkomend de volgende clausule ondertekenen:

1. *‘verklaart zich, bij de uitvoering van empirisch onderzoek in het kader van zijn/haar masterproef, te verbinden tot naleving van de discretie, vertrouwelijkheids- en geheimhoudingsplicht van uitspraken, dossiers, gegevens en inlichtingen, waarvan zij/hij tijdens de uitvoering van het empirisch onderzoek kennis krijgt.*
2. *verbindt zich ertoe om de uitspraken, Excel files en andere documenten verkregen als onderdeel van het empirisch onderzoek op geen enkele manier en aan geen enkele persoon verder te verspreiden, om welke reden dan ook, behoudens toestemming van professor Verbeke of dr. Goossens.*
3. *verbindt zich ertoe om, na kennisname van het bericht dat hij/zij geslaagd is voor de masterproef, de uitspraken, Excel files en andere documenten in zijn/haar bezit verkregen als onderdeel van het empirisch onderzoek te verwijderen van alle elektronische dragers (laptop, tablet, harde schijf, GSM etc.) en alle fysieke kopieën te vernietigen.’*

De meeste deelnemende rechtscolleges namen genoegen met deze werkwijze. Drie rechtscolleges verkozen om zelf rechtspraak te selecteren en deze geanonimiseerd te bezorgen, in plaats van studenten toe te laten ter griffie. Een aantal rechtscolleges vereiste bijkomend dat de studenten een vertrouwelijkheidsclausule ondertekenden vooraleer zij de rechtspraak mochten inkijken. Van de rechtscolleges die hun medewerking weigerden, gaf één rechtbank expliciet gegevensbescherming als reden (en meer in het bijzonder de overeenstemming met de Algemene verordening gegevensbescherming).³⁶

5.2.3 Rol van studenten

Het onderzoek werd in de eerste plaats uitgevoerd door masterstudenten, weliswaar onder ons toezicht. De grote betrokkenheid van studenten behoorde tot de essentie van dit onderzoek; ons project had immers ook een pedagogische doelstelling (*supra*). De grote verantwoordelijkheid van studenten vormde echter ook een uitdaging voor het welslagen van het onderzoek. Het ging immers om studenten die nog niet afgestudeerd waren en wier kennis van het familiaal vermogensrecht nog zeer pril was. Toch werden zij verondersteld om te beslissen of een vonnis of arrest al dan niet moest worden opgenomen in het onderzoek, en moesten zij het vonnis of arrest vervolgens analyseren en coderen. De goede resultaten op het vlak van de interbeoordelaarsbetrouwbaarheid (zie *supra*) tonen aan dat ook een grootschalig onderzoek met studenten tot solide onderzoeksresultaten kan leiden. De betrokken masterstudenten gaven gedurende het hele project blijk van een grote motivatie en onderzoeksmaturiteit.

De open communicatie met en veelvuldige terugkoppeling naar de studenten was essentieel om dit project te laten slagen. We besteedden veel tijd aan het uitleggen van het belang van het onderzoek en aan het enthousiasmeren van studenten. De studenten werden betrokken als volwaardige onderzoekers, naar wier input werd geluisterd. Hiervoor lichtten we reeds toe dat we drie plenaire colleges organiseerden met de volledige groep studenten, waarin de balans van het onderzoek werd opgemaakt en waar nodig werd bijgestuurd (bijv. aanpassen codeerschema). De groepen studenten moesten tweemaal schriftelijk verslag

³⁶ De spanning tussen de principiële openbaarheid van rechtspraak, zoals bepaald in art. 6 EVRM en art. 149 Gw., en de privacy van partijen is het voorwerp van een klassiek debat, dat de laatste jaren bijzonder actueel is geworden door de Algemene verordening gegevensbescherming (AVG). Zie hierover onder meer S. De Somer & C. Fornoville, 'Rechtspraakpublicatie anno 2011: historische beschouwingen, hedendaagse bedenkingen en toekomstperspectieven', *RW* 2011-12, nr. 4, p. 206 (evenwel daterend van vóór de invoering van de AVG). Zie voor het Nederlandse debat over de invloed van de AVG op de mogelijkheden tot openbaarmaking van jurisprudentie: M. van Opijnen, "Alle rechtspraak online? Hoe dan?" Noodzakelijke ingrediënten voor een wettelijke regeling', *Ars Aequi* 2021/127, alsmede Peeraer en Van Gestel in deze bundel.

doen van hun bevindingen (tussentijds verslag en groepsmemo). De onderzoekers van het RRDI waren tussendoor steeds beschikbaar voor verder overleg, via e-mail of na de lesmomenten. Veel studenten maakten hier gebruik van. De studenten moesten de bijdragen van de groepsgenoten en van zichzelf ook tweemaal evalueren in een *peer assessment*: bij het tussentijds verslag en na de indiening van de groepsmemo. De bedoeling hiervan was dat de studenten elkaar zouden aanmoedigen en ook scherp houden, alsmede om *free rider* gedrag snel in te dijken. Dit heeft wellicht goed gewerkt. Bij de eerste *peer assessment* gaf deze werkwijze de studenten de mogelijkheid om waar nodig hun inspanningen op te drijven.

Tot slot kregen de studenten voor hun gedegen inspanningen ook een bonuspunt op het examen Familiaal vermogensrecht. Hoewel dit bonuspunt wellicht mede een rol heeft gespeeld bij het motiveren van de studenten om bij te dragen aan het onderzoek, denken wij dat dit geen determinerende aansporing was voor de grote inzet, motivatie en engagement van de studenten. Tijdens de colleges was er een grote betrokkenheid en veel interactie en ook uit de *peer review* konden we afleiden dat er goed en intens werd samengewerkt tussen de studenten. We menen dat het project zelf en het opzet van onze colleges hiertoe hebben bijgedragen. Dit sluit aan bij de zelfdeterminatietheorie van Ryan en Deci.³⁷ In deze motivatietheorie gaat men ervan uit dat men het meest en meest duurzaam wordt gemotiveerd wanneer men intrinsiek gemotiveerd is. Men kan dit bekomen wanneer er rekening wordt gehouden met drie elementen: autonomie, betrokkenheid en het gevoel van competentie. Aan de hand van de lessen en de discussies in groepsverband werd er gewerkt aan hun niveau van competentie, door het interactieve groepswerk werd er gewerkt aan betrokkenheid en daarnaast kregen de studenten autonomie om de dataverzameling te regelen. Ze kregen als het ware een rugzakje mee om samen en autonoom op weg te gaan.

Als pedagogisch experiment gaven we in maart 2019 ook aan 25 bachelorstudenten van het vak Personen-, familie- en familiaal vermogensrecht de mogelijkheid om in groepsverband een vijftal arresten te coderen. Dit gebeurde onder begeleiding van een onderzoeker van het RRDI. De studenten bleken evenwel onvoldoende juridisch inzicht te hebben om tot een gedegen codering te komen, zo bleek uit de vragen die zij tijdens dit college stelden. De codering van deze studenten werd dan ook niet meegenomen in het onderzoek. Er zijn diverse verklaringen denkbaar voor dit gebrek aan expertise van de bachelorstudenten. Er is vooreerst de algemene juridische vorming, die bij deze studenten in het tweede semester van het tweede jaar van de opleiding nog erg beperkt is. Dit in tegenstelling tot de studenten

37 E.L. Deci & R.M. Ryan, 'Self-determination theory: A macrotheory of human motivation, development, and health', *Canadian psychology/Psychologie canadienne* 2008/49, afl. 3, p. 182.

van de master, in het vierde of vijfde jaar, die al een bachelor in de rechten behaalden en dus alle technische juridische vakken in de voorbije drie jaar met succes volgden. Daarnaast is er het gebrek aan specifieke expertise in de technische materie van het familiaal vermogensrecht omdat deze bachelorstudenten op het moment van de coderingen nog in het prille begin van het tweede semester het basisvak familiaal vermogensrecht aan het volgen waren en er mogelijk nog niet echt op gestudeerd hadden en er ook niet met succes een examen over hadden afgelegd. Tot slot is er de mogelijk positieve selectie van de masterstudenten die het klemtoonvak familiaal vermogensrecht kozen. Dit zijn studenten die bijzonder gemotiveerd en geïnteresseerd zijn in het privaatrecht (dat zij kozen als maior of minor) en die daarenboven specifiek geïnteresseerd zijn in het familiaal vermogensrecht (dat een vak is met een zware reputatie en gemiddeld lagere examenscores dan andere vakken privaatrecht).

5.3 *Beperkingen*

We benadrukten hiervoor al dat ons onderzoek voornamelijk kwantitatief van aard is waarbij het de bedoeling is conclusies te trekken over de types van conflicten tussen feitelijke samenwoners en de betrokken rechtsgronden. Derhalve komen vragen zoals waarom samenwoners niet trouwen, welke verwachtingen zij hebben over hun relatie, en in welke mate zij op de hoogte zijn van de juridische implicaties van de gekozen relatievorm, niet aan bod in dit onderzoek. Vervolgonderzoek zou dieper kunnen ingaan op deze thema's.

Voorts is voorzichtigheid geboden bij de interpretatie van de onderzoeksresultaten. Het onderzoek laat enkel toe om bepaalde vaststellingen en patronen over conflicten tussen feitelijke samenwoners te beschrijven. Mogelijke verklaringen waarom de resultaten in de ene of de andere zin zijn, vloeien niet rechtstreeks voort uit het onderzoek en zijn dus eerder speculatief. Het is daarom niet evident om aan het onderzoek ook normatieve conclusies te verbinden.

Daarnaast betreft het onderzoek weliswaar een aanzienlijke hoeveelheid vonnissen en arresten over de bestudeerde thematiek, maar omvat het niet alle relevante beslissingen uit de betrokken periode (september 2014-september 2018). Zo deden niet alle aangeschreven rechtbanken mee aan het onderzoek, verzamelden de studenten bij sommige deelnemende rechtbanken niet alle relevante rechtspraak vanwege de hoge werklast, en kozen sommige hoven en rechtbanken ervoor om zelf een selectie op te sturen.

Tot slot omvat het onderzoek enkel conflicten die voor de rechtbank komen. Het is niet duidelijk of deze ook representatief zijn voor de conflicten die feitelijke samenwoners

doorgaans ervaren na beëindiging van de relatie, dan wel of men alleen voor bepaalde conflicten de stap naar de rechtbank wil maken. Zo kan men zich afvragen of het grote aantal zaken die handelen over de verdeling van een huis, er ook effectief op wijzen dat dit het meest voorkomend probleem is bij een relatiebreuk na een feitelijke samenwoning. Een mogelijke andere verklaring is dat feitelijke samenwoners voor deze conflicten sneller naar de rechter stappen, bijvoorbeeld omdat de gezinswoning het grootste vermogensbestanddeel is, of de grootste emotionele waarde heeft, of nog omdat men deze vordering de grootste kans op slagen toedicht.

6 CONCLUSIE

Ondanks de beperkingen en uitdagingen heeft dit onderzoek ons een aantal nuttige zaken bijgebracht.

Zo hebben we geleerd dat het ondanks de moeilijke toegankelijkheid van onze rechtspraak mogelijk is om bijna duizend uitspraken over een korte periode en een specifiek onderwerp bijeen te krijgen, dank zij de goede wil van de actoren van Justitie en het vele werk van een grote groep vrijwilligers. Desalniettemin heeft deze grote inspanning voor ons alleen nog maar benadrukt hoe essentieel het voor de burger in het algemeen en de juristen in het bijzonder is om alle rechterlijke uitspraken op een vlotte en digitale manier te kunnen consulteren. Vooral voor academische juristen is dit cruciaal: willen we ooit op een volwassen manier aan kwantitatief empirisch onderzoek kunnen doen, dan zal onze basisgrondstof toch gemakkelijker ontginbaar moeten zijn.

Daarnaast bevestigde dit project dat kwantitatief empirisch onderzoek het best wordt gevoerd in multidisciplinair verband.³⁸ Ook al kozen wij wegens de beperkte voorkennis van de studenten voor een toegankelijk codeerschema en een gegevensverwerking met Excel, dan nog bleek de samenwerking tussen juristen en een psycholoog essentieel voor het welslagen van het project. De juristen en masterstudenten reichten stonden in voor het inhoudelijke luik: het bepalen van de variabelen in het codeerschema, de eigenlijke analyse van de rechtspraak en de vertaling van de uitspraken naar het codeerschema. De psycholoog bewaakte de methodologische component. Zij zorgde voor de training van de studenten (volgens het principe van *Assess-Discuss-Review*; *supra*) en controleerde de interbeoordelaarsbetrouwbaarheid.

38 Zie ook Vols in deze bundel.

Ook hebben we geleerd dat het kwantitatief onderzoek van de rechtspraak bijna noodzakelijk moet worden aangevuld met een kwalitatief luik. Interviews met de rechters en raadsheren over de gevonden patronen en bevindingen zou het onderzoek van een descriptief naar een voorzichtig normatief niveau kunnen brengen.

Tevens hebben we geleerd dat we aan de universiteit bevoorrecht zijn, met een groep intelligente, hard werkende en goed gemotiveerde studenten. Maar dat wisten we natuurlijk al. Heel in het bijzonder hebben we geleerd dat de interactie met de studenten als volwaardige onderzoekspartners een grote toegevoegde waarde oplevert voor het denkproces en verfijning van het onderzoeksdesign. In de rechtenfaculteit bestaat geen traditie om masterproefonderzoek in te bedden in het ruimere onderzoek van een instituut of een departement. De ervaring van dit project toont aan dat we daar nuttig mee zouden kunnen experimenteren, zoals bijvoorbeeld wel gebeurt aan de faculteit psychologie.

De pedagogische toegevoegde waarde van dit project is groot voor onze studenten, zo bevestigen ook de studenten zelf. Het anders kijken naar rechtspraak, patronen leren zien, daarover met elkaar goed gestructureerd in discussie gaan, erover rapporteren, schriftelijk en mondeling in plenaire colleges, elkaar aanmoedigen en evalueren via de *peer assessments*, het zijn boeiende aspecten van een gevarieerd leerproces. Zelf hebben wij kunnen vaststellen dat de interacties tijdens de discussiecolleges van een hoog niveau waren, hoger dan bij andere discussiecolleges, zo denken wij. Nochtans zijn onze studenten daar ook gemotiveerd en goed voorbereid. We veronderstellen dat een belangrijke factor hier was dat de voorbereiding veel verder ging dan het louter theoretische, en sterk doorleefd was door terreinwerk op de griffies en met elkaar. Het feit dat de studenten een actieve rol kregen, en medeactor waren in het denkproces, creëert een gevoel van autonomie en betrokkenheid, dat wellicht een positieve rol speelde. We nemen hieruit voor toekomstige discussiecolleges mee dat het nuttig is om de studenten zelf actiever bij het voorbereidingsproces van dergelijke colleges te betrekken.

Tot slot hopen we dat dit onderzoek ook andere collega's kan inspireren en aanzetten tot experimentele vormen van onderzoek en dialoog met masterstudenten. We kunnen er langs de beide kanten veel van leren.



6 METHODOLOGISCHE KANSEN EN BEPERKINGEN VAN GEBRUIK VAN STATISTIEK EN MACHINE LEARNING BIJ DE BESTUDERING VAN RECHTSPRAAK

*Michel Vols**

INLEIDING

Dit hoofdstuk gaat over de kwantitatieve analyse van jurisprudentie. Het vormt een aanvulling en vervolg op de publicatie van Vols en Jacobs uit 2017.¹ In die publicatie wordt een onderscheid gemaakt tussen het tellen *van* jurisprudentie en het tellen *in* jurisprudentie. Het tellen van jurisprudentie gaat over het achterhalen en rapporteren van het aantal uitspraken dat rechters over een rechtsvraag hebben gedaan. Het tellen in jurisprudentie slaat op het verkrijgen van kwantitatieve data over rechtspraak door het systematisch verzamelen en analyseren van gerechtelijke uitspraken.²

In onderliggend hoofdstuk gaat het, net als in de eerdere publicatie, over het tellen *in* jurisprudentie. Dit noem ik in de rest van het hoofdstuk kwantitatieve jurisprudentieanalyse. De publicatie van Vols en Jacobs richt zich op de algemene eisen die gelden voor kwantitatieve jurisprudentieanalyses. Het doel van dit hoofdstuk is door een bespreking van drie onderzoeksprojecten meer details te geven over hoe de methode van kwantitatieve jurisprudentieanalyse in juridische onderzoek kan worden gebruikt.³ Ik bespreek methodologische eisen die voor kwantitatieve jurisprudentieanalyses gelden. Bovendien bespreek ik

* Prof. mr. dr. M. Vols is hoogleraar Openbare-orderecht aan de Rijksuniversiteit Groningen. Ik bedank Marije Eding voor de nuttige opmerkingen bij een eerdere versie van dit hoofdstuk.

1 M. Vols & J.P.A.M Jacobs, 'Juristen als rekenmeesters: over de kwantitatieve analyse van jurisprudentie', in: P.A.J. van den Berg & G. Molier (red.), *In dienst van het recht*, Den Haag: Boom juridisch 2017.

2 Vols & Jacobs 2017, p. 90.

3 Zie M. Medvedeva, M. Vols & M. Wieling, 'Using machine learning to predict decisions of the European Court of Human Rights', *Artificial Intelligence and Law* 2020[a], p. 237-266, op p. 239-242 (<https://doi.org/10.1007/s10506-019-09255-y>) en L. Epstein & N.W. Waterbury, 'Judicial behaviour', in: *Oxford Research Encyclopedia of Politics* 2020 voor overzichten van andere studies met kwantitatieve jurisprudentieanalyses.

– aan de hand van een van de onderzoeksprojecten – hoe computers een rol kunnen spelen in de automatische verzameling en analyse van jurisprudentie binnen de kwantitatieve jurisprudentieanalyse.

In de eerste paragraaf komen de vragen aan bod die in de drie te bespreken projecten centraal staan. Ook bespreek ik welke rol juridische en niet-juridische theorieën spelen in de projecten. De tweede paragraaf gaat over de data die gebruikt zijn in de drie projecten: welke en hoeveel uitspraken zijn verzameld en geanalyseerd? In de derde paragraaf komen de methoden van dataverzameling, -selectie (par. 3.1) en -analyse (par. 3.2) aan bod. Paragraaf 4 bespreekt de lessen die kunnen worden getrokken uit de verschillende projecten, en verkent wat de kansen en beperkingen zijn van de methode. Ik sluit af met een korte vooruitblik op nieuwe projecten.

1 VRAGEN EN THEORIEËN

In dit hoofdstuk staan drie onderzoeksprojecten centraal. Tabel 1 geeft een kort overzicht van de projecten. Het eerste project gaat over ontbindings- en ontruimingsvorderingen binnen het Nederlandse huurrecht. Dit project is uitgevoerd in het kader van een VENI-subsidie van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO), en wordt in de rest van het hoofdstuk aangeduid als het VENI-project. Het tweede project betreft de toepassing van artikel 13b Opiumwet. Deze bepaling geeft de burgemeester de bevoegdheid om panden te sluiten vanwege drugshandel. Artikel 13b Opiumwet wordt ook wel de Wet Damocles genoemd, en daarom noem ik het project in de rest van het hoofdstuk het Damocles-project. Het derde project draait om het automatisch classificeren van uitspraken van het Europees Hof voor de Rechten van de Mens (EHRM), en wordt in het vervolg het EHRM-project genoemd. In de rest van deze paragraaf zal ik de drie projecten nader beschrijven, en de centrale onderzoeksvragen en de rol van theorieën in de projecten belichten.

Tabel 1 Projecten

Naam project	Belangrijkste bepaling	Rechtsgebied	Niveau jurisprudentie
VENI	Artikel 6:265 BW	Privaatrecht	Rechtbank
Damocles	Artikel 13b Opiumwet	Bestuursrecht	Rechtbank
EHRM	Titel 1 EVRM	Mensenrechten	EHRM

1.1 VENI-project

Het VENI-project betreft een privaatrechtelijk project. Het draait om huurrechtelijke huisuitzettingen. Als een huurder tekortschiet in de nakoming van zijn verplichtingen (denk aan huurachterstand of overlast), kan dat gezien artikel 6:265 Burgerlijk Wetboek (BW) leiden tot ontbinding van de huurovereenkomst. Anders dan in deze reguliere ontbindingsregeling, vereist artikel 7:231 BW dat de rechter de huurovereenkomst ontbindt.

De hoofdregel is dat een tekortkoming ontbinding van de huurovereenkomst rechtvaardigt. Op grond van de tenzij-regel in artikel 6:265 BW kan de rechter echter de vordering afwijzen, als de tekortkoming, gezien haar bijzondere aard of geringe betekenis, de ontbinding met haar gevolgen niet rechtvaardigt.⁴ Jaarlijks buigen rechters zich over meer dan 15.000 vorderingen van verhuurders, al wordt maar een relatief gering deel van de vonnissen op rechtspraak.nl gepubliceerd.⁵ In het VENI-project zijn de op rechtspraak.nl gepubliceerde vonnissen geanalyseerd (zie par. 2). In de bestudeerde vonnissen moet de rechtbank de vraag beantwoorden of de tekortkoming van de huurder de ontbinding van de huurovereenkomst rechtvaardigt. De centrale vraag van het VENI-project is hoe rechters omgaan met deze ruime ontbindingsbevoegdheid, en welke verweren van huurders invloed lijken te hebben op de beslissing van de rechter om wel of niet tot ontbinding te besluiten.

Binnen het VENI-project worden grofweg twee typen theorie gebruikt.⁶ Het eerste type theorie betreft klassiek-juridische theorieën (of leerstukken): het onderzoek maakt gebruik van systematische beschrijvingen van het positieve Nederlandse (huur)recht. De analyse van de vonnissen vindt plaats binnen het wettelijke kader van artikel 6:265 BW, en daarom spelen klassiek-juridische begrippen zoals ontbinding, ontruiming en tenzij-formule een belangrijke rol. Hoewel het VENI-project zelf niet tot doel heeft om bij te dragen aan de nadere systematisering van het huurrecht, hoopt het wel dieper inzicht in de uitleg van belangrijke juridische begrippen te geven.

Het tweede type theorie dat gebruikt is, betreft microlevel theorieën. Dat zijn hypothesen, of in andere woorden, 'stellige antwoorden op een gesloten vraag en zijn bedoeld om de

4 Zie HR 28 september 2018, ECLI:NL:HR:2018:1810.

5 Het gaat om ontbindingsvorderingen en/of ontruimingsvorderingen. In kort geding kan alleen de ontruiming van de huurwoning worden gevorderd, maar is de door de rechter gebruikte belangrijkste wettelijke maatstaf gelijk aan die in de bodemprocedure: art. 6:265 BW. Zie M. Vols, P.G. Tassenaar & J.P.A.M. Jacobs, 'Dutch Courts and Housing Related Anti-social Behaviour. A first statistical analysis of legal protection against eviction', *International Journal of Law in the Built Environment* 2015, 7, p. 148-161.

6 Zie M. Vols, *Juridisch onderzoek*, Den Haag: Boom juridisch 2020, p. 69-70 voor verschillende typen theorie die kunnen worden onderscheiden.

reactie ‘juist’ of ‘onjuist’ uit te lokken’.⁷ Microlevel theorieën staan veelal in juridisch onderzoek niet los van klassiek-juridische theorieën. Leerstukken contextualiseren door gaans de hypothesen in juridische zin, en dat is ook te zien bij de in het VENI-project gebruikte *micro-level* theorieën.⁸

- a. Als de tekortkoming druggerelateerde overlast betreft, is de kans op ontbinding groter.
- b. Als de huurder zich laat bijstaan door een gemachtigde, dan is de kans op ontbinding kleiner.
- c. Als de huurder een proportionaliteitsverweer voert, dan is de kans op ontbinding kleiner.

1.2 Damocles-project

Het Damocles-project is een bestuursrechtelijk project. Het richt zich op uitspraken van rechtbanken in zaken over de toepassing van artikel 13b Opiumwet. De burgemeester past de in deze bepaling neergelegde bevoegdheid tot het sluiten van panden meer dan duizend keer per jaar toe. Het is gezien de gevolgen van een sluiting niet vreemd dat bestuursrechters zeer geregeld moeten beoordelen of een sluiting rechtmatig is. Een belangrijke vraag in het Damocles-project is welke factoren de succeskans van de procespartijen beïnvloeden. Meer concreet gaat het bijvoorbeeld over de vraag of het voeren van een proportionaliteitsverweer door belanghebbenden de succeskans van de belanghebbenden bij de rechter vergroot.⁹ Daarnaast is onderzocht in welke mate kenmerken van procespartijen (bijv. grootte van de gemeente) de succeskans van de partijen beïnvloeden.

Er zijn klassiek-juridische en *microlevel* theorieën gebruikt in het project. Bij de dataverzameling en -analyse is aangesloten bij bestuursrechtelijke concepten en leerstukken, zoals het evenredigheidsbegrip als bedoeld in artikel 3:4 Algemene wet bestuursrecht (Awb) en artikel 8 van het Europees Verdrag van de Rechten van de Mens (EVRM). Daarnaast zijn microlevel theorieën gebruikt, onder andere ontleend aan de publicaties van Galanter.¹⁰ In deze theorieën zijn bestuursorganen als burgemeesters getypeerd als ‘upperdogs’ en burgers of bedrijven als ‘underdogs’. Voorbeelden van in het Damocles-project getoetste hypothesen zijn:

7 K. van den Bos, *Kijken naar het recht* (oratie Utrecht), 2014, p. 13.

8 Vols e.a. 2015; M. Vols, ‘European law and private evictions: property, proportionality and vulnerable people’, *European Review of Private Law* 2019, 29, p. 719-752.

9 L.M. Bruijn, M. Vols & J.G. Brouwer, ‘Home closure as a weapon in the Dutch war on drugs: Does judicial review function as a safety net?’, *International Journal of Drug Policy* 2018, p. 137-147.

10 Zie M. Galanter, ‘Why the “haves” come out ahead: Speculations on the limits of legal change’, *Law & Society Review* 1974, nr. 1, p. 95-160., zoals naar verwezen in: L.M. Bruijn & M. Vols, ‘Upperdogs Versus Underdogs’, *Recht der Werkelijkheid* 2020, nr. 1, p. 28-31.

- a. *Upperdogs* hebben een grotere kans om te winnen in de rechtszaal dan *underdogs*.
- b. Sterke *upperdogs* hebben een grotere kans om te winnen in de rechtszaal dan zwakke *upperdogs*.
- c. Sterke *underdogs* hebben een grotere kans om te winnen in de rechtszaal dan zwakke *underdogs*.

1.3 EHRM-project

Het EHRM-project is een mensenrechtelijk project, en focust op het automatisch classificeren van uitspraken van het EHRM. In deze uitspraken gaat het om de vraag of een recht uit titel 1 van het EVRM door een lidstaat is geschonden of niet. Het project heeft een sterk interdisciplinair karakter: samen met computationeel linguïsten zijn met computertechnieken alle arresten van het EHRM verzameld en geanalyseerd. Het doel van het project is om te onderzoeken of met behulp van *Natural Language Processing* (NLP)-technieken jurisprudentie zodanig kan worden geanalyseerd dat de uitspraken automatisch correct kunnen worden geclassificeerd. Dat wil zeggen dat een computermodel uitspraken op basis van de tekstuele informatie uit die uitspraken correct aanmerkt als (1) een uitspraak waarin het EHRM een schending van het EVRM constateert of (2) een uitspraak waarin het EHRM oordeelt dat er geen schending is.

In het project spelen klassiek-juridische theorieën over de werkwijze van het EHRM, de uitleg van bepalingen van het EVRM en de structuur van uitspraken van het EHRM een rol. Daarnaast zijn belangrijke concepten en theorieën uit de computationele linguïstiek gebruikt. Het gaat dan bijvoorbeeld om het uitgangspunt dat niet alleen de gehele arresten, maar ook de aparte woorden en combinaties van woorden in het arrest als data moeten worden aangemerkt ('text as data').¹¹ Ook komen theorieën aangaande training en test sets, *machine learning* en *k-fold cross validation* aan bod. De betekenis van al de gebruikte theorieën zijn hopelijk op een voor juristen te begrijpen manier in de publicaties beschreven, maar zullen verderop in dit hoofdstuk ook kort worden aangestipt.¹² Ten slotte zijn in het EHRM-project enkele *microlevel* theorieën gebruikt. Zo zijn de volgende hypothesen getoetst:

- a. Het classificeren van uitspraken is minder succesvol met een dataset die bestaat uit oudere uitspraken (van bijvoorbeeld tien jaar geleden) dan met een dataset die bestaat uit zeer recente uitspraken.
- b. De naam van de rechter is een voorspeller voor de uitspraak van het EHRM.

11 Zie J. Grimmer & B.M. Stewart, 'Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts', *Political Analysis* 2013, p. 1-31.

12 Zie voor de uitleg: Medvedeva, Vols & Wieling 2020a.

2 DATA

In de drie projecten is steeds systematisch jurisprudentie verzameld en geanalyseerd. Tabel 2 geeft een overzicht van het type jurisprudentie, de hoeveelheid uitspraken en de vindplaats van de data.

Tabel 2 Data

Naam project	Type uitspraken	Hoeveelheid uitspraken	Vindplaats uitspraken
VENI	Vonnissen rechtbank	483	www.rechtspraak.nl/
Damocles	Uitspraken rechtbank	217	www.rechtspraak.nl/
EHRM	Arresten EHRM	11.532	https://hudoc.echr.coe.int/

In het VENI-project en het Damocles-project zijn uitspraken van rechters in eerste aanleg verzameld. Het gaat om enkele honderden online gepubliceerde vonnissen of uitspraken. Het EHRM-project gebruikt Engelstalige arresten van het EHRM als data. Deze arresten zijn allemaal te downloaden op de website van het EHRM, meer specifiek de Hudoc-database.

3 METHODEN

In deze paragraaf bespreek ik hoe de voornoemde data zijn verzameld en geselecteerd. Vervolgens komt aan de orde welke analysemethoden zijn aangewend om de data te onderzoeken.

3.1 *Dataverzameling en -selectie*

Tabel 3 geeft een overzicht van kenmerken van de dataverzameling en -selectie in de drie verschillende projecten. In de publicaties in het kader van het VENI- en het Damocles-project is steeds expliciet genoemd waar de uitspraken zijn gevonden, wat de tijdsperiode is waarbinnen de uitspraken zijn gepubliceerd/gedaan en welke zoektermen zijn gebruikt bij het zoeken van de uitspraken. Zo is de reproduceerbaarheid van de dataverzameling gewaarborgd.¹³ Voor het EHRM-project zijn geen vaste zoektermen gebruikt, omdat alle in september 2017 beschikbare Engelstalige arresten in de Hudoc-database zijn verzameld.¹⁴

¹³ Zie bijvoorbeeld Bruijn e.a. 2018, p. 141; Vols 2019, p. 743; Bruijn & Vols 2020, p. 32.

