

rechtstreeks 2014 nr 2

rechtstreeks 2014 - nr 2



de Rechtspraak

Recht in Nederland | Naar de rechter | Organisatie | Actieschets | Uitvoeren en registreren | Procedure | Werken bij

Uitspraken en besluiten | Uitspraak | Zaken in uitvoering

ECLI:NL:RBNHO:2014:9376

Instantie	Rechtbank Noord-Holland
Datum uitspraak	09-10-2014
Datum publicatie	09-10-2014
Zaaknummer	14-001123
Rechtsgebieden	Strafrecht
Bijzondere kenmerken	Raadkamer
Inhoudsindicatie	Officier van justitie geeft gedragsaanwijzing art. 509h Sv. aan lid motorclub: uitgaansverbod in centrum Alkmaar. Versichte stelt daartegen beroep in. Openbaar Ministerie is niet in staat voor de zitting dossier aan de rechtbank te doen toekomen. Rechtbank verklaart beroep gegrond.

Vindplaatsen: Rechtspraak.nl

Uitspraak

RECHTBANK NOORD-HOLLAND
Afdeling Publiekrecht, Sectie Straf

Locatie Alkmaar
Eensvoudige raadkamer

Registratienummer: 14-001123
Parketnummer: 15/00044-14
Uitspraakdatum: 9 oktober 2014

Beschikking (art. 509h Wetboek van Strafvordering)



de Rechtspraak

Raad voor de
rechtspraak

Recent verschenen

- 2014 – nr 1 ‘Collectief vakmanschap’
Rechtspraak en digitalisering (deel 1)
- 2013 – nr 4 Rechters, ga experimenteren!
Regie in de rechtspraak (deel 4)
- 2013 – nr 3 Grondrechten in de Nederlandse rechtspraak
Regie in de rechtspraak (deel 3)

Hoofdartikel Rechtspraakdata: open, linked
en big

Serie Rechtspraak en digitalisering (deel 2):
De acht meest gestelde vragen over KEI IT

Rechtspraakdata: open, linked en big

Marc van Opijnen

1 Inleiding

1.1 Handige knoppen?

‘Stel, u worstelt met een ingewikkelde casus en hebt eindelijk een relevant arrest gevonden. Nu wilt u alle jurisprudentie en rechtswetenschappelijke artikelen vinden die op dit arrest betrekking hebben. Dat kan, met één druk op de knop.’

Met de beschrijving van dergelijke, nog niet bestaande, zoekfuncties opende de Rechtstreeks 2010/1, getiteld ‘Rechtspraak en digitale rechtsbronnen: nieuwe kansen, nieuwe plichten’.¹ In die Rechtstreeks heb ik getracht uit te leggen waarom dergelijke ‘handige knoppen’ makkelijk ogen maar moeilijk zijn te realiseren, en welke verantwoordelijkheden de Rechtspraak zelf heeft in de daarvoor noodzakelijke digitalisering van het juridische informatieaanbod.

We zijn inmiddels vier jaar verder, en nog steeds heeft u de beschreven functies nog niet tot uw beschikking. Is er in die vier jaar dan niets gebeurd? Jazeker wel, maar we zijn er nog lang niet. De benodigde data (zoals rechterlijke uitspraken, wetgeving en bestuurlijke kerncijfers) zijn in steeds ruimere mate beschikbaar, kunnen steeds beter in onderlinge samenhang worden beschouwd en er kan steeds meer kennis uit deze data worden gedistilleerd. Dat gaat overigens ook om kennis die u misschien liever niet direct in de openbaarheid ziet. Zo kan een computervaardige advocaat of burger zich voorbereiden op de zitting met behulp van een analyse van uw eerdere uitspraken, uw uit Facebook, LinkedIn en Twitter blijkende hobby's, politieke opvattingen en sociale netwerk. En de minder computervaardige advocaat of burger kan deze informatie gewoon kopen.

Drie begrippen spelen in al deze ontwikkelingen een cruciale rol: *open data*, *linked data* en *big data*. Hierover gaat dit artikel. Het is een vervolg op de genoemde Rechtstreeks 2010/1, maar kan ook zelfstandig worden gelezen. Beide bijdragen zijn nauw gerelateerd aan mijn proefschrift *Op en in het web. Hoe de toegankelijkheid van rechterlijke uitspraken kan*

1 Van Opijnen 2010.

worden verbeterd, het fysieke eindresultaat van een onderzoek dat mede mogelijk was dankzij een bijdrage uit het wetenschapsprogramma van de Raad voor de rechtspraak.² De primaire onderzoeksvraag luidde: is het mogelijk een *Model for Automated Rating of Case Law* (MARC) te ontwikkelen, waarmee de computer rechterlijke uitspraken kan classificeren naar hun jurisprudentiële waarde? Op de weg daar naartoe moesten eerst enkele andere vraagstukken worden geadresseerd. *Op en in het web* begint met een uitgebreide analyse van het juridisch kader rond de (internet)publicatie van rechterlijke uitspraken, alsmede met een toetsing van de Nederlandse praktijk aan dat juridisch kader. Dit onderwerp laat ik hier verder rusten, maar beknopt samengevat luidt de conclusie dat het juridisch kader zich ontwikkelt in de richting van een plicht tot publicatie van in ieder geval de belangrijkste uitspraken. De eis tot anonimisering van in uitspraken voorkomende persoonsgegevens staat tegenwoordig niet meer ter discussie en met de groei van de uitsprakendatabanken nemen ook de eisen toe aan de toegankelijkheid van de erin opgeslagen informatie. De Nederlandse rechtspraak doet het – in een internationale vergelijking – heel behoorlijk, maar de ontwikkelingen gaan zo snel dat er geen enkele reden is om genoegzaam achterover te leunen.

Een belangrijker stap nog die moest worden gezet om tot MARC te komen was het bouwen van een databank met linked data: de verbanden tussen uitspraken, wetgeving en rechtswetenschappelijke literatuur. Dit had nogal wat voeten in aarde, en heeft bij elkaar enkele jaren geduurd. De problemen waren niet alleen technisch van aard, maar hadden ook te maken met de beschikbaarheid en standaardisatie van (open) data.

Over de recente ontwikkelingen rond open data en linked data gaat hoofdstuk 2. Hoofdstuk 3 bevat een beknopte beschrijving van het Model for Automated Rating of Case Law. In hoofdstuk 4 worden enkele beleidsmatige vragen opgeworpen met betrekking tot de onstuitbare ontwikkelingen rond open, linked en big data, begrippen die allereerst een korte uitleg vereisen.

1.2 Open, linked, big

Open data, linked data, big data. Drie termen die steeds vaker in media en beleidsagenda's opduiken, maar niet altijd even zorgvuldig worden gebruikt. De begrippen hangen onderling nauw samen, maar duiden toch drie verschillende concepten aan.

Met open data duiden we de vrije beschikbaarheid van ruwe data aan, in het bijzonder (semi-)overheidsdata. Het begrip ziet niet op de mate waarin, en de manier waarop informatie van een overheidsorganisatie doorzoekbaar wordt gemaakt op de eigen website, maar op de beschikbaarheid van deze data voor hergebruik door anderen.

2 Van Opijnen 2014.

De juridische grondslag voor open data is te vinden in de Europese Hergebruikrichtlijn, in Nederland gematerialiseerd in hoofdstuk V-A van de Wet openbaarheid van bestuur.³ In artikel 5 lid 1 van de Europese richtlijn werd alleen de eis gesteld van een licentie tot hergebruik: ‘Openbare lichamen stellen hun documenten beschikbaar in eender welk reeds bestaand formaat of eender welke taal, indien mogelijk en passend langs elektronische weg. (...)’

In de praktijk vormen de technische eigenschappen van de data een minstens even belangrijke factor bij het bepalen van de feitelijke herbruikbaarheid. Drie aspecten spelen daarbij een belangrijke rol.⁴

In de eerste plaats dienen de data machineleesbaar te zijn. Ze dienen daartoe goed te zijn gestructureerd, bijvoorbeeld als *eXtensible Markup Language* (XML) of *Character Separated Values* (CSV). Pdf en Word zijn niet of moeilijk computerleesbaar en verdienen dus niet de voorkeur.

Vervolgens dienen gegevens uniek en persistent geadresseerd te kunnen worden met behulp van specifieke nummers of codes, aangeduid als *uniform resource identifiers* (URI's). Juist omdat URI's uniek en persistent zijn kan iedereen ze altijd en overal gebruiken om naar hetzelfde ding te verwijzen. Burgerservicenummers en autokentekens zijn URI's, maar ook ECLI's en de documentnummers van Europese Verordeningen of Richtlijnen. Ook internet-adressen zijn uniek, maar sommige veranderen regelmatig. Om als betrouwbare URI te kunnen worden gebruikt moet de eigenaar persistentie garanderen.

Ten slotte bevatten data bij voorkeur ook links naar andere data. Die links – die kunnen verwijzen naar gegevens in de eigen dataset, maar ook naar die van anderen – moeten worden gemaakt met behulp van de URI's van die gerelateerde data.

Data die aan al deze drie voorwaarden – machineleesbaarheid, unieke en persistente URI's en links naar andere data – voldoen, noemen we *linked open data*. Daarbij is het goed voor ogen te houden dat linked data niet altijd ook open data zijn. Zo zijn veel data bij overheids- en uitvoeringsorganisaties wel linked – omdat ze met behulp van URI's als burgerservicenummers, KvK-nummers, postcodes, kentekenbewijzen en bankrekeningnummers aan elkaar worden geknoopt – maar niet open; ze worden niet gepubliceerd op internet, en zijn niet voor algemeen hergebruik beschikbaar.

De technische eisen die aan de herbruikbaarheid van open data worden gesteld hebben inmiddels ook juridisch hun weerslag gekregen: in de herziene Hergebruikrichtlijn,⁵ waar-

3 Richtlijn 2003/98/EG van het Europees Parlement en de Raad van 17 november 2003 inzake het hergebruik van overheidsinformatie.

