

# M. van Opijnen, *Op en in het web. Hoe de toegankelijkheid van rechterlijke uitspraken kan worden verbeterd*

*Diss. Amsterdam UvA,  
promotor prof. dr. T.M. van  
Engers, Den Haag: Boom  
Juridische uitgevers 2014, 656  
p., ISBN 978-90-8974-874-4.*

## 1. Inleiding

Op 20 februari 2014 promoveerde Marc van Opijnen op het proefschrift, getiteld *Op en in het web*. De ondertitel luidt: *Hoe de toegankelijkheid van rechterlijke uitspraken kan worden verbeterd*. Het is een dik boek geworden dat hoge eisen stelt aan de lezers. Zij moeten juristen zijn en ze moeten iets van ICT weten. Beter gezegd, ze moeten een intrinsieke belangstelling voor ICT-toepassingen hebben, want de auteur laat hen alle details van zijn mechanisme MARC zien (MARC staat voor: Model for Automated Rating of Case law). Voordat het zover is, vertelt de auteur enthousiast over de lange weg naar MARC. Een adequate bespreking van deze vuistdikke dissertatie moet mijns inziens, (1) de essentie weergeven, (2) niet te lang zijn en (3) een waardeoordeel bevatten.

Om een goed inzicht te krijgen in de feiten en vragen die de auteur tot zijn onderzoek gemotiveerd hebben, reproduceer ik hieronder de eerste alinea van de samenvatting met dien verstande dat ik de lay-out en interpunctie enigszins aanpas en een nummering aan de vragen toevoeg (dit doe ik ter verhoging van de duidelijkheid).

‘De publicatie van rechterlijke uitspraken heeft door de opkomst van internet een metamorfose ondergaan. Tot diep in de jaren negentig van de twintigste eeuw hadden commerciële uitgeverij een monopolie op het verzamelen, selecteren, publiceren en verrijken van jurisprudentie. Internet(technologie) heeft echter voor de rechterlijke macht de mogelijkheden gecreëerd om de publicatie van uitspraken in eigen hand te nemen,

zowel voor intern gebruik als om te voorzien in de behoeften van rechtspraktijk, wetenschap en de maatschappij. De verstrekking van uitspraken aan het publiek heeft twee functies: de legitimatiefunctie en de rechtsvormende functie. De nieuwe mogelijkheden roepen echter tal van vragen op:

1. Moeten alle uitspraken worden gepubliceerd, of volstaat een selectie?
2. Hoe ziet zo’n eventuele selectie er dan uit?
3. In hoeverre moeten uitspraken worden geanonimiseerd?
4. Volstaat een blote plaatsing van documenten, of moeten ontsluitingsfaciliteiten worden aangeboden waarmee de gebruiker effectief en efficiënt informatie kan vinden in een ras groeiende databank?’

Dit is een interessante alinea en toch zien we hier dat de vier vragen relatief ver verwijderd zijn van de vraag: ‘Hoe kan de toegankelijkheid van rechterlijke uitspraken worden verbeterd?’ Laat ik allereerst vaststellen dat de auteur alle vier de vragen voortreffelijk (en uitvoerig) beantwoordt.

Hierbij teken ik aan dat Van Opijnen zich in de voetsporen beweegt van een stroming die vaak en luid van zich liet horen tussen 1990 en 2010: ‘Alles wat offline geldt, geldt ook online (...)’.<sup>1</sup> Twee prominente voorbeelden zijn Hans Franken en Aernout Schmidt<sup>2</sup> (2003): ‘What holds good offline, also applies online’, en Henk Sniijders en Stephen Weatherill<sup>3</sup> (2003): ‘What holds offline, holds online, but (...)’.