¹⁴ Zie Medvedeva e.a. 2020a, p. 246.

Alle in september 2017 beschikbare arresten zijn eenmalig gedownload met behulp van een online beschikbaar computerscript.¹⁵ Er is in het EHRM-project dus geen sprake van een *data pipeline*, zoals besproken in de bijdrage van Tjong Tjin Tai.¹⁶

Tabel 3 Dataverzameling en selectie

Naam project	Reproduceerbaarheid dataverzameling en -selectie	Representativiteit	Codeurs	Getest codeboek
VENI	Ja	Nee, maar erkend	3	Ja
Damocles	Ja	Nee, maar erkend	2	Ja
EHRM	Ja	Ja	computer	Nee

Bij alle projecten is in de publicaties uitgelegd hoe een selectie is gemaakt van de gevonden uitspraken. Zo zijn bij het VENI-project van de zoekresultaten met de vaste zoektermen op rechtspraak.nl alleen de privaatrechtelijke zaken geselecteerd, en alleen de uitspraken van de rechtbanken. Al de overgebleven zoekresultaten zijn vervolgens handmatig doorgenomen, en op die manier zijn alleen de zaken over ontbinding en/of ontruimingsvorderingen inzake woonruimte geselecteerd.¹⁷ Bij het Damocles-project is dezelfde selectiemethode gehanteerd, alleen was daar sprake van bestuursrechtelijke uitspraken van de rechtbank.¹⁸ Een nadeel van deze handmatige selectie is dat deze tamelijk arbeidsintensief is, en dus – afhankelijk van het aantal variabelen dat wordt onderscheiden – alleen bij relatief kleine datasets kan worden gebruikt.¹⁹

In de publicatie binnen het EHRM-project is ook duidelijk gemaakt welke selectie is gemaakt van alle beschikbare uitspraken. Zo zijn alleen uitspraken geselecteerd waarin de klacht ontvankelijk is verklaard. Daarnaast zijn uitspraken uitgesloten waarin er sprake is van meerdere klachten, en het EHRM een klacht gegrond en een andere klacht ongegrond heeft verklaard.²⁰ Hiervoor is gekozen omdat we het model willen trainen om een uitspraak waar het draait om één klacht binair te classificeren: ‘schending’ of ‘geen schending’. Bij uitspraken waar het draait om meerdere klachten, is een dergelijke binaire classificatie moeilijker te maken: de eerste klacht kan bijvoorbeeld gegrond worden verklaard en de tweede ongegrond. Ten slotte zijn uitspraken over artikelen waar weinig jurisprudentie

¹⁵ Zie Medvedeva e.a.2020a, p. 246 (voetnoot 11).

¹⁶ Zie de bijdrage van Tjong Tjin Tai in deze bundel.

¹⁷ Vols 2019, p. 743.

¹⁸ Zie Bruijn & Vols 2020, p. 32.

¹⁹ Vgl. Verbruggen in deze bundel. Een voorbeeld van een grotere dataset (ca. 4000 uitspraken die handmatig verzameld zijn) wordt beschreven in de bijdrage van Wijntjens in deze bundel.

²⁰ Zie Medvedeva e.a. 2020a, p. 247.

over bestaat (bijv. art. 4 EVRM en art. 12 EVRM) niet geselecteerd. Het was niet noodzakelijk om een handmatige selectie te maken, omdat de metadata in de Hudoc-database een automatische selectie relatief eenvoudig maken. Het begrip metadata slaat op data over data: in de Hudoc-database is bij elke uitspraak bijvoorbeeld te vinden welk artikel een rol speelt in de uitspraak, en of de klacht ontvankelijk is verklaard.²¹

Een belangrijke vraag bij het verzamelen van data is voorts of de geselecteerde uitspraken representatief zijn voor de hele populatie uitspraken. Bij het VENI- en het Damocles-project is gebruikgemaakt van op www.rechtspraak.nl gepubliceerde jurisprudentie. Gezien het Besluit selectiecriteria uitsprakendatabank Rechtspraak.nl worden niet alle uitspraken van rechtbanken gepubliceerd.²² De rechterlijke macht beslist zelf welke uitspraken voor het publiek toegankelijk worden gemaakt door middel van opname in de online databank.²³ De gepubliceerde jurisprudentie wordt bovendien niet willekeurig gekozen, voor het publiceren gelden verschillende selectiecriteria. Vooral het criterium dat het een uitspraak betreft die een jurisprudentievormend karakter heeft, is van belang.²⁴ Kortom, er is er geen sprake van een aselechte steekproef van alle vonnissen en uitspraken.²⁵ In de publicaties binnen het VENI- en het Damocles-project is dus ook steeds aangegeven dat de geselecteerde jurisprudentie niet representatief is voor de hele populatie jurisprudentie op het gebied van ontbinding/ontruiming en woningsluitingen. Er is in de publicaties gewaarschuwd de resultaten van de analyse te lezen met deze constatering in het achterhoofd. Ook wordt in de publicaties gereflecteerd op de gevolgen van deze selectiebias.²⁶

Het voordeel van de Hudoc-database is dat die alle arresten van het EHRM bevat. Het is dus in beginsel mogelijk om de hele populatie van EHRM-uitspraken te onderzoeken, waardoor de vraag van representativiteit niet of minder van belang is. Hoewel door de selectiecriteria en de wijze van downloaden enkele relevante zaken verloren gegaan (kunnen) zijn, is een dusdanig grote dataset verzameld (zie tabel 2) dat er sprake is van een representatieve steekproef van de EHRM-arresten.²⁷

21 Zie het tabblad 'case details' bij de uitspraak op <https://hudoc.echr.coe.int/>.

22 Zie www.rechtspraak.nl/Uitspraken/Paginas/Selectiecriteria.aspx. Zie ook A.T. Marseille & M. Wever, 'Het verstrekken van niet-gepubliceerde uitspraken door de Rechtspraak: Gunst of recht?', *Nederlands Juristenblad* 2019, p. 123-124.

23 Art. 1 Besluit selectiecriteria uitsprakendatabank Rechtspraak.nl 2012.

24 Art. 5 Besluit selectiecriteria uitsprakendatabank Rechtspraak.nl 2012.

25 Zie www.rechtspraak.nl/Uitspraken/Paginas/Selectiecriteria.aspx. Zie ook Marseille & Wever 2019.

26 Zie Vols 2019, p. 743; Bruijn & Vols 2020, p. 32-33 en 45-46.

27 Zie Medvedeva e.a. 2020a, p. 246.

3.2 Coderen

Nadat de jurisprudentie verzameld en geselecteerd is, begint het codeerproces. Coderen is het systematisch noteren van de relevante kenmerken van elke uitspraak (de variabelen) op een dusdanige manier dat systematische analyse mogelijk wordt.²⁸ Voordat ik de wijze van coderen nader bespreek, wil ik eerst de typering van coderen in het onderzoeksproces kort aanstippen. Het coderen kan zowel als dataverzameling als data-analyse worden aangemerkt, afhankelijk van wat de onderzoeker als data beschouwt. Als hij de uitspraken op zichzelf als data ziet, dan is het coderen van de uitspraak een vorm van data-analyse.²⁹ Men kan het coderen ook als dataverzameling zien: door het systematisch lezen van de uitspraken genereert men data, die later met behulp van statistische methoden worden geanalyseerd. In dit hoofdstuk kies ik ervoor om het coderen te zien als onderdeel van de dataverzameling. Het is echter niet zo dat het coderen anders had plaatsgevonden als het proces als data-analyse was aangemerkt.

Tabel 3 laat zien dat binnen het VENI- en het Damocles-project de jurisprudentie door meerdere mensen – codeurs – is gecodeerd. Het coderen vond plaats aan de hand van een codeboek. In een codeboek staan de verschillende variabelen die bij het lezen van elke uitspraak worden bekeken en genoteerd, met de verschillende waarden die aan een variabele kunnen worden toegekend. Een codeboek wordt gebruikt om het codeerproces te structureren en de betrouwbaarheid en reproduceerbaarheid van het onderzoek te vergroten.³⁰

In beide projecten is het codeboek in een aantal fasen ontwikkeld. In de eerste fase is op basis van eerder onderzoek en de te beantwoorden onderzoeksvragen een aantal variabelen in het codeboek opgenomen, en is in het codeboek omschreven hoe de variabele geïnterpreteerd moet worden en in welke gevallen wat genoteerd dient te worden. Vervolgens is een tiental uitspraken aan de hand van dit voorlopige codeboek gecodeerd. In de tweede fase is door de codeurs besproken welke variabelen moeten worden toegevoegd, samengevoegd of verwijderd, en tegen welke onduidelijkheden zij zijn aangelopen. Het codeboek is vervolgens aangepast. Uitgangspunt was dat liever te veel variabelen in het codeboek worden opgenomen: het is eenvoudiger later iets te verwijderen of samen te voegen dan om alle uitspraken weer opnieuw te moeten bekijken omdat een nieuwe variabele is toege-

28 Zie L. Epstein & A.D. Martin, 'Quantitative approaches to empirical legal research', in: P. Cane & H.M. Kritzer, *Oxford Handbook of Empirical Legal Research*, Oxford: Oxford University Press 2013; M. Salehijam, 'The Value of Systematic Content Analysis in Legal Research', *Tilburg Law Review* 2018, nr. 1, p. 34-42.

29 Zie bijvoorbeeld Verbruggen in deze bundel.

30 Zie L. Epstein & A.D. Martin, *An introduction to empirical legal research*, Oxford: Oxford University Press 2014, p. 95-116; Vols & Jacobs 2017, p. 95-96.

voegd. In de derde fase hebben verschillende codeurs aan de hand van het codeboek een aantal dezelfde zaken gecodeerd. Met behulp van *inter-coder reliability* testen is vervolgens de betrouwbaarheid van het codeboek getoetst.³¹ Als uit de test bleek dat de score voor een bepaalde variabele te laag was, dan zijn de verschillen tussen de codeurs besproken en is waar nodig het codeboek aangepast. Vervolgens zijn wederom enkele dezelfde zaken door meerdere codeurs gecodeerd, maar is alleen gefocust op de zwakker scorende variabelen. Als de variabele na deze ronde wel voldoende scoorde, is deze opgenomen in het codeboek. Bij een onvoldoende score is wederom dezelfde procedure gevoerd als in de derde fase of is de variabele uiteindelijk geschrapt uit het codeboek en database.

In de vierde fase is met behulp van het geteste betrouwbare codeboek de hele set van uitspraken gecodeerd. Zo zijn voor alle uitspraken data verzameld in een database opgesteld in het computerprogramma *Microsoft Access*. In deze database zijn verschillende typen variabelen opgenomen. Veelal gaat het om nominale variabelen: 'heeft huurder een advocaat: ja/nee', 'naam gemeente van burgemeester die sluit', 'bewoner voert een proportionaliteitsverweer: ja/nee', 'uitspraak: ontruiming of afwijzing'. Er zijn ook variabelen op ratio-niveau opgenomen: denk bijvoorbeeld aan het aantal dagen dat de huurovereenkomst van kracht is, de duur van de sluiting en de hoogte van een proceskostenvergoeding. Veel van de nominale variabelen zijn binair, ook wel dummyvariabelen genoemd. Dat heeft gevolgen voor welke statistische analysetechnieken gebruikt kunnen worden, hetgeen later nader aan bod komt.

Binnen het EHRM-project is geen sprake van een handmatig codeerproces: de computer heeft daar automatisch (tekstuele) data verzameld. Ook hier is in meerdere fasen gewerkt. In de eerste fase zijn de teksten van de arresten van de Hudoc-database gedownload, en zijn sommige delen van de tekst automatisch verwijderd. De verwijderde delen betreffen de onderdelen 'Law', 'Judgement' en de *dissenting/concurring opinions*. Waarom deze delen verwijderd zijn, is uitvoerig uitgelegd in de publicatie over het EHRM-project.³² Kort gezegd gaat het erom dat deze onderdelen de uitspraak van het EHRM bevatten, en dat het daarom weinig interessant is om met behulp van deze informatie de uitspraak te classificeren.

In de tweede fase zijn de in de Hudoc-database beschikbare metadata verzameld en aan elke gedownloade tekst gekoppeld. Van elke tekst is bijvoorbeeld bekend welk artikel van het EVRM centraal staat, en wat de uitspraak van het EHRM is. Binnen het project is de uitspraak van het EHRM het *label*, en is het betrokken artikel een *feature*. Een label is

31 Zie Vols 2019, p. 743-744; Bruijn & Vols 2020, p. 32-33.

32 Zie Medvedeva e.a. 2020a, p. 248-249.

datgene dat met behulp van *machine learning* geclassificeerd/voorspeld wordt (vergelijkbaar met de afhankelijke variabele), en de *features* zijn andere observeerbare kenmerken ('*observable characteristics*'³³), die vergelijkbaar zijn met onafhankelijke variabelen.

De derde fase bestond uit de identificatie van *features* voor elke uitspraak. Er zijn verschillende technieken uit de *Natural Language Processing* (NLP) gebruikt. Zo zijn voor elke tekst *n-grams* verzameld. Een *n-gram* is een aansluitende reeks van woorden in een tekst. Het kan gaan om één woord: een *unigram* (bijvoorbeeld 'hondje'). Er kan ook sprake zijn van een *bigram* ('hondje loopt'), een *trigram* ('hondje loopt buiten') en zo verder. In het EHRM-project zijn *unigrams* tot en met *fourgrams* gebruikt. Er is niet slechts genoteerd hoe vaak een *n-gram* in een uitspraak voorkomt. Als dat zal gebeuren dan zullen bijvoorbeeld in dit kader weinigzeggende veel voorkomende woorden (bijvoorbeeld *article of law*) als belangrijk worden aangemerkt. Er is juist gezocht naar de meest voor een zaak kenmerkende *n-grams* door gebruik te maken van *term frequency-inverse document frequency (tf-idf)*.³⁴ Ook is er gebruikgemaakt van *data pre-processing*, door bijvoorbeeld stopwoorden te verwijderen en hoofdletters om te zetten naar kleine letters.³⁵ Op deze wijze kan het model de classificeerders ontwikkelen waarmee de uitspraken automatisch worden geclassificeerd als 'schending' of 'niet schending'.

3.3 Data-analyse

Nadat de data zijn verzameld, zijn zij aan een analyse onderworpen. In de voornoemde publicatie van Vols en Jacobs is op het gebied van analyse een onderscheid gemaakt tussen het gebruiken van methoden uit de beschrijvende statistiek en de toepassing van geavanceerdere statistische technieken om bijvoorbeeld het verband tussen twee of meer variabelen te onderzoeken.³⁶

Tabel 4 laat zien dat in alle in dit hoofdstuk besproken projecten gebruik is gemaakt van beschrijvende statistiek. In de diverse publicaties is bijvoorbeeld in tabellen informatie te vinden over het aantal keer dat een bepaald verweer wordt gevoerd, of hoe vaak geklaagd is over de schending van recht neergelegd in het EVRM.³⁷ Ook is bijvoorbeeld gerapporteerd

33 Medvedeva e.a. 2020a, p. 249.

34 Medvedeva e.a. 2020a, p. 250.

35 Medvedeva e.a. 2020a, p. 251-252.

36 Zie Vols & Jacobs 2017, p. 96.

37 Zie bijvoorbeeld Bruijn e.a. 2018, p. 142-143; Vols 2019, p. 744; Medvedeva e.a. 2020a, p. 246.

hoelang de sluitingen op basis van artikel 13b Opiumwet in de bestudeerde jurisprudentie gemiddeld duren, en wat de mediaan en standaarddeviatie is.³⁸

Tabel 4 Analysetechnieken

Naam project	Beschrijvende statistiek	Geavanceerdere analyse
VENI	Ja	Regressieanalyse
Damocles	Ja	Regressieanalyse
EHRM	Ja	<i>Machine learning</i>

In de drie projecten is ook gebruikgemaakt van meer geavanceerde statistische analysemethoden. Het is niet de bedoeling om in dit hoofdstuk de betekenis van al deze methoden te bespreken, maar meer om de lezer een indruk te geven welke technieken kunnen worden gebruikt.³⁹ Zo is – vanwege een relatief kleine dataset – bijvoorbeeld het verband tussen variabelen getoetst met behulp van de zogenoemde ‘Fisher’s Exact Test’.⁴⁰ In andere publicaties is vanwege de binaire (nominale) afhankelijke variabele (ontruiming of niet; burgemeester wint of niet) een logistische regressie uitgevoerd.⁴¹ In één publicatie is ook de betekenis van (logistische) regressie uitgebreider besproken.⁴²

In het EHRM-project is met behulp van *machine learning* technieken onderzocht of en in welke mate de uitspraken succesvol worden geclassificeerd. Classificeren ziet in dit verband op het automatisch correct aanduiden van de beslissing van het EHRM in een uitspraak (het *label*, dat wil zeggen schending of niet schending) aan de hand van data verzameld uit de uitspraak zelf (de *features*), zonder dat in die data de beslissing wordt genoemd. Om te classificeren is met behulp van een *balanced training set* een algoritme getraind. Dat houdt in dat er in de training set evenveel zaken over schending als niet-schending zijn opgenomen. Dit voorkomt dat het systeem eenvoudigweg leert welke soort beslissing (verreweg) het meeste voorkomt, en dus de uitspraken steeds als zodanig classificeert.⁴³ Het algoritme dat is getraind betreft een *Support Vector Machine (SVM) Linear Classifier*.

38 Bruijn & Vols 2020, p. 36.

39 Zie Epstein & Martin 2014, p. 143-221 voor een goed overzicht van de verschillende technieken die gebruikt kunnen worden.

40 Zie M. Vols & L.M. Bruijn, ‘De strijd van de burgemeester tegen drugscriminaliteit: Een eerste statistische analyse van de toepassing van artikel 13b Opiumwet’, *Netherlands Administrative Law Library*, oktober 2015, p. 1-23; L.M. Bruijn & M. Vols, ‘Ondermijning, drugscriminaliteit en vertrouwen in de rechtstaat: Een analyse van de toepassing van de Wet Damocles’, in: H.D. Tolsma & P. de Winter (red.), *De wisselwerking tussen recht en vertrouwen bij toezicht en handhaving* (Serie Governance & Recht deel 15), Den Haag: Boom juridisch 2017, p. 189-205; Bruijn e.a. 2018.

41 Zie Vols 2019, p. 745-747; Bruijn & Vols 2020, p. 37-40.

42 Zie Bruijn & Vols 2020, p. 33-34.

43 Zie Medvedeva e.a. 2020a, p. 247-248.

In een publicatie over het EHRM-project is uitvoerig – en hopelijk ook op een voor juristen begrijpelijke wijze – uitgelegd wat een SVM voor algoritme is.⁴⁴ Bij de training van het algoritme is gebruikgemaakt van *threefold cross-validation*.⁴⁵ Dat betekent, kort gezegd, dat het algoritme driemaal is getraind en getest, met steeds een ander gedeelte van de *training set*.⁴⁶ Na het optimaliseren van het algoritme, is met een *test set* van niet bij de training betrokken uitspraken onderzocht hoe succesvol het algoritme uitspraken classificeert.⁴⁷ Het in het EHRM-project ontwikkelde model classificeert ten aanzien van de gebalanceerde test set van uitspraken in 75% de uitspraak juist als ‘schending’ of ‘niet schending’. Dat is fors hoger dan de gokkans van 50% (de test set bestaat voor de helft uit zaken waarin het EHRM oordeelt dat er een schending is en voor de helft uit zaken waarin het EHRM oordeelt dat er geen schending is). Het percentage ligt iets lager dan een model ontwikkeld door Aletras en anderen, maar in de literatuur zijn ernstige methodologische bezwaren tegen in die studie gebruikte data opgeworpen.⁴⁸

4 RESULTATEN EN LESSEN

In deze paragraaf bespreek ik niet zo zeer de inhoudelijke resultaten van de diverse projecten (zie daarvoor de publicaties), maar veel meer de lessen die getrokken kunnen worden over de toepassing van de methode van kwantitatieve jurisprudentieanalyse. Ik bespreek vijf lessen.

De eerste les betreft het gebruik van theorie bij de kwantitatieve jurisprudentieanalyse. Theorie speelt een belangrijke, maar ook veelal wat impliciete rol bij de toepassing van deze methode. De les is dat de rol van de theorie in het onderzoek door de juridische onderzoeker steeds moet worden geëxpliciteerd. Dat maakt duidelijk in welke mate het onderzoek aansluit bij bestaande kennis, wat voor soort theorie wordt gebruikt, en in hoeverre het juridisch onderzoek bijdraagt aan theorie. Bovendien kan de onderzoeker door aandacht te besteden aan theorie voorkomen dat in het wilde weg data worden verzameld. Dataverzameling om de dataverzameling is weinig zinvol: het is noodzakelijk dat de verzamelde data ingebed zijn in een theoretisch kader, en ook helpen om de theorie verder te vormen en verfijnen.

44 Zie Medvedeva e.a. 2020a, p. 243.

45 Zie Medvedeva e.a. 2020a, p. 252.

46 Zie Medvedeva e.a. 2020a, p. 244.

47 Zie Medvedeva e.a. 2020a, p. 248.

48 Zie N. Aletras, D. Tsarapatsanis, D. Preotiu-Pietro & V. Lampos, ‘Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective’, *PeerJ Computer Science* 2016, p. 1-19; Medvedeva e.a. 2020a, p. 262.

Concreet betekent dit dat de onderzoeker bij het opstellen van een plan om jurisprudentie te verzamelen en selecteren, aansluit bij (klassiek-juridische) theorieën. Zo moet bijvoorbeeld worden achterhaald welke wetsartikelen, procedures en juridische concepten relevant zijn om de relevante gerechtelijke uitspraken te vinden. Ook bij het opstellen van een codeboek dient de onderzoeker zo veel mogelijk aan te sluiten bij deze theorie. Wat voor soort verweren kan een procespartij bijvoorbeeld voeren? Welke maatstaven hanteert de rechter bij de beoordeling van een vordering, verweer of besluit?

Daarnaast kunnen *microlevel* theorieën worden gebruikt, en toetsbare hypothesen opgesteld worden. Dat kunnen meer klassiek-juridisch georiënteerde hypothesen zijn (bijvoorbeeld ‘het voeren van een proportionaliteitsverweer vergroot de winkans van de huurder’), maar ook meer meta-juridische hypothesen (bijvoorbeeld ‘rechters die al langer dan tien jaar rechter zijn, wijzen minder vorderingen toe dan rechters die korter rechter zijn’). Als de onderzoeker normatieve (in de zin van prescriptieve of evaluatieve) theorieën wil gebruiken, dan moeten ook deze worden geëxpliciteerd.⁴⁹ Het is dan van belang om te weten dat de verzamelde data zelf niet als een normatieve standaard functioneren: de maatstaf moet buiten de data worden gevonden (bijvoorbeeld in een mensenrechtenverdrag of in de politieke filosofie).⁵⁰

De tweede les ziet op de noodzaak om te reflecteren op de beschikbaarheid en kwaliteit van de data. Onderzoekers die een kwantitatieve jurisprudentieanalyse uitvoeren, moeten steeds nagaan in hoeverre de data die zij onderzoeken representatief zijn voor de hele populatie uitspraken. Een eerste groep jurisprudentie betreft zaken waarvan de hele populatie uitspraken gepubliceerd is. De tweede groep betreft zaken waarvan slechts een deel van de uitspraken wordt gepubliceerd.

Voorbeelden van de eerste groep betreft jurisprudentie van het EHRM, het Hof van Justitie van de Europese Unie, de Hoge Raad en de Afdeling Bestuursrechtspraak van de Raad van State.⁵¹ Als onderzoekers slechts – bijvoorbeeld vanwege een gebrek aan tijd – een deel van dergelijke uitspraken kunnen onderzoeken, zullen zij een representatieve steekproef moeten trekken en uit moeten leggen hoe de steekproef is getrokken.⁵² Voorbeelden van de tweede groep zijn uitspraken van de Nederlandse rechtbanken en gerechtshoven. Op

49 Zie H.S.T. Taekema, ‘Theoretical and Normative Frameworks for Legal Research: Putting Theory into Practice’, *Law and Method* 2018, p. 11 (doi:10.5553/REM/000031).

50 Zie Vols 2020, p. 75-76.

51 Zie ook K. Weinshall & L. Epstein, ‘Developing High-Quality Data Infrastructure for Legal Analytics: Introducing the Israeli Supreme Court Database’, *Journal of Empirical Legal Studies* 2020, nr. 2, p. 416-434, over een nieuwe set data met (data over) uitspraken van het Hooggerechtshof van Israël.

52 Zie Vols & Jacobs 2017, p. 94-95 & Vols 2020, p. 140-141 over het trekken van een representatieve steekproef.

rechtspraak.nl wordt slechts een deel van de uitspraken gepubliceerd, en is geen sprake van een aselechte steekproef van de uitspraken. Zo worden bijvoorbeeld verstekvonnissen nagenoeg niet gepubliceerd, al vormen zij juist in de hele populatie van vonnissen een omvangrijk deel.⁵³ Ook in veel andere landen wordt slechts een beperkt deel van (lagere) rechtspraak gepubliceerd.⁵⁴ Onderzoekers die kwantitatieve jurisprudentieanalyses gebruiken om deze tweede groep jurisprudentie te bestuderen, dienen steeds weer te waarschuwen voor deze *selection bias*, en te reflecteren op wat de mogelijke effecten zijn van deze *bias* voor de resultaten van hun onderzoek. In welke mate zijn de resultaten van de analyse vertekend door de *selection bias*, en in welke mate kan (n)iets worden geconcludeerd aangaande de hele populatie uitspraken?

Naast representativiteit verdient ook de kwaliteit van de data aandacht.⁵⁵ Natuurlijk kunnen databases als rechtspraak.nl en Hudoc fouten bevatten. De uitspraken kunnen typefouten bevatten of de metadata in databases kunnen onvolledig zijn.⁵⁶ Daarnaast kunnen fouten ontstaan bij het coderen van de uitspraken. Uit eerder onderzoek blijkt dat in tamelijk veel projecten onduidelijk is hoe de uitspraken gecodeerd zijn, door wie dat is gebeurd en of een (gevalideerd) codeboek is gebruikt.⁵⁷ Het is noodzakelijk om in de publicaties duidelijkheid te verschaffen over het codeerproces en de wijze waarop een codeboek tot stand is gekomen. Alleen zo wordt duidelijk dat een onderzoeksproject betrouwbare data en resultaten genereert, en dat de onderzoeksvragen op een gedegen wijze beantwoord worden.

Een derde les ziet op juridische koudwatervrees op het gebied van het gebruik van binnen de rechtswetenschap minder gehanteerde onderzoeksmethoden. Veel rechtswetenschappers achten zich onvoldoende opgeleid om gebruik te maken van bijvoorbeeld statistische technieken. Hoewel methodologische onvoorzichtigheid en overmoedigheid natuurlijk onverstandig zijn, moet de juridische terughoudendheid niet leiden tot onnodige stilstand. Deyevre en anderen wijzen er terecht op dat in andere disciplines voortdurend nieuwe methoden worden aangeleerd: 'Social scientists thus have no choice but to upgrade their

53 Zie daaromtrent De Rechtspraak, *Kengetallen 2019, 2020*, www.rechtspraak.nl/SiteCollectionDocuments/kengetallen-2019.pdf.

54 Zie C.L. Boyd, P.T. Kim & M. Schlanger, 'Mapping the iceberg: the impact of data sources on the study of district courts', *Journal of Empirical Legal Studies* 2020, nr. 3, p. 466-492 voor de bespreking van de publicatie van jurisprudentie van lagere rechtspraak in de Verenigde Staten, en de gevolgen van het publicatiebeleid voor de representativiteit van bestudeerde datasets.

55 Porton en anderen laten zien dat er in de Amerikaanse juridische archieven bijvoorbeeld veel fouten zitten in de data omtrent huissuitzettingen. Zie A. Porton, A. Gromis & M. Desmond, 'Inaccuracies in eviction records: implications for renters and researchers', *Housing Policy Debate* 2020 (<https://doi.org/10.1080/10511482.2020.1748084>).

56 Zie de bijdrage van Tjong Tjin Tai in deze bundel.

57 Zie L. Epstein & G. King, 'The Rules of Inference', *University of Chicago Law Review* 2002, 69, p. 1-133; Vols & Jacobs 2017.

methodological know-how in order to keep pace with the evolution of their discipline. This contrasts with the legal academy, where the prevailing wisdom seems to be that the old legal dog cannot learn new tricks'.⁵⁸ Ook juristen zijn goed in staat om statistische methoden te leren gebruiken, al moet niet worden onderschat dat dit wel een tijdsinvestering vergt. Met een middag naar instructievideo's op YouTube kijken ben je er niet. Het voorgaande betekent niet dat juristen zich allemaal gelijk moeten richten op het leren van geavanceerde statistische analysemethoden. Toepassing van beschrijvende statistiek in kwantitatieve jurisprudentieanalyses is een goed begin, en is ook vanuit wetenschappelijk oogpunt zeer nuttig.

De vierde les ziet op de mogelijkheden die samenwerking met andere disciplines biedt.⁵⁹ In de hierboven besproken projecten heeft samenwerking met econometristen en computationeel taalkundigen er bijvoorbeeld voor gezorgd dat methodologische fouten zijn voorkomen. Zonder samenwerking was het gebruik van *machine learning* technieken bij de analyse van jurisprudentie niet goed mogelijk geweest. Daarnaast biedt deze interdisciplinaire samenwerking ook mogelijkheden om het tijdsintensieve proces van handmatige dataverzameling, -selectie en -codering te verkorten door automatisering.⁶⁰ Het EHRM-project laat zien dat automatisering van deze processen mogelijk is, en dat ook het verzamelen en analyseren van wat grotere (*legal big*) datasets mogelijk is.

De vijfde les betreft het inzicht dat een kwantitatieve jurisprudentieanalyse niet los kan staan van een meer klassiek-juridische (kwalitatieve) duiding van de onderzoeksresultaten.⁶¹ Zoals gezegd vormt klassiek-juridisch onderzoek doorgaans de basis voor de identificatie van variabelen die een rol spelen in de analyse. Bij de uitleg van de resultaten van de kwantitatieve analyse (denk aan het aantal keren dat een verweer succesvol is of een aangetroffen verband tussen variabelen) is het ook goed om gebruik te maken van inzichten vanuit klassiek-juridisch onderzoek. Ook bij de duiding van met behulp van *machine learning* gevonden classificeerders kan goed worden teruggegrepen naar inzichten uit klassiek-juridisch onderzoek. Kwantitatieve jurisprudentieanalyses en klassiek-juridisch onderzoek zijn dus geen concurrenten van elkaar, maar vullen elkaar juist goed aan.