4 Vergelijk *5 star open data* van WWW-uitvinder Tim Berners-Lee: <http://5stardata.info>.

5 Richtlijn 2013/37/EU van het Europees Parlement en de Raad van 26 juni 2013 tot wijziging van Richtlijn 2003/98/EG inzake het hergebruik van overheidsinformatie.

over in de zomer van 2013 is besloten en die in juli 2015 in nationale wetgeving moet zijn omgezet, leest artikel 5 lid 1: ‘Openbare lichamen stellen hun documenten ter beschikking in de reeds bestaande formaten of talen en, indien mogelijk en passend, in een open en machinaal leesbaar formaat, samen met hun metadata. Zowel het formaat als de metadata dient voor zover mogelijk aan formele open standaarden te voldoen.’

Het tempo waarin elektronische data worden geproduceerd groeit exponentieel. Daarbij komt al snel het buzzword big data om de hoek kijken, maar om big te zijn, moeten data meer zijn dan alleen maar volumineus. Een bruikbare definitie levert onderzoeksbureau Gartner: *‘Big data is high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making.’*⁶

Belangrijke elementen in deze definitie zijn:

1. Kenmerken van de data: volume, variëteit en variabiliteit. Ontbreekt een van de drie dan is geen sprake van big data.
2. Door deze eigenschappen zijn de data niet met traditionele hardware, software en analysemethoden verwerkbaar.
3. Analyse van de data kan nieuwe kennis opleveren.

De verzamelwoede van de NSA valt zonder twijfel binnen de definitie van big data, maar ook Google gebruikt big data om zoekresultaten aan te passen aan fysieke locatie, sociaal netwerk en persoonlijke voorkeuren van de gebruiker. Netflix gebruikt big data om filmsuggesties te doen aan individuele gebruikers en om de optimale ingrediënten te bepalen van de volgende dramaserie, en supermarkten en banken gebruiken big data om op basis van aankoop- en betaaldegedrag klanten persoonlijke aanbiedingen te doen.

Met deze voorbeelden wordt ook meteen duidelijk dat veel big data wel linked zijn, maar niet open. De informatieverzamelaars willen de data niet open maken omdat ze waardevolle bedrijfsinformatie bevatten, of ze mogen dat niet omdat er gevoelige persoonsgegevens in zitten.

2 De datatrein rijdt

2.1 Open data

Open data zijn onmisbare ingrediënten bij het vormgeven van een transparante overheid.⁷ Bovendien groeit het besef dat de overheid zelf een van de grootste hergebruikers van de

⁶ www.gartner.com/it-glossary/big-data.

⁷ Zie bijvoorbeeld de Beleidsvisie Open Overheid en het Actieplan Open Overheid (*Kamerstukken II 2013/14*, 32 802, nr. 5 (bijlagen)).

eigen data is, waardoor een goed open-databeleid ook een belangrijke bijdrage kan leveren aan een efficiëntere overheid. Daarnaast maken ook steeds meer bedrijven gebruik van open overheidsdata. Om al deze open-datahergebruikers te helpen bij het vinden van de juiste datasets is het portaal Data.overheid.nl ingericht, waarop inmiddels zesduizend datasets zijn aangemeld.

Ook de Rechtspraak heeft een aantal gegevensverzamelingen als open data beschikbaar gesteld – al zijn die (nog) niet via Data.overheid.nl te vinden. De belangrijkste is natuurlijk de uitsprakendatabank, die al sinds 2004 om niet voor uitgevers, content-integrators, advocatenkantoren, wetenschappers en bedrijven beschikbaar is.

Ook de gegevens uit het centraal insolventieregister kunnen worden hergebruikt, zij het dat belangstellenden eerst moeten verklaren dat zij de informatie alleen zullen gebruiken binnen de grenzen van de Wet bescherming persoonsgegevens.

2.2 URI's

In paragraaf 1.2 hebben we gezien dat open data, om verwerkbaar en betrouwbaar te zijn, moeten zijn voorzien van URI's: nummers of codes waarmee gegevens uniek en persistent worden geïdentificeerd. Ik bespreek hier eerst de belangrijkste ontwikkelingen met betrekking tot gestandaardiseerde URI's voor juridische bronnen.

European Case Law Identifier

Voor rechterlijke uitspraken heeft de Rechtspraak in juni 2013 een belangrijke stap gezet met de vervanging van het Landelijk Jurisprudentienummer (LJN) door de *European Case Law Identifier* (ECLI).⁸ ECLI is een (niet-verplichte) standaard, in 2010 vastgesteld door de Raad van ministers van de EU,⁹ die als doel heeft om het computerleesbaar citeren van rechterlijke uitspraken te vergemakkelijken. Daardoor kunnen, zowel in nationale als Europese context, uitspraken beter vindbaar worden, ook met behulp van bijvoorbeeld citatie-indexen. De ECLI-systematiek bestaat uit vier onderdelen: de ECLI-code, gestandaardiseerde metadata, de organisatiestructuur en het Europees ECLI-zoekportaal.

De ECLI-code mag inmiddels bekend worden verondersteld: hij bestaat altijd uit vijf delen, gescheiden door dubbele punten: de aanduiding 'ECLI', de landcode, de gerechtscodex, het jaartal van uitspraak en een unieke code, die mag bestaan uit letters, cijfers en punten.

De gestandaardiseerde metadata zijn voor de gebruiker niet direct zichtbaar. Ze dienen ertoe om zoekvelden en filters overal vergelijkbare resultaten te laten geven, onder meer in het Europees ECLI-zoekportaal.

⁸ Van Opijnen 2013.

⁹ Raadsconclusies waarin de invoering wordt aanbevolen van een Europese identificatiecode voor jurisprudentie (ECLI), en van een minimumaantal uniforme metagegevens betreffende jurisprudentie (CELEX:52011XG0429(01)).

Alhoewel ECLI-code en metadata zijn gestandaardiseerd, hebben de lidstaten een grote mate van vrijheid bij het toepassen van de standaard. Per land besluit de nationale ECLI-coördinator over de wijze van introductie – geleidelijk of, zoals in Nederland, met een *big bang* – en het vijfde deel van de ECLI-code. Deze coördinator is ook verantwoordelijk voor documentatie van de nationale regels op het Europese e-Justice-portaal, zodat juristen en informatici daarvan kennis kunnen nemen.¹⁰

Op dat e-Justice-portaal zal straks ook de ECLI-zoekinterface te vinden zijn. Alle partijen die uitspraken (her)publiceren, zowel gerechten als commerciële of non-profit databanken – zoals de Jurifast-databank van de Associatie van Raden van State¹¹ – kunnen deze doorzoekbaar laten maken door deze Europese zoekmachine. Nationale uitspraken krijgen daarvoor veel meer Europese exposure, onder meer omdat een (buitenlandse) gebruiker direct te zien krijgt waar eventuele vertalingen of samenvattingen te vinden zijn. Met dit ECLI-portaal wordt dus eindelijk praktisch invulling gegeven aan de eis die het Europees Hof van Justitie in 1982 in het Cilfit-arrest formuleerde: bij het besluit dat de juiste interpretatie van het gemeenschapsrecht zo evident is dat daarover aan het Hof geen vraag gesteld behoeft te worden, moet de nationale rechter ervan overtuigd zijn dat de gekozen interpretatie even evident is voor de rechterlijke instanties van de andere lidstaten.¹²

Naar verwachting gaat het Europese ECLI-zoekportaal nog in 2014 live. Natuurlijk zullen alle op Rechtspraak.nl gepubliceerde uitspraken daarin te vinden zijn, maar ook jurisprudentie van gerechten uit andere landen die ECLI inmiddels in hun publieke databanken hebben geïntroduceerd: Duitsland (*Bundesverwaltungsgericht*), Frankrijk (*Cour de Cassation* en *Conseil d'État*), Tsjechië (Hooggerichtshof), Slovenië (Hooggerichtshof en gerechtshoven) en Oostenrijk (diverse gerechten). Daarnaast zijn ook Spanje, Finland, Italië en Malta bezig met de introductie van ECLI. Het Europees Hof van Justitie heeft ECLI op 24 maart 2014 ingevoerd, en gebruikt ECLI ook voor de citaties van de eigen uitspraken. Daarbij wordt overigens het eerste deel van de ECLI-code – het woordje 'ECLI' – weggelaten. Dat is jammer, want daardoor zijn de ECLI's voor mensen moeilijker, en – belangrijker nog – voor computers helemaal niet meer te herkennen, hetgeen de vindbaarheid van deze citaties bemoeilijkt. De *Board of appeal* van het European Patent Office heeft al zijn beslissingen inmiddels van een ECLI voorzien en ten slotte is ook het Europees Hof voor de Rechten van de Mens vergoederd met de invoering.

Identificatie van wet- en regelgeving

Wetgeving neemt een niet te onderschatten rol in als grondslag voor, en 'kapstok' van allerlei soorten overheidsinformatie, zoals lagere en decentrale regelgeving, rechterlijke

¹⁰ <http://e-justice.europa.eu> > NL > rechtspraak > De Europese identificatiecode voor jurisprudentie (ECLI).

¹¹ <http://www.juradmin.eu> > Case law > Jurifast.

¹² ECLI:EU:C:1982:335.

uitspraken, beleidsinstructies, uitvoeringsvoorschriften, Kamerstukken en publieksfolders. Maar ook juridische vakliteratuur, wetenschappelijk onderzoek en voorlichtingsmateriaal van consumentenorganisaties is aan de wet te relateren. Omdat op eenduidige URI's gebaseerde links tussen al die informatie tot voor kort ontbraken, was het verzamelen van relevante informatie over bijvoorbeeld het boetebesluit ex artikel 17c AOW een moedeloos makende onderneming. De komst van *content-integrators* als Rechtsorde en Legal Intelligence bracht geen volledige verlichting, want je kunt er niet op vertrouwen dat die zoekmachines alle schrijfwijzen – officiële titel, citeertitel, afkorting of alias (zoals 'de Koppelingwet') – herkennen. Titels zijn bovendien niet uniek en de variëteit aan afkortingen en aliases is schier oneindig.¹³ Het gebruik van publicatienummers (in bijvoorbeeld Staatsblad, Tractatenblad of Staatscourant) is onhandig omdat daarmee niet naar geconsolideerde regelgeving wordt verwezen.