In par. 2 (Wat gaan we ontsluiten?) behandel ik kort de eerste twee

hierboven genoemde vragen. Daarna wordt het tijd om een overzicht te geven van de rijke inhoud van het proefschrift (par. 3: Recht, informatica en kunstmatige intelligentie). Vervolgens sta ik stil bij de bijdrage van de auteur aan de wetenschap (par. 4: De betekenis van MARC). In par. 5 (Wat kan de computer bieden?) geef ik een slotbeschouwing, waarbij de laatste twee vragen (3 en 4) eveneens beantwoord worden. Tevens wordt in deze paragraaf op basis van het proefschrift een toekomstperspectief gegeven.

## 2. Wat gaan we ontsluiten?

In hoofdstuk 5 getiteld ‘Selectie en relevantie’ behandelt de auteur het selectievraagstuk, dat gebeurt vanzelfsprekend vanuit het offlineperspectief. Natuurlijk komt ook hier (en al eerder) de Aanbeveling R (95) 11 van het Comité van Ministers van de Raad van Europa ter sprake, evenals de (oude en nieuwe) selectiecriteria van Rechtspraak.nl (p. 328). Het is aardig om te zien hoe de ontwikkeling van selectiecriteria in Nederland heeft plaatsgevonden. Ik neem als voorbeeld de selectiecriteria van de *NJ*. Zij kennen zes criteria (zie p. 331 van het proefschrift), te weten:

1. juridisch belang van de zaak;
2. actualiteitswaarde;
3. vindt rechtsvorming plaats;
4. wordt open rechtsvraag ingevuld;
5. wordt een ‘zeker kleurengamma ten aanzien van feitelijke oordelen’ geboden;
6. in hoeverre is de gegeven beslissing ook buiten de gegeven casuspositie van belang.

In dit opzicht is het interessant om op te merken dat Henk Sniijders al in 1978 wees op het feit dat ogen-

schijnlijk oninteressante gevallen (het ging over omgangsregelingen) toch van belang konden zijn. Van Opijnen (p. 330) geeft het volgende citaat van Snijders:

‘Vast staat thans dat de gepubliceerde jurisprudentie niet een doorsnee van de praktijk der rechtsvinding maar een selectie van interessante gevallen betreft. Terecht worden vooral die gevallen gepubliceerd. Ten onrechte gaat ook de rechtstheorie voornamelijk uit van de gepubliceerde rechtspraak met als gevolg een kloof tussen theorie en praktijk der rechtsvinding. Ergo, ik bepleit meer aandacht voor de ‘oninteressante casus in de rechtsvinding.’

Met dit citaat had Van Opijnen zijn eerste vraag als volgt kunnen beantwoorden: ‘Alle uitspraken publiceren, een selectie volstaat niet.’ Natuurlijk zou dit voor een promovendus een te snelle stap zijn geweest, maar het zou duidelijkheid verschaffen, te weten: online kan er veel meer dan offline en er zijn ook andere technieken beschikbaar. Een verwijzing naar big-data-technieken is hier op zijn plaats.<sup>4</sup> Van Opijnen doet het heel voorzichtig. Hij gaat over op het begrip ‘relevantie’ en start daarmee de tweede vraag: ‘Hoe ziet zo’n eventuele selectie er dan uit?’ De auteur beschrijft vervolgens een generieke theorie, gaat over op het meten van relevantie, behandelt traditionele statistiek versus netwerk-analyse en beschrijft toepassingen op jurisprudentiedatabanken. Hij onderscheidt vijf typen van relevantie, te weten (1) systeemrelevantie, (2) onderwerpsrelevantie, (3) cognitieve relevantie, (4) situationele relevantie, en (5) domeinrelevantie. Op p. 357 verklaart de promovendus dat hij zich richt op de domeinrelevantie. In hoofdstuk 6 wordt daarna het verzamelen van onderzoeksgegevens besproken. Het gaat om ‘de samenstelling van de datacollecties en de wijze waarop allerlei voor het model belangrijke gegevens aan rechterlijke en wetenschappelijke teksten zijn onttrokken’ (zie p. 358). In hoofdstuk 7 wordt dan MARC besproken.