58 A. Dyevre, W. Wijtvliet & N. Lampach, 'The Future of European Legal Scholarship: Empirical Jurisprudence', *Maastricht Journal of Comparative and International Law* 2019, 26(3), p. 348-371, op p. 370.

59 Zie daarover ook de bijdrage van Goossens e.a. in deze bundel.

60 Zie Epstein & Waterbury 2020, p. 7.

61 Zie ook Bruijn & Vols 2020, p. 43-46.

5 CONCLUSIE

Dit hoofdstuk heeft hopelijk inzicht gegeven in de potentie van kwantitatieve jurisprudentieanalyses binnen de rechtswetenschap. De drie besproken projecten zijn slechts enkele voorbeelden van de toepassing van deze techniek. Inmiddels zijn nieuwe projecten gestart waarin verder wordt geëxperimenteerd met kwantitatieve analyse van rechtspraak. Zo verzamelt en analyseert Eding in het kader van haar promotieonderzoek naar beschermingsbewind (een civielrechtelijke procedure die jaarlijks tot vele tienduizenden beschikkingen leidt) jurisprudentie die in het niet-openbare *e-archief* van de Raad van de Rechtspraak is gepubliceerd.⁶² Zo wordt onder andere hopelijk duidelijk hoe representatief www.rechtspraak.nl is voor de hele populatie jurisprudentie op dat specifieke rechtsgebied. Ook Bruijn en anderen verzamelen en analyseren inmiddels in dit *e-archief* gepubliceerde jurisprudentie in het kader van een evaluatie van de toepassing van artikel 13b Opiumwet.⁶³ Medvedeva en anderen hebben op basis van documenten die ver voor de uitspraak van het EHRM verschijnen (de zogenoemde *communicated cases*) uitspraken van het EHRM voorspeld.⁶⁴ Ten slotte zal Vols in het EVICT-project experimenteren met het toepassen van *machine learning* op Nederlandstalige rechtspraak over huisuitzettingen, en dat combineren met technieken uit de *citation network analysis*.⁶⁵ Hopelijk leiden al deze nieuwe projecten tot verdere inzichten in de voor- en nadelen van de kwantitatieve analyse van rechtspraak, en kan daar in een volgende publicatie nader op worden gereflecteerd.

62 Zie over de thematiek van dit onderzoek M. Eding, 'Beschermingsbewind op maat', *WPNR* 2019/7248, p. 554-560.

63 Zie www.wodc.nl/onderzoeksdatabase/3135-monitoring-inzake-artikel-13b-van-de-opiumwet.aspx.

64 Zie M. Medvedeva, X. Xu, M. Wieling & M. Vols, 'JURI SAYS: an automatic judgement prediction system for the European Court of Human Rights', *Conference Proceedings JURIX 2020*[b], www.martijnwieling.nl/files/JURIX2020.pdf; M. Medvedeva, A. Üstün, X. Xu, M. Vols, & M. Wieling, 'Automatic judgement forecasting for pending applications of the European Court of Human Rights', *Proceedings of the Fifth Workshop on Automatec Semantic Analysis of Information in Legal Text (ASAIL 2021)*, p. 12-23.

65 Zie www.eviction.eu.



7 EXPLORING AND SEARCHING JUDICIAL OPINIONS WITH TOP2VEC

*Arthur Dyevre**

1 INTRODUCTION

Searching, exploring and classifying judicial opinions form key legal research tasks, in legal practice as much as in academia. Whether it is assessing legal risk, writing a legal brief or commenting the latest ruling, relevant cases must be retrieved, analyzed and sorted out. All lawyers are familiar with the pain and frustration that often accompany the quest for information. How do we know what cases are out there? How do we go about searching for them? Can we be sure we have retrieved all the rulings corresponding to our concept, doctrine or legal issue of interest?

To be sure, most legal databases provide some functionalities to help retrieve and classify legal information. First is keyword search. Type one or more keywords and the search engine will retrieve the rulings in which these keywords appear. What is typically presented as ‘advanced’ search then allows you to combine keywords with Boolean logic (AND, OR, etc.) on top of specifying a time range or additional metadata. In addition, some judicial bodies have developed their classification schemes, with classification codes and extra metadata. The Court of Justice of the European Union (CJEU), for example, has evolved a classification system based on the summaries of judgments and orders published in the European Court Reports (ECR).¹ The search engine integrated in the Court’s online database, Curia, allows users to search cases via this classification scheme.²

Yet every experienced lawyer knows the limitations of these search methods. Keyword search is based on exact matching. However, lawyers are not so much interested in words as in concepts. Words are there to express legal concepts or fact patterns. But everyone knows that there are many ways to express the same concept or idea, which nobody can

* Prof. dr. A. Dyevre is Professor of Empirical Jurisprudence at the KU Leuven Centre for Empirical Jurisprudence. I am grateful to Nusret Ipek for invaluable research assistance. I acknowledge financial support from European Research Council Horizon 2020 Starting Grant #638154 (EUTHORITY).

1 See <https://curia.europa.eu/common/juris/en/aideGlobale.pdf#page=20> (accessed 16 May 2021).

2 <https://curia.europa.eu/juris/recherche.jsf?language=en> (accessed 16 May 2021).

be sure to know *a priori*. Exact matching, even combined with Boolean algebra, cannot address the challenges arising from synonymy, polysemy and co-reference. Likewise, the classification schemes elaborated by judges and their staff do not fare much better. For one thing, any manual classification scheme must be informed by an understanding of what information is most important or relevant. Yet which feature of a case is important or relevant depends on one's research question. A feature – fact pattern or legal issue – deemed secondary may later become crucial as new judges and litigants seek to distinguish new cases from older ones. For the same reasons, manual classification schemes do not always age well. As new issues rise into prominence while others recede into obscurity, they come to reflect the law and preoccupations of the past. In 2010, the CJEU introduced a new classification scheme devised to reflect the evolution of EU law following the coming into force of the Lisbon Treaty (the old classification scheme still applies to the case law prior to 2010).³

Furthermore, where courts have bothered to develop such a classification scheme, there are often strong reasons to doubt their reliability and consistency. The process of annotating decisions does not habitually follow a systematic procedure ensuring inter-annotator reliability. In fact, we have to imagine that, for a relatively busy court like the CJEU, some member of the court staff is instructed to prepare a summary of the case in accordance with the classification scheme. However, when a classification system contains dozens of headings and hundreds of entries, inconsistencies easily multiply if no provision is made for training annotators, monitoring quality and addressing mistakes.⁴ Faced with a cognitively straining task, annotators have the tendency to overly opt for the more general classification terms (almost any legal case is in some sense about 'procedure'). Nor is there much guaranty that two independent annotators – especially if they are jurists trained in distinct jurisdictions – will spontaneously associate the same ideas and concepts with classificatory headings such as 'vested rights' or 'price controls'.

Whoever has made an earnest attempt at a comprehensive case law search – *e.g.* retrieving all cases pertaining to a particular doctrine, fact pattern or legal question – knows these limitations. Such attempts often end in frustration, leaving the legal researchers with finds that can be aptly characterized as false positives.

³ See note 2.

⁴ To be sure, no moderately complex annotation project ever achieves 100% inter-annotator agreement. But putting such procedures in place (1) reduces discrepancies and, more important still, (2) discrepancies and the areas where they arise can be identified and quantified.

Computer scientists and computer linguistics have designed a whole family of methods and techniques, broadly falling under the heading of Natural Language Processing (NLP), to automate the exploration and classification of large collections of documents. NLP has been making rapid progress. So too have legal applications. Yet the so-called ‘bag-of-words’ approach, until recently dominant in NLP research, meant that these methods failed to satisfactorily address the practical problems I have just described. NLP, however, has been experiencing a major paradigm shift, moving away from the bag-of-words paradigm to the distributed linguistics paradigm – often summarized by the mantra ‘words are defined by the company they keep’.⁵ At the heart of this paradigm shift is a fundamental change in the way computers learn and construct word representations. These advances open up new possibilities for legal research.

In this chapter, I introduce a novel NLP technique developed by Dimo Angelov called ‘Top2Vec’.⁶ I illustrate its potential for legal research using a corpus of Belgian high court rulings. Combining unsupervised document clustering with distributed word representation generated via a neural network architecture, Top2Vec, provides a tool to classify, explore and search judicial decisions all at once. To explain how Top2Vec works, I begin by discussing the bag-of-words paradigm on which NLP approaches to document retrieval and classification used to rely. Next I present the new ‘distributed’ paradigm underpinning Top2Vec. In the section that follows I then go on to present the Top2Vec algorithm. Finally, I showcase a legal application with a corpus of Belgian decisions and examples of search and exploration tasks.

To dispel any lingering doubt, I feel obliged to state, at the outset, that this chapter has not been written with computer scientists or legal academics with a strong interdisciplinary bent, such as law & economics and empirical legal scholars, in mind. Instead, my target audience is very much the typical jurist, law professor or practitioner, for whom law is synonymous with doctrinal analysis. Accordingly, I have adapted the prose to avoid off-putting mathematical notation (with very limited exceptions).

5 The phrase is commonly attributed to J.R. Firth, ‘Synopsis of Linguistic Studies 1930-1955’, reprinted in F. Palmer (ed.), *Selected Papers of J.R. Firth*, Longman (1968).

6 D. Angelov, ‘Top2Vec: Distributed Representations of Topics’, (2020) arXiv preprint arXiv:2008.09470.

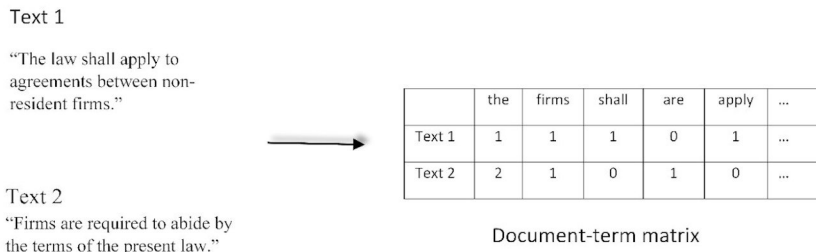
2 **LIMITATIONS OF BAG-OF-WORDS APPROACHES TO DOCUMENT CLUSTERING
AND DOCUMENT RETRIEVAL**

When trying to explain how NLP methods work, it is useful to consider first how language is processed by the human brain. When we read a text our brain parses its meaning applying our knowledge of semantics, syntax and context. In any language, the stock of words is finite, but syntactic rules allow the construction of infinitely many sentences from this finite stock of words. Not only that. Humans are able to communicate more than they say or write by taking the context into account. It is why we ascribe different meanings to the sentence 'I would like a table' when uttered in a restaurant and when uttered in a furniture shop.⁷

The language of judicial opinions occasionally diverges, sometimes significantly, from everyday language. It also tends to be more self-contained with less reliance on the non-linguistic context. Yet the basic principles of linguistic cognition and interpersonal communication I just described hold true.

Now, NLP methods do not parse texts quite the same way. All NLP techniques – even the most advanced language models trained on billions of pages of text – are ultimately based on the same basic statistical principle: First find a way to turn texts into numbers and then look for patterns in these numbers. While the ability to detect patterns depends on the amount of textual data and the sophistication of the algorithm, the basic principle remains the same.

⁷ D. Sperber & D. Wilson, *Relevance: Communication and Cognition*, Wiley (1996).

Figure 1 Illustrating the Bag-of-Words Paradigm: Converting Texts into Document-Term Matrix

Traditionally, NLP techniques have involved a good deal of complexity reduction. The bag-of-words paradigm that used to dominate the discipline pushed complexity reduction very far. To see why, let us assume, for the sake of simplicity, that we have a corpus with just two texts, which I shall call Text 1 and Text 2, as in Figure 1. The bag-of-words approach involves converting these texts into sequences of word counts. We call the sequence of word counts representing a text a ‘vector’. This vector contains counts of all the words occurring in that text and zeros for the words occurring in the other text but not in that particular text. For Text 1 the zeros will represent all the words that appear in Text 2 but not in Text 1, and vice-versa. In a large corpus spanning a vocabulary of millions of words, the vector of word counts representing a text will contain mostly zeros – accounting for all the words that occur in other texts but not in the one under consideration.

It is easy to see all the information that gets lost when texts are forced into this procrustean bed. The document-term matrix does not encode any information about syntax nor about word order. A text thus truly becomes a bag of words.

This modus operandi has indisputably produced useful methods and results. Probabilistic topic modelling,⁸ for example, was a popular tool to explore all sorts of document collections: from publications in scientific journals⁹ to law review articles,¹⁰ legislative acts¹¹ and

8 D.M. Blei, ‘Probabilistic Topic Models’, (2012) *Communications of the ACM* 55, p. 77-84.

9 *Ibid.*

10 A. Dyevre *et al.*, ‘The Voices of European Law: Legislators, Judges and Law Professors’, (2021) *German Law Journal* 22(6), p. 1-14.

11 *Ibid.*

judicial decisions.¹² Probabilistic topic modelling looks for patterns of word usage in the document-term matrix and defines topics to produce a document classification. While the researcher decides how many topics he/she wants (*i.e.* how granular the summary of the corpus should be), it is unsupervised: the topics are defined by the algorithm (not by the researcher), which also reports topic proportion for each document.

Topic modelling offers a quick way to classify large document collections, such as the entire corpus of CJEU rulings.¹³ However, it can produce topics that are difficult to interpret (so-called ‘garbage’ topics). It also ignores synonymy, just as polysemy. It is indubitably useful for many exploratory purposes when one wishes to have a general idea of the content of a large corpus: What range of issues are addressed in preliminary rulings?¹⁴ How prevalent are internal market issues in infringement cases, compared to, say, environmental questions?¹⁵ Which subjects enjoy the most attention in the *Common Market Law Review*?¹⁶ Still, the fact that, as with bag-of-words methods in general, it does not encode any information about the context of words limits its ability to capture the conceptual structure expressing the doctrines and fact patterns of interest to legal researchers.

3 FROM THE BAG-OF-WORD TO THE DISTRIBUTED LINGUISTICS PARADIGM

At the heart of the shift from the bag-of-words to the distributed linguistics paradigm is a shift in what information is encoded about the words occurring in a corpus of texts. Instead of simply encoding information about word counts across documents to build a document-term matrix, the basic idea is to encode information about the words appearing around the target word. Consider, for example, the words likely to occur around the word ‘contract’ in a contract or a judicial opinion. Terms springing to mind are ‘legal’ (as in ‘legal contract’), ‘parties’ (as in ‘parties’ to the contract), ‘clause’, ‘terms’ and so on. Now, if we think about the word ‘agreement’, we can see that it will often be found in the company of the same words (‘legal agreement’, ‘parties to the agreement’, etc.). This is the intuition behind the mantra ‘words are defined by the company they keep’. Words that tend to share the company of similar words are likely to share a similar meaning, just like ‘contract’ and ‘agreement’ in legal discourse.

12 A. Dyevre & N. Lampach, ‘Issue Attention on International Courts: Evidence from the European Court of Justice’, (2020) *The Review of International Organizations*, p. 1-23.

13 *Ibid.*

14 *Ibid.*

15 *Ibid.*

16 Dyevre *et al.*, *The Voices of European Law* (2021).

Building off from this intuition, a major breakthrough came in 2013 when researchers at Google released Word2Vec.¹⁷ Word2Vec uses a shallow, neural network architecture to produce vector representations of words.¹⁸ The neural network is trained to predict words appearing around the target word or to predict the surrounding words given the target word. The vectors representing words are positioned in a high-dimensional space in such a way that vectors representing words sharing the same contexts are closer to one another. Word2Vec models trained on large corpora of English texts have the vector for *man* close to the vector for *woman* and *king* close to *queen*. Vector operations, such as vector addition and vector subtraction, can be performed on the word vectors to produce new vectors. Interestingly, the operation *king* – *man* yields the word vector for *queen*, which goes to show that Word2Vec models do capture conceptual structures – at least when trained on a sufficiently large corpus.¹⁹

After Word2Vec came Doc2Vec.²⁰ Doc2Vec rests on the same neural network architecture and learning strategy but learns document vectors on top of the word vectors. Like word vectors, document vectors are positioned in the high-dimensional space so that documents that are more similar have more proximate locations. The distance between document vectors thus becomes a measure of textual similarity.

Top2Vec builds on both Word2Vec and Doc2Vec which it combines with a clustering procedure to construct classificatory topics.

4 THE TOP2VEC ALGORITHM

The ‘Top’ in Top2Vec is for topic modelling, while the ‘2Vec’ underlines its reliance on vector representation. Dimo Angelov presented the algorithm as ‘distributed representation of topics’.²¹ A breakdown of the Top2Vec algorithm in six steps looks as follows:

1. Build word and document vectors from the corpus. This step applies to the Doc2Vec and Word2Vec techniques described above.

17 T. Mikolov, I. Sutskever, K. Chen, G.S. Corrado and J. Dean, J., ‘Distributed representations of words and phrases and their compositionality’, *Advances in Neural Information Processing Systems* (2013), p. 3111-3119.

18 While neural networks trained for deep learning have many layers of neurons, ‘shallow’ networks like Word2Vec have only two layers.

19 For a more extensive discussion of the potential of word embeddings for the analysis of legal language see E. Ash & D.L. Chen, ‘Case Vectors: Spatial Representations of the Law Using Document Embeddings’, in M. Livermore and D. Rockmore (eds.), *Law as Data*, Santa Fe Institute Press (2019).

20 Q. Le and T. Mikolov, ‘Distributed representations of sentences and documents’, (2014) *International conference on machine learning PMLR*, p. 1188-1196.

21 Angelov, Top2Vec (2020).

2. Map the document vectors into a lower dimensional space (this is done with a technique called Uniform Manifold Approximation and Projection or ‘UMAP’).
3. Find areas of the lower dimensional space with dense document clusters (for this another technique called ‘Hierarchical Density-Based Spatial Clustering of Applications with Noise’, or ‘HDBSCAN for short, is used).
4. For each dense area compute the centroid of the document vectors. The centroid (which is calculated using the original high-dimensional space to ensure greater precision) defines the topic vector for the dense area.
5. Find the word vectors closest to the topic vector. These word vectors become the words associated with the topic vector.

The output of the Top2Vec algorithm thus consists of three categories of vectors: (1) word vectors, (2) document vectors and (3) topic vectors.

The possibility to use these three vectors in combination is what makes Top2Vec a powerful tool to explore and search large repositories of judicial opinions.

In the next section, I showcase its application to a corpus of Belgian high court rulings.

5 APPLICATION: BELGIAN HIGH COURT DECISIONS

This section illustrates how Top2Vec can be used to explore and search a corpus of judicial rulings. The model I trained for the purpose of the present chapter is a prototype in the sense that, while good enough to illustrate the potential of Top2Vec for case law research, it leaves room for fine-tuning. Also, I do not provide systematic validation. Neither do I offer a benchmarking of performance compared to competing methods.²² Instead, I assess results informally.

5.1 *Corpus*

Belgian high court decisions – Court of Cassation, Council of State and Constitutional Court – were extracted from Jure-Juridat, the official case law repository of the Belgian ministry of justice,²³ using Python scripts. Decisions are usually available in both French and Dutch. In this chapter, I only consider decisions published in the Dutch language.

²² This would require human-labelled decisions.

²³ Accessed 15 April 2020. In the meantime, Jure-Juridat has been replaced by the JUPORTAL (<https://juportal.be/home/welkom>).

More than 150,000 rulings were scraped. The raw corpus includes all judgments of the Constitutional Court from 1985 up to 2019. Council of State and Court of Cassation rulings span the period 2004-2019. To ensure faster training, the corpus was further reduced to a random sample of 10,896 rulings. In some instances, the brief (*verzoekschrift*) is attached to the text of the decision of the court. While removing briefs from such a corpus is always a possibility (which can be easily performed using regular expressions), the argument can also be made that they offer additional information about the context of the corresponding ruling.

The amount of texts influences the quality of word embeddings. The algorithm learns from the repetition of phrasal patterns. So if there are fewer and/or shorter texts, the embeddings will usually be less accurate. In general, the adage ‘the more, the better’ holds. However, the algorithm is designed in such a way that embeddings should normally be reasonably accurate even for rare words when training with a small corpus of less than 3,500 rulings.²⁴

5.2 *Training*

After loading the raw texts to a Google Colaboratory²⁵ session, some cleaning was performed to remove carriage returns and html tags. I did not fix typos or words spanning line and page breaks. For that reason, some words appear in stunted form (e.g. ‘*ichtlijn*’ for ‘*richtlijn*’) in the examples below. As we will see below, this Top2Vec is implemented in a Python library. I used version 1.0.24. I trained the model using both the ‘fast’ and ‘deep-learn’ options. Using Google Colaboratory’s Graphic Processing Unit (GPU), training took about two hours with the fast option; with the deep-learn option training took close to twelve hours. The latter has the advantage of producing higher quality vectors, although the difference is not immediately apparent. Deep-learn training produced a lower number of topics (98) than fast training, which yielded 108 topics.

5.3 *Results*

What I want to show here is how Top2Vec can aid lawyers in exploring a corpus of judicial opinions and in searching for decisions relating to a specific concept. Remember that Top2Vec generates three categories of vectors: (1) word vectors, (2) document vectors and (3) topic vectors. These vectors can all serve to explore and retrieve decisions in the corpus.

²⁴ This is because, similar to the so-called ‘skip-gram’ approach for training Word2Vec models, Top2Vec is trained to predict surrounding words in the context window from the target word, see Angelov, Top2Vec (2020).

²⁵ Google Colab is a free programming environment run in the clouds.

Because these vectors are positioned in the same high-dimensional space, we can relate word vectors to document and topic vectors. Because word vectors capture conceptual structures, we can use them to search for decisions relevant to a particular concept by locating the document vectors closest to the word vector associated with the concept in question. Similarly, we can search for the topic vectors closest to a particular word vector; or for the rulings closest to a specific topic vector.²⁶

At an exploratory level, the topic vectors give us a sense of what issues are addressed by rulings in the corpus. Topics vary in size, with the largest topic comprising 2,679 decisions and the smallest 22.²⁷ Figure 2 to 4 illustrate some topics (simply numbered in decreasing size from largest to smallest) in the form of word clouds.

Larger words in the word cloud denote closer proximity to the topic vector. In short, this means larger words are more characteristic of the topic. For example, the largest word for topic 6 is ‘*onderwijs*’ (education) followed by ‘*gesubsidieerd*’ (subsidized). The other words in the word cloud (‘*leerkrachten*’, ‘*hogescholen*’) suggest litigation relating to funding for education (which in Belgium is managed by the linguistic communities – hence the prominent ‘*franse*’).

Figure 2 Topic 6: Education



Figure 3 Topic 35: Divorce



26 Of course, document vectors and cosine similarity scores can also serve to determine case similarity. Topic vectors offer additional help (more similar cases should also exhibit topical overlap).

27 Note that documents outside dense areas in the lower-dimensional projection (UMAP) are treated as outliers and are not assigned to a topic.

Figure 4 Topic 79: Adoption

Topics generated via Top2Vec are generally cohesive and easily interpretable – more so than the topics from a bag-of-words probabilistic topic model.²⁸ Some topics, though, occasionally capture information of orthogonal interest to the task of mapping the legal areas covered by the cases included in one’s corpus. The second largest topic, Topic 1, for example, captures the names of the registrar (Johan Pafenols) and function of judges (‘*sectievoorzitter*’, section president). This information usually appears at the top of a judgment (see Figure 5).

Topic vectors are useful on their own for corpus exploration – that is, to get some idea of the broad content of a corpus. Yet the added value of Top2Vec for information retrieval lies primarily in the possibility to specify word vectors manually and then identify documents and topics close to the specified vector.

How is that useful? Remember what we saw in the introduction about the way the search engines of legal databases typically work. The user specifies some keywords which return rulings where these keywords are exactly matched. The limitations of such an information retrieval approach were made plain. Now, Top2Vec allows researchers to conduct a keyword search, with the important difference that keywords become word vectors able to capture conceptual structures.

²⁸ This has been demonstrated by means of *systematic performance benchmarking*, see Angelov, Top2Vec (2020).

Figure 5 Topic 1: Information about sitting judges and registrar



Suppose we want to find Belgian rulings addressing EU law issues. We can specify one or add (or subtract) several word vectors. I have tried, by way of example, the vector *europese + unie + richtlijn* (in English that would translate into: *european + union + directive*). The decisions – that is the document vectors – closest to that word vector will be the decisions most likely to feature the corresponding concept and, therefore, the most relevant to our search.

Why should the results be better than for the exact-matching approach? Figure 6 helps understand what information is embedded in a word vector. Listed are the twenty words closest to that word vector *europese + unie + richtlijn* in the high-dimensional space. Numbers indicate the cosine similarity, with 1 indicating perfect synonymy and 0 the absence of shared context. The word most similar to our EU law vector is ‘*lidstaten*’ (member states) with cosine similarity of 0.79, followed by ‘*justitie*’ (justice as in ‘Court of Justice’), ‘*eeg*’ (EEC for European Economic Community), ‘*unierecht*’ (Union law) and so on. As we can see, all these terms are closely associated with EU law. This, in short, is how word vectors encode conceptual information. In Top2Vec, a keyword does not call a mere sequence of letters (as is the case with the exact-matching approach) but a whole set of lexical relations. These relations – encoded in the word vector – is what makes the algorithm a considerably more powerful search engine.

Note the stunted words such as ‘*pese*’, ‘*ichtlijn*’ and ‘*richtlij*’, caused by page and line breaks or (for older PDF files) poor digitization also appear. These are, in effect, treated like typos. When texts contain typos the algorithm will recognize that they are associated with their correctly spelled form because they appear in the company of the same words.

Figure 6 Words Closest to Vector *europese + unie + richtlijn*. Number Indicates Next to Word Cosine Similarity

```

lidstaten 0.7862719746265601
justitie 0.7201846720990392
eeg 0.6814893408394782
unierecht 0.6693159310212815
lidstaat 0.6664169772702708
vweu 0.6628048891307439
pese 0.6439236445047856
ichtlijn 0.6353940122979993
chtlijn 0.6280429889394502
grondrechten 0.62622076033376
ropese 0.6190540810488415
richtlij 0.5795002428041915
veu 0.5773079943698531
richtli 0.5748442792219819
hvj 0.572936134525861
verdrag 0.5720676718383884
nationale 0.567840581614723
richtl 0.5656647650967712
uitgelegd 0.5632335210235333
onderdanen 0.5608093776073136
europes 0.550423165491795

```

Before examining what decisions Top2Vec identifies as most relevant to our search, let us take a brief look at the topics closest to *europese + unie + richtlijn*. Intuitively, these topics should encompass the cases most likely to feature references to EU law. Figure 7 to 9 illustrate the three topics closest to our word vector. The closest, Topic 12, pertains to immigration. Next is Topic 17 which relates to energy (including renewable energy). Third is Topic 32 which addresses privacy (including data protection).

Figure 7 Topic 12: Immigration



Figure 8 Topic 17: Energy



Figure 9 Topic 32: Privacy



So now, what are the rulings closest to our *europese + unie + richtlijn* vector? The top three rulings are all rulings of the Constitutional Court featuring extensive discussion of the case law of the CJEU on EU citizenship.²⁹

1. Judgment 167/2013 of 19 December 2013 regarding a preliminary question from the Council for Immigration Affairs.
2. Judgment 123/2013 of 26 September 2013 regarding a preliminary question from the Council of State on Article 9 of Act of 8 July 2011 modifying Act of 15 December 1980 regulating the entry, sojourn, establishment and deportation of foreign nationals.

²⁹ K. Lenaerts, 'EU Citizenship and the European Court of Justice's "Stone-by-Stone" Approach', (2015) *International Comparative Jurisprudence* 1 (1), p. 1-10.

3. Judgment 121/2013 of 26 September 2013 on constitutional complaints against various legislative provisions regulating immigration, residence and family reunion rights.

As with all vectors (whether document, word or topic) distance is measured by cosine similarity, whereby documents with higher cosine similarity are closer to the specified word vector. Similar to a Google search, all documents in the corpus can thus be ranked from the most relevant to the least relevant to the conceptual search.

Whereas the first seven rulings in our EU law search are all opinions of the Constitutional Court, we find in eighth position – as measured by cosine similarity to our word vector – a decision of the Court of Cassation of 12 October 2017 (C.17.0278.N) on the right to compensation and assistance of airline passengers in case flight cancellation, delays and overbooking. Like the aforementioned constitutional rulings, the opinion contains extensive elaborations of EU legislation and CJEU decisions.

To give another example, let us suppose we want to find decisions addressing the scope of judicial review and, more specifically, whether it should be restricted to a reasonableness review, which in Dutch and Belgian law is often discussed under the heading of ‘*marginale toetsing*’ (literally ‘marginal review’). An obvious way to express this concept is the word vector *marginale + toetsing*. Figure 10 displays the words most closely associated with this vector. Cosine similarity is lower than for the vector *european + union + directive*, which may possibly reflect the fact that these words appear in several unrelated contexts aside from the discussion of the concept of ‘*marginale toetsing*’. Still, words such as ‘*opportuniteit*’ (opportunity/expediency), ‘*beoordelen*’ (evaluate), ‘*wettigheid*’ (legality/lawfulness), ‘*bevoegheid*’ (authority/competence), ‘*grondwettigheid*’ (constitutionality), ‘*discretionaire*’ (discretionary) are familiar terms for jurists working in administrative law.

Figure 11 Words Closest to Vector *marginale* + *toetsing*. Number Next to Word Indicates Cosine Similarity

```
marginaal 0.5994762690074329
toetsen 0.572991989048841
getoetst 0.548422615692167
opportunititeit 0.5198787994091496
beoordelen 0.5040501786967213
wettigheid 0.5015270449818503
merkt 0.4684215240589723
nagaan 0.4660236678059979
biedt 0.45928996187124727
bevoegdheid 0.4495330597885385
jurisdictionele 0.448121657100567
ziet 0.44687846280911325
overeenstemming 0.44667443352776426
geding 0.4437775389174424
grondwettigheid 0.44146980120438184
discretionaire 0.43851984718324816
onevenredig 0.436700612414454
mening 0.434893661643183
wetskrachtige 0.4339238655549861
voorts 0.4324679680120572
ruime 0.43219273298188626
onderzocht 0.431481944328483
dezen 0.4270582428118396
aangelegenheid 0.42355982243583856
formulering 0.4197206951604433
leent 0.4191192198903091
rechtspraak 0.41699824894068405
bovendien 0.4163596850901381
```

In the order of relevance defined by cosine similarity, the ten documents closest to our word vector are a mixture of Constitutional Court and Court of Cassation opinions.