Om deze problemen op te lossen is in het publiek-private standaardisatieplatform Juriconnect, waarin ook de Rechtspraak participeert, de 'BWB-standaard' ontwikkeld. Dit is een systematiek waarmee gestandaardiseerd verwezen kan worden naar een specifiek artikel (of ander onderdeel) van een regeling die is opgenomen op Wetten.nl. Optioneel kan daarbij nog worden vermeld dat de verwijzing is bedoeld naar het artikel zoals geldend op moment T1, en zoals gezien op moment T2 (hetgeen vooral van belang is voor regelingen die met terugwerkende kracht zijn gewijzigd). Alhoewel de standaard ook voor mensen leesbaar is, is deze niet bedoeld om – zoals bij een ECLI – de tekst van de verwijzing te vervangen.¹⁴ Ook voor decentrale regelgeving is er een Juriconnect-standaard (de 'JCDR'), maar deze kent – nog – geen verwijzingen naar het artikelniveau.¹⁵ In november 2013 zijn de Juriconnect BWB-standaard, de JCDR en het Nederlandse ECLI-profiel opgenomen op de pas-toe-of-leg-uit-lijst van het Forum Standaardisatie. Gebruik van deze standaarden is daarmee verplicht voor overheden (bijvoorbeeld in specificaties voor aanbestedingen); niet-gebruik moet worden verklaard.

In navolging van ECLI is in 2012 door de Raad van ministers de *European Legislation Identifier* (ELI) vastgesteld als (eveneens vrijwillige) URI-standaard voor nationale en Europese regelgeving.¹⁶ ELI kan worden gecombineerd met nationale standaarden en zou een belangrijke rol kunnen spelen bij, onder meer, het inzichtelijk maken van de verbanden tussen Europese richtlijnen en nationale implementatieregelgeving. Verschillende landen zijn begonnen met de invoering van ELI, maar dit is nog niet zo vergevorderd als bij ECLI.

13 Zo bestaan er bijvoorbeeld wel acht regelingen met de naam 'Aanwijzing opsporingsambtenaren'.

14 Een voorbeeld: de Juriconnect-BWB-URI voor artikel 17c van de AOW zoals geldend op 10 juni 2014 luidt: `1.0:c:BWBR0002221&artikel=17c&g=2014-01-06`.

15 Op de website Juriconnect.nl zijn de standaarden tot in detail beschreven.

16 Council conclusions inviting the introduction of the European Legislation Identifier (ELI) (CELEX:52012XG1026(01)).

Overige bronnen

De parlementaire geschiedenis – vanaf 1995 – heeft op [Zoek.officielebekendmakingen.nl](http://zoek.officielebekendmakingen.nl) goede URI's, maar dat geldt alleen voor de documenten als geheel; het is nog niet mogelijk om te verwijzen naar een specifieke paragraaf van een memorie van toelichting. Ook rechtswetenschappelijke literatuur heeft nog geen URI-systematiek. Dit zijn wel zaken waaraan wordt gewerkt, onder meer in het kader van Juriconnect.

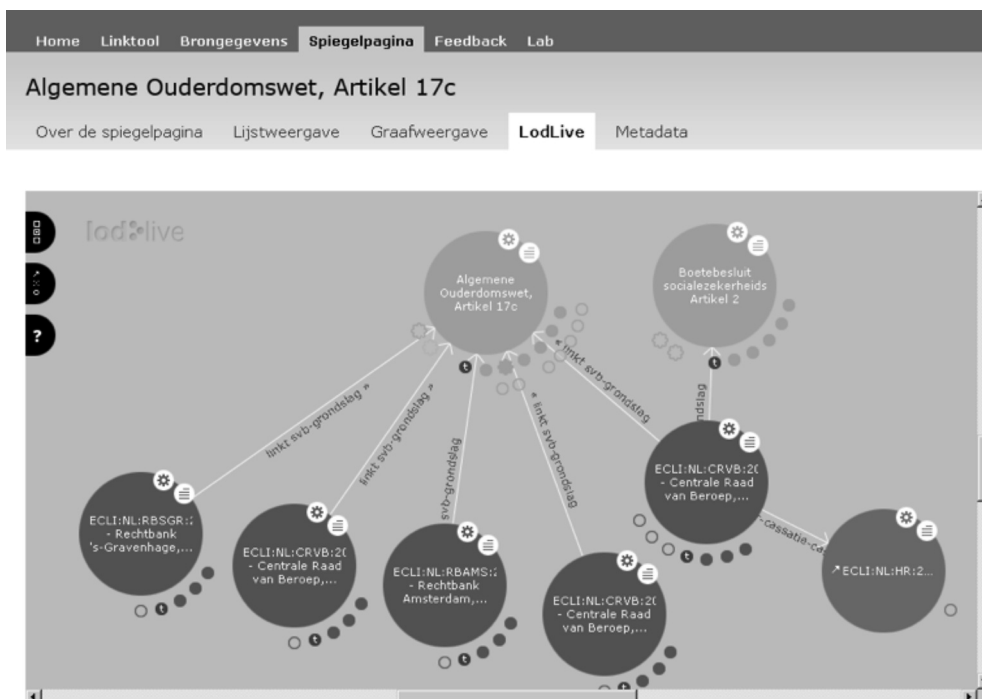
2.3 Linked data

De in de vorige paragraaf besproken URI's zijn er niet alleen om gegevens uniek te adresseren en citeren, vooral ook zijn ze een *condicio sine qua non* voor het creëren van linked data. Alvorens nader in te gaan op het creëren van links, kijken we eerst naar een project dat tot doel heeft om aan wet- en regelgeving gerelateerde overheidsdata als linked data aan te bieden, vooral ten behoeve van de overheid zelf.

Linked Data Overheid

Het project Linked Data Overheid (LiDO) heeft als doel om de verbanden tussen allerlei overheidsinformatie beter inzichtelijk te maken. Op de website linkeddata.overheid.nl zijn inmiddels de eerste resultaten zichtbaar. Naast de database waarin alle links worden verzameld, bestaat de LiDO-infrastructuur uit:

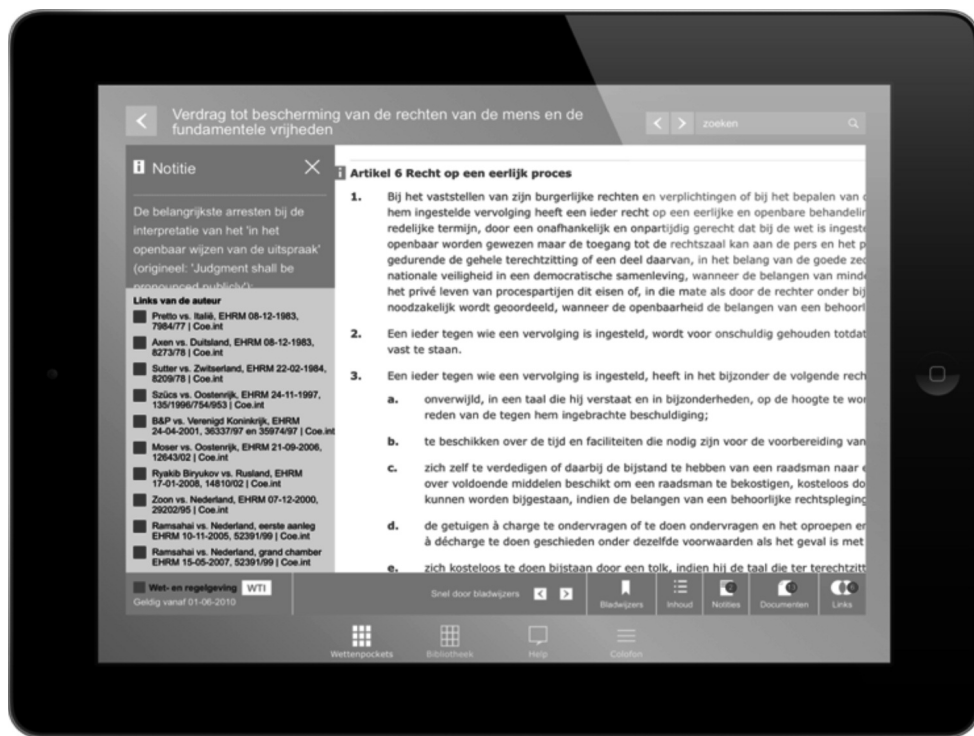
- een *linktool*, waarmee iedereen gestandaardiseerde verwijzingen kan maken. Deze linktool kan direct op de website worden gebruikt, maar ook worden geïntegreerd in redactie-omgevingen van overheidsorganisaties;
- visualisatie van gevonden links (zie het voorbeeld in figuur 1). Op basis van de beschikbare linked data kan iedereen – in zijn eigen IT-omgeving – zijn eigen soort visualisaties maken of de data op andere wijze integreren;
- een attenderingsservice, waarmee gebruikers met grote nauwkeurigheid kunnen worden gewaarschuwd indien er wijzigingen optreden in informatie die is gerelateerd aan de gegevens waarvoor zij zelf verantwoordelijkheid dragen.



Figuur 1. linkeddata.overheid.nl: visualisatie van (enkele) relaties tussen regelgeving en jurisprudentie inzake het boetebesluit ex artikel 17c AOW.

Steeds meer databronnen worden op LiDO aangesloten, zodat de waarde van de erin opgeslagen informatie gestaag groeit. Voor een willekeurige jurist mag de informatie op de website overvloedig, technisch en verwarrend overkomen, maar bedacht moet worden dat de waarde van deze informatie ontstaat als deze – na eventuele selectie – wordt geïntegreerd in de eigen vertrouwde omgeving. Omdat alles is gestandaardiseerd (URI's, metadata en technische interfaces) én omdat iedereen zijn kennis aan het 'ecosysteem' kan toevoegen, kunnen tegen relatief lage kosten hoogwaardige maatwerkproducten worden ontwikkeld. Een voorbeeld van zo'n linked-data-toepassing is de Wettenpocket.¹⁷

17 <http://wettenpocket.overheid.nl>.