Hiermee geeft Van Opijnen een afdoende antwoord op de tweede onderzoeksvraag. In het kort luidt zijn

antwoord: het is afhankelijk van de domeinrelevantie.

Het is wetenschappelijk allemaal keurig in orde (mijn complimenten). Toch eist ook hier de offline-online-gedachte zijn tol. Online kan zoveel meer en doet het zoveel beter (in precisie en in snelheid). Waarom daar dan geen gebruik van maken? Natuurlijk is Van Opijnen dit pad opgegaan, getuige zijn alinea (p. 339): ‘Derhalve moeten we de aandacht verleggen van de inputzijde naar de outputzijde: welke – voor hem relevante – selectie uit het totale bestand kan de gebruiker verkrijgen? Daarmee wordt het selectievraagstuk een relevantievraagstuk dat we kennen uit de wereld van zoekmachines en *information retrieval*. Daarop zullen we de aandacht dan ook nu gaan richten.’ Van Opijnen slaat de spijker op z’n kop. Zo is het, en zo zal het verder moeten gaan. Ik beschrijf nu eerst het pad dat Marc op zoek naar MARC heeft gevolgd.

### 3. Recht, informatica en kunstmatige intelligentie

De vraag van Marc is: ‘Hoe kan de toegankelijkheid van rechterlijke uitspraken worden verbeterd?’ Dit is een procedurevraag (een hoe-vraag), waarvan de beantwoording ten slotte leidt tot de wetenschappelijke bijdrage van de promovendus. Hieraan vooraf gaat, zoals we reeds gezien hebben, de vraag: ‘Wat gaan we eigenlijk ontsluiten?’ (Zie nogmaals de vier vragen in Marcs samenvatting.) De wat-vraag krijgt veel aandacht in het proefschrift, naast de vragen waarom, voor wie en onder welke voorwaarden. Hieronder geef ik derhalve eerst de opbouw van het proefschrift; het bevat 8 hoofdstukken en 21 bijlagen. De hoe-vraag komt helemaal aan het einde ter sprake. Voor een goed inzicht en adequaat overzicht bespreek ik daarom eerst de hoofdstukken die voorafgaan aan de hoe-vraag, zij het zo kort mogelijk.

#### 3.1. Hoofdstuk 1: Inleiding

De Inleiding bestaat uit een onderzoeksdoelstelling (1.1), onderzoeksvragen en een beschrijving van de opzet van het boek (1.2) alsmede een leidraad bij het lezen (1.3). Wat mij betreft raakt het begin van de samenvatting (zie citaat hierboven) de kern van het onderzoek (afgezien van de

andere kern: MARC) beter en is daarmee voor de lezer ook aantrekkelijker.

#### 3.2. Hoofdstuk 2: Juridisch kader

De bespreking van het juridisch kader is diepgaand en veelomvattend. Ik noem: (i) openbaarheid van de uitspraak, (ii) verstrekking van de uitspraken, (iii) toegankelijkheid, (iv) anonimisering, (v) rechtsvergelijkend perspectief, en (vi) doctrines over de eerste vier onderwerpen.

#### 3.3. Hoofdstuk 3: Uitsprakenpublicatie van overheidswege in Nederland

De feitelijke ontwikkelingen in Nederland worden uitvoerig (mogelijk volledig) beschreven. We zien een heldere analyse van een corpus dat door uitgever is gepubliceerd. De problematiek wordt geschetst aan de hand van zaakafhandelingsstatistieken. De onderzoeker concludeert dat de in 2013 gepubliceerde selectie als voldoende representatief gekarakteriseerd kan worden. Hierna gaat het over de toegankelijkheid van de juridische databanken.