Judgment 101/2015 of 2 July 2015 of the Constitutional Court deals with a preliminary question from the Court of First Instance of Antwerp regarding the constitutionality of Article 329bis, § 2, third paragraph of the Civil Code. The question submitted to the Constitutional Court pertained precisely to the scope of judicial review:

‘Does Article 329bis, § 2, third paragraph of the Civil Code, where this article states that the court can refuse recognition if it is manifestly contrary to the best interests of the child, violate Article 22bis, paragraph 4 of the Constitution when interpreted as allowing the court only a marginal review [*marginale toetsing*] of the best interests of the child?’ [Translation is mine.]

The first Court of Cassation opinion on the list, a judgment of 17 September 2017 (Nr. S.16.0057.N), rejected an appeal against the decision of a labour court which refused to go beyond a marginal review of determinations made by the Federal Social Security Agency. The decision itself is rather terse and the discussion of marginal review somewhat cryptic. The judges restrict themselves to asserting that the Federal Social Security Agency acted as the law required. Yet the party briefs contains extensive elaborations about judicial review powers, with the plaintiff insisting that the labour court should have applied a full review.

Aside from showing why it might be useful to include, where available, party briefs in the analysis, this decision is also interesting in that neither the judges nor the parties use the phrase ‘*marginale toetsing*’. This illustrates the fundamental difference with approaches to information retrieval relying on exact matching: relevant documents can be retrieved even if they do not feature the exact word or phrase used to specify our concept of interest.

Finally, a look at the topic vectors closest to our *marginale + toetsing* vector suggests decisions involving discussing of marginal review often relate to child protection and filiation, and construction and urban planning. Topic 31 (the closest to *marginale + toetsing* by cosine similarity score) relates to the planning instruments (WWROSP and ‘*gewestplan*’) of the Walloon and Flemish regions (see Figure 12). Topic 34 regards paternity (‘*vaderschap*’) and ancestry (‘*afstamming*’).

Figure 12 Topic 31: Construction and Urban Planning



Figure 13 Topic 34: Filiation and Paternity



6 CONCLUSION

In this chapter I have tried to explain and showcase the potential of the Top2Vec algorithm for legal research. To that end, I trained a prototype using a corpus of several thousand Belgian high court decisions and walked the reader through concrete examples of search and exploration tasks.

The Top2Vec algorithm provides a geometric representations of words, decisions and topics, accounting for the contextualized nature of linguistic meaning. This rich representation makes it a powerful tool to find the proverbial needle in the haystack of legal documents.

In my experience, Top2Vec and embedding-based methods vastly outperform exact matching approaches (even with clever use of Boolean operators).³⁰ To ensure that it does, I recommend to check, as I did in my examples, what information is encoded in the specified word vector. If the specified word vector captures relevant conceptual associations, it should consistently deliver better result than an exact matching search. Cosine similarity effectively works as a measure of search relevance and the documents with the highest cosine similarity to the word vector should not contain false positives.

The prototype I presented here can be improved in several ways. First, a larger corpus would help generate better vector embeddings. With more texts, the algorithm can learn more nuances about the contexts in which words occur. Second, Top2Vec can be fed word embeddings generated by transformer language models like Google's BERT.³¹ The word vectors produced by Word2Vec are 'static' and type-based. As such, they can handle synonymy (if 'car' and 'vehicle' are used to mean the same thing they will have high cosine similarity score) but not polysemy or co-reference resolution (to determine what a pronoun refers to). Yet the vector representing the word 'party', for instance, will not differentiate between party as in 'political party' and the party to a legal case. In a large and relatively diverse corpus, the Word2Vec is thus liable to assign high cosine similarity to words associated with both contexts. By contrast, transformer models, which are trained to predict surrounding words as well as surrounding sentences, generate word vectors that go beyond generalizing across contexts. These vectors represent words as dynamic, token-based vectors, thereby coming much closer to capturing the particular, sentence-specific context of

30 For an evaluation of the performance of vector-based methods for information retrieval tasks, see F. Diaz *et al.*, 'Query Expansion with Locally-Trained Word Embeddings', (2016) arXiv preprint arXiv:1605.07891.

31 J. Devlin *et al.*, 'Bert: Pre-Training of Deep Bidirectional Transformers for Language Understanding', (2018) arXiv preprint arXiv:1810.04805.

occurrence of a word. These word vectors are thus able to handle polysemy and co-reference resolution on top of synonymy.

Third, a transformer model for Dutch legal language has yet to be developed and training one from scratch requires significant resources (a very large corpus and vast computing power). Existing transformer models for the Dutch language have been trained on non-legal texts,³² making their word embeddings less suitable for legal tasks. This is the reason why I opted for training my word embeddings locally using the corpus of Belgian decisions. The development of transformer models for legal Dutch, however, may bring significant improvements.

32 P. Delobelle *et al.*, 'Robbert: A Dutch Roberta-Based Language Model', (2020) arXiv preprint arXiv:2001.06286.



8 METHODOLOGISCHE ASPECTEN VAN EEN DATA PIPELINE VOOR RECHTSPRAAKANALYSE

*Eric Tjong Tjin Tai**

1 INLEIDING

Sinds enkele jaren is er veel aandacht voor het gebruik van data science-methoden in het juridisch onderzoek.¹ Hiervoor zijn twee oorzaken aan te wijzen: enerzijds de grotere beschikbaarheid van rechtspraakdata, anderzijds de opkomst van nieuwe analysetechnieken, in het bijzonder de AI-techniek die wordt aangeduid als *deep learning*. Het gebruik van zulke technieken belooft veel, maar tegelijk is er enige aarzeling: hoe moeten zulke technieken worden toegepast op een wijze die wetenschappelijk verantwoord is?² Dit is met name van belang nu veel juristen niet bekend zijn met methoden voor onderzoek anders dan het traditionele doctrinaire of dogmatische onderzoek aan de hand van op papier gedrukte teksten.

In dit hoofdstuk zal ik proberen tot enige methodologische aanwijzingen te komen voor het gebruik van data science-technieken bij rechtspraakanalyse. Deze aanwijzingen zijn gebaseerd op een experimenteel systeem binnen Tilburg Law School. Dit systeem bestaat uit een zogenoemde *data pipeline*, een geautomatiseerde stapsgewijze verwerking van data teneinde deze geschikt te maken voor analyse. Dit hoofdstuk beschrijft de opzet van dit systeem, enige ervaringen die zijn opgedaan, en enkele voorlopige methodologische bevindingen en aanbevelingen.

Enkele waarschuwingen vooraf zijn op zijn plaats. Ik zal in de tekst veelvuldig IT-jargon en Engelse termen gebruiken. Ik ga geen moeite doen deze door Nederlandse woorden te

* Prof. mr. dr. ir. T.F.E. Tjong Tjin Tai is Hoogleraar Privaatrecht aan Tilburg University en onderzoeker bij het Tilburg Institute for Private law (TIP).

1 Een (niet geheel compleet) overzicht van beschikbare databanken en researchgroepen is te vinden op <https://github.com/Liquid-Legal-Institute/Legal-Text-Analytics>.

2 Zie bijvoorbeeld kritisch over het trekken van onjuiste conclusies uit AI-voorspellingen van rechtspraak: F. Bex & H. Prakken, 'De juridische voorspelindustrie: onzinnige hype of nuttige ontwikkeling?', *Ars Aequi* 2020/3, p. 255-259.

vervangen: zulke vertalingen zouden gekunsteld zijn en het lastiger maken om thuis te raken in dit gebied. De lezer die dergelijk onderzoek ook wil uitvoeren zal uiteindelijk meestal veel Engelse literatuur en websites moeten raadplegen en het is dan gemakkelijker om al bekend te zijn met de daar gebruikte termen en jargon.

2 BESCHRIJVING VAN HET SYSTEEM

Het systeem waar deze bijdrage over gaat, is niet zozeer voortgekomen uit een hang naar *methode*, als wel uit experimenteerdrijf. Sinds 2017 heb ik af en aan gewerkt aan een experimentele omgeving³ waarbinnen verschillende data science- en AI-technieken konden worden uitgetoetst op rechtspraakdata. Het oogmerk was om beter te begrijpen hoe deze technieken werken, wat hun mogelijkheden en beperkingen zijn, en hopelijk ook daarbij nieuwe inzichten te genereren. Een nevenuitkomst hiervan is evenwel ook geweest dat ik beter zicht kreeg op de waarborgen die noodzakelijk zijn om betrouwbare uitkomsten te verkrijgen.

Ik zal het systeem in enig detail beschrijven. Deze beschrijving lijkt voor juristen op het eerste gezicht misschien nodeloos gedetailleerd. Ik zal later beargumenteren waarom wij, als wij data science-methoden willen inzetten, toch enigszins vertrouwd moeten raken met deze details.

2.1 De infrastructuur

De *hardware* van het systeem bestaat uit een Dell PC met een betrekkelijk bescheiden CPU,⁴ 32 Gb RAM geheugen en een GPU (Graphics Processing Unit, grafische kaart).⁵ Een GPU is een aparte processor die gespecialiseerd is in het snel en efficiënt uitvoeren van berekeningen voor grafische output. GPU's zijn oorspronkelijk in de jaren negentig ontstaan om geavanceerde *graphics* mogelijk te maken voor spelsoftware (*games*). Maar GPU's zijn ook bijzonder geschikt voor het soort berekeningen (matrixbewerkingen) die nodig zijn voor het trainen van algoritmen. Veel belangrijke *deep learning libraries* ondersteunen *GPU acceleration*,⁶ en GPU-fabrikanten richten zich expliciet op de AI-markt.⁷

3 Deze omgeving is gefinancierd door Tilburg Law School en het Department of Private Law.

4 Intel Core i7-7700K (Quad Core 4.2GHz, 4.5Ghz Turbo).

5 NVIDIA Quadro P4000 8 Gb.

6 Bijvoorbeeld Tensorflow (www.tensorflow.org/guide/gpu).

7 Bijvoorbeeld www.nvidia.com/en-us/deep-learning-ai/products/solutions/.

Verder zijn belangrijk een Solid State Disk drive (SSD) die aanzienlijk lagere toegangstijden heeft dan de traditionele harde schijf (HDD), wat een belangrijke versnelling van verwerking oplevert bij de grote hoeveelheden data die aan de orde zijn bij AI.

Als *operating systeem* is gekozen voor Ubuntu, een variant van Linux, dat weer een *open source* variant van Unix is. Unix-achtige systemen (zoals ook MacOX) worden bij data science-ontwikkeling meestal geprefereerd boven een andersoortig OS als Windows. De meeste ervaring is opgedaan binnen Unix-achtige omgevingen, en het beheer van zo'n omgeving is ook gemakkelijker binnen een Unix-variant.⁸

De *software* waarmee de pipeline is opgezet is grotendeels gebaseerd op *open source libraries*. Dit is inmiddels gebruikelijk. De belangrijkste programma's en bibliotheken zijn als *open source software* beschikbaar gesteld.⁹ Aanpassingen van software worden ook publiekelijk bediscussieerd; kennis over gebruik en het oplossen van fouten wordt breed gedeeld en is een belangrijke aanjager van de ontwikkelingen op dit gebied. Dit geldt zowel voor de basissoftware (programmeertalen als Python, Javascript, databasesystemen als MySQL, NoSQL, MongoDB, Elasticsearch, interface en visualisatiesoftware als NodeJS, Cytoscape e.d.) als specifieke AI-bibliotheken (Gensim, Tensorflow). Het werk zit vooral in het aan elkaar knopen van de losse bestanddelen en het aanpassen aan specifieke behoeften (*customizen*).

2.2 De data

Bij het systeem is gewerkt met Nederlandse en met Engelse¹⁰ rechtspraak. Omdat de meeste ervaring is opgedaan met de Nederlandse data zal daar in het vervolg op worden ingegaan.

Mede in het licht van Europese richtlijnen die verplichten tot beschikbaar maken van overheidsdata¹¹ maakt de Raad voor de rechtspraak alle uitspraken die via rechtspraak.nl zijn te vinden ook beschikbaar als compleet pakket.¹² Dit bestaat uit een Zip-bestand van

8 Installatie en versiebeheer van tools is meestal geënt op het bevoegdheidssysteem en de *directory*-structuur van Unix, handleidingen en voorbeelden bevatten vaak stukjes Unix-script (bash o.i.d.). Zie ook zoekresultaten met termen 'data science unix windows'.

9 Dat wil zeggen in essentie gratis beschikbaar.

10 Dit was als proef teneinde de haalbaarheid hiervan te beoordelen. Hier is verder nog geen uitgebreide analyse op uitgevoerd.

11 Richtlijn 2003/98/EG en Richtlijn 2013/37/EG. Algemeen over open data van rechtspraak: M. van Opijnen, *Op en in het web. Hoe de toegankelijkheid van rechterlijke uitspraken kan worden verbeterd*, (diss. Amsterdam UvA), Den Haag: Boom Juridische uitgevers 2014. Kritisch over ongebreideld openbaar maken van data: M. van Eechoud, *De lokroep van open data* (oratie Amsterdam UvA), 2014.

12 www.rechtspraak.nl/Uitspraken-en-nieuws/Uitspraken/Paginas/Open-Data.aspx.

5,18 Gb (1 december 2020). Uitgepakt komt dit neer op 2,8 miljoen bestanden van totaal 23,7 Gb.¹³ Elk bestand bevat één uitspraak of conclusie P-G.¹⁴ Verder kunnen nieuwe uitspraken worden gedownload door verzoeken aan een speciaal daarvoor opgezette server.¹⁵

2.3 De data pipeline

Het systeem dat op de hierboven genoemde hardware, software en data is geconstrueerd is een zogenaamde *data pipeline* voor rechtspraakdata. Bij projecten waarin data worden gebruikt (*big data, data science, deep learning*) wordt vaak gewerkt met een dergelijke data pipeline.¹⁶ Dit is een structuur waarbij ruwe data stapsgewijs, door opeenvolgende bewerkingen, geschikt worden gemaakt voor gebruik. Er is niet één unieke structuur, het gaat om verschillende stappen, afhankelijk van het doel en de aard en kwaliteit van de data. Voor zover hier relevant gaat het met name om de volgende bewerkingen:¹⁷

- *verzamelen* van data (verschillende bestanden);
- *integreren* van data in één bestand, wat meestal uniformeren van format van data vereist;
- *scrubbing*: opschonen van ruwe data;
- *transformaties*: wijzigen van de structuur van de data;
- *toepassingen*: gebruik van de data.

De achtergrond van deze stappen is dat data die afkomstig zijn van een externe leverancier over het algemeen niet geschikt zijn om direct gebruikt te worden.

13 In XML-formaat.

14 Bij oude uitspraken, zoals het *Kelderluik*-arrest, ontbreekt meestal de volledige tekst en zijn alleen zaaksgegevens opgenomen.

15 Het betreft een zogenaamde *RESTful web service*. REST (REpresentational State Transfer) is een bepaalde stijl van netwerkkarchitectuur die geschikt is om via internet grote hoeveelheden data beschikbaar te stellen: een computer kan dan de server bevragen om precies die data te krijgen die nodig zijn. Zie: https://en.wikipedia.org/wiki/Representational_state_transfer, verwijzend naar R. Fielding, *Architectural Styles and the Design of Network-based Software Architectures*, Ph.D. thesis University of California (Irvine), 2000.

16 Op internet zijn hier talloze beschrijvingen over te vinden (bijv. <https://towardsdatascience.com/a-beginners-guide-to-the-data-science-pipeline-a4904b2d8ad3>, <https://analyticsweek.com/content/data-science-pipeline/>, www.ibm.com/developerworks/library/ba-intro-data-science-1/index.html, <https://medium.com/analytics-vidhya/what-is-a-pipeline-in-machine-learning-how-to-create-one-bda91d0ceaca>). Deze zijn vooral praktisch en niet wetenschappelijk onderbouwd. Bij wetenschappelijk onderzoek wordt hier ook gebruik van gemaakt, maar dit wordt methodisch meestal niet onderbouwd. Ik kom hier nog op terug.

17 Vergelijk ook C.S. Alexander e.a., 'Using Text Analytics to Predict Litigation Outcomes', in: M.A. Livermore & D.N. Rockmore (red.), *Law as Data: Computation, Text and the Future of Legal Analysis*, New Mexico: Santa Fe Institute Press 2019, p. 275-311, die wel vrij gedetailleerd het proces beschrijven om een bruikbare database te verkrijgen (p. 278-305).

2.3.1 Verzamelen

Het verzamelen van data kan op verschillende manieren gebeuren. Zoals hierboven is aangegeven zijn voor de rechtspraak.nl-data twee wegen beschikbaar: een periodieke integrale download, en updates via een server. De eerste manier is handig bij het opstarten (omdat je dan alles in één keer hebt), maar niet voor updaten, omdat je dan steeds een enorme hoeveelheid opnieuw downloadt.

Voor de *pipeline* is een aparte *web scraper* gemaakt¹⁸ die de data van onder andere rechtspraak.nl verzamelt en aan de database toevoegt.

2.3.2 Integreeren

De data uit verschillende bronnen moeten worden samengevoegd tot één bestand.¹⁹ In het geval van rechtspraak.nl-data is dat gemakkelijk omdat het allemaal in hetzelfde format (bestandsformaat) is, maar bij data uit verschillende bronnen is dat meestal niet zo. Dan moeten de data in eenzelfde structuur worden gebracht om gezamenlijk verwerkt te kunnen worden. In het hier beschreven lab is het wel zo dat de data die door de webscraper zijn verkregen moeten worden geïntegreerd met de data die in bulk (het grote zip-bestand) zijn gedownload, maar dat is voor de hand liggend en verschilt niet met het gewone bijwerken van de data in de totale database.

Toch kunnen al in dit stadium problemen voorkomen. Een voorbeeld is dat een uitspraak wordt gedownload die al in de database aanwezig lijkt te zijn. Dit zou kunnen omdat de uitspraak is verbeterd: het lijkt erop dat rechtspraak.nl fouten in de data periodiek wijzigt zonder dit expliciet aan te geven.

Dat ligt anders voor de data die uit bijvoorbeeld Engelse bronnen is verkregen. Daarbij geldt dat de structuur van de rechtspraak anders is dan in Nederland: er zijn andere gerechten, de vorm van de uitspraak is anders (de rechters staan aan het begin genoemd), en dergelijke. In de mij bekende literatuur wordt meestal niet geprobeerd zulke data samen te voegen: men gebruikt data uit een enkele jurisdictie.²⁰

18 De *web scraper* is door een extern bedrijf gemaakt.

19 Als er volledig gescheiden bestanden zijn die niet meer gecombineerd worden, gaat het in feite om gescheiden applicaties.

20 Een uitzondering zijn mogelijk de Verenigde Staten, voor zover de verschillende Amerikaanse staten en gerechten als afzonderlijke jurisdicties moeten worden beschouwd. Deze hebben namelijk wel grotendeels eenzelfde structuur, ook al zijn zij formeel misschien onafhankelijk. In E. Ash & D.L. Chen, 'Case Vectors: Spatial Representations of the Law using Document Embeddings', in: Livermore & Rockmore 2019, p. 313-337, is de rechtspraak van de federale appelrechters bestudeerd, maar dat lijkt te berusten op een databank waarin de rechtspraak al genormaliseerd was (de databank wordt overigens niet geïdentificeerd). In M. Dumas, 'Detecting Ideology in Judicial Language', in: Livermore & Rockmore 2019, p. 383-405, wordt

2.3.3 Scrubbing

De data zijn, zoals meestal bij *big data*, niet schoon. Zij bevatten fouten.²¹ Dit komt doordat de brondata fouten of lacunes bevatten. Dit kan komen door invoerfouten, doordat sensoren tijdelijk defect waren, of door fouten in de software waarmee de data werden gegenereerd. Een voorbeeld van een invoerfout is de uitspraak ECLI:NL:GHAMS:1982:BH6225: in het veld datum uitspraak staat '31 december 1982', maar de uitspraak is blijkens de tekst van 23 december 2008. Een ander voorbeeld betreft de uitspraak in ECLI:NL:RBARN:2008:BD6872, die geen dictum bevat. Waarschijnlijk is het slot van de uitspraak weggevallen of niet ingevoerd.

Ook kan er sprake zijn van wijziging in het format van de data: bijvoorbeeld doordat tot een bepaalde datum gegevens in de ene vorm werden opgenomen en vanaf die datum in een andere vorm (zoals in het verleden LJV werd gebruikt en vanaf bepaalde datum ECLI). Een ander probleem is niet-uniform gebruik van het format. Zo is de inhoud van het veld 'uitspraak' (in de XML-structuur) niet uniform. De ene keer begint in dat veld gewoon de tekst van de uitspraak, de andere keer is dit veld onderverdeeld in subvelden.²² Voor een computerprogramma is dit lastiger om te verwerken omdat er met verschillende mogelijkheden rekening moet worden gehouden. Het ligt voor de hand om de onderverdeling te negeren en de tekst van de uitspraak gewoon samen te stellen door al de subvelden aan elkaar te plakken, maar het is mogelijk dat hiermee relevante informatie verloren gaat.

Voor veel toepassingen van big data is het niet erg als de data kleine fouten bevatten: die zijn statistisch niet significant en verdwijnen bij analyses over grote hoeveelheden data.²³ Maar sommige fouten kunnen ertoe leiden dat de analyse te veel vertekend wordt. Denk aan een tikfout waardoor het inkomen van een persoon als 1 triljard euro is opgegeven: dit zou het gemiddelde inkomen significant vertekenen.

Ook is het mogelijk dat fouten moeten worden aangepakt omdat andere delen van de software anders niet weten hoe hiermee om te gaan. Dan zou in de database de datum van de bovenvermelde uitspraak moeten worden aangepast. Het bezwaar is dat dan de data

gebruikgemaakt van data afkomstig uit Westlaw, evenzo in A. Feldman, 'Opinion Clarity in State and Federal Trial Courts', in: Livermore & Rockmore 2019, p. 407-430. In het onderzoek beschreven in V. Eidelman e.a., 'Predicting Legislative Floor Action', in: Livermore & Rockmore 2019, p. 117-150 worden parlementaire stukken van verschillende Amerikaanse staten samengevoegd; zij gaan kort in op het proces van normalisatie om dit mogelijk te maken (p. 125-126).

21 Zie ook de beschrijving in M. van Opijnen, 'Legal(ly) Linked Data', *Computerrecht* 2018/2, p. 64-69.

22 Bijvoorbeeld respectievelijk ECLI:NL:RBHAA:2006:AV9107 (het veld 'uitspraak' begint gewoon met tekst) en ECLI:NL:GHARL:2016:498 (het veld uitspraak begint met een subveld 'uitspraak.info' waarin pas de tekst is opgenomen).

23 O.a. V. Mayer-Schönberger & K. Cukier, *Big Data*, Mariner Books 2014.

wordt gewijzigd, wat wetenschappelijk gezien niet zomaar is toegestaan. Een alternatief is dat de software wordt aangepast om met zulke fouten rekening te houden. Het nadeel van die aanpak is dat de software daardoor veel groter en onoverzichtelijker wordt (iedere uitzondering vereist extra code) wat de kans op fouten in de hand werkt. Bovendien wordt de software daardoor ook steeds trager.

2.3.4 Transformaties

Het is vaak nodig om de data ook te transformeren, dat wil zeggen een andere database-structuur te gebruiken of het format aan te passen. Het kan zo zijn dat de structuur niet gelijk of niet handig is voor de beoogde toepassing. De data van rechtspraak.nl worden geleverd in losse XML-bestanden, die zeer inefficiënt zijn om te doorzoeken. Het is zinvol om de data te transformeren naar een efficiëntere database zoals Elasticsearch.²⁴ Daarnaast is het handig als de data vooraf worden geanalyseerd om bijvoorbeeld de namen van rechters of het dictum apart te identificeren: als dit in de pipeline is gedaan, hoeft dit niet bij onderzoek naar de data nog een keer gedaan worden, en dat scheelt veel tijd.

Transformatie: data-elementen onderscheiden

Een belangrijke stap is dat uit de uitspraken relevante elementen worden gehaald.²⁵ Deze stap is niet altijd nodig, het hangt ervan af wat je wilt onderzoeken. Als je bijvoorbeeld wilt zoeken naar uitspraken waarin een bepaalde advocaat is opgetreden lijkt het voldoende om *full-text* te zoeken op een bepaalde naam. Een nadeel hiervan is dat tikfouten in de naam dan niet worden opgepikt.

Voor het hier beschreven project is gewerkt aan het onderscheiden van de namen van de rechters die de uitspraken hebben gewezen, en het onderscheiden van de verschillende onderdelen van een uitspraak: de kop, feiten, vorderingen, motivering en dictum. Dit is alleen geprobeerd voor civielrechtelijke uitspraken omdat deze normaal gesproken een vrij uniforme structuur hebben die afwijkt van wat bij strafrecht en bestuursrecht gebruikelijk is. Ook is gewerkt aan het onderscheiden van NJ- en ECLI-verwijzingen. Het is mij bekend dat ook elders hieraan is gewerkt: diverse bedrijven claimen dat zij rechters en advocaten uit uitspraken hebben gehaald; rechtspraak.nl heeft een groot project opgezet om de uitspraken te verrijken met kloppende NJ- en ECLI-verwijzingen (zie de knop

24 www.elastic.co. Deze database is zeer efficiënt voor *full text search* en heeft ook goed ontwikkelde interfaces (zoals mogelijkheden voor opzetten van een dashboard).

25 Vergelijk H. Laqueur & A. Venancio, 'A Computational Analysis of California Parole Suitability Hearings', in: Livermore & Rockmore 2019, p. 193-231, die beschrijven hoe zij pdf's van *hearings* omzetten in gestructureerde data (p. 204-207). Zij beschrijven dat zij een *parser* hebben gemaakt om de tekst van de gevangene uit de gehele tekst te halen, maar geven niet aan hoe dit is gedaan en hoe verzekerd is dat dit proces correct is uitgevoerd.

‘verrijkte uitspraak’ op rechtspraak.nl bij de uitspraken).²⁶ De precieze werkwijze van deze initiatieven is echter niet openbaar: de code is niet beschikbaar en het is dan ook niet mogelijk om te controleren of deze processen daadwerkelijk correct zijn.

Bij het vervaardigen van *code* om de uitspraken te ontleden traden verschillende moeilijkheden op, die ik hier kort zal schetsen.

De uitspraken zijn niet uniform qua structuur. Hierboven is al aangestipt dat de manier waarop het veld ‘uitspraak’ wordt benut niet uniform is. Daarnaast is ook de tekst niet uniform gestructureerd. Het dictum, bijvoorbeeld, wordt op veel verschillende manieren aangegeven in de tekst, zoals ‘De beslissing’, ‘beslissende’, ‘de uitspraak’, ‘Rechtdoende’, ‘Beslissing’, ‘De rechtbank:’, ‘Het gerechtshof’, om maar een paar voorbeelden te noemen. Als je alleen op bepaalde zaken zoekt, zoals de voorwaarden voor ontbinding van een arbeidsovereenkomst, wordt het gemakkelijker omdat je dan alleen hoeft te kijken naar de beslissing over een bepaald type vordering, waar waarschijnlijk specifieke termen in zijn gebruikt. De resultaten van zo’n onderzoek zijn echter niet zomaar te generaliseren naar alle soorten dicta.

Daarnaast bevatten de data, zoals hierboven al aangegeven, regelmatig fouten. Veel van deze fouten zijn voor menselijke lezers onproblematisch, maar een computer kan hier alleen met een specifieke instructie mee omgaan. Het bepalen van de rechters die de uitspraak hebben gedaan is bijvoorbeeld lastig om goed te doen, aangezien het in wezen kennis vereist van het soort namen dat kan voorkomen. Hoe moet de computer bijvoorbeeld bepalen dat in ‘mrs. Van Leuven, Mos-Verstraten en van der Zanden’ de derde rechter Van der Zanden heet, en in ‘mrs Valk, Wesseling-Lubberink en Van Osch en de leden mr ing Jansens van Gellicum en ir Rogaar’ de woorden mr ing en ir verwijzen naar titels? Hoewel de meeste rechters een achternaam hebben van één woord zijn er ook veel rechters met namen die uit meer onderdelen bestaan. Voor correcte verwerking van de data moet er derhalve veel kennis in het programma worden gestopt, niet alleen van de juiste betekenis, maar ook van varianten (zoals het vermelden van een titel zonder ‘.’).

Een voorbeeld kan dit punt goed illustreren: in de uitspraak HR 21 januari 2000, ECLI:NL:HR:2000:AA4428 staat: ‘Deze beschikking is gegeven door de vice-president Mijnsen als voorzitter en de raadsheren Heemskerk, Fleers, De Savornin Lohman en Hammerstein (...)’. Dan moet de computer achterhalen dat De Savornin Lohman de naam van een raadsheer is, waarbij geen initialen zijn vermeld. Een jurist weet dan dat dit B.C. de Savorning Lohman is, en niet O. of A.H. (naamgenoten die ook rechter waren), omdat

26 Zie hierover: M. van Opijnen, ‘Legal(ly) Linked Data’, *Computerrecht* 2018/2, p. 64-69.

B.C. in die tijd raadsheer in de Hoge Raad was. Dergelijke kennis moet expliciet worden opgenomen om de data te verrijken. Voor raadsheren in de Hoge Raad lijkt dat nog te overzien. Maar voor lagere gerechten is dat praktisch ondoenlijk: zie bijvoorbeeld Hof 's-Gravenhage 13 november 2002, ECLI:NL:GHSGR:2002:AF2536: 'Deze beschikking is gegeven door mrs. Fockema Andreae-Hartsuiker, Stille en Jansen, (...)'. De kans is groot dat rechter Jansen een vaste raadsheer was bij het hof in die tijd, maar het is zonder nader onderzoek niet uit te sluiten dat het een rechter-plaatsvervanger is met dezelfde achternaam.²⁷

Om dit alles softwarematig op te lossen is een behoorlijk complex programma nodig dat gegevens uit verschillende bronnen moet combineren.²⁸ Er is veel kennis nodig over de structuur van gerechten en de werkwijze in de rechtspraak, naast feitelijke kennis over werkzame juristen, om de data te completeren en zo'n programma te vervaardigen. Overigens zijn deze ervaringen gebaseerd op Nederlandse rechtspraak; daaruit kan niet worden afgeleid of deze problemen zich ook voordoen bij andere jurisdicties, zoals de rechtspraak van het Europees Hof voor de Rechten van de Mens (EHRM).²⁹

Deze overwegingen maken duidelijk dat verwerken van uitspraken daadwerkelijk een *menselijke* interpretatie veronderstelt die achtergrondkennis behoeft.³⁰ Geautomatiseerde verwerking is vooralsnog lastig doordat de data daar niet voor bedoeld waren en ook niet zonder meer geschikt voor zijn. Zo werd in het verleden naar *NJ*-publicaties verwezen in

27 Ik ga dan nog voorbij aan de aanwezigheid van tikfouten die voor mensen vaak evident zijn maar door een computer vrijwel alleen te corrigeren door een handmatige lijst van varianten op te stellen, waarbij moet worden bedacht dat Nederlandse namen ook regelmatig bestaan uit woorden die volgens de moderne Nederlandse regels verkeerd gespeld zijn (bijv. Dijkstra, Fortuijn).