Figuur 2. De Wettenpocket, met wettekst, commentaar en links naar jurisprudentie.

De Wettenpocket (zie de schermafbeelding in figuur 2) is een app voor tablets waarmee men wettenverzamelingen op specifieke rechtsgebieden altijd in de meest actuele versie bij de hand kan hebben. Deze verzamelingen zijn er in twee smaken: bundels en pockets. ‘Bundels’ bevatten automatisch verzamelde wetteksten van een bepaalde wetsfamilie, bijvoorbeeld de ‘Bundel BW Boek 7’ of de ‘Bundel Wet inkomstenbelasting 2001’. Naast de hoofdwet bevatten de bundels ook de gerelateerde AMvB’s, beleidsregels en ministeriële regelingen. ‘Pockets’ daarentegen worden niet automatisch gegenereerd, maar samengesteld door bij overheidsinstanties werkzame auteurs. Deze auteurs kunnen zelf een selectie maken van relevante (onderdelen van) regelingen, hyperlinks toevoegen en verklarende notities schrijven. Ook kunnen ze documenten die (nog) niet in de LiDO-omgeving zijn opgenomen, als pdf aan de pocket toevoegen. Voorbeelden van reeds gemaakte wettenpockets illustreren dat deze kunnen worden gemaakt op zeer specifieke aandachtsgebieden: ‘Slachtoffer in het strafprocesrecht’, ‘Elektronische overheidspublicaties’ of ‘Energie-investeringsaftrek’.

Gebruikers kunnen in gedownloade pockets en bundels zelf aantekeningen en bladwijzers aanbrengen. Bovendien – en daarin zit het grote voordeel van de verbinding met de *linked data cloud* – kan de gebruiker bij elk wetsartikel gerelateerde links zien, en kan hij direct doorklikken naar de LiDO-webpagina, zodat hij het hele informatienetwerk verder kan doorzoeken.

Links: extractie, kwaliteit en betekenis

Belangrijke vragen in linked-data-projecten zijn: hoe gaan we al die linkjes maken? Hoe controleren we de kwaliteit daarvan? En wat betekent het, inhoudelijk-juridisch, eigenlijk als twee informatieobjecten aan elkaar gelinkt zijn?

Idealiter worden links gemaakt door de citerende auteurs: wetgevingsjuristen brengen in wet en memorie van toelichting links aan naar wetsartikelen en parlementaire documenten, rechters maken links naar regelgeving en jurisprudentie. Zover zijn we echter nog niet: het is ingewikkeld en tijdrovend, en de door juristen gebruikte tekstverwerkers ondersteunen het nog niet. ‘Redactioneel linken’ gebeurt dan ook nog slechts voor een zeer beperkte groep documenten. Bovendien zitten we met miljoenen historische documenten die we ook graag van linkjes zouden willen voorzien.

Het alternatief is het gebruik van *link extractors*: software waarmee de computer verwijzingen in juridische teksten kan herkennen en omzetten naar de URI-gebaseerde links die ook door andere computers kunnen worden begrepen. Bij voorkeur maken we het deze link extractors zo eenvoudig mogelijk. Als een rechter een ECLI gebruikt om te verwijzen naar een arrest van de Hoge Raad, dan is een link extractor, dankzij het vaste patroon van de ECLI, in staat om daarvan een stabiele link af te leiden. Als de rechter niet een ECLI, maar een LJN, een of meerdere vindplaatsaanduidingen of een combinatie van zaaknummer, instantienaam en datum gebruikt, wordt het al een stuk lastiger, maar link extractors kunnen ook hieruit – dankzij alle gegevens die in het ECLI-register zijn opgeslagen – de correcte ECLI's afleiden. Ook verwijzingen naar nationale en Europese regelgeving – waarvoor juristen zo'n vijftig verschillende schrijfwijzen hebben verzonnen – kunnen door een link extractor vrijwel foutloos naar een nette URI worden vertaald.

Een tweede aspect van de link is de kwalificatie ervan: wat is de reden dat juist deze twee objecten met elkaar zijn verbonden? Dit kan interpretatie vragen, en daarom minder gemakkelijk door een link extractor te bepalen. Stel: een hofarrest (H) bevat verwijzingen naar twee rechtbankvonnissen (R1 en R2). R1 wordt aangehaald omdat het de aangevallen beslissing is. R2 heeft betrekking op een geheel andere zaak, maar wordt aangehaald om redenen van materiële analogie. Als de link extractor beide verwijzingen herkent, zal het moeten weten wat de aard is van de relatie tussen enerzijds H en R1 en anderzijds H en R2. Een verkeerde interpretatie zou er immers toe kunnen leiden dat bij R2 wordt vermeld dat deze in hoger beroep door H is vernietigd, hetgeen natuurlijk onjuist is.

Nu is het onderscheid tussen ‘formele’ (op de wet gebaseerde) en ‘materiële’ (inhoudelijk relevante) jurisprudentieverwijzingen nog objectief te omschrijven, maar dat wordt al lastiger bij wetsverwijzingen. Naar veel wetsartikelen wordt alleen om procesrechtelijke of motive-ringstechnische redenen verwezen (zoals de artikelen 8:75, 8:77 en 8:81 Abw of 47 en 57 Sr), zonder dat de verwijzende beslissingen enige uitspraak doen over die artikelen zelf. Als we alle door link extractors aangebrachte linkjes naar deze wetsartikelen zouden gebruiken om te zoeken naar jurisprudentie die voor de interpretatie van deze wetsartikelen relevant is, dan komen er tienduizenden irrelevante uitspraken terug. Dit voorbeeld wordt vaak aan-gehaald om het belang van automatische linkextractie te bagatelliseren, maar daarbij wordt vergeten dat een link extractor ook de verwijzingen naar tienduizenden andere wetsartikelen opspoort, die bij gebruik in een zoekvraag wel degelijk naar de inhoudelijk relevante juris-prudentie zullen leiden – een zoekactie die anders uren in beslag kan nemen.

Tegenovergesteld aan het vraagstuk van de ‘nutteloze’ procesrechtelijke verwijzingen is het probleem van impliciete links: een situatie die zich voordoet indien een uitspraak betrek-king heeft op een bepaald wetsartikel zonder het expliciet te noemen. Het aanbrenge van een dergelijke link is niet alleen subjectief, maar vergt ook aanzienlijke inhoudelijke kennis, van mens dan wel computer. Dat geldt evenzeer voor het linken naar specifieke (al dan niet letterlijk in een uitspraak benoemde) thema’s, of het specificeren van een gevonden mate-riële relatie tussen uitspraken en/of wetten (analogie, inhoudelijke tegenspraak, bevestiging, uitwerking, inperking, et cetera).

Het moge duidelijk zijn dat, net als gestandaardiseerde URI’s, link extractors weliswaar een onmisbare schakel zijn voor de structurering van juridische data, maar op zichzelf nog niet afdoende om de (juridische) gebruiker optimaal te faciliteren bij al zijn informatievragen. Maar ook de geschetste problemen rond het duiden van de betekenis van een link en de impliciete links zijn oplosbaar. Het gaat stap voor stap en de volgende heuvel wordt pas zichtbaar als de vorige is beklommen, maar we bewegen onmiskenbaar in de goede richting.

3 Rekenen met het recht

Vroeger, toen jurisprudentiecollecties nog uitsluitend op papier werden uitgegeven, was er regelmatig discussie over de selectie van de op te nemen uitspraken. In de beginjaren van Rechtspraak.nl werd deze discussie ook over deze online uitsprakendatabank gevoerd, maar de laatste jaren wordt steeds duidelijker dat het selectieprobleem zich verplaatst van de inputzijde naar de outputzijde: op Rechtspraak.nl staan nu ruim driehonderdduizend uit-spraken, en wekelijks komen er honderden bij. Hoe kunnen we in die enorme verzameling effectief zoeken en het kaf nog van het koren scheiden?

Natuurlijk zou het al een enorme winst zijn als er gezocht kan gaan worden met behulp van linked data: geef mij alle uitspraken waarin Verordening (EG) nr. 1408/71 wordt aangehaald,

of waarin wordt voortgebouwd op ECLI:NL:HR:2007:AX0771 – ongeacht natuurlijk de vraag hoe die verordening of dat arrest feitelijk is aangehaald.

Met behulp van de in paragraaf 2.3 beschreven link extractors is dit technisch goed realiseerbaar, maar dan weten we nog niet welke van de gevonden zoekresultaten echt belangrijk zijn: wat zijn de bagateluitspraken die louter om anekdotische of publicitaire redenen zijn geselecteerd, en wat zijn de beslissingen die werkelijk richting geven aan jurisprudentie en wetenschappelijk debat?

Om daarbij behulpzaam te zijn heb ik het eerder genoemde Model for Automated Rating of Case Law (MARC) ontwikkeld, waarvan de hoofdlijnen hierna worden samengevat. Daaruit zal duidelijk worden dat de ontwikkeling en toepassing van MARC een big-data-vraagstuk is, zoals gedefinieerd in paragraaf 1.2: het gaat om grote hoeveelheden data van allerlei (inhoudelijke en technische) aard, waarin continu mutaties optreden, die met de traditionele middelen niet verwerkbaar zijn, en waaruit nuttige kennis kan worden gedistilleerd die voorheen niet beschikbaar was.

3.1 Het theoretisch fundament: crowdsourcing

De bedoeling van MARC is om iedere uitspraak van een ‘MARC-indicator’ te voorzien die het juridisch belang ervan aanduidt. Ter illustratie kan worden verwezen naar de HUDOC-jurisprudentiedatabank van het Europees Hof voor de Rechten van de Mens.¹⁸ Een van de zoekfilters daar is de *importance level*; er zijn vier niveaus, waarbij de hoogste de in – om redenen van jurisprudentieel belang – de *case reports* opgenomen uitspraken omvat, en de laagste met name de kennelijk niet-ontvankelijk/ongegrondverklaringen.