#### 3.4. Hoofdstuk 4: Identificatie en citatie van rechterlijke uitspraken

Rechters, advocaten, en onderzoekers liepen in het verleden geregeld tegen hetzelfde probleem aan: over welke uitspraak hebben we het? Hoe kunnen we die vinden? En waar kunnen we hem vinden? De identificatie van uitspraken is het centrale punt van hoofdstuk 4. Het gaat daarbij om *uniciteit* en *persistentie*. In Nederland waren lange tijd diverse identificatiesystemen in gebruik (nog steeds trouwens), maar de wens om te uniformeren is aanwezig en wordt boeiend beschreven door de auteur. In de eerste helft van het hoofdstuk worden het theoretisch kader, toetsingskaders, en de citatievoorschriften tot in detail besproken. Ze worden vervolgens vergeleken met drie buiten-continentale landen (Canada, Zuidelijk Afrika, en Verenigde Staten), alsook met drie Europese landen (Spanje, Noorwegen, Finland) en de Europese Unie. In par. 4.5 is de auteur gereed voor het beschrijven van de grote stap: de Europese Case Law Identifier (ECLI). Op 29 april 2011 werden de ECLI-conclusies gepubli-

ceerd. De auteur heeft de tekst opgenomen als bijlage 6. In genoemde paragraaf (4.5) is ECLI *mean and lean* beschreven, precies, ingekort en toch volledig (opnieuw complimenten). Daarna volgen nog veel wetenswaardigheden, waarbij het vergelijkingsoverzicht (fig. 4-6, p. 309) zeer illustratief is.

### 3.5. Hoofdstuk 5: Selectie en relevantie

De essentie van hoofdstuk 5 is in mijn par. 2 (Wat gaan we ontsluiten?) beschreven om de lezer enthousiast te maken voor het onderwerp.

### 3.6. Hoofdstuk 6: Uitspraken in context

Recht, informatica en kunstmatige intelligentie worden in een verstrengelde positie gebracht en besproken. Hoe kunnen we een relationele databank met behulp van metadata toegankelijk maken voor het vinden van de juiste uitspraken? Dit wil zeggen dat de uitspraken de juiste context dienen te hebben. Het proefschrift beschrijft een normalisatieproces van citatiestrings. Om tot canonieke strings te komen wordt een beroep gedaan op – wat de auteur noemt – het canonicalisatieproces. Daarna laat de auteur zien dat het systeem geschikt is om wetswijzigingen te verwerken.

### 3.7. Hoofdstuk 7: Een maat voor domeinrelevantie

Eigenlijk hebben hoofdstuk 4, 5 en 6 tot doel om de auteur in hoofdstuk 7 te laten schitteren. Ik vind het ook een mooie vondst om het systeem MARC te noemen. MARC is een model dat de domeinrelevantie aan geeft. Dit is 'de mate waarin een individuele rechterlijke uitspraak bijdraagt aan de algemene rechtsopvatting'. Marc gebruikt daarvoor een kennisinfrastructuur van vijf componenten: (1) publicatie in jurisprudentieperiodieken, (2) publicatie van een annotatie, (3) inkomende citaties uit eenmalige literatuur, (4) inkomende citaties uit continue literatuur en (5) inkomende citaties uit uitspraken.

Er worden vervolgens drie fasen onderscheiden: de publicatiefase, de transitiefase en de citatiefase. Elke fase heeft zijn eigen model en ook zijn eigen MARC-indicator. Omdat in de citatiefase gebruikgemaakt kan

worden van *crowd*variabelen begint de auteur met deze fase (de citatiefase). Hij ontwikkelt voor elke variabele van de betreffende component van de infrastructuur een weefactor. Dat is een intensieve taak waarbij computers een goed hulpmiddel zijn. Vervolgens vindt een uitvoerige netwerkanalyse plaats, waarin onder andere aandacht wordt geschonken aan het verschil in distributies en de rol die dit verschil kan spelen. Ten slotte bereikt de auteur zijn doel en noemt het MARC-CF. Voor de publicatiefase en de transitiefase kan de auteur (nog) niet van dezelfde technieken gebruikmaken – er is nog niet voldoende materiaal voor handen. Daarom heeft hij voor de publicatiefase twee competentieklassen (klasse 1: overige gerechten, en klasse 5: Europese gerechten) buiten beschouwing gelaten. Vervolgens wordt op een iets andere wijze opnieuw een MARC-indicator geformuleerd (MARC-PF). Analoog met weer andere redeneringen wordt ook een MARC-TF ontwikkeld. We hebben nu drie waarden (MARC-PF, MARC-TF, MARC-CF) en met behulp van een vijfpunts Likert-schaal (zie p. 451-452) wordt nu de MARC-indicator vastgesteld. Er volgen nog enkele bladzijden met testresultaten die aangeven hoe beloftevol dit systeem is.