28 C.S. Alexander e.a., 'Using Text Analytics to Predict Litigation Outcomes', in: Livermore & Rockmore 2019, p. 275-311, geven aan dat de taak om de beslissingen van rechters en appèlrechters te classificeren te lastig bleek ('this approach was stymied by the wide variation in the language...'), p. 303.

29 Zoals het project beschreven in de bijdrage van Vols, dat overigens ook niet afhankelijk lijkt te zijn van correcte identificatie van rechters. Bij het EHRM zijn minder rechters dan in de volledige Nederlandse rechtspraak, en er komen bij mijn weten geen dubbele namen voor. Het is mij niet bekend of de EHRM-rechtspraak ook tikfouten in de namen van rechters bevat. Hetzelfde kan gelden bij de rechtspraak van het U.S. Supreme Court, zoals bestudeerd in K. Carlson e.a., 'Style and substance on the US Supreme Court', in: Livermore & Rockmore 2019, p. 83-115, of het ECJ, waarover J. Frankenreiter, 'Writing Style and Legal Traditions', in: Livermore & Rockmore 2019, p. 153-191. In het onderzoek van E. Ash & D.L. Chen, 'Case Vectors: Spatial Representations of the Law using Document Embeddings', in: Livermore & Rockmore 2019, p. 313-337 is blijkbaar een niet nader geïdentificeerde databank gebruikt waarin de rechters al correct geïdentificeerd waren.

30 Zie ook de combinatie van menselijke classificatie en softwarematige analyse beschreven in C.S. Alexander e.a., 'Using Text Analytics to Predict Litigation Outcomes', in: Livermore & Rockmore 2019, p. 275-311.

de vorm *NJ* 1985, 613, maar dat bleek voor computers lastig te verwerken, reden waarom dat tegenwoordig vervangen is door de vorm *NJ* 1985/613.³¹

Er zijn overigens ook andere problemen: de bestanden bevatten niet alleen rechterlijke uitspraken, maar ook tuchtrechtelijke uitspraken, processen-verbaal,³² herstelvonnissen.³³ Zulke uitspraken zijn niet geschikt voor analyses die uitgaan van gewone rechterlijke uitspraken op basis van een regulier verzoek of vordering. Ook op dit punt is het nodig om keuzes te maken voor de selectie van te verwerken uitspraken.

Transformatie: databasestructuur

Voor het hier beschreven systeem is gewerkt met diverse databasestructuren en afzonderlijke databases, omdat elke toepassing weer andere eisen had. Bij de grote hoeveelheden data die gebruikt werden, zou het nodeloos vertragend werken als alle toepassingen steeds op één database moesten worden uitgevoerd. Het is veel sneller en efficiënter om een specifieke aangepaste versie van een database te maken waar dan de toepassing op kan worden uitgevoerd.

Ik geef twee voorbeelden ter illustratie. Ten eerste, de telling van *NJ*-verwijzingen kon worden gedaan door een bestand te maken waarin elke *NJ*- of ECLI-verwijzing was opgenomen. Die database was veel kleiner (en daarmee ook sneller te gebruiken) dan de database met complete uitspraakdata. Natuurlijk is het, indien nodig, vrij eenvoudig om die complete database te koppelen aan de verwijzingsdatabase. Het tweede voorbeeld betreft een database waarin per uitspraak de namen van de rechters waren opgenomen. Deze transformaties waren echter, zoals hierboven ook al aangegeven, niet rechttoe rechtaan te verrichten. Het uiteindelijke programma om deze verwijzingen en rechtersnamen eruit te halen is vrij complex en nog steeds niet perfect.

Bij deze transformaties moet je goed opletten of de getransformeerde versie ook is afgeleid van de juiste hoofdversie. Zoals eerder aangegeven verandert het hoofdbestand voortdurend doordat er steeds nieuwe uitspraken bijkomen. Als je een afgeleide database maakt, moet je goed bijhouden van welke versie van de hoofddatabase deze is afgeleid, anders is de kans aanwezig dat zij uit de pas lopen en er fouten ontstaan als je afleidingen maakt (bijvoorbeeld over aantallen en relatieve proporties). Daarnaast worden er door rechtspraak.nl ook soms

31 Eenzelfde reden is er voor het gebruik van ECLI: dat is een verwijzingsvorm die vooral bedoeld is om het computers gemakkelijker te maken: voor mensen is de ECLI-notatie juist lastiger te lezen en onthouden dat de vroegere verwijzingsvormen.

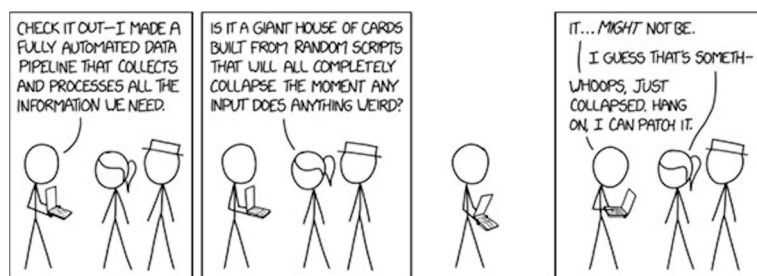
32 Bijv. ECLI:NL:RBHAA:2008:BD2954.

33 Bijv. ECLI:NL:RBROT:2008:BG8407.

fouten in de data verbeterd: het is niet helemaal duidelijk of goed wordt doorgegeven dat de data veranderd zijn en hoe dit dan moet worden verwerkt in het totale bestand.

Dit versiebeheer van databases is complex. Een goed opgezette pipeline zou er automatisch voor moeten zorgen dat de afgeleide databases ook steeds worden bijgewerkt, maar dat vergt een behoorlijk complexe set-up die in de praktijk al snel omvalt (doordat er iets in de omgeving of de bronnen verandert) en dan weer door een deskundig programmeur aan de praat moet worden gekregen (zie figuur 1). Deze handmatige werkzaamheden kunnen, bij onoplettendheid, gemakkelijk ertoe leiden dat bijvoorbeeld de verschillende transformaties niet op elkaar aansluiten of dat een verkeerde versie van een transformatie wordt gebruikt, waardoor de data gecorrumpereerd raken en de resultaten van analyse van deze data niet in overeenstemming zijn met de oorspronkelijke data.

Figuur 1 De complexiteit van data pipelines



Bron: <https://xkcd.com/2054/>

Deze moeilijkheden kunnen ten dele worden vermeden door niet, zoals in het bedrijfsleven, te werken met een steeds uitdijende dataset. In plaats daarvan kan – zoals bij veel wetenschappelijk onderzoek – worden gewerkt met een statische, gefixeerde dataset. Dat neemt niet weg dat het waarschijnlijk toch wenselijk is dat het onderzoek gebaseerd wordt op vrij recente, actuele data. Dan is er misschien geen dagelijks bijgewerkte pipeline nodig, maar zal wel een pipeline worden opgesteld waarmee bijvoorbeeld eens per jaar de meest actuele dataset van rechtspraak.nl wordt verwerkt. Zo'n geautomatiseerde verwerking is een reden waarom het onwenselijk is dat er handmatig correcties in de data worden aangebracht: die correcties gaan verloren als dezelfde data nogmaals door de pipeline worden verwerkt.³⁴

³⁴ Dat kan gebeuren als rechtspraak.nl een oude uitspraak zou corrigeren en die correctie als aangepaste data wordt aangeboden aan de pipeline.

2.4 Conclusie

De opzet van een data pipeline voor rechtspraakdata laat zien dat het misleidend is om te spreken over ‘de’ data. Er zijn verschillende bestanden aanwezig die allemaal gebaseerd zijn op de oorspronkelijke data. Het resultaat van het gebruik van een data pipeline is dat de ruwe data zijn veranderd, bewerkt. De processen waarmee die bewerkingen zijn uitgevoerd bepalen wat eruitkomt.

3 TOEPASSINGEN

De output van de pipeline is een geüniformeerde, opgeschoonde database in een gemakkelijk bruikbaar format (of verschillende varianten van dezelfde data, die voor verschillende doeleinden geschikt zijn). Die database kan gebruikt worden als basis voor diverse soorten toepassingen en experimenten.³⁵ Een paar voorbeelden.

3.1 Netwerkanalyse

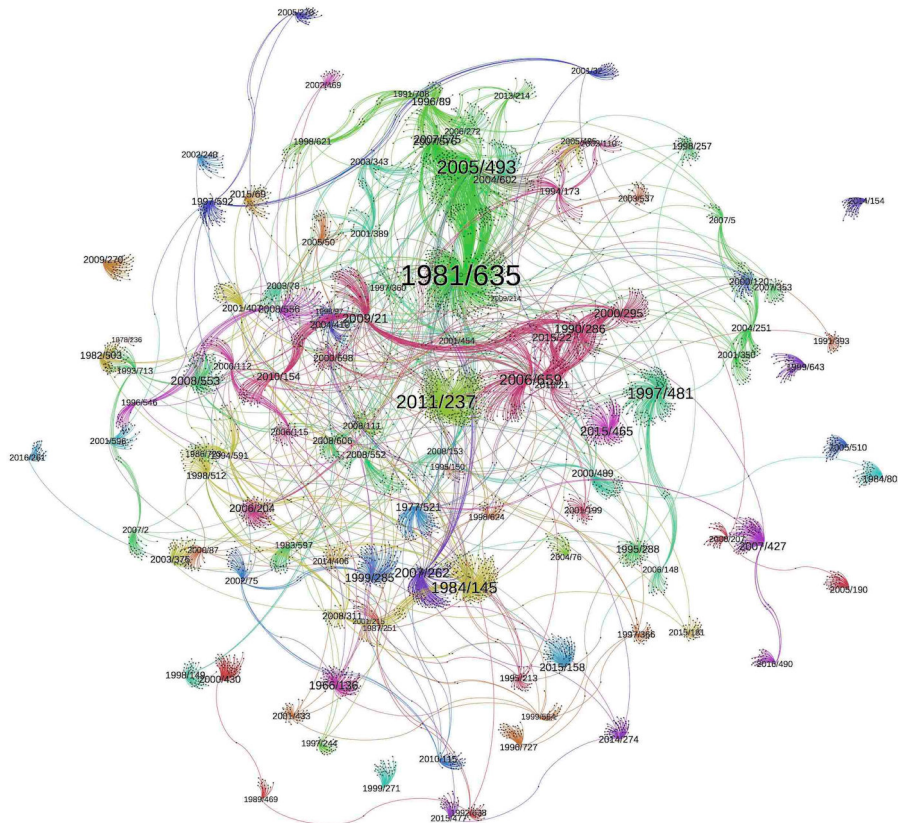
In de juridische literatuur is er al enkele jaren aandacht voor het gebruik van netwerkanalyse bij juridisch onderzoek, in het bijzonder om verwijzingen tussen rechterlijke uitspraken aan elkaar te verbinden.³⁶ Voor Nederlandse rechtspraak betekent dit bijvoorbeeld dat de expliciete verwijzingen naar andere rechtspraak in de vorm van *NJ*- of ECLI-verwijzingen in kaart worden gebracht. Dit kan met specifieke software of aan de hand van ondersteunende software als Gephi of Cytoscape. Figuur 2 geeft een voorbeeld van het resultaat van een dergelijke analyse, waarbij in kaart is gebracht wanneer een rechterlijke uitspraak of conclusie A-G naar een uitspraak in de *NJ* verwijst.³⁷

35 Vgl. B. Custers, ‘Methods of data research for law’, in: V. Mak, T.F.E. Tjong Tjin Tai & A. Berlee, *Research Handbook in Data Science*, Cheltenham: Edward Elgar 2018, p. 355-377.

36 Uitvoerig: C. Coupette, *Juristische Netzwerkforschung*, Tübingen: Mohr Siebeck 2019. In Nederland zie G. van Dijck & D. van Kuppevelt, ‘Netwerkanalyse van werkgeversaansprakelijkheid: Deelonderwerpen en precedentes in het kader van werkgeversaansprakelijkheid ex artikel 7:658/7:661 BW’, *NJB* 2019/527, afl. 10, p. 652-661, G. van Dijck & J. Cleuters, ‘Plas/Valburg na CBB/JPO? Een netwerkanalyse van rechtspraak over afgebroken onderhandelingen in de precontractuele fase’, *Contracteren* 2020, p. 11-19.

37 De kleine stippen geven de uitspraken aan die een verwijzing bevatten en de lijn is getrokken naar de uitspraak waarnaar is verwezen (die met de *NJ*-verwijzing is geïdentificeerd). Ter wille van de overzichtelijkheid zijn alleen de uitspraken waarnaar vaker wordt verwezen (dertig keer of meer) opgenomen.

Figuur 2 Verwijzingsnetwerk



Hierin is te zien dat bijvoorbeeld heel vaak naar *Haviltex* (NJ 1981/635) wordt verwezen. Een aantrekkelijke kant van dit onderzoek is dat het fraai ogende plaatjes oplevert. Er is echter twijfel aan de wetenschappelijke opbrengst van zulk onderzoek: het levert slechts beperkt nieuwe inzichten op.³⁸ Veel onderzoek beperkt zich tot het toepassen van netwerkanalyse op juridische data, zonder rekening te houden met wat juridisch relevant is.³⁹ Methodologisch is bijvoorbeeld een beperking dat impliciete verwijzingen zoals het gebruik

38 Zie Coupette 2019, p. 324-327, en het overzichtsartikel van R. Whalen, 'Legal Networks: The Promises and Challenges of Legal Network Analysis', *Michigan State Law Review* 2016, p. 539. Verder kritisch N. Peterson & E.V. Towfigh, 'Network Analysis and Legal Scholarship', *German Law Journal* 2017, p. 695-700, op p. 697-699, ook J. Frankenreiter, 'Network Analysis and the Use of Precedent in the Case Law of the CJEU – A Reply to Derlén and Lindholm', *German Law Journal* 2017, p. 687-694, ook de slotopmerkingen van Van Dijk & Kuppevelt, *NJB* 2019/527.

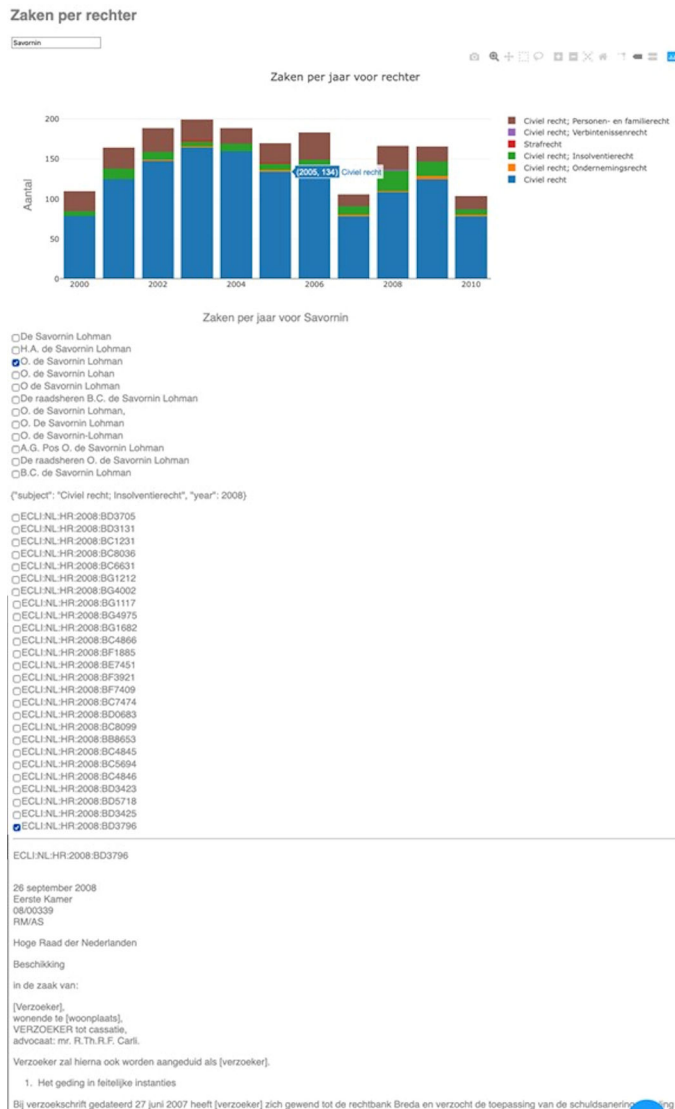
39 Coupette 2019, p. 325-326, vgl. p. 315-316.

van een *Haviltex*-formule of de *Kelderluik*-criteria zonder *NJ*-verwijzing buiten beeld blijven.

3.2 *Namen van rechters in uitspraken*

Een ander voorbeeld is het onderzoeken van de uitspraken die een rechter heeft gedaan. Op basis van een *parser* (software die tekst in onderdelen splitst) zijn de namen van de rechters die uitspraken hebben gewezen opgenomen in een database. Dit proces is niet geheel foutloos (zoals hierboven beschreven). De database is in een experimentele set-up toegankelijk gemaakt in een programma met browser-interface waarmee relatief snel voor een naam de bijbehorende uitspraken kunnen worden gezien in soort en jaar, en kunnen worden geraadpleegd.

Figuur 3 Tool om de door een rechter gewezen zaken grafisch in beeld te brengen en snel te raadplegen: de selectie van de zaak in de lijst van zaken per jaar, met daaronder direct de tekst van de uitspraak



Dit resultaat heeft niet direct wetenschappelijk nut. Weliswaar kan men proberen om verschillen tussen rechters te achterhalen aan de hand van de door hen gewezen uitspraken, maar zoals aangegeven is op dit moment de kwaliteit van de herkenning van de rechters

nog onvoldoende om er wetenschappelijk betrouwbare conclusies aan te verbinden.⁴⁰ Wel heeft dit deelproject nuttige ervaring opgeleverd, en een *proof of concept* van een mogelijke webinterface om data te analyseren.

De ervaring is enerzijds dat het voor de (Nederlandse)⁴¹ databestanden waarmee is gewerkt lastig is om de namen 100% correct uit de data te halen als gevolg van de vele tikfouten en varianten die in de data aanwezig zijn (er is veel werk nodig om een perfecte *parser* te maken), anderzijds dat een *web-based tool* het daadwerkelijk gemakkelijker maakt om de data te ontsluiten (zie ook par. 3.4).

3.3 *Natural language processing*

In de literatuur wordt *natural language processing* (NLP) vaak besproken als mogelijke toepassing van AI, algoritmes of deep learning op rechtspraak. NLP houdt in dat tekstuele data worden verwerkt met technieken die geschikt zijn om natuurlijke taal te verwerken.⁴² Een voorbeeld is het classificeren van woorden in categorieën naar onderwerp (*topic modeling*). Figuur 4 geeft een voorbeeld.

40 Overigens zal bij dergelijk onderzoek ook moeten worden gelet op juridische beperkingen als het recht op privacy en specifieke regelgeving (zoals in Frankrijk, zie G.M. de Vries, 'Rechtswetenschappelijk onderzoek naar rechters door kwantitatieve analyse van jurisprudentie: Waar ligt de ethische grens bij het profileren van rechters?', *RMThemis* 2021, afl. 2, p. 50-68).

41 Zoals boven al aangegeven zijn deze ervaringen uitsluitend gebaseerd op de Nederlandse rechtspraakdata.

42 Zie voor een toepassing van NLP de bijdragen van Vols en Dyevre in deze bundel.

Figuur 4 Een voorbeeld van het resultaat van topic modeling: ieder topic bestaat uit een lijst woorden die volgens het algoritme behoren bij één samenhangend onderwerp

topic 65	topic 66	topic 67	topic 68	topic 69
aan	eiser	of	gemeente	voor
door	schade	worden	aan	met
koper	vordering	omstandigheden	met	die
die	door	kan	voor	als
bedrijf	voor	vraag	door	onder
deze	geleden	redelijkheid	tot	aan
verkoop	tot	naar	grond	worden
met	vergoeding	die	om	nieuwe
was	jegens	zijn	dit	tot
verkoper	gevorderde	geval	besluit	zijn
koopovereenkomst	afgewezen	billijkheid	deze	uit
als	worden	dient	achmea	om
onderneming	schadevergoeding	moet	bouwvergunning	gebouw
levering	onrechtmatige	voor	aanvraag	eerste
voor	zijn	bij	worden	over
zaken	rechtbank	aan	woningen	door
project	aansprakelijk	belang	bij	bij
verkocht	ten	indien	verband	meer
koop	vorderingen	zo	hebben	aantal
bij	deze	oordeel	college	oude

Zoals hieruit blijkt is het mogelijk om woorden te identificeren die vaak in elkaars nabijheid te vinden zijn. Dat geeft inderdaad betekenisvolle clusters van woorden. Door het instellen van bepaalde parameters (keuzes zoals het aantal topics, hoe lang het algoritme wordt getraind op basis van de uitspraken) kan het resultaat worden beïnvloed. De keuze van de parameterwaarden is vooral een kwestie van uitproberen (*trial and error*).

Bij nadere bestudering zijn deze clusters echter veel te algemeen en helpen zij weinig voor juridisch gebruik. Dat een uitspraak over aansprakelijkheid en schadevergoeding gaat, weten we meestal wel en kon ook al op basis van simpele zoekopdrachten worden vastgesteld. Het zou vooral handig zijn als een algoritme zou kunnen zoeken op zeer specifieke rechtsvragen. Om dat te bereiken zou echter sowieso een verdeling in zeer veel topics nodig zijn,⁴³ en de techniek werkt vooral goed bij een beperkt aantal topics. Het is overigens mogelijk dat met andere technieken of een combinatie van technieken betere resultaten

⁴³ Ter toelichting: een onderverdeling in zo'n 300 topics van alle civiele uitspraken over een jaar (zie tabel 1) gaf bij topic 66 een woordcluster dat duidt op schadevergoeding wegens aansprakelijkheid. Dat is echter alleen een hoofdcategorie, daadwerkelijk onderzoek vindt plaats op het niveau van gedetailleerdere subsub-categorieën als soorten aansprakelijkheid, schade en schadevergoeding. Om tot dat niveau te komen zul je in de orde van duizenden onderwerpen moeten categoriseren, of andere technieken gebruiken.

kunnen worden bereikt.⁴⁴ Het gaat het bestek van dit artikel te buiten om hier nader op in te gaan.

3.4 Combinatie van technieken

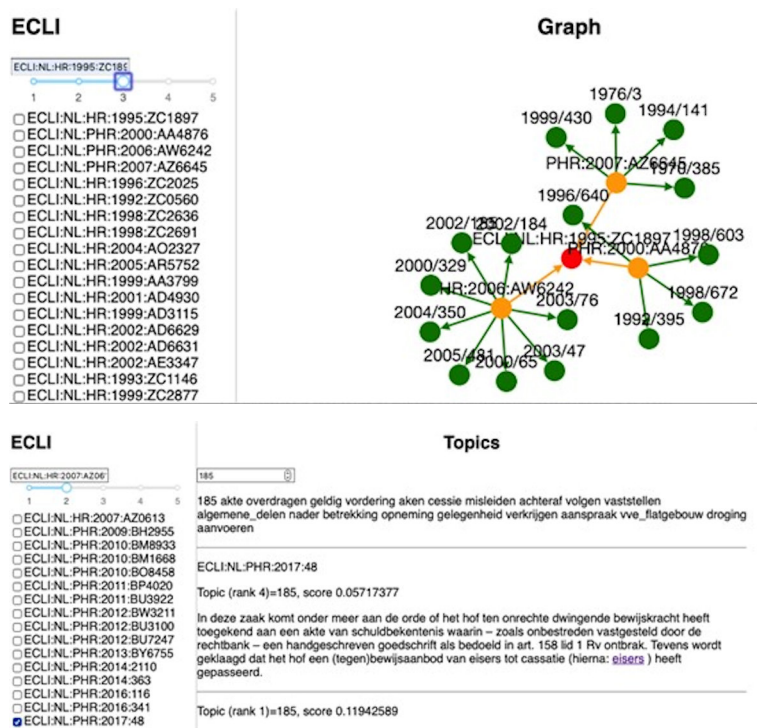
Een mogelijkheid die ontstaat doordat alle data op een eigen server zijn geplaatst is dat diverse technieken kunnen worden gecombineerd. Zo kan bijvoorbeeld een webinterface worden geconstrueerd waarmee enerzijds op basis van netwerkanalyse kan worden doorzocht, anderzijds delen van de gevonden uitspraken aan de hand van topic modeling worden geclassificeerd. Het nut van zo'n benadering zou kunnen zijn dat dit voor onderzoekers de mogelijkheid biedt geavanceerder door de data te zoeken, wat misschien voor bepaalde vragen sneller of vollediger zou kunnen werken.⁴⁵ Daarnaast kan zo'n benadering een geschikte tool vormen om de gebruikte technieken te verfijnen (bijvoorbeeld eenvoudiger uit te vinden welke parameters het beste werken bij topic modeling).

Ter illustratie volgen twee *screenshots* van een dergelijke tool met zo'n interface. In de eerste stap kan aan de hand van eenvoudige netwerkanalyse (verwijzingen tussen uitspraken) een indicatie worden verkregen of bepaalde uitspraken over dezelfde onderwerpen gaan. Dit maakt het mogelijk om betrekkelijk snel en simpel de relevante uitspraken te selecteren voor verdere analyse. In de tweede stap kan vervolgens een van de geselecteerde uitspraken worden geanalyseerd aan de hand van een lijst van *topics* verkregen uit *topic modeling*. Aan de bovenzijde wordt dan een topic geselecteerd, en de passages uit de uitspraak die het beste passen bij dat topic worden op volgorde van beste *match* gepresenteerd.

44 Geavanceerde toepassingen van AI berusten vaak op ingenieuze aanpassingen van het basismodel, door tussenvoegen van extra lagen in het neurale netwerk, verschillende vormen van terugkoppeling, combinatie van verschillende netwerken of trainingsmodellen.

45 In afwezigheid van zo'n tool is een onderzoeker aangewezen op de omslachtiger werkende interface van rechtspraak.nl. Zie bijvoorbeeld de ervaringen gedocumenteerd in de bijdragen van Wijntjens en Verbruggen in deze bundel.

Figuur 5 Twee screenshots van tool om uitspraken op samenhang te selecteren en binnen uitspraak op topic



4 METHODOLOGISCHE OPMERKINGEN

Deze vrij uitvoerige beschrijving geeft aanleiding tot enkele methodologische opmerkingen. Wetenschappelijke kennis wordt voortgebracht door een methode die de betrouwbaarheid en reproduceerbaarheid van de resultaten waarborgt. Voor doctrinair juridisch onderzoek is dit van oudsher niet zo'n probleem, doordat dit onderzoek meestal is gebaseerd op openbaar toegankelijke, gepubliceerde rechtspraak en literatuur, zodat de lezer de brongegevens zelf kan raadplegen en de argumentatie ook kan volgen (die is immers in de publicatie neergelegd). Als we overgaan naar onderzoek gebaseerd op grote hoeveelheden rechtspraakdata ligt dat anders. Ik ga hier niet in op onderzoek door handmatige analyse van de data. Daarover gaan enkele andere bijdragen in deze bundel. Ik richt mij op onderzoek dat is gebaseerd op een data pipeline, dat wil zeggen de geautomatiseerde verwerking van grote hoeveelheden data, zoals de in paragraaf 3 genoemde voorbeelden.

Betrouwbaarheid is niet vanzelfsprekend, aangezien de data te omvangrijk zijn om te controleren voor de lezer, zelfs als de data publiek beschikbaar zijn. Er vinden verschillende bewerkingen plaats op de data waarbij vergissingen kunnen worden gemaakt.

Daarnaast zijn de data zelf niet vrij van fouten en is een methodologische vraag hoe met zulke fouten moet worden omgegaan. Dit zal ook afhangen van de toepassing. Voor het trainen van een algoritme om de uitkomst te voorspellen op basis van de feiten is het misschien niet van belang dat spelfouten in de namen van rechters worden gecorrigeerd,⁴⁶ maar voor een toepassing waarbij wordt gezocht naar patronen in de wijze waarop individuele rechters uitspraken doen is dat wel van belang. Als er wetenschappelijke claims worden gedaan over verschillen tussen rechters is van belang dat controleerbaar is hoe de rechters aan de uitspraken zijn gekoppeld.

De *reproduceerbaarheid* is eveneens niet vanzelfsprekend. Om de uitkomsten daadwerkelijk reproduceerbaar te maken zou het nodig zijn om de experimenten in veel meer detail te beschrijven dan op dit moment gebruikelijk is. De mij bekende artikelen beschrijven slechts in vrij algemene termen de gebruikte software en parameters, en zwijgen nagenoeg geheel over de hierboven genoemde keuzes en problemen. Dit kan erop wijzen dat de onderzoekers deze keuzes en problemen niet als zodanig ervaren. De zwijgzaamheid op dit punt kan ook daaruit voortkomen dat in andere disciplines ook niet alle details worden vermeld. Dat kan ten dele verklaard worden doordat er in de meeste disciplines veel *knowhow* aanwezig is waardoor een ervaren onderzoeker de niet vermelde aspecten van het onderzoek kan invullen: deze zijn geïmpliceerd. Zo zal bij een verslag van (bio)chemisch onderzoek niet worden vermeld dat alle apparatuur naar behoren is schoongemaakt en dat *samples* ook zodanig zijn behandeld dat zij niet verontreinigd zijn: dat is vanzelfsprekend.⁴⁷ Maar voor data science-technieken is er nog geen algemeen gedeelde *knowhow* of *best practice*.

Bij bepaalde disciplines is zulke kennis uitbesteed aan bijvoorbeeld laboranten (*lab technicians*). Programmeurs in data science zijn echter niet vanzelfsprekend opgeleid in een wetenschappelijk verantwoorde manier van softwareontwikkeling en databeheer.

46 Vgl. het onderzoek beschreven in de bijdrage van Vols in deze bundel.

47 Als het toch mis blijkt te zijn is dat aanleiding voor terugtrekking van het artikel. Zie de geruchtmakende terugtrekking van een artikel door Science, www.sciencemag.org/news/2011/12/updated-rare-move-science-without-authors-consent-retracts-paper-tied-mouse-virus.