Het gaat dus niet om de relevantie van een bepaalde uitspraak in relatie tot de letterlijke vraag die aan de zoekmachine is gesteld (de ‘systeemrelevantie’), of voor de oplossing van het probleem dat de gebruiker tot zijn zoekvraag aanzette (de ‘situationele relevantie’), maar uitsluitend om de juridische relevantie van een uitspraak in het algemeen (de ‘domeinrelevantie’).

Door de beperkte omvang van de HUDOC-databank en het zeer specialistische rechtsgebied kan een klein team van specialisten bij het Hof nieuwe uitspraken van een *ranking* voorzien, en deze eventueel bij oudere uitspraken wijzigen. Bij omvangrijker databanken, zoals Rechtspraak.nl of het E-archief, stuit dit op organisatorische en financiële grenzen.

We zouden dit kunnen oplossen als we het oordeel van de gehele juridische gemeenschap – in internettermen ook wel de *crowd* genoemd – over een bepaalde uitspraak zouden kunnen meten. Linked data bieden hier uitkomst. Veel links immers vormen de weerslag van de opinie van de juridische crowd: mensen die expert zijn op een bepaald rechtsgebied selecteren uitspraken – als redactielid – ter publicatie in een jurisprudentieperiodiek, of ze

18 <http://hudoc.echr.coe.int>.

halen deze – als wetenschapper – aan in een rechtsgeleerd artikel of proefschrift, of ze citeren deze – als rechter – in een latere uitspraak.

3.2 De onderzoeksdatabank: big en linked

De onderzoeksdatabank werd gevuld met alles waarop de digitale hand kon worden gelegd:

- de volledige uitsprakendatabank van Rechtspraak.nl, het E-archief en alle jurisprudentietijdschriften van Sdu en Kluwer. Na verwijdering van dubbele voorkomens bleven 850.000 uitspraken over;¹⁹
- alle rechtswetenschappelijke literatuur – annotaties, tijdschriftartikelen, handboeken en commentaaredities – uit de portalen van Sdu en Kluwer, alsmede de volledige inhoud van Wiki Juridica, bij elkaar 557.000 bestanden. Erg ‘hergebruikvriendelijk’ bieden de uitgevers de data trouwens niet aan; er waren vele weken nodig om de bestanden bij elkaar te schrapen en met speciaal ontwikkelde software ‘schoon te wassen’.
- 11.000 actualiteiten van Rechtspraak.nl;
- als hulpbestanden werd ook alle citeertitels en afkortingen van Wetten.nl, EUR-Lex, de Centrale Voorzieningen Decentrale Regelgeving en Rechtspraak-interne databanken ingelezen, bij elkaar 184.000 records. Ook dit vergde de nodige harmonisatie-inspanningen.

Al deze bestanden bevatten geen of – onvoldoende betrouwbare – linked data. De volgende stap was daarom het ontwikkelen van link extractors (zie par. 2.3), om jurisprudentiecitaties in uitspraken en rechtswetenschappelijke literatuur op te sporen, alsmede wetsverwijzingen in uitspraken. Door deze link extractors werd mijn onderzoeksdatabank gevuld met:

- 444.000 jurisprudentieverwijzingen in rechterlijke uitspraken;
- 680.000 jurisprudentieverwijzingen in rechtswetenschappelijke literatuur;
- 5.659.000 wetsverwijzingen in rechterlijke uitspraken naar 57.000 verschillende wetsartikelen.

Daarnaast werden nog tal van andere data verzameld: vindplaatsen van uitspraken in jurisprudentieperiodieken, het tijdsverloop tussen uitspraakdatum en latere aanhaling, de lengte van de uitspraak, enkel- of meervoudigheid van de kamer, rechtsgebied, et cetera.

De schat aan gegevens die dit opleverde, was niet alleen relevant voor de ontwikkeling van MARC, maar leverde ook andere intrigerende kennis op, bijvoorbeeld over selectiebeleid en publicatiesnelheid van jurisprudentieperiodieken en Rechtspraak.nl, citatiepraktijk en kruisverwijzingen tussen verschillende gerechten,²⁰ en over Europeesrechtelijke verwijzingen in rechtspraak en literatuur.²¹

¹⁹ Getallen zijn afgerond op duizendtallen.

²⁰ Van Opijnen 2011a.

²¹ Van Opijnen 2011b.

3.3 Het model

Met al deze miljoenen gegevens in de databank kon de ontwikkeling van MARC beginnen. Na alle ellende die reeds bij het verzamelen van de data was ondervonden, begonnen hier de big-data-nachtmerries pas echt. Voor linked big data kan ‘sociaalnetwerkanalyse’ worden gebruikt om betekenisvolle relaties op te sporen en in kaart te brengen. De software die bedrijven als Google, Facebook en LinkedIn daarvoor hebben ontwikkeld is echter niet commercieel verkrijgbaar en de in de academische wereld ontwikkelde programma’s bleken nauwelijks geschikt voor echt grote datasets – begeeft de software het niet uit zichzelf, dan zit er wel een geheugenlek in waardoor je hardware crasht. Door de zeldzame combinatie van disciplines waren helpende handen nauwelijks te vinden.

Overigens bleek sociaalnetwerkanalyse uiteindelijk weinig nuttige indicatoren op te leveren. Statistisch significante resultaten kwamen er wel uit een geavanceerde variant van regressieanalyse. Eenvoudig gezegd kan met regressieanalyse worden voorspeld hoe groot de kans is dat een bepaald verschijnsel (de ‘regressor’) zich voordoet, gegeven een of meerdere variabelen (de ‘predictors’). Regressieanalyse kan worden geïllustreerd met een voorbeeld uit de medische wetenschap: hoe groot is de kans dat iemand een bepaalde ziekte (de regressor) zal ontwikkelen, gegeven geslacht, genetische achtergrond, leeftijd, ziektegeschiedenis en leefpatroon (de predictors)?

Voor rechterlijke uitspraken kan niet met één regressiemodel worden volstaan, omdat veel van de variabelen die we als predictor zouden willen gebruiken, ten tijde van de publicatie van de uitspraak (vaak de dag van uitspraak zelf) nog niet bestaan: de uitspraak is dan nog niet gepubliceerd, nog niet geannoteerd en nog niet geciteerd. Om dit probleem te onderwerpen is een onderscheid gemaakt in drie perioden. De kenmerken van deze perioden en de daarvoor gebruikte variabelen zijn samengevat in figuur 3 en worden hier besproken.

	Publicatieperiode	Transitieperiode	Citatieperiode
Karakter	Uitspraak ziet levenslicht	Studie en commentaar	Voortdurende roem of geleidelijke vergetelheid
Duur	Een week	100 dagen	Oneindig
Regressor	Publicatie m.u.v. Rechtspraak.nl	Gewogen gemiddelde van: <ul style="list-style-type: none"> • MARC-publicatieperiode • MARC-citatieperiode, afhankelijk van dag in transitieperiode 	Citatie in rechtspraak en eenmalige literatuur in komende drie jaar
Predictors	<ul style="list-style-type: none"> • Uitgaande jurisprudentiecitaties • Uitgaande wetsverwijzingen • Publicatie op Rechtspraak.nl • Nieuwsbericht op Rechtspraak.nl • Enkel-/meervoudige kamer • Lengte • Hiërarchische positie gerecht • Rechtsgebied 		<ul style="list-style-type: none"> • Publicatie (gewogen, logaritmisch) • Annotatie (idem) • Citatie in continue literatuur (logaritmisch) • Citatie in eenmalige literatuur (logaritmisch en met gewogen voortschrijdend gemiddelde) • Citatie in rechtspraak (idem) • Leeftijd • Hiërarchische positie gerecht • Rechtsgebied

Figuur 3. Kenmerken, regressors en predictors van de drie opeenvolgende perioden in het leven van een rechterlijke uitspraak.

Periodes

Wanneer een uitspraak door de rechter wordt gedaan, breekt de eerste periode aan: de publicatieperiode. Misschien komt de uitspraak niet verder dan het E-archief, maar over het algemeen worden belangrijke beslissingen in deze periode op Rechtspraak.nl gepubliceerd. In de transitieperiode, die na een week begint en honderd dagen duurt, wordt de uitspraak door de juridische crowd gelezen en op zijn belang voor de algemene rechtsopvatting beoordeeld: door selectie voor jurisprudentieperiodieken, door vermelding in vakliteratuur en misschien reeds door citatie in andere uitspraken. In de citatieperiode die vervolgens aanbreekt zal blijken in hoeverre de uitspraak blijvend een plek heeft veroverd in het collectief juridisch geheugen. Ik begin de bespreking bij deze laatste, in beginsel oneindige periode. Eerst evenwel nog een terminologische verheldering. Er is onderscheid gemaakt in twee soorten rechtswetenschappelijke literatuur. In de eerste plaats is er ‘eenmalige literatuur’: werk dat slechts eenmaal wordt geschreven en gepubliceerd, en vervolgens nooit meer wordt herzien. Tijdschriftartikelen, noten in jurisprudentieperiodieken en proefschriften zijn sprekende voorbeelden. Daartegenover staat ‘continue literatuur’: publicaties – vroeger losbladig, tegenwoordig vooral online – die beogen om continu de actuele staat van het recht te beschrijven en daartoe frequent worden bijgewerkt.

Citatieperiode

Voor de citatieperiode is als regressor gekozen: de mate waarin een uitspraak in de daaropvolgende drie jaar geciteerd zal gaan worden in rechtspraak en eenmalige literatuur. De volgende variabelen blijken daarvoor een voorspellende waarde te hebben.