### 3.8. Hoofdstuk 8: Slotbeschouwing

De slotbeschouwing bevat de belangrijkste bevindingen, de kortetermijnperspectieven (relatie recht en informatica) en de lonkende perspectieven. De laatste zin is het waard om geciteerd te worden: 'Een schat aan intrigerende bevindingen ligt te wachten.'

## 4. De betekenis van MARC

De betekenis van MARC voor de wetenschap is zonder meer groot te noemen. De auteur maakt waar wat rechtsinformatici wisten maar nooit implementeerden, en opent de ogen van praktijkjuristen en onderzoekers voor welk juridisch deelgebied ook maar. Het proefschrift is multidisciplinair en dat betekent dat sommige delen kennis van een specifieke discipline vereisen. De auteur heeft evenwel zelf voor tijdige disseminatie gezorgd door reeds in 2010 over zijn plannen te publiceren.<sup>5</sup> Na de

verschijning van het proefschrift heeft hij in *Rechtstreeks* een 'algemeen toegankelijke' bijdrage geschreven als vervolg op het artikel uit 2010.<sup>6</sup> Ook in de tussentijd (2011-2013) heeft hij van zich laten horen.<sup>7, 8, 9</sup>

Het artikel in *Rechtstreeks* 2014 heeft geleid tot een reactie van Peter van der Jagt en Simon van der Linde<sup>10</sup> onder de titel 'Juridische uitgevers omarmen MARC'. Uit deze titel is duidelijk dat zij enthousiast zijn. Toch zetten ze ook twee kanttekeningen.

Van der Jagt en Van der Linde schrijven op p. 41-42:

'Een kanttekening die wij willen plaatsen is dat MARC vanuit de rechterlijke macht is ontwikkeld. Hierdoor zijn (impliciete) keuzes gemaakt vanuit het perspectief van de rechterlijke macht. Een rechterlijke macht die keuzes maakt over de selectie van uitspraken die opgenomen moeten worden in [www.rechtspraak.nl](http://www.rechtspraak.nl). Daar komt dan ook het selectiebeleid van de uitgever bij. De uitgever omringt zich met redacturen en selecteurs die niet alleen afkomstig zijn uit de rechterlijke macht, maar ook uit de wetenschap, advocatuur, het notariaat, etc. Zij bepalen op basis van op de rechtspraktijk toegesneden normen de waarde van uitspraken voor de juridische gemeenschap en geven zo een extra smaak en keuzemogelijkheid voor de juridische professional die zicht wil houden op de juridische realiteit.

(...)

Van Opijnens model geeft handvaten voor het geautomatiseerd classificeren van juridisch belangrijke uitspraken. Het publieke belang dat wordt gediend, aldus Van Opijnens, met het selecteren, kwalificeren, koppelen en ontsluiten van rechtspraak is in de huidige vorm naar onze mening slechts het beperkte belang van de juridische professional en niet het brede belang van de gemiddelde burger die geïnteresseerd is in rechterlijke beslissingen.'