Daarnaast ligt een deel van de keuzes besloten in software. Er is de mogelijkheid dat de software programmeerfouten bevat, waardoor er onjuiste uitkomsten kunnen resulteren.⁴⁸ Dit pleit ervoor dat de gebruikte software beschikbaar wordt gesteld. Bij de onderzoeken beschreven in het boek Livermore en Rockmore getiteld *Law as Data* lijkt dat niet het geval te zijn. Overigens is het in het bredere data science-onderzoek al wel gangbaar dat veel data science-software op open source-basis beschikbaar wordt gesteld. Dat is echter nog niet genoeg als in een publicatie niet ook duidelijk wordt gemaakt welke versie van de software er gebruikt is. Om eenzelfde reden valt te bepleiten dat de gebruikte data, afhankelijk van het soort onderzoek, ook beschikbaar moet zijn voor controle. De universiteiten werken op dit moment ook aan data *repositories* waarbij in de toekomst standaard de gebruikte data kunnen worden opgeslagen. Het is hierbij echter ook weer van belang dat regels omtrent dataopslag met voldoende oog voor de details van het onderzoek worden toegepast. Voor dit artikel is het bijvoorbeeld niet van belang dat de concrete data en software toegankelijk is, omdat er geen conclusies aan de data zelf worden verbonden.⁴⁹ Het is ook niet nodig dat voor ieder onderzoek over rechtspraakdata de complete dataset van vele gigabytes wordt opgeslagen.

5 METHODOLOGISCHE AANBEVELINGEN

Gelet op dit alles lijkt het van belang dat er een methodologisch debat plaatsvindt waarbij de hierboven genoemde punten worden geadresseerd. Ik denk hierbij met name aan de volgende punten:

- Hoe gaan we om met ‘fouten’ in de brondata? Mogen deze worden gecorrigeerd of niet? Hoe geef je aan dat je fouten hebt gecorrigeerd?
- Hoe verzekeren we integriteit van de gebruikte data: juistheid, volledigheid, representativiteit?
- Wat zijn *best practices* voor versiebeheer van software en data(bestanden)?
- Wat zijn *best practices* voor openbaarheid van software en data(bestanden)?
- Is het nodig dat alle auteurs van een artikel kennis hebben van de gebruikte software en data science-technieken? Het is niet altijd duidelijk of de combinatie van juridische en technische kennis daadwerkelijk tot betrouwbare resultaten leidt: controleert de jurist de data scientist daadwerkelijk inhoudelijk?

48 Zie voor een voorbeeld hiervan in de chemie: J. Bhandari Neupane *et al.*, ‘Characterization of Leptazolines A-D, Polar Oxazolines from the Cyanobacterium *Leptolyngbya* sp., Reveals a Glitch with the “Willoughby-Hoye” Scripts for Calculating NMR Chemical Shifts’, *Organic Letters* 2019 (21) 8449.

49 De fouten in de data die hier zijn besproken zijn alle onderbouwd met een ECLI-verwijzing waardoor deze op rechtspraak.nl kunnen worden gecontroleerd.

ERIC Tjong Tjin Tai

- Zou er een opleiding voor wetenschappelijke data science-ontwikkelaars nodig zijn, analoog aan opleidingen voor laborant, en welke eisen zouden daaraan gesteld moeten worden?
- Weten we wanneer rechtspraak.nl de data heeft veranderd? Wat *zijn* de data waarmee we werken, zodat de analyse reproduceerbaar wordt.

Op dit moment lijkt de methodologie van juridisch data science-onderzoek nog in de kinderschoenen te staan. Dit artikel is een poging om de discussie een stap verder te brengen.

9 SYSTEMATISCHE JURISPRUDENTIEANALYSE ALS UITDAGING VOOR ONDERWIJS EN ONDERZOEK

Frederik Peeraer en Rob van Gestel

1 INLEIDING

Rechtswetenschappers die via juridisch-dogmatisch onderzoek analyseren wat het geldende recht is, of wat het idealiter zou moeten zijn,¹ worden in toenemende mate gedwongen hun methoden en technieken te expliciteren.² Dit houdt onder andere verband met het feit dat multidisciplinair samengestelde commissies die onderzoekssubsidies toekennen van alle wetenschappers verlangen dat wordt uitgelegd hoe zij denken hun onderzoeksvragen op een betrouwbare en systematische wijze te kunnen beantwoorden. Daar komt bij dat momenteel het empirisch-juridische onderzoek zowel in de lage landen als daarbuiten in opkomst is.³ Bij dit type onderzoek is het traditiegetrouw zo dat onderzoekers veel meer oog hebben voor zaken als systematische dataverzameling, selectie van bronnen ('*sampling*'), verantwoording van methoden en technieken en (on)mogelijkheden met betrekking tot de generalisatie van onderzoeksbevindingen.

Tegelijkertijd is er ook kritiek op het empirisch-juridische onderzoek, mede omdat het te weinig zou aansluiten bij de normatieve vragen waarmee in de rechtspraktijk geworsteld

* Prof. dr. F. Peeraer is doctor-assistent Rechtsmethodologie aan de UGent en docent Juridische argumentatieleer aan de Universiteit Antwerpen. Prof. dr. R.A.J. van Gestel is hoogleraar theorie en methode van wetgeving aan de Universiteit van Tilburg en hoogleraar methodologie van rechtswetenschappelijk onderzoek aan de KU Leuven.

1 Volgens J. Smits, *Omstreden rechtswetenschap*, Den Haag: Boom Juridische uitgevers 2009, p. 28, is de kern van de juridisch-dogmatische methode dat de beoefenaars ervan zich niet alleen bezighouden met de beschrijving van het geldende recht, maar ook met hoe dat recht behoort te luiden.

2 Zie daarover uitgebreid R.A.J. van Gestel, 'Revitalisering van juridisch-dogmatisch onderzoek: expliciteren van het impliciete', *NJB* 2021, p. 286 e.v.

3 W.H. van Boom, P. Desmet & P. Mascini (red.), *Empirical Legal Research in Action. Reflections on Methods and their Applications*, Cheltenham: Edward Elgar 2018, p. 2. Zie echter ook: G. van Dijck, S. Sverdlov & G. Buck, 'Empirical Legal Research in Europe: Prevalence, Obstacles, and Interventions', *Erasmus Law Review* 2018, p. 105 e.v.

wordt en bovendien te veel gericht zou zijn op het meten van effecten van juridische interventies zonder oog te hebben voor theorievorming die de kennis over de ontwikkeling van het recht verder brengt. Zo is het goed om te weten dat de rechterlijke macht evenals nagenoeg iedere beroepsgroep te kampen heeft met allerlei ‘biases’ waar het gaat om de eigen oordeelsvorming,⁴ maar de vraag is uiteindelijk wat de rechtspraak hier precies mee kan, hoe vooringenomenheid voorkomen of beperkt kan worden in individuele zaken en wat dit zou moeten betekenen voor bijvoorbeeld de opleiding van rechters of de verantwoording van hun methoden van rechtsvinding.

Voor het antwoord op laatstgenoemde vragen blijft de verbinding met juridisch-dogmatisch onderzoek essentieel. Daar komt bij dat zonder degelijke handboeken, annotaties, grondwetscommentaren en kronieken, de ontsluiting van het geldende nationale recht dat in toenemende mate verweven is met Europees en internationaal recht achterop zal geraken. De vraag is bovendien of het vruchtbaar is om empirisch-juridisch en juridisch dogmatisch onderzoek tegenover elkaar te plaatsen en te vergelijken. Davies stelt terecht dat goed juridisch-dogmatisch onderzoek onmisbaar is voor empiristen, alleen al om de juiste vragen te kunnen stellen waarop de praktijk graag antwoorden zou willen zien. Omgekeerd is het de vraag of klassiek opgeleide doctrinalisten niet wat zouden kunnen leren van de systematische wijze waarop empirici vaak hun onderzoek verrichten en vooral van de wijze waarop zij hun methoden en technieken verantwoorden.

Deze bijdrage gaat over dit laatste en betreft één specifieke vorm van jurisprudentieanalyse die als een poging kan worden gezien om het goede van beide typen van onderzoek (juridisch-dogmatisch en empirisch) met elkaar te verenigen, namelijk gestructureerde inhoudsanalyse toegepast op doorgaans grotere hoeveelheden rechterlijke uitspraken in plaats van, of in aanvulling op, meer klassieke vormen van tekstuele interpretatie van individuele arresten.

2 OPZET

Het gebruik van ‘*structured content analysis*’ in het kader van de bestudering van jurisprudentie is een vrij recent fenomeen. De invloedrijke bijdrage van Hall en Wright⁵ is voor

4 Vgl. o.a. C. Guthrie, J.J. Rachlinski & A.J. Wistrich, ‘Judging by Heuristic: Cognitive Illusions in Judicial Decision Making’, *Judicature* 2002, p. 44-50; B. Barry, *How Judges Judge. Empirical Insights into Judicial Decision-Making*, Oxon: Informa law from Routledge 2021.

5 M.A. Hall & R.F. Wright, ‘Systematic Content Analysis of Judicial Opinions’, *California Law Review* 2008, p. 63-122. Deze auteurs beschrijven systematische jurisprudentieanalyse op p. 64 als: ‘a scholar collects a

veel auteurs van deze bundel een belangrijke gids geweest,⁶ omdat het artikel een duidelijk overzicht biedt van zowel de kansen als de beperkingen van deze methode en bovendien een reeks praktische tips geeft waar het gaat om keuzes rond zaken als: selectie van uitspraken, coderen, testen van de betrouwbaarheid van het codeerproces en het gebruik van statistische analysetechnieken om bijvoorbeeld trends in de jurisprudentie vast te stellen. In deze bijdrage focussen we echter op meer kwalitatieve⁷ systematische jurisprudentieanalyse, aangezien die het dichtst staat bij traditioneel (interpretatief) juridisch-dogmatisch onderzoek.⁸ We leggen eerst kort uit wat deze vorm van jurisprudentieanalyse inhoudt en waarin die verschilt van traditioneel jurisprudentieonderzoek (par. 3). Daarbij verdedigen we dat beide vormen van jurisprudentieonderzoek niet als tegenpolen moeten worden gezien, maar dat het gaat om verschillende typen onderzoek, waarbij de ene vorm niet beter is dan de andere, maar geschikter is voor een ander doel. Systematische jurisprudentieanalyse stelt de onderzoeker kortgezegd in staat om heel gestructureerd en met precisie patronen te ontdekken in grote hoeveelheden jurisprudentie, terwijl traditionele interpretatieve vormen van jurisprudentieanalyse meer ruimte bieden om bij individuele uitspraken de diepte in te gaan en daarbij vrijer te associëren.

Na de uiteenzetting van dit onderscheid gaan we in op de meer algemene kwaliteitsmaatstaven die hebben te gelden voor juridisch-dogmatisch onderzoek (par. 4) om te laten zien dat systematische jurisprudentieanalyse uit de aard van de methode wat gemakkelijker aan eisen van verantwoording, evenwichtig brongebruik en neutraliteit kan voldoen, juist vanwege het feit dat deze onderzoekers dwingt keuzes in het onderzoek nadrukkelijker te expliciteren. De vraag is of doctrinalisten die traditioneel jurisprudentieonderzoek doen daar wellicht niet iets van zouden kunnen leren. Omgekeerd zullen we laten zien dat systematische jurisprudentieanalyse zeker geen Haarlemmerolie is die iedere vorm van traditioneel interpretatief jurisprudentieonderzoek overbodig maakt.

set of documents, such as judicial opinions on a particular subject, and systematically reads them, recording consistent features of each and drawing inferences about their use and meaning’.

- 6 Inhoudsanalyse zelf als methode dateert al van de jaren vijftig en vindt haar oorsprong in onderzoek naar massacommunicatie: B. Berelson, *Content Analysis in Communication Research*, Glencoe (Illinois): The Free Press 1952. Ook in het (Amerikaanse) recht werd inhoudsanalyse al meermaals toegepast in de twintigste eeuw, maar voor het artikel van Hall en Wright ontbrak een ruimer overzicht.
- 7 Zie voor de bespreking van enkele technieken voor kwantitatieve systematische jurisprudentieanalyse de bijdragen van Dyevre, Vols en Wolters in deze bundel. De mogelijkheden en beperkingen van kwantitatieve analyse van jurisprudentie (althans voor de bestudering van geldend recht) staan centraal in Wolters’ bijdrage. Voor de detectie van wijzigingen in jurisprudentie in het bijzonder, zie A. Dyevre, ‘The promise and pitfall of automated text-scaling techniques for the analysis of jurisprudential change’, *Artificial Intelligence and Law* 2020, p. 1 e.v.
- 8 Zie Hall & Wright 2008, p. 64: ‘This method comes naturally to legal scholars because it resembles the classic scholarly exercise of reading a collection of cases, finding common threads that link the opinions, and commenting on their significance.’

Hierna schakelen we naar de mogelijke betekenis van systematische jurisprudentieanalyse voor het rechtenonderwijs (par. 5). Daarvoor geldt dat de opkomst van empirisch-juridisch onderzoek een ‘boost’ lijkt te kunnen gaan geven aan het methodenonderwijs in met name onderzoeksmasters en PhD-trainingsprogramma’s. Het gevaar daarvan is echter dat er geleidelijk een tweestromenland ontstaat waarbij het onderwijs voor deze onderzoekersopleidingen loszingt van het reguliere vaardighedenonderwijs. Wij denken juist dat systematische jurisprudentieanalyse mogelijk een brug kan slaan tussen methoden van juridisch-dogmatisch onderzoek en empirisch onderzoek. Daarmee houdt het aspirant-wetenschappers in contact met normatief onderzoek naar het geldende recht en laat het juristen die voorbereid worden op de rechtspraak juist eens kennismaken met een andere methode van onderzoek waardoor ze ook bepaalde basiskennis kunnen opdoen van een quasi-empirische methode, die inzicht biedt in het nut van het expliciteren van keuzes en verantwoording van de eigen werkwijze.

Om systematische jurisprudentieanalyse tot een succes te maken in het onderzoek zal er echter niet enkel iets moeten gebeuren aan het (methoden)onderwijs. Er zijn ook uitdagingen voor het onderzoek zelf. Op één daarvan, ontsluiting van data en aandacht voor databeheer, gaan we daarom wat dieper in (par. 6). Op de lange duur zal systematische jurisprudentieanalyse enkel een succes kunnen blijven wanneer bijvoorbeeld datasets ook beschikbaar blijven voor nieuw onderzoek en ook door derden controles kunnen worden uitgevoerd op het primaire onderzoek om te kijken of zich daarbij geen grove fouten of nalatigheden hebben voorgedaan. We ronden af met het schetsen van enkele denkbare toekomstscenario’s met betrekking tot de wijze waarop systematische jurisprudentieanalyse zich in de toekomst zou kunnen gaan ontwikkelen en hoe deze scenario’s passen binnen de bredere ontwikkeling van het juridisch-dogmatische onderzoek van de toekomst (par. 7).

3 WAT IS SYSTEMATISCHE JURISPRUDENTIEANALYSE?

‘*Structured content analysis*’ is een methode waarbij men systematisch en op repliceerbare wijze de inhoud van diverse tekstvormen, zoals wetgeving, jurisprudentie en transcripties van interviews analyseert.⁹ Het is een methode die, zoals de bijdragen in deze bundel illustreren, gebruikt kan worden voor zowel kwantitatief als kwalitatief onderzoek.¹⁰ Het is een

9 S. Stemler, ‘An Overview of Content Analysis’, *Practical Assessment, Research, and Evaluation* 2000, p. 1 e.v.; M.A. Hall, *Coding Case Law for Public Health Law Evaluation*, Robert Wood Johnson Foundation, Princeton NJ, 2011, p. 3. Zie ook de inleiding van Verbruggen in deze bundel.

10 Voor inleidende werken op inhoudsanalyse, zie o.m. K. Krippendorff, *Content Analysis: An Introduction to Its Methodology*, Los Angeles: Sage 2004; C.M. Oldfather, J.P. Bockhorst & B.P. Dimmer, ‘Triangulating Judicial Responsiveness: Automated Content Analysis, Judicial Opinions, and the Methodology of Legal

aanvulling op – en dus geen vervanging van – traditionele vormen van jurisprudentieonderzoek, waarbij tekstanalytische interpretatiemethoden worden gebruikt en vaak ook naar ‘lijnen’ in de jurisprudentie wordt gezocht met behulp van een doorgaans beperktere selectie van uitspraken.¹¹

Door gestructureerd met behulp van coderingstechnieken en computersoftware te analyseren wat rechters doen en dat systematisch in kaart te brengen, focust systematische jurisprudentieanalyse op de ‘breedte’ van het onderzoeksobject en niet zozeer op zijn ‘diepte’.¹² In die zin is het enigszins vergelijkbaar met het onderscheid tussen enquêteonderzoek en het gebruik van interviews. Beide zijn op zichzelf geschikte methoden om informatie op te halen die in het bezit van ‘respondenten’ is, maar met enquêtes is het mogelijk om grote hoeveelheden opinies op een relatief oppervlakkige manier te analyseren, terwijl men bij het afnemen van interviews juist individueel de diepte in kan gaan. Afhankelijk van het doel van het onderzoek is de ene of de andere methode meer geschikt. Wie bijvoorbeeld wil weten of rechtswetenschappers de voorkeur geven aan *peer review* bij het beoordelen van hun eigen onderzoek boven het gebruik van citatieanalyse is gebaat bij een elektronische enquête die per e-mail aan alle rechtswetenschappers in het land kan worden gestuurd.¹³ Wie daarentegen te weten wil komen wat voor type dilemma’s rechtswetenschappers in verschillende rechtsgebieden ervaren bij het zoeken, selecteren, analyseren en synthetiseren van bronnen en hoe zij daarmee omgaan, zal eerder bij interviews uitkomen.¹⁴

Zo is het ook met systematische jurisprudentieanalyse versus het gebruik van klassieke juridische interpretatiemethoden. Eerstgenoemde methode is vooral geschikt om door middel van codering van tekstfragmenten op systematische wijze hypothesen te toetsen of (on)verwachte patronen te ontdekken – vaak met behulp van computersoftware, zoals NVivo, MAXqda of Atlas.ti – in grote hoeveelheden jurisprudentie. Daarbij wordt dan zoveel mogelijk voorkomen dat onderzoekers door cognitieve beperkingen, zoals geheugenstoornis, preconcepties of bepaalde vormen van *bias*, zaken over het hoofd zien. Niet-

Scholarship’, *Florida Law Review* 2012, p. 1189 e.v.; J. Saldaña, *The Coding Manual for Qualitative Researchers*, Los Angeles: Sage 2013; J. Drisko & T. Maschi, *Content Analysis*, Oxford: Oxford University Press 2015; K. Neuendorf, *The Content Analysis Guidebook*, Los Angeles: Sage 2017 en U. Kuckartz, ‘Qualitative Text Analysis: A Systematic Approach’, in: G. Kaiser & N. Presmeg, *Compendium for Early Career Researchers in Mathematics Education*, Cham: Springer International Publishing 2019, p. 181 e.v. en de literatuurlijst aan het einde van die bijdrage.

11 Hall & Wright 2008, p. 66; Hall 2011, p. 4.

12 Hall 2011, p. 6.

13 W.H. van Boom & R.A.J. van Gestel, ‘Rechtswetenschappelijk onderzoek: uitkomsten van een landelijke enquête’, *NJB* 2015, p. 1340. Zie voor de volledige enquête: www.njb.nl/umbraco/uploads/2015/5/Van-Boom-en-Van-Gestel-2015.pdf.

14 M. Snel, *Meesters over bronnen*, Den Haag: Boom juridisch 2016.

temin zal men voor het verklaren van wat er achter de gevonden patronen in de jurisprudentie zit toch vaak deels moeten terugvallen op een theoretisch kader en op klassieke juridische interpretatiemethoden waarvoor dan bijvoorbeeld sleuteluitspraken kunnen worden ontleed die een trendbreuk inluiden met betrekking tot de wijze waarop de rechter met bepaalde juridische vragen omgaat. Aanvullende methoden en andere invalshoeken, zoals rechtsfilosofische, rechtssociologische, rechtshistorische, rechtseconomische, rechtsvergelijkende... invalshoeken, kunnen daarbij worden gecombineerd met pure tekstanalytische interpretatie van rechterlijke uitspraken.

Het is bijvoorbeeld één zaak om vast te stellen dat rechters vaststaande rechtspraak plots wijzigen. Zoiets zou met behulp van systematische jurisprudentieanalyse concreet kunnen worden benoemd en gekwantificeerd. Voor een verklaring en een beoordeling van die vaststelling is echter meer nodig. Waarom sommige uitspraken terugkomen op eerdere rechtspraak en andere uitspraken niet, kan immers niet worden nagegaan door het louter in kaart brengen van de rechterlijke praktijk. Daarvoor is het nodig om een theoretisch kader uit te werken, de toonaangevende uitspraken van hoogste gerechtshoven in de diepte te analyseren en de waarden te expliciteren waarmee rechters rekening moeten houden indien zij wensen terug te komen op eerdere rechtspraak (zoals ‘rechtszekerheid’ – wat pleit voor het ongewijzigd continueren van de eerdere rechtspraak – en ‘respect voor regelgevend optreden’ – wat pleit voor de verandering ten opzichte van de eerdere rechtspraak, aangezien de regelgever de nieuwe regel inhoudelijk ‘beter’ acht). Dat alles kan mede geschieden door middel van rechtsvergelijking. Wanneer de in kaart gebrachte rechtspraak tegen die achtergrond wordt geplaatst, kan meer duidelijkheid worden gegeven over de argumenten die een wijziging van vaststaande rechtspraak kunnen verantwoorden, en in welke omstandigheden zulks aanvaardbaar is.¹⁵

3.1 *Behoeftte aan normatieve inbedding*

Systematische jurisprudentieanalyse is geen empirisch onderzoek in de klassieke zin, aangezien het niet gaat om het bestuderen van de invloed van het recht op een fysiek observeerbare werkelijkheid, wat wel gebeurt bij interviews, observaties en experimenten. Het blijft – hoe systematisch en repliceerbaar ook – tekstanalyse.¹⁶ Desalniettemin loopt

15 Zie daarover S. Verstraelen, *Rechterlijk overgangsrecht*, Antwerpen: Intersentia 2015.

16 Wie ‘empirisch’ ruimer interpreteert en bijvoorbeeld slechts vereist om van empirisch onderzoek te spreken dat de onderzoeker gebruikmaakt van data, beschouwt systematische jurisprudentieanalyse mogelijk wel als een empirische onderzoeksmethode. Voor die interpretatie van ‘empirisch’, zie bijv. L. Epstein & A.D. Martin, *An Introduction to Empirical Legal Research*, Oxford: Oxford University Press 2014, p. 3 e.v. In zijn inleidende bijdrage in deze bundel bespreekt Verbruggen bondig in welke mate systematische

deze onderzoeksmethode het risico soortgelijke verwijten te krijgen als die jegens empirisch juridisch onderzoek in het algemeen vaak worden geuit, namelijk: dat het onderzoek te instrumenteel en beleidsgeoriënteerd is en een normatief kader ontbeert.¹⁷

Aangezien systematische jurisprudentieanalyse ‘slechts’ een methode¹⁸ is, staat ze los van enig theoretisch raamwerk of perspectief.¹⁹ Daarom is het temeer van belang dat deze methode niet *l’art pour l’art* wordt gehanteerd, maar wordt ingebed in debatten die gaande zijn in de rechtswetenschappelijke literatuur en bijdraagt aan het dichten van aanwijsbare lacunes in bestaande kennis. Met andere woorden, het is van belang dat systematische jurisprudentieanalyse wordt voorafgegaan door een probleemstelling die duidelijk maakt wat de onderzoeker nu exact wil toevoegen aan de bestaande kennis, welke relevantie dat heeft en waarom deze onderzoeksmethode het meest geschikt is om de onderzoeksvragen te beantwoorden.²⁰ Dat dit mogelijk is, blijkt uit de bijdragen in deze bundel en uit andere voorbeelden, waarin systematische jurisprudentieanalyse kadert binnen een ruimer juridisch-dogmatisch onderzoek.

3.2 Voorbeelden

Wat de bijdragen in deze bundel betreft, verwijzen we in de eerste plaats naar het promotieonderzoek van Wijntjens over de rol die excuses spelen in het medische aansprakelijkheidsrecht. Op grond van de literatuur en klassieke jurisprudentieanalyses werd gevreesd dat het aanbieden van excuses voor medische fouten door artsen door de rechter zou kunnen worden uitgelegd als het erkennen van schuld en aansprakelijkheid. Wijntjens heeft gebruikgemaakt van onder meer systematisch jurisprudentieonderzoek om die hypothese te testen. Uit de jurisprudentieanalyse bleek dat het nauwelijks voorkomt dat aanbieden van excuses leidt tot erkenning van aansprakelijkheid. Daarmee is de juridische drempel voor bijvoorbeeld artsen om dat vaker dan thans te doen verwaarloosbaar. Voor

jurisprudentieanalyse als een ‘empirische’ methode kan worden beschouwd, maar voegt daar (terecht) aan toe dat die discussie niet meteen erg relevant is.

17 Zie G. Davies, ‘The Relationship between Empirical Legal Studies and Doctrinal Legal Research’, *Erasmus Law Review* 2020, p. 6 e.v.

18 Of preciezer, zoals Verbruggen in zijn inleidende bijdrage in deze bundel stelt, een verzamelnaam voor de toepassing van diverse sociaalwetenschappelijke methoden.

19 ‘Content analysis is a method that is independent of theoretical perspective or framework (e.g., grounded theory, phenomenology)’, H. Julien, ‘Content Analysis’, in: L. Given, *The SAGE Encyclopedia of Qualitative Research Methods*, Los Angeles: Sage 2008, p. 120. Vgl. Hall & Wright 2008, p. 88: ‘Because of these methodological limitations, very few opinion-coding projects explicitly aim to evaluate the legal correctness of judicial opinions.’

20 Vols wijst daar ook expliciet op in zijn bijdrage in deze bundel. Zie ook de bijdrage van Tjong Tjin Tai: hij is in soortgelijke zin kritisch over de wetenschappelijke opbrengst van bijvoorbeeld netwerkanalyse in een juridische context.

zover je ook de kleine kans zou willen wegnemen dat excuses aangeboden door een professionele dienstverlener, zoals een verzekeraar, onbedoeld tot erkenning van aansprakelijkheid kunnen leiden, stelt Wijntjens de introductie van een wettelijke uitlegregel voor die expliciet bepaalt dat excuses niet mogen worden opgevat als een erkenning van aansprakelijkheid. Dat zou er volgens haar niet alleen toe kunnen leiden dat vaker excuses worden aangeboden, maar ook dat dit een de-escalerend effect heeft en onnodig juridische procedures kan voorkomen. Op die manier heeft Wijntjens systematische jurisprudentieanalyse dus gebruikt als middel en tussenstap om een ruimere onderzoeksvraag naar de rol en betekenis van excuses in het aansprakelijkheidsrecht te beantwoorden.

Vols doet in zijn bijdrage verslag van meerdere onderzoeksprojecten. Bij die projecten beoogt de systematische jurisprudentieanalyse bij te dragen aan zowel klassiek-juridische theorieën als de toetsing van meer concrete hypothesen. Het 'VENI-project' tracht een dieper inzicht te verschaffen in hoe rechters gebruikmaken van hun ontbindingsbevoegdheid en in juridische begrippen die een rol spelen in het positieve Nederlandse (huur)recht. Dat project tracht tevens te beoordelen of stellingen als 'De kans op ontbinding is groter, als de tekortkoming druggerelateerde overlast betreft' juist of onjuist zijn. Het 'Damocles-project' streeft er onder meer naar te verduidelijken hoe rechters de rechtmatigheid inschatten van de beslissingen van burgemeesters om panden te sluiten en hoe het evenredigheidsbegrip (zoals bedoeld in art. 3:4 Algemene wet bestuursrecht (Awb)) in de praktijk wordt ingevuld. Dat project toetst daarnaast concrete hypothesen die ontleend zijn aan publicaties van Galanter, zoals 'Upperdogs are more likely to win in court than underdogs' (waarbij burgemeesters dan zijn getypeerd als 'upperdogs' en burgers of bedrijven als 'underdogs'). Het 'EHRM-project' draagt op zijn beurt dan weer bij tot een beter inzicht in de factoren die bepalend zijn voor de uitleg van bepalingen van het Europees Verdrag tot bescherming van de rechten van de mens en de fundamentele vrijheden (EVRM). Inzicht in die factoren maakt het mogelijk om in aanzienlijke mate te voorspellen hoe het Europees Hof voor de rechten van de mens (EHRM) zal oordelen in een bepaalde zaak. Dergelijk inzicht heeft vooral praktische relevantie: het zijn met name advocatenkantoren die dat inzicht kunnen verzilveren door aan de hand van dat inzicht sneller en accurater dan voorheen te beslissen of het vanuit financieel oogpunt een goede keuze is om een bepaalde rechtszaak al dan niet aan te spannen. Onrechtstreeks kan dergelijk inzicht ook van grote theoretische waarde zijn, doordat het toelaat de rol en onmisbaarheid van menselijke waardeoordelen bij juridische beslissingen nauwkeuriger te benoemen en te onderbouwen.²¹

21 Bepaalde auteurs beklemtonen terecht de onvervangbaarheid van het menselijke waardeoordeel in dit verband. Zie bijvoorbeeld C. Devins e.a., 'The Law and Big Data', *Cornell Journal of Law and Public Policy* 2017, p. 357 e.v.

Het onderzoek dat Verbruggen in deze bundel beschrijft, kadert binnen ruimer onderzoek naar private regulering. De concrete leemte die dit onderzoek probeert op te vullen, betreft de voorwaarden die het privaatrecht stelt aan handelingen van sportbonden. Talrijke open normen (zoals ‘redelijk en billijkheid’ en ‘maatschappelijke betamelijkheid’) geven burgerlijke rechters weinig concrete richtsnoeren om dergelijke handelingen te beoordelen op hun rechtmatigheid. Hoe dergelijke rechters die handelingen in de praktijk daadwerkelijk beoordelen, kan een systematische jurisprudentieanalyse verhelderen. De resultaten van die analyse kunnen op hun beurt dan weer aanknopingspunten vormen voor verdere theorievorming over de concretisering van open normen.

Ook Goossens, Bollen en Verbeke hebben niet ‘zomaar’ rechtspraak systematisch geanalyseerd. Momenteel wordt de juridische verhouding tussen Belgische feitelijke samenwoners slechts beheerst door het gemene verbintenissen- en vermogensrecht. Of bijzondere regelgeving nodig is en hoe die dan zou moeten luiden, zijn normatieve vragen die men slechts gefundeerd kan beantwoorden door te steunen op empirische gegevens. Dankzij een systematische jurisprudentieanalyse zoals deze komen dergelijke gegevens aan het licht en kan het maatschappelijk debat met meer kennis van zaken worden gevoerd.