1. Publicatie in jurisprudentieverzamelingen, zowel commercieel als Rechtspraak.nl
Omdat selectie voor de Nederlandse Jurisprudentie (NJ) een ander gewicht geeft aan een uitspraak dan selectie voor de Praktijkids, is de ‘JurisprudentiePeriodiekWegings-Factor’ (JPWF) ontwikkeld. De JPWF, die dus de status van een tijdschrift uitdrukt, is berekend op basis van het aantal malen dat de erin gepubliceerde uitspraken, gemiddeld genomen, later worden geciteerd. Voor alle periodieken en databanken is de JPWF berekend. De NJ heeft een JPWF van 3,16, de Praktijkids scoort 0,94 en Rechtspraak.nl 0,85.
2. Annotatie
Over het algemeen worden alleen de belangrijkste uitspraken in jurisprudentietijdschriften van een noot voorzien. Dit gegeven is dan ook als een afzonderlijke voorspellende variabele opgenomen. Voor de berekening van het gewicht van de annotatie is de JPWF van het publicerende tijdschrift gebruikt.
3. Citatie in continue literatuur
Uitspraken die belangrijk zijn voor de juridische doctrine zullen worden vermeld in een of meer van de ‘continue’ commentaaredities, zoals Tekst & Commentaar, OpMaat of Groene Serie.
4. Citatie in eenmalige literatuur
Anders dan bij continue literatuur – waar een verwijzing gewoon zal worden verwijderd als deze z’n belang verloren heeft – moet bij analyse van citaties in eenmalige literatuur rekening worden gehouden met de factor ‘tijd’. Stel: twee uitspraken van acht jaar oud zijn twintig keer in wetenschappelijke publicaties aangehaald, maar bij uitspraak A hebben al die citaties vrijwel direct na de uitspraakdatum plaatsgevonden, terwijl bij uitspraak B die citaties verspreid over de tijd, tot aan een maand geleden, zijn gemaakt. Verondersteld mag dan worden dat, heden ten dage, uitspraak B belangrijker is dan uitspraak A. Om dit tijdsaspect tot uitdrukking te brengen is gebruikgemaakt van een gewogen voortschrijdend gemiddelde: hoe recenter een citatie, des te zwaarder deze meetelt.

5. Citatie in rechtspraak

Eenzelfde rekenmethodiek werd gebruikt voor weging van de citaties van een uitspraak in latere jurisprudentie. Een onderscheid naar citerende instantie (mede in relatie tot de geciteerde instantie) is daarbij overigens niet gemaakt.

6. Leeftijd

Los van de onder de vorige twee punten genoemde actualiteit van eventuele citaties is ook de leeftijd van een uitspraak zelf een factor waarmee rekening moet worden gehouden; de kansen van een ongeciteerde uitspraak om alsnog te worden aangehaald zullen dalen naarmate de uitspraak ouder is, en dat geldt ook voor wel-geciteerde uitspraken.

7. Rechterlijke hiërarchie

De hiërarchische positie van het uitsprekende gerecht is natuurlijk ook een factor om rekening mee te houden. Om het model niet overgecompliceerd te maken zijn vijf categorieën gebruikt: internationale/Europese gerechten, hoogste nationale gerechten (Hoge Raad, Centrale Raad van Beroep, College van Beroep voor het bedrijfsleven en Afdeling bestuursrechtspraak van de Raad van State), appelinstanties (gerechtshoven), gerechten van eerste aanleg (rechtbanken en oude kantongerechten) en als laatste de niet-rechterlijke en buitenlandse instanties.

8. Rechtsgebied

Ten slotte mag ook van het rechtsgebied invloed worden verwacht. Een eenvoudige indeling in drie rechtsgebieden is gehanteerd: civiel recht, strafrecht en bestuursrecht.

Al deze factoren blijken statistisch significant te zijn, maar de mate waarin verschilt nogal. In figuur 4 is een overzicht opgenomen met de *Wald statistic*, een statistische maat waarmee het onderlinge gewicht van de variabelen kan worden uitgedrukt. Hieruit blijkt dat citatie in latere rechtspraak de belangrijkste variabele is; maar ook de andere variabelen leggen voldoende gewicht in de schaal.

De statistische maat R^2 voor de citatieperiode is 0,47. Dat wil zeggen dat 47% van de variantie in de regressor (dus de mate waarin een uitspraak de komende drie jaar zal worden geciteerd in rechtspraak en eenmalige literatuur) kan worden verklaard uit deze acht predictor-variabelen.

Variabele	Wald statistic
Citatie in rechtspraak	454806
Citatie in continue literatuur	27191
Hiërarchische positie gerecht	20738
Leeftijd	19944
Publicatie in periodieken	6067
Rechtsgebied: strafrecht	
Rechtsgebied: civiel recht	4701 ²²
Rechtsgebied: bestuursrecht	
Citatie in eenmalige literatuur	2615
Annotaties	1513

Figuur 4. Wald statistic voor de variabelen in de citatieperiode.

Publicatieperiode

Indien een uitspraak belangstelling krijgt in literatuur en rechtspraak, gaat daaraan over het algemeen publicatie in een of meerdere tijdschriften vooraf; redacties zijn de eerste die interessante uitspraken signaleren. Daarom is voor de publicatieperiode, de eerste week na de uitspraak, de selectie voor tijdschriften als regressor gebruikt. Anders geformuleerd: hoe groot is de kans dat een uitspraak in (belangrijke) tijdschriften wordt gepubliceerd? De eerder besproken JurisprudentiePeriodiekWegingsFactor (JWPF) ligt aan deze regressor ten grondslag, maar dan zonder Rechtspraak.nl, die we in deze fase juist als predictor gebruiken.

Omdat de uitspraak in de publicatieperiode door de juridische crowd nog niet is bestudeerd, kunnen we ‘exogene variabelen’ als citatie en annotatie niet gebruiken, en moeten we ons in plaats daarvan verlaten op ‘endogene variabelen’: gegevens in de uitspraak zelf. Ook hier zitten veel linked data bij:

1. Uitgaande jurisprudentieverwijzingen

Als een uitspraak veel wordt geciteerd in latere rechtspraak heet zo’n uitspraak – in sociaalnetwerkterminologie – een *authority*. Als een uitspraak daarentegen zelf veel naar eerdere uitspraken verwijst, noemen we het een *hub*. Ook dat blijkt een indicatie voor juridische relevantie. Gemiddeld genomen zal een rechter immers meer jurisprudentie aanhalen als de casuïstiek lastig en de zaak juridisch ingewikkeld is; daarmee wordt de kans vergroot dat de uitspraak zelf later een *authority* zal worden. Uitgaande jurisprudentiecitatie blijken inderdaad significant meer voor te komen bij uitspraken die later in de citatieperiode goed scoren.

22 Civiele uitspraken bleken de zwaarste weging te krijgen, gevolgd door bestuursrecht en strafrecht.

2. Uitgaande wetsverwijzingen

Op vergelijkbare wijze kan ook het aantal uitgaande wetsverwijzingen worden gebruikt: hoe meer wetsartikelen worden aangehaald, hoe ingewikkelder de juridische casus geacht kan worden te zijn. Hierbij is een correctie toegepast voor de in paragraaf 2.3 besproken problematiek van wetsartikelen die vooral om procesrechtelijke of motiveer- ringstechnische redenen worden aangehaald: alle verwijzingen naar de ruim 250 wets- artikelen die meer dan 2.000 keer werden aangehaald zijn buiten beschouwing gelaten.

3. Publicatie op Rechtspraak.nl

Het oordeel van de rechter over het belang van haar eigen vonnis, dat tot uiting komt in publicatie op Rechtspraak.nl, vormt vanzelfsprekend een belangrijke voorspellende variabele.

4. Actualiteit op Rechtspraak.nl

Ook de plaatsing van een nieuwsbericht op Rechtspraak.nl over de uitspraak kan worden geacht iets te zeggen over het juridisch gewicht van de uitspraak.

5. Enkel- of meervoudige kamer

Juridisch complexe zaken worden over het algemeen door de meervoudige kamer afge- handeld, eenvoudiger zaken door de enkelvoudige kamer. Hoe complexer, hoe groter de kans op een juridisch belangrijke uitspraak, dus ook dit aspect wordt in ogenschouw genomen.

6. Lengte van de uitspraak

Ook de lengte van een uitspraak kan iets zeggen over het gewicht ervan. Hoe meer woorden er nodig zijn voor een uitspraak, hoe groter de kans dat het een belangrijke uitspraak betreft.

7. Rechterlijke hiërarchie

De hiërarchische positie van het uitsprekende gerecht kan, net als in de citatieperiode, in de publicatieperiode als voorspellende variabele worden gebruikt.

8. Rechtsgebied

Hetzelfde geldt voor het rechtsgebied, hierbij is dezelfde indeling gebruikt als voor de citatieperiode.

Analyse leert dat al deze variabelen statistisch significant zijn, met uitzondering van de actualiteit op Rechtspraak.nl. Vervolgens kan ook hier de Wald statistic worden gebruikt om het belang van de verschillende predictors te vergelijken (zie figuur 5).

Variabele	Wald statistic
Hiërarchische positie gerecht	12130
Rechtsgebied: strafrecht	} 5447 ²³
Rechtsgebied: civiel recht	
Rechtsgebied: bestuursrecht	
Enkel-/meervoudige kamer	5367
Uitgaande jurisprudentieverwijzingen	4654
Publicatie op Rechtspraak.nl	1278
Uitgaande wetsverwijzingen	1095
Lengte van de uitspraak	358

Figuur 5. Wald statistic voor de variabelen in de publicatieperiode.

Hierbij valt ten eerste op dat de verschillen in het relatieve belang van de variabelen in de publicatieperiode minder groot zijn dan in de citatieperiode. Van die variabelen blijken de aard van het gerecht en de omvang van de behandelende kamer van groot belang. Opvallend is verder de rol van uitgaande jurisprudentiecitaties; wetsverwijzingen daarentegen zijn van minder gewicht. R^2 komt voor de publicatieperiode uit op 0,28, dat wil zeggen dat 28% van de variantie in de regressor (de publicatie in jurisprudentieperiodieken) verklaard kan worden uit deze zeven variabelen.

Transitieperiode

Als we de citatieperiode direct op de publicatieperiode zouden laten aansluiten, dan zou de MARC-indicator gedurende de eerste maanden grote fluctuaties kunnen vertonen. Een nieuwe publicatie in een tijdschrift, vermelding in verschillende werken van continue literatuur, annotaties: vrijwel dagelijks zou de rating van de uitspraak grote schommelingen ondervinden, hetgeen de gebruiker in verwarring zou kunnen brengen. Teneinde deze overgang meer geleidelijk te laten verlopen is de transitieperiode geïntroduceerd: de periode van honderd dagen waarin het belang van de uitspraak door de juridische crowd wordt gewikt en gewogen.