Deze kanttekeningen sluiten aan bij de eerdere discussie in par. 2 over (1) selectie (alleen de interessante uitspraken) en (2) de offlinebenadering. Zie hier, niet de juristen, niet

de wetenschap en ook niet de uitgevers moeten beslissen over de selectie, maar de gebruiker in samenwerking met een computerprogramma. Het zoekprogramma moet op aanwijzingen van de gebruiker bepalen welke uitspraken relevant zijn voor de taak die de betreffende gebruiker voor ogen staat. De wetenschap moet daartoe nog een aantal stappen maken in samenwerking met juristen, uitgevers en *agile software engineers*. Ik ben het met Marc eens, een schat aan intrigerende bevindingen ligt op ons te wachten.

### 5. Wat kan de computer bieden?

Het tijdperk van de offlinebenadering nadert zijn einde. Computers kunnen wat mensen niet kunnen en mensen kunnen wat computers niet kunnen. Maar computers leren snel (en 'diep', zoals in *deep learning*). De derde vraag – 'In hoeverre moeten uitspraken worden geanonimiseerd?' – kunnen we nu gemakkelijk beantwoorden. We houden ons aan de redelijk stringente eisen van ECLI en laten de computerprogramma's het werk doen. Na enkele aanloopfouten zal blijken dat ze minder over het hoofd zien dan mensen.

Dan rest nog de vierde vraag: 'Volstaat een blote plaatsing van documenten, of moeten ontsluitingsfaciliteiten worden aangeboden waarmee de gebruiker effectief en efficiënt informatie kan vinden in een ras groeiende databank?' Het antwoord hierop is mijns inziens een volmondig 'ja'. Natuurlijk moet blote plaatsing goed afgeschermd zijn, maar vanuit die positie kunnen intelligente computerprogramma's zorg dragen voor de juiste ontsluiting van alle informatie in de ras groeiende databank. Vooralsnog is dit toekomstmuziek. Voor de stap die Marc van Opijnen gezet heeft verdient hij evenwel alle hulde.

*H.J. van den Herik*<sup>11</sup>  
Leiden, april 2015

- 
1. *Kamerstukken II* 1997/98, 25880, 1, p. 1.
  2. H. Franken & A. Schmidt, 'Law as code, code as law – general remarks on legal requirements engineering', in: H.J. Snijders & S. Weatherill (red.), *E-commerce Law. National and Transnational Topics and Perspectives* (Meijers-reeks, nr. 54), The Hague: Kluwer Law International 2003, p. 117-139.
  3. H.J. Snijders & S. Weatherill, 'What holds offline, holds online, but...', in: H.J. Snijders & S. Weatherill (red.), *E-commerce Law. National and Transnational Topics and Perspectives* (Meijers-reeks, nr. 54), The Hague: Kluwer Law International 2003, p. 1-4.
  4. V. Mayer-Schönberger & K. Cukier, *De big data revolutie*, Amsterdam: Maven Publishing 2013.
  5. M. van Opijnen, 'Rechtspraak en digitale rechtsbronnen: nieuwe kansen, nieuwe plichten', *Rechtstreeks* 2010, afl. 1.
  6. M. van Opijnen, 'Rechtspraakdata: open, linked en big', *Rechtstreeks* 2014, afl. 2, p. 12-39.
  7. M. van Opijnen, 'Jurisprudentiestatistiek. Over de publicatie, annotatie en citatie van rechterlijke uitspraken', *NJB* 2011/1658.
  8. M. van Opijnen, 'Het communautaire geteld – een kwantitatief onderzoek naar Europeesrechtelijke verwijzingen in rechterlijke uitspraken en rechtswenschappelijke literatuur', *SEW* 2011, afl. 6.
  9. M. van Opijnen, 'Van LJN naar ECLI. Vijf prangende vragen', *NJB* 2013/1607.
  10. P. van der Jagt & S. van der Linde, 'Juridische uitgevers omarmen MARC', *Rechtstreeks* 2014, afl. 2, p. 40-42.
  11. **Prof. dr. H.J. van den Herik is hoogleraar Recht en Informatica aan de Universiteit Leiden en voorzitter van Board of Directors van het Leiden Centre of Data Science (LCDS).**