Onderzoek dat Vandaele in het kader van haar proefschrift heeft verricht, maakt duidelijk dat systematische jurisprudentieanalyse ook buiten deze bundel al succesvol is toegepast. Zij heeft onderzocht of de invulling van het belangvereiste in het annulatiecontentieux van de Belgische Raad van State compatibel is met het recht op toegang tot een rechter. Deze probleemstelling is ingegeven door literatuur en rechtspraak van het EHRM, waaruit blijkt dat beperkingen op het uit artikel 6 EVRM voortvloeiende recht op toegang tot een rechter duidelijk en voorspelbaar moeten zijn. Dit impliceert concreet dat van (rechts)personen die een bestuursrechtelijke procedure starten niet mag worden verwacht dat zij voldoen aan ontvankelijkheidseisen waarvan men de toepassing niet kan begrijpen of welke men onmogelijk heeft kunnen voorzien bij de start van het geding.

Met het oog op het beantwoorden van de vraag of daarvan in de jurisprudentie van de Raad van State sprake is, heeft Vandaele een systematische jurisprudentieanalyse van zo’n duizend arresten uitgevoerd. Zij komt daarbij tot de volgende bevindingen. Ten eerste krijgen de facetten die de rechtsgeleerde doctrine toeschrijft aan het belangvereiste als ontvankelijkheidsvoorwaarde – namelijk dat men een rechtstreeks, zeker, persoonlijk, legitiem en actueel belang moet hebben – in de praktijk vaak nauwelijks concrete betekenis, hetgeen afbreuk doet aan de voorzienbaarheidseis. Ten tweede blijkt de Raad van State als voorwaarde voor het voldoen aan de eis van belanghebbendheid te verlangen dat er een oorzakelijk verband is tussen de bestreden bestuurshandeling en het nadeel dat de verzoeker ondervindt en dat verzoekers niet reeds genoegdoening hebben verkregen buiten de

bestuursrechtelijke procedure om. Ook deze beide eisen stuiten volgens onderzoekster op problemen waar het gaat om de uitleg die het EHRM geeft aan artikel 6 EVRM, met name op het punt van de duidelijkheid en voorzienbaarheid van ontvankelijkheidseisen.²²

Bovenstaande voorbeelden illustreren dat systematische jurisprudentieanalyse een geschikte methode kan zijn voor hypothesetoetsing. Wijntjens heeft de aanname weerlegd – of minstens in belangrijke mate genuanceerd – dat het aanbieden van excuses voor medische fouten door artsen door de rechter zou kunnen worden uitgelegd als het erkennen van schuld en aansprakelijkheid en daardoor soms onbedoeld kunnen nopen tot schadevergoeding, terwijl een van de bedoelingen van het geven van een prominentere rol aan excuses in het aansprakelijkheidsrecht juist is om juridische conflicten te de-escaleren en vaker in der minne te schikken. In elk van de drie projecten die Vols beschrijft, zijn concrete hypothesen geformuleerd en getoetst. Vandaele heeft op haar beurt aangetoond dat de diverse kenmerken van het belangvereiste die in literatuur worden aangetroffen in de praktijk vaak weinig tot geen betekenis hebben, waarop ze beschrijft hoe die vereisten dan wel worden gehanteerd.

Uit de vermelde voorbeelden blijkt voorts dat op grond van systematische jurisprudentieanalyse alleen nog geen normatieve aanbevelingen kunnen worden gedaan. In het geval van Wijntjens zou een dergelijke aanbeveling kunnen zien op de wijze waarop de rechter in de toekomst zou moeten omgaan met excuses in het aansprakelijkheidsrecht, bijvoorbeeld door de genoegdoeningsfunctie van dit recht op te rekken en niet langer te beperken tot financiële genoegdoening. Hoe rechters gebruik zouden moeten maken van hun ruime ontbindingsbevoegdheid en of rechterlijke beoordelingen van sluitingsbeslissingen van burgemeesters wel naar recht verantwoord zijn, zijn vragen die het ‘VENI’- respectievelijk ‘Damocles-project’ van Vols op zich nog niet beantwoordt. Hetzelfde geldt voor Verbruggens onderzoek naar de normatieve waardering van de rechterlijke toetsing van handelingen van sportbonden: in kaart brengen van wat rechters in de praktijk doen is één ding; of die praktijk strookt met het Nederlandse recht een ander. Goossens, Bollen en Verbeke geven uitdrukkelijk aan dat hun onderzoek de basis vormt voor een normatief debat over de regulering van de rechtsverhouding tussen feitelijke samenwoners. Om van de resultaten van hun onderzoek naar dat debat te gaan, moet nog een ‘normatieve sprong’ worden gezet en moeten die resultaten nog worden gewaardeerd. Vandaele kan op grond van haar analyse weliswaar aantonen dat onderscheidingen die in de doctrine worden gemaakt in de rechtspraak een veel minder prominente rol spelen dan door menigeen wordt gedacht, maar daarmee kan nog niet de vraag beantwoord worden of de wijze waarop de Raad van

22 A.-S. Vandaele, *Procesbelang bij de Raad van State en het recht op toegang tot een rechter*, Brugge: die Keure 2019.

State met de eisen van belanghebbendheid omgaat uiteindelijk ook in overeenstemming zijn met de eisen die het EHRM daaraan stelt op grond van artikel 6 EVRM. Om die reden kadert Vandaele haar systematische jurisprudentieanalyse dan ook binnen meer traditioneel jurisprudentieonderzoek naar wat de sleutelarresten van het EHRM op dat punt precies vereisen.

3.3 *(Te?) hoge eisen op onderzoeksgebied*

De gegeven voorbeelden kunnen de indruk wekken dat systematische jurisprudentieanalyse de lat wel erg hoog legt. De inbedding ervan in de actuele wetenschappelijke discussies en de combinatie ervan met een grondig inzicht in het positieve recht, vergt tijd en middelen die niet steeds beschikbaar zullen zijn.²³ Toch kan systematische jurisprudentieanalyse ook een waardevolle bijdrage leveren wanneer zij niet volledig strookt met dat ideaalbeeld.

In de eerste plaats kan een dergelijke analyse van belang zijn voor vakpublicaties die niet meteen de intentie hebben om iets nieuws toe te voegen aan de stand van het wetenschappelijk onderzoek, maar veeleer beogen de praktijk te informeren over nieuwe inzichten en actuele tendensen, zoals: hoe vaak komt het eigenlijk voor dat feitenrechters een rechtsvormende uitspraak doen of afwijken van jurisprudentie van de hoogste rechter en gebeurt dat thans meer dan in het verleden of op bepaalde rechtsgebieden meer dan gemiddeld? Door te focussen op patronen, tendensen en relaties tussen data, kunnen Goossens, Bollen en Verbeke in eerste instantie conclusies trekken over de soorten conflicten die zich voordoen tussen feitelijke samenwoners en over de rechtsgronden die de rechtspraak daarbij hanteert. Dankzij Verbruggens analyse wordt duidelijk welke argumenten en beginselen zoal gehoor vinden bij rechters die handelingen van sportbonden moeten beoordelen.

In de tweede plaats kunnen dit soort analyses heel goed vragen opleveren die vervolgens om fundamenteel (vervolg)onderzoek schreeuwen. In de zojuist genoemde voorbeelden: hoe komt het bijvoorbeeld dat feitenrechters meer dan in het verleden zelf ook aan rechtsvorming zijn gaan doen en kijken Nederlandse rechters op dit punt af van rechters in omliggende landen? Of is het wellicht zo dat het feit dat lagere rechters de afgelopen decennia steeds vaker de weg naar het Luxemburgse Hof hebben weten te vinden, hen ook in staat heeft gesteld onwelgevallige arresten van de hoogste nationale rechter bij te sturen? Het onderzoek van Goossens, Bollen en Verbeke wijst uit dat de meeste zaken door vrouwen

23 In zijn bijdrage benadrukt Wolters dit als een belangrijke beperking van (kwantitatieve) systematische jurisprudentieanalyse.

worden ingeleid (361 door vrouwen vs. 244 door mannen). Waarom is dat zo? Daarnaast hebben zij vastgesteld dat de feitelijke situatie van partijen (zoals leeftijd van de samenwoners, hun inkomen...) verrassend genoeg amper wordt meegenomen in de rechterlijke beoordeling van de rechtsverhouding tussen feitelijke samenwoners. Welke motieven daarachter schuilgaan, kunnen interviews met de betrokken rechters mogelijk uitwijzen. Verbruggen signaleert dat Nederlandse rechters handelingen van sportbonden veelvuldig toetsen aan rechtsbeginselen zoals 'hoor en wederhoor' en 'motivering'. In welke mate dergelijke – van oorsprong publiekrechtelijke – rechtsbeginselen rechten en plichten kunnen verantwoorden in privaatrechtelijke verhoudingen, blijft een open vraag.

Systematische jurisprudentieanalyse kan met andere woorden ook heel goed een opstapje zijn voor vervolgonderzoek, zeker wanneer er onverwachte resultaten uit naar voren zouden komen.

4 MEERWAARDE VAN SYSTEMATISCHE JURISPRUDENTIEANALYSE VOOR DE KWALITEIT VAN JURIDISCH-DOGMATISCH ONDERZOEK

4.1 *Meerwaarde voor wetenschappelijk onderzoek*

Bijkomend voordeel van systematische jurisprudentieanalyse is dat het kan bijdragen aan het beter expliciteren van de wijze waarop juridisch-dogmatisch onderzoek wordt uitgevoerd. Momenteel treffen we bijvoorbeeld nog heel vaak weinig informatie aan in onderzoek waarbij jurisprudentie wordt geanalyseerd met betrekking tot hoe de onderzochte uitspraken zijn geselecteerd. Of het enkel om gepubliceerde of ook om ongepubliceerde uitspraken gaat, maar ook wat voor analysekader bijvoorbeeld is gebruikt om uitspraken te bestuderen zodat de consistentie van de beoordeling is gegarandeerd.²⁴ Vaak lijken juristen die juridisch-dogmatische methoden toepassen zich nauwelijks bewust van het bestaan van kwaliteitsmaatstaven die voor deze vorm van onderzoek hebben te gelden.

In een recente internationale studie naar kwaliteitsmaatstaven voor juridisch-dogmatisch onderzoek komt Snel op basis van literatuur en empirisch onderzoek niettemin tot drie categorieën van kwaliteitsmaatstaven met negen criteria die hieronder schematisch (vertaald) zijn weergegeven.²⁵

²⁴ Vgl. bijv. H.T.M. Kloosterhuis, *Juridische methoden: casus oplossen, jurisprudentie-analyse en argumenteren*, Den Haag: Boom juridisch 2020 waar het gaat om het gebruik van analyseschema's.

²⁵ Overgenomen uit: M. Snel, 'Making the implicit quality standards and performance expectations for traditional legal scholarship explicit', *German Law Journal* 2019, p. 10.

Tabel 1 Categorieën van kwaliteitsmaatstaven

Onderzoeksdesign	Beantwoording onderzoeksvragen	Publicatie als geheel
Precieze onderzoeksvraag: concrete en precies geformuleerde probleemstelling	Verantwoording: precieze en correcte verwijzing naar bronnen en explicitering en rechtvaardiging van belangrijke keuzes in het onderzoek	Leesbaar: grammaticaal correct, duidelijk en met een logische opbouw en structuur
Originaliteit: vraag indiceert wat het onderzoek beoogt toe te voegen aan de stand van de wetenschap	Accuratesse: argumenten en bronnen moeten daadwerkelijk bijdragen aan het beantwoorden van de onderzoeksvraag(en)	Overtuigend: de argumenten zijn ter zake, logisch opgebouwd, gefocust, logisch, valide en consistent
Relevantie: vraag maakt duidelijk wat de theoretische en maatschappelijke relevantie van het onderzoek is	Evenwichtig: compleet en representatief gebruik van bronnenmateriaal dat recht doet aan het gewicht van verschillende bronnen	Neutraal/eerlijk: de inhoud mag niet beïnvloed worden door niet verantwoorde preconcepties of aannames van de onderzoeker
Methoden: zijn geschikt om de onderzoeksvraag te beantwoorden	Geloofwaardig: alle claims en beweringen van de auteur moeten correct of verdedigbaar zijn	

Op punten als ‘verantwoording’, ‘evenwichtigheid’ en ‘neutraliteit/eerlijkheid’ lijkt systematische jurisprudentieanalyse iets extra’s te bieden voor juridisch-dogmatisch onderzoek, omdat het onderzoekers dwingt aan de buitenwereld duidelijk te maken hoe de analyse precies in z’n werk gaat. Door bijvoorbeeld alle te bestuderen uitspraken te coderen maken onderzoekers duidelijk in welke informatie zij geïnteresseerd zijn en door die keuzes in een codeerboek vast te leggen wordt controleerbaar welke keuzes zijn gemaakt bij het bestuderen van jurisprudentie.²⁶ Verder kan de betrouwbaarheid van de inhoudsanalyse – zeker wanneer bij die analyse meerdere onderzoekers betrokken zijn – worden vergroot door bijvoorbeeld een ‘*intercoder reliability*’ test²⁷ uit te voeren waarmee wordt duidelijk gemaakt in hoeverre zij op dezelfde wijze de teksten coderen en vervolgens door dezelfde lens naar het object van onderzoek kijken.²⁸ Samenwerking tussen codeurs draagt er bovendien toe bij dat onderzoekers elkaar scherp houden.²⁹

26 Wolters verduidelijkt in zijn bijdrage welke problemen (deels) kunnen worden ondervangen door gebruik te maken van een uitgewerkt codeerboek. Beklemtoond moet worden dat het expliciteren van de gemaakte keuzes niet zaligmakend is: een zekere dosis interpretatieve subjectiviteit kan daarmee worden geobjectiveerd, maar volledig ‘objectief’ wordt het nooit: zie C. O’Connor & H. Joffe, ‘Intercoder Reliability in Qualitative Research: Debates and Practical Guidelines’, *International Journal of Qualitative Methods* 2020, p. 4 e.v.

27 Voor een overzicht van de argumenten *pro* en *contra intercoder reliability*, zie O’Connor & Joffe 2020, p. 3 e.v.

28 O’Connor & Joffe 2020, p. 2.

29 C. MacPhail e.a., ‘Process guidelines for establishing Intercoder Reliability in qualitative studies’, *Qualitative Research* 2015, p. 198 e.v.

Systematische jurisprudentieanalyse kan ervoor zorgen dat juridisch-dogmatisch onderzoek zijn normatieve ambities beter kan waarmaken.³⁰ Zo laat het onderzoek van Wijntjens duidelijk zien dat een oproep aan artsen om vaker excuses aan te bieden bij medische fouten niet gehinderd hoeft te worden door angst van ziekenhuizen of zorgverzekeraars voor het erkennen van aansprakelijkheid en het daarmee samenhangende risico van een plicht tot schadevergoeding. De data die Goossens, Bollen en Verbeke hebben vergaard, vormen dan weer de fundamenten voor een normatief debat over de regulering van de rechtsverhouding tussen feitelijke samenwoners.

Dit laatste betekent niet dat de resultaten van dergelijk onderzoek niet langer aangevallen kunnen worden op zowel normatieve als empirische gronden. Zo hebben Geeraets en Maas onafhankelijk van Wijntjens verdedigd dat het loskoppelen van excuses en aansprakelijkheid kan leiden tot een instrumentalisering van (onoprechte) spijtbetuigingen met het oog op de voor daders mogelijk positieve effecten en dat er bovendien geen overtuigend empirisch bewijs zou zijn dat vaker aanbieden van excuses binnen de Nederlandse context leidt tot minder juridische procedures, gelet op onder andere ons afwijkende zorgstelsel en het ontbreken van een claimcultuur.³¹ Dergelijke kritiek zal ongetwijfeld tot vervolgonderzoek leiden, maar het voordeel daarvan ten opzichte van traditioneel juridisch-dogmatisch onderzoek is dat het minder gemakkelijk wordt voor onderzoek om al dan niet bewust vooral op zoek te gaan naar bronnen die het eigen waardeoordeel bevestigen en daarbij tegenargumenten en tegenbewijs dat onwelgevallig is weg te laten.³² Daarbij helpt meer evenwichtig bronnenonderzoek mogelijk vooringenomen standpunten terug te dringen.

4.2 Meerwaarde voor praktijkgericht onderzoek

Gestructureerde en systematische inhoudsanalyse van grote hoeveelheden jurisprudentie kan ook een meerwaarde bieden voor bijvoorbeeld het schrijven van kronieken of overzichten van rechtspraak. In de eerste plaats kan deze methode dergelijke vakpublicaties inhoudelijk rijker en meer divers maken.

30 Zie uitgebreid H. Dagan, R. Kreitner & T. Kricheli-Katz, 'Legal Theory for Legal Empiricists', *Law & Social Inquiry* 2018, p. 292 e.v.; M. Salehijam, 'The Value of Systematic Content Analysis in Legal Research', *Tilburg Law Review* 2018, p. 5; Davies 2020, p. 9 e.v.

31 V. Geeraets & R. Maas, 'Excuses belicht vanuit multidisciplinair perspectief', *NJB* 2020, p. 2463 e.v.

32 Zie daarover ook J.B.M. Vranken, 'Als een arrest niet bevalt, laten wij het gewoon weg', in: A.G. Castermans, J. Hijma, K.J.O. Jansen, P. Memelink, H.J. Snijders & C.J.J.M. Stolker, *Ex libris Hans Nieuwenhuis*, Deventer: Kluwer 2009, p. 77 e.v.

Weten of een besproken uitspraak een anekdotische uitzondering is dan wel exemplarisch is voor de rechtspraak van bepaalde rechters kan voor de praktijk bijzonder relevant zijn en dat geldt temeer wanneer op deze manier bijvoorbeeld onvermoede patronen uit de jurisprudentie naar boven worden gehaald, zoals dat bepaalde rechtscolleges of individuele rechters minder geneigd zijn precedenten te volgen dan andere.³³ Tegelijk kan de meer expliciete methodische benadering ervoor zorgen dat wanneer er gebruik van wordt gemaakt in vakpublicaties, de grondigheid en diepgang van de publicaties meer gewicht krijgen en beter op waarde kunnen worden geschat bij onderzoeksbeoordelingen in de rechtswetenschap.³⁴ Door te verduidelijken wat men precies doet, wordt de uitgevoerde analyse controleerbaar, wat op zijn beurt een indicatie voor de kwaliteit van die analyse kan zijn.

Systematische jurisprudentieanalyse kan echter traditionele interpretatieve analyse niet vervangen.³⁵ Ze kan helpen patronen te identificeren, maar voor het duiden en verklaren van die patronen schiet ze op zichzelf te kort. Daarvoor zijn immers vaak bronnen en inzichten nodig die niet te vinden zijn in de onderzochte uitspraken zelf, maar bijvoorbeeld in de literatuur, waardoor ze nooit met behulp van codering van tekstfragmenten uit arresten naar boven kunnen worden gehaald. Een andere beperking van deze methode is dat zij tijdrovend en arbeidsintensief is en bovendien enige training en oefening vereist, zowel om de data te genereren volgens de regelen der kunst als om diezelfde vergaarde data nadien te duiden met behulp van specifieke softwarepakketten die de patroonherkenning vergemakkelijken.

5 BETEKENIS VOOR HET RECHTENONDERWIJS

Gebruik van systematische jurisprudentieanalyse als onderzoeksmethode zal alleen op grotere schaal ingang vinden indien jonge onderzoekers hierin worden opgeleid. Problematisch is hier dat Europese rechtenfaculteiten traditioneel weinig belangstelling hebben voor onderzoeksmethodologie die verder gaat dan bronnenonderzoek en tekstuele inter-

33 Vgl. het onderzoek van Goossens, Bollen en Verbeke in deze bundel en de vaststelling dat rechters de feitelijke situatie van partijen amper meenemen in hun beoordeling van de rechtsverhouding tussen feitelijke samenwoners.

34 Voor een pleidooi voor de rehabilitatie van de vakpublicatie, zie R. van Gestel & M. Snel, 'Rehabilitatie van de vakpublicatie', *NJB* 2019, p. 3152 e.v. Ook buiten het domein van het recht wordt benadrukt dat inhoudsanalyse zeker niet alleen een onderzoeksfinaliteit heeft: zie Drisko & Maschi 2015, p. 131 e.v.; Neuendorf 2017 in 'Chapter 1', bij 'The Myths of Content Analysis' en specifiek 'Myth #4: Content analysis is for academic use only'.

35 Vols wijst daar ook op, in de 'vijfde les' van zijn bijdrage in deze bundel.

pretatie,³⁶ waarbij bovendien doorgaans het perspectief van de praktizijn (rechter, advocaat, notaris enz.) dominant is.³⁷ Interessant is niettemin dat door de opkomst van empirisch-juridisch onderzoek recentelijk het methodenonderwijs in met name onderzoeksmasters en PhD-trainingprogramma's wel meer aandacht lijkt te krijgen.³⁸

De combinatie van weinig geëxpliciteerde juridisch-dogmatische onderzoeksmethoden voor het reguliere rechtenonderwijs en een afzonderlijke meer empirisch-juridisch georiënteerde focus voor de opleiding van aspirant-rechtswetenschappers kan er evenwel gemakkelijk toe leiden dat er, net zoals eerder in de VS op sommige faculteiten wel is gebeurd, een kloof ontstaat tussen het onderwijs aan studenten en onderzoek door docenten.³⁹ Dat is niet zonder risico, want zonder degelijk positiefrechtelijk onderzoek zal het onderwijs dat nog altijd sterk is gericht op de voorbereiding van studenten op de rechtspraktijk ongetwijfeld verschromelen.⁴⁰ De keerzijde daarvan zou bovendien kunnen zijn dat het rechtswetenschappelijke onderzoek in vooral de meer onderzoeksgelateerde faculteiten steeds verder loszingt van ontwikkelingen in de rechtspraktijk waarvoor de president van het Amerikaanse hooggerechtshof al eens waarschuwde in een interview met de woorden:

'Pick up a copy of any law review that you see, and the first article is likely to be, you know, the influence of Immanuel Kant on evidentiary approaches in 18th Century Bulgaria, or something, which I'm sure was of great interest to the academic that wrote it, but isn't of much help to the bar.'⁴¹

Los van een dergelijke potentiële kloof tussen onderwijs en onderzoek, waarvan voorlopig in de lage landen nog allerminst sprake lijkt, is misschien nog wel belangrijker dat goed empirisch-juridisch onderzoek ook niet zonder input uit juridisch-dogmatisch onderzoek kan. Dat is bijvoorbeeld het geval waar het gaat om het vaststellen van wat de belangrijkste

36 Daarover o.a.: A. Dyevre, 'Fixing European Law Schools', *ERPL* 2017, p. 151 e.v.

37 Zie reeds: R.C. Van Caenegem, *Judges, Legislators, and Professors: Chapters in European Legal History*, Cambridge: Cambridge University Press 1992.

38 E.S. Pannebakker e.a., *Onderweg naar morgen – een inventarisatie van empirische onderzoeksmethoden in juridisch onderwijs*, Platform Empirical Legal Studies juli 2020.

39 B. Newton, 'Preaching What They Don't Practice: Why Law Faculties' Preoccupation with Impractical Scholarship and Devaluation of Practical Competencies Obstruct Reform in the Legal Academy', *South Carolina Law Review* 2010, p. 108 e.v.

40 Zie daarover voor de VS, R. Posner, *Diverging paths: The academy and the judiciary*, Cambridge (Ma.): Harvard University Press 2016 en M.J.A.M. Ahsmann, 'Civiel effect: nut en noodzaak. Met enkele rechtsvergelijkende beschouwingen over de initiële en postinitiële opleiding', in: *Handelingen NJV*, Deventer: Wolters Kluwer 2018, p. 15 e.v.

41 Zie: www.abajournal.com/news/article/law_prof_responds_after_chief_justice_roberts_disses_legal_scholarship.

problemen zijn in de rechtspraak en welke systematische kennis bijvoorbeeld nodig is om bepaalde op normatief onderzoek gebaseerde aannames van rechters of wetgevers onderuit te halen.

Een voorbeeld op dit punt is de (impliciete) aanname bij het Europese Hof van Justitie dat nationale rechters die prejudiciële vragen stellen aan het Hof vervolgens de beslissingen van het Luxemburgse rechtcollege ook automatisch naleven. Eerder systematisch jurisprudentieonderzoek laat evenwel zien dat dit zeker niet automatisch het geval is en dat het Hof daarop nauwelijks toezicht houdt.⁴²

Systematische jurisprudentieanalyse kan fungeren als een brug tussen empirisch-juridisch en juridisch-dogmatisch onderzoek.⁴³ Voor empiristen heeft zij tot voordeel dat zij ‘slechts’ een methode is waarmee rechtspraak kan worden geanalyseerd en waarbij derhalve de normatieve uitspraken van rechters leidend zijn. Het helpt empiristen met andere woorden om geworteld te blijven in het geldende recht.⁴⁴ Voor doctrinalisten die later de rechtspraak in gaan, vormt systematische jurisprudentieanalyse een toegangspoort tot andere manieren om het recht te bestuderen dan de traditionele dogmatische methoden.⁴⁵ Het maakt juristen duidelijk waarom het van belang is om keuzes te expliciteren en de eigen werkwijze te verantwoorden: als een systematische jurisprudentieanalyse niet uitdrukkelijk stapsgewijs wordt toegelicht, kan men totaal andere resultaten krijgen zonder dat verklaarbaar is waarom.

6 DATATOEGANG EN DATADELING: UITDAGINGEN VOOR SYSTEMATISCHE JURISPRUDENTIEANALYSE

Zelfs als we onderzoekers vertrouwd kunnen maken met het correct uitvoeren van systematische jurisprudentieanalyse, zijn er (in ieder geval momenteel nog) praktische obstakels die verhinderen dat deze methode haar volle potentieel kan ontplooiën. We beperken ons

42 R.A.J. van Gestel & J.C.A. de Poorter, *In the Court we Trust: Cooperation, Coordination and Collaboration between the European Court of Justice and Supreme Administrative Courts*, Cambridge: Cambridge University Press 2019, p. 177-178.

43 Hoe die brug tussen empirisch-juridisch en juridisch-dogmatisch onderzoek in het rechtenonderwijs moet worden geïmplementeerd, spreekt niet voor zich: dat hangt af van fundamentele beleidskeuzes met betrekking tot de doelstelling van rechtenonderwijs en de concrete vormgeving ervan. In par. 7 van dit hoofdstuk schetsen we enkele toekomstscenario's.

44 Zie voor de verschillende doelstellingen van juridisch-dogmatisch onderzoek J. Smits, 'What is Legal Doctrine? On The Aims and Methods of Legal-Dogmatic Research', in: R. van Gestel, H. Micklitz & E. Rubin, *Rethinking Legal Scholarship: A Transatlantic Dialogue*, New York: Cambridge University Press 2017, p. 213 e.v.

45 Zie Davies 2020, p. 9 e.v.

hier tot de uitdagingen op het gebied van datatoegang in Nederland (par. 6.1) en in België (par. 6.2) en op het gebied van datadeling in het algemeen (par. 6.3). Wat betreft de uitdagingen op het gebied van datakwaliteit (in het kader van AI-technieken), verwijzen we naar de bijdrage van Tjong Tjin Tai in deze bundel.

6.1 *Datatoegang in Nederland*

In een ideale wereld is alle jurisprudentie online toegankelijk. In de echte wereld wordt slechts een beperkt deel van alle uitspraken gepubliceerd. In Nederland zijn er jaarlijks rond de 1,7 miljoen uitspraken.⁴⁶ De criteria die bepalen welke daarvan worden gepubliceerd op de online databank rechtspraak.nl zijn neergelegd in het Besluit selectiecriteria uitspraken databank Rechtspraak.nl uit 2012.⁴⁷ Dat besluit gaat uit van het in beginsel publiceren van alle uitspraken van de hoogste rechters. Voor de lagere rechters geldt dat publicatie afhankelijk is van kort gezegd het (potentiële) belang van uitspraken. Bepaalde categorieën van uitspraken van lagere rechters, zoals die waarin een beroep wordt gedaan op Europees recht, worden standaard gepubliceerd. Daarbij is anonimisering voorgeschreven, hetgeen nog altijd handmatig dient te geschieden en daardoor arbeidsintensief is.⁴⁸ Uitspraken waarvoor mediabelangstelling bestaat, worden doorgaans gepubliceerd op de dag van de uitspraak, andere 'zo veel mogelijk' binnen een maand.

Ondanks het feit dat behoorlijk veel uitspraken online worden gepubliceerd is er veel kritiek op het beleid inzake openbaarmaking.⁴⁹ De kritiek betreft niet enkel het feit dat er meer gedaan kan worden om onder andere kenbaarheid, vindbaarheid en duidelijkheid van de informatie te bevorderen.⁵⁰ Van meer principiële aard is de kritiek dat het uiteindelijk de rechterlijke macht zelf is die de selectie van te publiceren rechterlijke uitspraken uitvoert en dat wie toegang tot ongepubliceerde jurisprudentie wenst daarvoor medewerking van de Raad voor de rechtspraak nodig heeft die onderzoekers vervolgens aan een toetsingsprocedure onderwerpt waarin allerlei vragen worden gesteld omtrent zaken als de institutionele inbedding van het onderzoek, de vraagstelling, de theoretische en/of beleidsmatige context, de methoden van onderzoek en de relevantie van het onderzoek voor de recht-

46 Deze informatie is ontleend aan: www.njb.nl/blogs/wettelijke-regeling-publicatie-rechterlijke-uitspraken/.

47 De databank is te vinden via: www.rechtspraak.nl/. Voor het Besluit selectiecriteria, zie: www.rechtspraak.nl/Uitspraken/Paginas/Selectiecriteria.aspx.

48 Voor de anonimiseringsrichtlijnen. Zie: www.rechtspraak.nl/Uitspraken/Paginas/Anonimiseringsrichtlijnen.aspx.

49 Zie W.J. Zwalve & C.J.H. Jansen, *Publiciteit van Jurisprudentie*, Deventer: Kluwer 2013, p. 222 e.v. en M. van Opijnen, *Op en in het Web. Hoe de toegankelijkheid van rechterlijke uitspraken kan worden verbeterd*, Den Haag: Boom Juridische uitgevers 2014.