De rating in de transitieperiode wordt berekend op basis van het gewogen gemiddelde van publicatieperiode en citatieperiode, waarbij de weging is gebaseerd op het tijdstip binnen de transitieperiode. Omdat de transitieperiode honderd dagen lang is, zal op bijvoorbeeld de 25^{ste} dag ervan de rating uit de publicatieperiode voor 75% meetellen, en die van de citatieperiode voor 25%.

Modelmatige vereenvoudiging

De rating die het model voor iedere uitspraak oplevert is een positief of negatief getal met heel veel cijfers achter de komma, dat de gebruiker van een uitsprakendatabank niet veel zal

23 Het relatief gewicht van de respectieve rechtsgebieden bleek gelijk als in de citatieperiode (zie voetnoot 20): civiel, bestuur, straf.

zeggen. In een volgende stap worden de ratings daarom gegroepeerd in vijf categorieën – van MARC-1 tot en met MARC-5 – gevisualiseerd met sterren, vergelijkbaar met de rating van hotels op Booking.com, boeken op Amazon.com, en films of theatervoorstellingen in een krantenrecensie.

Terwijl alle voorgaande stappen zijn gebaseerd op statistische berekeningen, is het trekken van een grens tussen de verschillende categorieën een subjectieve aangelegenheid, die onder meer afhangt van de aard en de omvang van de databank waarvan de uitspraken worden geclassificeerd: een databank waarin alleen de uitspraken van de tien meestgelezen jurisprudentieperiodieken zitten, zal verhoudingsgewijs veel meer MARC-5-uitspraken kennen dan het volledige E-archief. Voor de onderzoeksdatabase (zie de beschrijving in par. 3.2), waarin zowel gepubliceerde als ongepubliceerde uitspraken zaten, is gekozen voor een verdeling van MARC-1 tot MARC-5 van respectievelijk: 60%, 25%, 11%, 3,5% en 0,5%.

Verhouding tussen publicatieperiode en citatieperiode: de proef op de som

Het model voor de citatieperiode is met name gebaseerd op het oordeel van de juridische crowd, voor de publicatieperiode zijn met name endogene variabelen gebruikt. De vraag rijst of predictors uit de publicatieperiode, zoals wetsverwijzingen en uitgaande jurisprudentiecitaties, nog een extra voorspellende waarde zouden kunnen hebben in de citatieperiode. Het antwoord op die vraag blijkt ontkennend te luiden, en daarmee lijken de voorspellende variabelen van de publicatieperiode een goede indicator voor het oordeel dat in de citatieperiode door de domeindeskundigen wordt geveld.

Dat blijkt overigens ook op een andere manier. In figuur 6 is te zien in hoeverre uitspraken in zowel publicatieperiode als citatieperiode in dezelfde MARC-categorie terecht komen. Dit blijkt voor 87,5% van de uitspraken het geval. 11,9% schuift in de citatieperiode een categorie omhoog of omlaag, terwijl bij slechts 0,6% van de uitspraken het verschil twee categorieën of meer is.

Citatiefase	Publicatiefase					Totaal
	MARC-1	MARC-2	MARC-3	MARC-4	MARC-5	
MARC-1	71,1	0,1	0,0	0,0	0,0	71,2
MARC-2	3,9	11,1	0,9	0,0	0,0	15,8
MARC-3	0,0	4,8	4,8	1,2	0,0	10,9
MARC-4	0,0	0,5	0,7	0,4	0,2	1,7
MARC-5	0,0	0,0	0,1	0,1	0,1	0,3
Totaal	75,0	16,5	6,5	1,7	0,3	100,0

Figuur 6. Kruistabel met procentuele verdeling van de uitspraken uit 2006 over de vijf MARC-categorieën in publicatiefase en citatiefase. N=81.710.

4 De toekomstige rol van open, linked en big data binnen de rechtspraak

In hoofdstuk 1 is beschreven wat open, linked en big data zijn, in hoofdstuk 2 zijn de belangrijkste ontwikkelingen beschreven op het gebied van standaardisatie en beleid. Hoofdstuk 3 bevatte een van statistische details ontdane samenvatting van het Model for Automated Rating of Case Law, dat zou kunnen bijdragen aan een betere toegankelijkheid van grote jurisprudentiedatabanken zoals E-archief en Rechtspraak.nl.

In dit afsluitende hoofdstuk worden enkele prangende vragen inzake de relatie tussen rechtspraak en open, linked, en big data geïnventariseerd: allereerst enkele veelgestelde vragen inzake het MARC-systeem, vervolgens meer generieke vraagstukken.

Is de MARC-rating eenvoudig te manipuleren, bijvoorbeeld door je eigen uitspraken veel te citeren?

Niets is onmogelijk, maar eenvoudig is het niet. Een eerste hindernis wordt gevormd door het statistisch model, dat slechts een beperkte waarde toekent aan citatie. Om een uitspraak effectief te 'boosten' zal deze ook moeten worden gepubliceerd en geannoteerd in jurisprudentieperiodieken en geciteerd in wetenschappelijke literatuur. Een tweede obstakel wordt gevormd door de grote getallen: een enkel citatietje valt weg in de grote massa. En mocht uit nadere analyse blijken dat (sommige) rechters hun eigen uitspraken onevenredig vaak citeren, dan passen we het model gewoon zo aan dat dergelijke zelfcitaties worden genegeerd.

Ik ken korte rechtbankuitspraken zonder wets- en jurisprudentieverwijzingen die toch baanbrekend genoemd kunnen worden. In MARC zouden ze laag scores, dan klopt het model toch niet?

Zowel in het recht als in de statistiek geldt dat het enkele bestaan van een uitzondering geen reden mag zijn de regel van onwaarde te verklaren. Enigszins gesimplificeerd gezegd ontstaan statistische modellen door in een 'puntenwolk' die rechte lijn te trekken waarbij de gemiddelde afstand van alle punten tot de lijn het kleinst is. Maar hoe je die lijn ook trekt, er zullen altijd punten zijn die ver van de lijn af liggen.

De kracht van het MARC-systeem is nu juist dat een kennelijk onjuiste plaats van een uitspraak in die puntenwolk door de juridische gemeenschap zelf wordt gecorrigeerd: in de publicatieperiode zal de in de vraag opgevoerde uitspraak slecht scoren, maar als deze inderdaad baanbrekend is, dan zullen publicatie, annotatie en citatie ervan binnen enkele weken tot een hogere rating leiden. MARC is weliswaar een computermodel, maar het is gebaseerd op het inzicht van duizenden deskundige juristen.

Zijn er niet meer variabelen die een rol zouden moeten spelen in MARC?

Alleen al vanuit statistisch perspectief moet die vraag bevestigend worden beantwoord. De R^2 – die een maximale waarde kent van 1 – van 0,47 voor de citatieperiode en zeker die van

0,28 voor de publicatieperiode is voor verbetering vatbaar. Er zijn nog genoeg variabelen waarvan aannemelijk is dat ze een rol spelen. In hoeverre is tegen een uitspraak beroep aangekondigd? Zo ja, wat is de uitslag van dat beroep (vernietigingen zijn relevanter dan bevestigingen)? Kunnen we nader specificeren naar rechtsgebied, onderwerp en/of behandelende kamer/sector/team? Wat is het financieel belang? Zoekt de rechter steun in doctrine of wetsgeschiedenis?

Het expliciteren van al deze gegevens (afgeleid uit de uitspraakteksten of door koppeling met registratieve systemen) vergt echter inspanningen die dit beperkte – doch nu reeds zeven jaar durende – onderzoek te boven gingen. Alhoewel verdere uitwerking ongetwijfeld tot verdere verfijning en verbetering van het model zal leiden, lijkt het huidige model reeds voldoende betrouwbaar om in praktijk te worden toegepast.

Waarom staat de MARC-indicator dan nog niet op Rechtspraak.nl?

De wetenschappelijke vraag luidde of het mogelijk was een dergelijk model te ontwikkelen, de implementatievraag is van geheel andere orde. MARC inbouwen op Rechtspraak.nl is geen *no brainer*, want we hebben het over big data: dagelijks moeten grote hoeveelheden gegevens worden verwerkt, afkomstig uit verschillende bronnen, van een grote variëteit en continu aan verandering onderhevig. Introductie vergt een weloverwogen besluit, gebaseerd op het besef dat dit geen standaard IT is. Bovendien zou de introductie van MARC geen doel op zich moeten zijn, maar onderdeel van een bredere strategie.

Ik lees dat er nu een databank is met miljoenen linkjes tussen uitspraken, wetgeving en rechtswetenschappelijke literatuur. Waarom beschik ik dan nog niet over al die 'handige knoppen' die daar kennelijk op zijn gebaseerd?

Het creëren van al die linkjes was een essentiële tussenstap voor MARC, maar ook zonder MARC-indicator is met die linkjes al veel nuttigs te doen, zoals het maken van citatie-indexen en het zoeken op aangehaalde wetsartikelen. Implementatie daarvan is minder ingewikkeld dan de MARC-indicator, maar ook een link extractor is niet echt eenvoudige software. Er bestaan geen standaardpakketten voor en het bouwen vergt diepgaande kennis van *natural language processing*, linked-data-technologie en juridische verwijzingstechnieken. Ook hier geldt dat niet zonder meer de voor het onderzoek ontwikkelde software kan worden gebruikt; een productieversie stelt aanvullende eisen aan snelheid, stabiliteit en onderhoudbaarheid. Het goede nieuws is evenwel dat met de 'intelligente kern' van de ontwikkelde link extractor thans een productieversie wordt ontwikkeld in de LiDO-omgeving. Met de miljoenen linkjes die hieruit voortkomen kan iedereen, ook de Rechtspraak, zijn eigen 'handige knoppen' gaan maken.

Moeten we als rechtspraak ons eigenlijk met dit soort dingen bezighouden? Is het niet aan commerciële uitgevers om dergelijke producten te ontwikkelen?