50 M.M. Groothuis, 'De digitale overheid en de Awb. Bestuursrechtelijke aspecten van elektronische communicatie', in: *VAR-Preadviezen* (Reeks 146), Den Haag: Boom Juridische uitgevers 2011, p. 7 e.v.

spraak, die direct aan de academische vrijheid raken.⁵¹ De vraag is bovendien hoe een dergelijke toetsingsprocedure zich verhoudt tot wettelijke openbaarheidsbepalingen, zoals die van artikel 8:79 lid 2 Awb, waaruit volgt dat anderen dan procespartijen afschriften of uittreksels van de uitspraak of van het proces-verbaal van de mondelinge uitspraak kunnen verkrijgen mits daarbij een wettelijke vergoeding voor de kosten van onder andere anoniemering wordt betaald.

Voor het doel van deze bijdrage is evenwel relevanter dat systematische jurisprudentieanalyse vaak niet goed mogelijk zal zijn juist vanwege het feit dat bij de publicatie van jurisprudentie een selectie wordt gemaakt. Marseille en Wever hebben er terecht op gewezen dat bijvoorbeeld verwacht mag worden dat gegronde beroepen vaker gepubliceerd zullen worden dan ongegronde, omdat gegronde beroepen doorgaans van groter praktisch belang zijn voor burgers.⁵² Tegelijkertijd maakt dit dat een systematische analyse van een representatieve dwarsdoorsnede van gepubliceerde uitspraken bij voorbaat onmogelijk is. Marseille en Van Ettekoven hebben bovendien terecht vastgesteld dat zolang niet alle jurisprudentie online gepubliceerd wordt elektronische ‘*data mining*’ en ‘*machine learning*’ eigenlijk niet mogelijk zijn.⁵³ Zoals onder meer Verbruggen, Vols, Wijntjens en Wolters in hun bijdragen aan deze bundel beklemtonen, geldt dat uiteraard ook voor ‘*structured content analysis*’. Door wijziging van de technische mogelijkheden zouden, volgens Marseille en Van Ettekoven, voorheen gebruikte argumenten voor beperkte publicatie, zoals kosten en privacyproblemen, achterhaald zijn. Barkhuysen heeft in het verlengde daarvan gepleit voor een regeling bij formele wet van de openbaarheid van rechterlijke uitspraken. Dat heeft ten eerste als voordeel dat de beslissing om uitspraken wel of niet te publiceren of ter beschikking te stellen, wordt weggehaald bij de rechterlijke macht die op dit punt geen rechter in eigen zaak zou moeten willen zijn. Ten tweede wordt op die manier ook een wettelijke garantie gecreëerd voor de privacybescherming van bij de uitspraken betrokken partijen.⁵⁴ Tegelijk kan bovendien ook een regeling worden getroffen met betrekking tot de openbaarheid van onderliggende processtukken die geen integraal onderdeel uitmaken van de uitspraak, maar desalniettemin buitengewoon interessant en relevant kunnen zijn voor fundamenteel onderzoek. Denk aan pleitnota’s, deskundigenadviezen, *amicus briefs* enzovoort.⁵⁵ Barkhuysen gaat zelf nog verder door te bepleiten dat ook metadata openbaar

51 Zie: A.T. Marseille & M. Wever, ‘Het verstrekken van niet-gepubliceerde uitspraken door de Rechtspraak: Gunst of recht?’, *NJB* 2019, p. 123-124.

52 Marseille & Wever 2019, p. 123.

53 A.T. Marseille & B.J. van Ettekoven, ‘Afscheid van de klassieke procedure in het bestuursrecht?’, in: *Afscheid van de klassieke procedure?* (Handelingen NJV), Deventer: Wolters Kluwer 2017, p. 259 e.v.

54 Zie daarover recent M. van Opijnen, ‘Alle rechtspraak online? Hoe dan? Noodzakelijke ingrediënten voor een wettelijke regeling’, *Ars Aequi* 2021/127.

55 T. Barkhuysen, ‘Een wettelijke regeling voor publicatie van rechterlijke uitspraken op rechtspraak.nl’, *NJB* 2017, p. 1649.

gemaakt zouden moeten worden en dat rechterlijke instanties bovendien verplicht zouden moeten worden uitspraken op een manier te classificeren die van belang is voor de rechtsontwikkeling en de rechtswetenschap. Daarbij gaat het bijvoorbeeld om de vraag of met een uitspraak een nieuwe lijn in de jurisprudentie is ingezet of wordt afgeweken van hogere rechtspraak. De behoefte aan een dergelijke classificatie zou enkel groter worden wanneer voortaan alle jurisprudentie online beschikbaar wordt gesteld, omdat het dan na verloop van tijd ook steeds belangrijker wordt om in de enorme berg data ook nog steeds relevante individuele uitspraken snel te kunnen vinden.⁵⁶

6.2 *Datatoegang in België*

Voor de situatie in België geldt dat deze nog aanzienlijk nijpender is. Daar wordt de jurisprudentie van het Belgische Grondwettelijk Hof en (sinds 1994) van de Belgische Raad van State openbaar gemaakt. De uitspraken van de rechterlijke orde (Hof van Cassatie en feitenrechtspraak) zijn daarentegen slechts zeer beperkt (vlot) toegankelijk.⁵⁷ Hoewel de Belgische wetgever met een wet van 5 mei 2019⁵⁸ heeft aangekondigd dat vonnissen en arresten voortaan op geanonimiseerde wijze integraal zullen worden opgenomen in een voor het publiek toegankelijke, elektronische databank van vonnissen en arresten van de rechterlijke orde, is de inwerkingtreding van die wet uitgesteld, tot in elk geval 1 september 2021.⁵⁹

Dit laatste maakt dat momenteel in België nog veel meer jurisprudentie dan in Nederland niet elektronisch toegankelijk is. Het onderzoek van Goossens, Bollen en Verbeke in deze bundel illustreert hoe moeilijk uitspraken van Belgische feitenrechters kunnen worden verkregen. Zij hebben eerst zelf 33 hoven en rechtbanken aangeschreven met het verzoek of die toegang wilden verlenen tot ongepubliceerde rechtspraak over de vermogensrechtelijke gevolgen van (affectieve) feitelijke samenwoning, uit de periode september 2014 tot en met september 2018. Uiteindelijk hebben alle vijf de hoven van beroep en vijftien rechtbanken toegang verleend. In de wetenschap dat dit verzoek mede afkomstig was van een gevestigde naam in het Belgische familiaal vermogensrecht als Verbeke, stemt dat

56 In die zin ook F. Bakker, 'Wettelijke regeling publicatie rechterlijke uitspraken', *NJB Blog* 2017.

57 Zie uitgebreid over de toegankelijkheid van Belgische jurisprudentie P. Taelman & M. Van Der Haegen, 'Recht op recht! Publicatie van rechtspraak', in: *De taal is gans het recht: liber amicorum Willy van Eeckhoutte*, Mechelen: Wolters Kluwer 2018, p. 257 e.v. Zie ook S. De Somer & C. Fornoville, 'Rechtspraakpublicatie anno 2011: historische beschouwingen, hedendaagse bedenkingen en toekomstperspectieven', *Rechtskundig Weekblad* 2011-12, p. 206 e.v.

58 Wet 5 mei 2019 tot wijziging van het Wetboek van strafvordering en het Gerechtelijk Wetboek wat de bekendmaking van de vonnissen en arresten betreft, *BS* 16 mei 2019.

59 Art. 73 wet 31 juli 2020 houdende diverse dringende bepalingen inzake justitie, *BS* 7 augustus 2020.

weinig hoopvol voor andere onderzoekers. Aan de andere kant is het wel veelbelovend dat op het niveau van het hof van beroep zonder onderscheid animo is om onderzoekers toegang tot die rechtspraak te verlenen. Maar zelfs als alle hoven en rechtbanken hun volledige medewerking verlenen, blijft het een nagenoeg onmogelijke opdracht om die jurisprudentie te vergaren. Voor het onderzoek van Goossens, Bollen en Verbeke zijn ongeveer 150 studenten ingeschakeld om de jurisprudentie ter griffie te kopiëren en te anonimiseren. Dat is niet alleen een monnikenwerk,⁶⁰ maar zal – zeker wanneer onderzoekers niet over een grote groep studenten kunnen beschikken – de nodige fondsen vergen om bijvoorbeeld kopieën te maken om die vervolgens in te scannen en om te zetten naar digitaal beschikbare data. De kopiekosten die gemaakt zijn in het onderzoek van Goossens, Bollen en Verbeke mogen dan wel afgedekt zijn door de werkingstoelage waarover Elise Goossens beschikt in het kader van haar postdoctoraal onderzoek (FWO), in talrijke andere gevallen zal zo'n toelage helaas niet voor handen zijn.

Zoals hierboven reeds betoogd voor de situatie in Nederland, zal het in de praktijk zeer lastig blijven om systematische jurisprudentieanalyse op grotere schaal te kunnen gebruiken voor fundamenteel rechtswetenschappelijk onderzoek zolang niet alle jurisprudentie behoorlijk digitaal ontsloten is. De Belgische wet van 5 mei 2019 belooft weliswaar beterschap, maar zolang er geen manier is om persoonsgegevens in jurisprudentie automatisch te anonimiseren, lijkt haar inwerkingtreding nog wel even op zich te zullen laten wachten, mede gelet op het feit dat de situatie in Nederland waar men al veel verder is, ook nog altijd niet wettelijk geregeld is.

6.3 *Datadeling*

Niet alleen de online beschikbaarheid van jurisprudentie lijkt een belangrijke randvoorwaarde voor een succesvolle ontwikkeling van systematische jurisprudentieanalyse. Ook de mate waarin de onderzoeksdata die door middel van deze methode vergaard zijn toegankelijk worden gemaakt voor derden vormt een achilleshiel.⁶¹ Een van de beperkingen op dit moment is dat er niet veel mogelijkheden zijn om datasets en codeerschema's die bij deze vorm van jurisprudentieanalyse worden gebruikt te controleren en te verifiëren, omdat in artikelen en boeken waarin resultaten van systematische jurisprudentieanalyse worden gepubliceerd vaak niet de onderliggende datasets en codeerschema's worden bijgevoegd.

60 Goossens, Bollen en Verbeke spreken in hun bijdrage niet onterecht van een 'huzarenwerk waarvoor vele handen en hoofden nodig zijn'.

61 In zijn bijdrage aan deze bundel wijst Eric Tjong Tjin Tai hier eveneens op.

Waar in andere wetenschappen regelmatig uitgebreide bijlagen bij een publicatie beschikbaar zijn via de website van bijvoorbeeld het tijdschrift waarin het onderzoek is gepubliceerd, lijkt dat op dit ogenblik nog niet echt het geval binnen de rechtswetenschap in de lage landen. Verder zijn er op dit ogenblik ook nog weinig tijdschriften en uitgevers die onderzoekers de mogelijkheid bieden om hun datasets en codeerschema's gratis te uploaden zodat de analyses daadwerkelijk gecontroleerd en geverifieerd kunnen worden. In Nederland kennen we gelukkig al wel een instituut voor permanente toegang tot digitale onderzoeksgegevens (DANS – Data Archiving and Networked Services), dat sinds 2011 ook NARCIS (National Academic Research and Collaborations Information System) omvat. Daardoor kunnen ook juridische onderzoekers hun datasets via narcis.nl ter beschikking stellen van het ruimere publiek. In de praktijk is het aantal juridische datasets op het gebied van de systematische jurisprudentieanalyse dat beschikbaar is via narcis.nl echter nog zeer beperkt. Dit heeft waarschijnlijk mede van doen met het feit dat zowel wetenschappers als hun lezerspubliek het vaak te veel rompslomp vinden om deze gegevens in dit soort databanken te gaan opslaan, terwijl ze er bovendien van uitgaan dat geïnteresseerden de primaire publicatiebron zullen raadplegen en wanneer zij daar geen achterliggende data vinden toch in eerste instantie naar de auteurs op zoek zullen gaan.

Een uitzonderlijk voorbeeld van onderzoek waar de bijbehorende data wel via narcis.nl beschikbaar zijn, is het onderzoek van Wolters dat aan de hand van een systematische analyse van 118 rechterlijke uitspraken in kaart brengt welke invloed de verhouding tussen schade en boete heeft op de matiging van boetebedingen.⁶² Via DANS⁶³ heeft Wolters de daarvoor gebruikte data, bestaande uit een bestand waarin de gegevens van de rechterlijke uitspraken worden beschreven en een codeerboek dat die gegevens toelicht, gedeeld. Een equivalent voor DANS of NARCIS is momenteel niet aanwezig in België. In theorie kunnen juridische onderzoekers in sommige gevallen eveneens algemene, internationale databanken gebruiken (zoals Zenodo en Open Science Framework),⁶⁴ maar in de praktijk lijken dergelijke online *repositories* vooral onbekend en mede daardoor onbemind.

Dat online databanken waarin onder andere systematisch jurisprudentieonderzoek zou kunnen worden opgeslagen momenteel nog niet echt ingeburgerd zijn in het juridische onderzoeksveld,⁶⁵ neemt niet weg dat onderzoekers ook los van die databanken ook zelf

62 P.J.T. Wolters, 'De invloed van de verhouding tussen schade en boete op de matiging van boetebedingen. Een empirische analyse', *AV&S* 2016, p. 109 e.v.

63 Zie <https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:64254/tab/2>.

64 Zie <https://zenodo.org> en <https://osf.io/>.

65 Op andere terreinen is dat deels wel anders. Denk bijvoorbeeld aan de beschikbaarheid van meta-evaluatie-studies op medisch gebied via de Cochrane foundation of op sociaalwetenschappelijk gebied via de Campbell

actie kunnen ondernemen om transparant te zijn over hun werkwijze. Zo zien we in het recente verleden steeds vaker dat juridische onderzoekers in zowel België⁶⁶ als Nederland⁶⁷ de gebruikte codeerschema's toevoegen als bijlage bij de gepubliceerde versie van hun dissertaties die ook steeds vaker online via de websites van universiteiten of universitaire bibliotheken ter beschikking worden gesteld. Diverse bijdragen in deze bundel tonen eveneens aan dat transparantie in talrijke gevallen wel degelijk doenbaar is, bijvoorbeeld door links op te nemen naar online plekken waar de data, codeerschema's en dergelijke meer ter beschikking worden gesteld. Bij gebrek aan gestandaardiseerde *best practices* qua datadeling zijn er echter nog wel een aantal stappen te zetten opdat systematische jurisprudentieanalyse optimaal kan bijdragen aan een betere publieke verantwoording van juridisch-dogmatisch onderzoek.

Uitgebreide en transparante datadeling mag de kwaliteit van onderzoek dan wel ten goede komen, in de praktijk kunnen geheimhoudingsclausules de mogelijkheden om data te delen in aanzienlijke mate beperken. Zelfs al kan door (manuele) anonimisering de privacy van de partijen die bij de rechtspraak zijn betrokken worden gegarandeerd, dan nog is het verre van evident om de geanalyseerde uitspraken ter beschikking te stellen van andere onderzoekers en van het ruime publiek. Het onderzoek van Goossens, Bollen en Verbeke toont aan dat vertrouwelijkheidsafspraken een belangrijke rol kunnen spelen om hoven en rechtbanken te overtuigen om (anonieme) rechtspraak ter beschikking te willen stellen. De uiteenlopende vereisten die rechtscolleges in het kader van dat onderzoek hebben gesteld om hun medewerking te verlenen, maken duidelijk dat er in dat verband – althans in België – momenteel nog geen uniforme richtlijnen zijn. Wellicht is het wachten tot de inwerkingtreding van de wet van 5 mei 2019 voor meer eenvormigheid.

7 ENKELE TOEKOMSTSCENARIO'S

'Voorspellen is lastig, vooral waar het de toekomst betreft' is een uitspraak die vaak aan de Deense natuurkundige Niels Bohr wordt toegedicht. Wij wagen ons daarom ook niet aan het voorspellen van wat de toekomst van systematische jurisprudentieanalyse is voor het onderzoek en onderwijs in de lage landen. Bezien van de juridisch-dogmatische traditie

Collaboration. Zie: www.cochranlibrary.com/ en <https://campbellcollaboration.org/> (en meer specifiek <https://onlinelibrary.wiley.com/journal/18911803>).

66 Zie bijvoorbeeld J. De Jaegere, *Judicial Review and Strategic Behaviour. An Empirical Case Law Analysis of the Belgian Constitutional Court*, Cambridge: Intersentia 2019, p. 339 e.v.

67 Zie bijvoorbeeld L.A.B.M. Wijntjens, *Als ik nu sorry zeg, beken ik dan schuld? Over het aanbieden van excuses in de civiele procedure en de medische tuchtprocedure*, Den Haag: Boom juridisch 2020, p. 347 e.v.

is het antwoord immers waarschijnlijk: dat hangt ervan af. Om die reden schetsen we drie mogelijke scenario's.

Het eerste scenario zouden we willen aanduiden als het *'business as usual scenario'*. Daarmee bedoelen we dat de verdere ontwikkeling van *'structured content analysis'* voor het bestuderen van jurisprudentie wordt overgelaten aan het vrije spel der academische krachten zonder noemenswaardige educatieve, organisatorische of financiële steun vanuit faculteiten, vanuit de rechtspraktijk, door de overheid of vanwege externe subsidieverstrekkingen. In deze situatie lijkt het toekomstige succes van deze methode van jurisprudentieonderzoek vooral afhankelijk van de voorbeeldfunctie die individuele studies kunnen bezitten. De geschiedenis van de (rechts)wetenschap laat immers zien dat veel onderzoekers in de praktijk leven volgens het motto 'beter goed gejat dan slecht bedacht'. Dat wil zeggen dat succesvolle individuele onderzoeken een doorbraakfunctie kunnen vervullen waar het gaat om de introductie van nieuwe methoden of theorieën. Denk recentelijk slechts aan de wereldwijde groei in populariteit van *'behavioural law and economics'* mede naar aanleiding van het boek over *nudging* van Thaler en Sunstein,⁶⁸ waarin zij inzichten en methoden uit onder andere de psychologie toepassen op het recht.

Dergelijke doorbraakstudies hebben we tot op heden op het gebied van systematische jurisprudentieanalyse nog niet gezien. De vraag is bovendien of we die in de toekomst zullen gaan zien. Daarmee is niet gezegd dat er op dit gebied geen uitstekende studies zijn met deels verrassende inzichten. We noemen hier bij wijze van voorbeeld slechts de studie van Cahill-O'Callaghan, waarin zij met behulp van *'structured content analysis'* aantoont dat en hoe rechters in het Britse Supreme court in *'hard cases'*, waarin sprake is van discretionaire ruimte, vaak gedreven worden door persoonlijke waarden in plaats van door beslis- en bewijsregels of precedenten.⁶⁹ In de lage landen mogen studies van onder andere Scholten,⁷⁰ Wiarda,⁷¹ Van Gerven⁷² en Vranken⁷³ dan wel reeds geruime tijd hebben vastgesteld dat persoonlijke waardeoordelen een rol kunnen spelen in rechterlijke besluitvorming, het bijzondere van studies zoals die van Cahill-O'Callaghan is vooral dat zij precies aantonen waar dat precies voorkomt in de jurisprudentie en op welke schaal. Daarmee

68 R.H. Thaler & C.R. Sunstein, *Nudge: Improving Decisions about Health, Wealth, and Happiness*, New Haven: Yale University Press 2008.

69 R. Cahill-O'Callaghan, *Values in the Supreme Court. Decisions, Division and Diversity*, Oxford: Hart 2020.

70 P. Scholten, *Handleiding tot de beoefening van het Nederlandsch burgerlijk recht. Algemeen deel*, Zwolle: Tjeenk Willink 1931.

71 G.J. Wiarda, *Drie typen van rechtsvinding*, Zwolle: Tjeenk Willink 1972.

72 W. Van Gerven, *Het beleid van de rechter*, Antwerpen: Standaard 1973.

73 J.B.M. Vranken, *Handleiding tot de beoefening van het Nederlands burgerlijk recht. Algemeen deel*, Zwolle: Tjeenk Willink 1995.

leggen zulke studies vooral patronen bloot; dieperliggende verklaringen kunnen zij op zichzelf niet aanreiken.

Nu systematische jurisprudentieanalyse vooral in de Verenigde Staten onder de naam 'jurimetrics' al decennialang wordt toegepast,⁷⁴ is het de vraag of er zonder nadere prikkels en investering in de opleiding van onderzoekers op dit punt op korte termijn grote doorbraken te verwachten zijn. Evengoed mogelijk is dat de belangstelling voor deze methode nadat het nieuwe eraf is weer wegebt.

Dit laatste zal wellicht anders worden wanneer systematische jurisprudentie-analyse kan meeliften op de groeiende belangstelling voor empirisch-juridisch onderzoek die bijvoorbeeld ook financieel wordt ondersteund door faculteiten en externe subsidieverstrekkers. Dit zouden we derhalve het 'ELS-slipstream scenario' kunnen noemen. Volgens de recent verschenen Nederlandse Encyclopedie Empirical Legal Studies zijn daarin drie pijlers te onderscheiden, namelijk: onderzoek naar empirische assumpties waarop het recht is gestoeld, onderzoek naar hoe het recht in de praktijk functioneert en onderzoek naar de effecten van het recht.⁷⁵ Systematische jurisprudentieanalyse kan gemakkelijk aansluiting vinden bij vooral de eerste twee pijlers, zoals de voorbeelden in deze bundel laten zien. Wat opvalt, is wel dat de nadruk thans nog sterk op de eerste pijler lijkt te liggen, namelijk het toetsen van aannames over met name de geldende rechtspraak die leven in de wetenschap of de praktijk of beide. De studies van Wijntjens, Vols, Wolters, Verbruggen, Vandaele, en Goossens, Bollen en Verbeke illustreren dat. Het feit dat empirisch-juridisch onderzoek zowel door faculteiten, subsidieverstrekkers als delen van de rechtspraak in de belangstelling staat, maakt het waarschijnlijk aantrekkelijk voor jonge onderzoekers om hun onderzoeksagenda daarop af te stemmen. Systematische jurisprudentieanalyse in vooral de kwalitatieve variant daarvan vertegenwoordigt bovendien misschien wel de meest natuurlijke stap voor studenten met een klassieke juridisch-dogmatische opleiding om kennis te maken met ook meer empirisch getint juridisch onderzoek dat momenteel de wind in de rug heeft. De vraag is alleen opnieuw hoe realistisch dit scenario is zolang er niet ook iets in het juridische onderwijs verandert en de rechtenopleiding vooral praktijkgeoriënteerd blijft en niet of nauwelijks wetenschappelijke onderzoeksgerichte finaliteit bezit.⁷⁶ In dat geval zal het vooral van onderzoeksmasters en PhD-opleidingen afhangen

74 Zie J. Jacobs & M. Vols, 'Juristen als rekenmeesters', in: P.A.J. van den Berg & G. Molier, *In dienst van het recht*, Den Haag: Boom juridisch 2017, p. 91 e.v.

75 M. Smit e.a., '25 jaar empirisch-juridisch onderzoek in Nederland – Een inleiding tot de Encyclopedie ELS', in: C. Bijleveld, A. Akkermans, M. Malsch, B. Marseille & M. Smit, *Nederlandse Encyclopedie Empirical Legal Studies*, Den Haag: Boom juridisch 2020, p. 13 e.v.

76 Dat is een algemene uitdaging voor empirisch onderzoek. Zie F. Bell, 'Empirical Research in Law', *Griffith Law Review* 2016, p. 268 e.v.

in hoeverre systematische jurisprudentieanalyse bredere ingang zal vinden.⁷⁷ Publicaties, zoals de onderhavige bundel, kunnen mogelijk de aantrekkingskracht van deze methode van jurisprudentieonderzoek vergroten, maar ook dan nog zal het afhangen van individuele wetenschappers die bereid en in staat zijn jonge onderzoekers op sleeptouw te nemen om hen vertrouwd te maken met deze methode, al was het maar omdat een of meer publicaties waarin de methoden van systematische jurisprudentieanalyse verkeerd blijken te zijn toepast ook een ‘*chilling effect*’ kunnen hebben.

Het derde en laatste scenario is het meest ambitieuze en veelomvattende, waarin degenen die juridisch-dogmatisch onderzoek verrichten zich langzaam maar zeker gaan realiseren dat hun normatieve onderzoek door de praktijk gedwongen zal zijn om mee te bewegen met moderne ontwikkelingen, zoals digitalisering, toenemende meergelaagdheid, opkomst van data science en artificiële intelligentie, in het besef dat die ontwikkelingen geen modegril zijn, maar ‘*here to stay*’. Dit zouden we misschien het *revitaleringsscenario* kunnen noemen. Wanneer de behoefte in de praktijk aan analyse van big data toeneemt zal automatisch ook de behoefte aan ‘*T-shaped lawyers*’ toenemen,⁷⁸ die niet enkel in staat zijn om klassieke tekstanalytische interpretatiemethoden toe te passen die een eeuwenlange traditie kennen. Tegelijk zal ook in toenemende mate behoefte zijn aan andere, meer empirisch georiënteerde methoden om bijvoorbeeld de slaagkansen van juridische procedures beter in te kunnen schatten, om foutieve normatieve of empirische assumpties van rechters, wetgever of advocaten omtrent de juiste uitleg van het geldende recht te kunnen weerleggen of om patronen te ontdekken in de toenemende hoeveelheid supranationale jurisprudentie die direct of indirect van belang is voor de uitleg van ons eigen nationale recht.

Dit scenario vergt in de eerste plaats druk vanuit de rechtspraktijk in de richting van juridische faculteiten om eindelijk eens werk te gaan maken van een broodnodige modernisering van het juridische onderwijs dat qua methodologische (zelf)reflectie mijlenver achterloopt op veel andere sociale wetenschappen en tot op heden nauwelijks aandacht besteedt aan zaken die verder gaan dan praktisch-juridische vaardigheden. Het feit dat er zowel in Nederland als België nog geen enkele voltijdse leerstoel methoden van juridisch onderzoek bestaat, spreekt op dit punt boekdelen. Een bijkomende impuls zou op dit punt kunnen

77 Uit een recente inventaris blijkt dat passieve en actieve beheersing van empirische methoden in bacheloropleidingen vooral in minoren en/of keuzevakken voorkomt, nog niet structureel aanwezig is in reguliere masterprogramma’s van Nederlandse rechtenopleidingen, maar dat empirisch-juridische vaardigheden wel meer kunnen worden uitgediept in veel van de onderzoeksmasters en PhD-opleidingen. Voor het rapport van die inventarisatie, zie Pannebakker e.a. 2020. Voor een bespreking van dat rapport, zie E. Pannebakker, H. Pluut, S. Voskamp en W. de Zanger, ‘Empirical Legal Studies in het juridisch onderwijs: waar staat Nederland en hoe nu verder?’, *RDW* 2021, p. 67 e.v.

78 E. Mak, *The T-shaped Lawyer and Beyond: rethinking legal professionalism and legal education for contemporary societies*, Den Haag: Eleven International Publishing 2017.

worden gegeven zodra bedrijven, overheden en advocatenkantoren steeds vaker zelf kennis en onderwijs gaan inhuren, welke door rechtenfaculteiten voorspeld niet wordt aangeboden.⁷⁹ Hierbij valt met name te denken aan experts op het gebied van recht en kunstmatige intelligentie en recht en empirische methoden meer in het algemeen. Zelfs wanneer dit scenario zich sneller voltrekt dan menigen mocht denken, zullen juridische faculteiten (maar ook bedrijven, overheden en advocatenkantoren) waar het gaat om de kansen die onder andere systematische jurisprudentieanalyse biedt nog altijd afhankelijk zijn van de mate waarin deze data op een toegankelijke manier elektronisch beschikbaar zijn. Op dat punt valt, zoals wij hierboven hebben betoogd nog een wereld te winnen, te beginnen bij de online publicatie van alle jurisprudentie die momenteel nog niet via digitale databases raadpleegbaar is.

Wij hopen in het voorgaande duidelijk te hebben gemaakt dat zo'n revitaliseringsscenario veel meer omvattend zal moeten zijn dan enkel het incorporeren van methoden van systematische jurisprudentieanalyse in de reguliere rechtenopleiding. Dit scenario zal alleen kans van slagen hebben wanneer er een algemeen besef ontstaat dat de rechtenopleidingen in Europa drastisch gemoderniseerd moeten worden, met name waar het gaat om de oriëntatie op de wetenschappelijke grondslagen van de discipline, het onderscheid tussen wetenschap en praktijk en de veel grotere nadruk die nodig is op het bekend maken van studenten met een breder palet van onderzoeksmethoden variërend van klassieke tekstinterpretatieve methoden tot meer empirische methoden, waaronder het gebruik van moderne datatechnologie.

79 Diverse wetenschappers hebben inmiddels voorspeld dat dit zal gaan gebeuren. Zie uit velen: C. Denvir, 'Scaling the Gap: Legal Education and Data Literacy', in: C. Denvir, *Modernising Legal Education*, Cambridge: Cambridge University Press 2020, p. 73-91; M. Livermore & D. Rockmore (red.), *Law as Data: Computation, Text, and the Future of Legal Analysis*, New Mexico: Santa Fe Institute Press 2019; J. Bosman & V. Cordo, *Data & Dialogue. A relationship redefined*, New York: TGO Consulting 2019; R. Susskind, *Tomorrow's Lawyers: An Introduction To Your Future*, Oxford: Oxford University Press 2017.



AUTEURSLIJST

Katalien Bollen

Senior Expert, Mediation, Communication, Coaching, Deloitte legal, & Family Business Dynamics-Governance en Onderzoeker, KU Leuven

Arthur Dyevre

Professor of Empirical Jurisprudence, KU Leuven

Rob van Gestel

Hoogleraar Theorie en methode van wetgeving, Tilburg University en hoogleraar Methodologie van rechtswetenschappelijk onderzoek, KU Leuven

Elise Goossens

Docent Privaatrecht, Vrije Universiteit Brussel, Postdoctoraal onderzoeker FWO Vlaanderen, KU Leuven en Gastprofessor, Universiteit Antwerpen

Frederik Peeraer

Doctor-assistent Rechtsmethodologie, Universiteit Gent en docent Juridische argumentatieleer, Universiteit Antwerpen

Eric Tjong Tjin Tai

Hoogleraar Privaatrecht, Tilburg University

Alain-Laurent Verbeke

Gewoon hoogleraar Privaatrecht & ADR, KU Leuven, Tilburg en Harvard

Paul Verbruggen

Universitair hoofddocent Privaatrecht, Tilburg University en TPR-Wisselleerstoelhouder verbonden aan de KU Leuven (2019-2021)

Michel Vols

Hoogleraar Openbare-orderecht, Rijksuniversiteit Groningen

Lianne Wijntjens

Universitair docent Privaatrecht, Tilburg University

METHODEN VAN SYSTEMATISCHE RECHTSPRAAKANALYSE

Pieter Wolters

Universitair hoofddocent Burgerlijk recht, Radboud Universiteit