Het antwoord is tweeledig. In de eerste plaats blijken uitgevers dit soort ontwikkelingen uit zichzelf nauwelijks op te pakken. Natuurlijk is de ene uitgever de andere niet, maar over het algemeen is het innovatief vermogen van juridische uitgevers – zowel in Nederland als in het buitenland – relatief gering; ze leunen op traditionele, inmiddels bijna ouderwetse bedrijfsmodellen, zijn weinig wendbaar en lijken geen heldere visie te hebben op de centrale rol die informatietechnologie zou moeten spelen in hun *core business*. Voor echte innovatie is nodig dat uitgevers meegaan in de internetmentaliteit van geven en nemen: een ruimhartiger ondersteuning van open (URI-)standaarden, en bereidheid om ook een deel van de eigen content – in ieder geval de metadata daarvan – voor derden open te stellen. Een tweede reden voor een ontkennend antwoord ligt in de erkenning dat het ook voor de rechtspraak gaat om core business. Het is illustratief om in dit verband nog even de ultieme ‘handige knop’ uit Rechtstreeks 2010/1 aan te halen: ‘Zou het niet handig zijn als u bij het begin van een niet-standaard zaak een overzicht krijgt van alle relevante literatuur, wetgeving en recente jurisprudentie? Het hoeft niet, maar ook dat kan. Zonder zelfs maar op een knop te drukken.’

Om te bepalen wat in een concrete casus onder ‘relevant’ moet worden verstaan, moet er niet alleen gebruik worden gemaakt van publieke databanken en commerciële kennisportalen, maar moet er ook toegang zijn tot de primair-processystemen (zoals Berber, Civiel en Compas), het E-archief en andere interne systemen, allemaal gevuld met privacygevoelige gegevens. Belangrijker nog is dat de selectie van de ‘relevante’ informatie raakt aan het hart van het rechterlijk bedrijf: zaken waarin eerder vergelijkbare verweren zijn gevoerd, waarin dezelfde *repeat player* een rol speelt, waarin bepaalde tekstblokken zijn gebruikt, die dezelfde concipiënt of auditeur hebben, et cetera. Daarbij gaat het niet alleen om de ‘bedrijfsgeheimen’ van de rechter, het betreft systemen die in het informatietechnologische hart staan van het rechterlijk bedrijf zoals zich dat de komende jaren en decennia zal gaan ontwikkelen. Dat mag en wil je niet uitbesteden.

Wordt het allemaal niet een beetje te veel IT? Hoe zorgen we ervoor dat we als rechter de regie houden?

Open data, linked data, big data: het zijn geen hypes, maar onomkeerbare trends in een steeds meer datagedreven maatschappij. Sterker nog, het zijn *disruptive technologies* waarvan de invloed moeilijk kan worden overschat. Er zijn inmiddels voorbeelden te over van sectoren die worden ‘overvallen’ door *game changers* die weten te profiteren van deze trends. Spotify maakte een eind aan de cd-winkels, Netflix is funest voor de video/dvd-shops, LinkedIn neemt de plaats in van klassieke arbeidsbemiddeling en Uber en Airbnb zetten een breekijzer in taxi- respectievelijk hotelmarkt.

De wereld van de rechtspraak zal veranderd gaan worden door *online dispute resolution*,

waarbij het niet vanzelfsprekend is dat de overheidsrechter het monopolie krijgt. Gelukkig worden in het KEI-programma thans stappen gezet om de rechtspraak in deze digitale ontwikkeling mee te laten gaan. Uit het antwoord op de vorige vraag bleek reeds dat daarbij de focus niet alleen zou moeten liggen op het digitaliseren van dossiers en het procedurele verkeer met burgers en juridische professionals, maar ook op het erkennen van alle daaruit voortvloeiende data als *asset*, als productiemiddel. Zo bezien zijn citatie-indexen en MARC-indicators geen doel op zich, maar slechts enkele van de nuttige uitvloeisels van een veel breder ‘databeleid’. Zo kan computeranalyse van (grote hoeveelheden) uitspraken, zeker indien deze reeds bij het concipiëren van een voor computers herkenbare structuur zijn voorzien, een schat aan informatie opleveren over de kwaliteit van het rechterlijk werk, de feitelijke keuzes die daarbij worden gemaakt en de juridische en maatschappelijke tendenzen die eruit zijn af te leiden.

Als rechter houd je de regie door al die data niet als het grote kwaad, maar als vriend en bron van kennis tegemoet te treden.

Zijn er meer datasets die de Rechtspraak als open data zou moeten aanbieden?

Onderdeel van zo’n ‘meebewegende’ houding is de erkenning dat je meer kennis uit data kunt distilleren als je ze deelt met anderen. Die anderen kunnen jouw gegevens combineren met andersoortige data, of onderzoek doen waar je zelf de tijd, middelen of expertise niet voor hebt. Bovendien is een goed open-databeleid onmisbaar om invulling te geven aan de beleidsdoelstelling ‘transparantie’.

Een eerste voorbeeld van een dataset die ‘open’ gemaakt zou kunnen worden is het Register beroepsgegevens en nevenbetrekkingen, dat weliswaar voor iedereen op Rechtspraak.nl raadpleegbaar, maar niet als open data beschikbaar is. Door de beperkingen in het zoekformulier kan de burger niet alle vragen beantwoord krijgen die hij van een transparante rechtspraak mogelijk verwacht. Zo kan men wel te weten komen welke nevenfuncties een specifieke rechter uitoefent, maar kan men geen antwoord krijgen op de vraag hoeveel rechters een (betaalde) nevenfunctie hebben bij een bepaalde juridische uitgeverij of verzekeringsmaatschappij. Beschikbaarstelling als open data zou dat probleem kunnen ondervangen.

Een tweede voorbeeld kan gevonden worden bij de Databank Consistente Straftoemeting, oorspronkelijk ontwikkeld om gehoor te geven aan de maatschappelijke roep om meer inzicht in de consistentie van de straftoemeting en de variabelen die daarbij een rol spelen. De databank is nooit een succes geworden, maar het als open data beschikbaar stellen van relevante – en natuurlijk geanonimiseerde – databestanden zou wetenschappers en journalisten de kans geven eigen onderzoek te doen.²⁴

24 Zie over de – voorlopige? – teloorgang van DCST en over de datasets die mogelijk interessant zouden kunnen zijn voor openstelling: Beaujean 2014.

In bredere zin is er geen reden te bedenken waarom de Rechtspraak niet de overheidsbrede trend zou volgen naar het publiekelijk aanbieden van meer data: datasets die ten grondslag liggen aan door of in opdracht van de Raad voor de rechtspraak uitgevoerd onderzoek, datasets die de basis vormen van publieke verslaglegging zoals instroom- en uitstroomgegevens, zaaksafhandelingsgegevens en financiële datasets.

Wat zijn de gevaren van al deze ontwikkelingen?

Tal van maatschappelijke sectoren hebben moeite te wennen aan de veranderingen die open, linked en big data teweegbrengen, en de Rechtspraak vormt daarop geen uitzondering. Ik noem hier drie risico's.

Een eerste gevaar is intern-organisatorisch van aard. Reeds nu klinken regelmatig klachten over de te veel op kwantitatieve indicatoren gerichte beoordeling van de individuele rechter. Hoe meer gegevens er beschikbaar komen, des te meer er gekwantificeerd kan worden en des te groter de kans dat de moeilijkst meetbare variabelen – de kwalitatieve – ondergesneeuwd raken.

Het tweede gevaar bestaat in de bijna onbeperkte en soms minder wenselijke mogelijkheden die open, linked en big rechtspraakdata bieden. In de inleiding werd het voorbeeld genoemd van de computervaardige advocaat die zich op de zitting voorbereidt met behulp van uw eerdere uitspraken en uw uit Facebook, LinkedIn en Twitter blijkende hobby's, politieke opvattingen en sociale netwerk. Een goede big data-analyse van al die gegevens zal uitwijzen welke verweren bij u al dan niet kans van slagen hebben. In de Verenigde Staten is handel in dit soort gegevens inmiddels *big business*, Nederland zal daarin ongetwijfeld volgen.

Op het eerste gezicht klinkt dat eng, maar waarom zou een rechter niet voorspelbaar en tot in detail controleerbaar mogen zijn? Echte 'transparantie' is meer dan heldere en actieve communicatie, het betekent ook het toelaten van kritische ogen en een analyserende computer.

Het derde gevaar schuilt in de eigen reactie op al deze gebeurtenissen. De kop in het zand is zinloos, verzet werkt contraproductief. De enige constructieve reactie is het serieus onderkennen van de ontwikkelingen, daar alert op reageren en zo mogelijk zelfs anticiperen. Daarvoor bestaat geen standaardrecept. Het is nieuw, het vergt een multidisciplinaire benadering, een brede verankering in beleid en bewustwording in de hele organisatie.

Literatuur

Beaujean 2014

Beaujean, R.G.A., 'Voorlopig geen GIDS voor consistente straftoemeting', *Trema Straftoemetingsbulletin* 2014-1, p. 3-9.

Van Opijnen 2010

Opijnen, M. van, 'Rechtspraak en digitale rechtsbronnen: nieuwe kansen, nieuwe plichten', *Rechtstreeks* 2010, nr. 1, Den Haag: Raad voor de rechtspraak, 2010.

Van Opijnen 2011a

Opijnen, M. van, 'Jurisprudentiestatistiek. Over de publicatie, annotatie en citatie van rechterlijke uitspraken', *NJB* 2011, 1658.

Van Opijnen 2011b

Opijnen, M. van, 'Het communautaire geteld – een kwantitatief onderzoek naar Europeesrechtelijke verwijzingen in rechterlijke uitspraken en rechtswetenschappelijke literatuur', *SEW, tijdschrift voor Europees en economisch recht* 2011, 98.

Van Opijnen 2013

Opijnen, M. van, 'Van LJN naar ECLI. Vijf prangende vragen', *NJB* 2013, 1607.

Van Opijnen 2014

Opijnen, M. van, *Op en in het web. Hoe de toegankelijkheid van rechterlijke uitspraken kan worden verbeterd* (diss. Amsterdam UvA), Den Haag: Boom 2014. Het proefschrift is in elektronische vorm beschikbaar op Rechtspraak.nl.