



**UvA-DARE (Digital Academic Repository)**

**Op en in het web: Hoe de toegankelijkheid van rechterlijke uitspraken kan worden verbeterd**

van Opijnen, M.

[Link to publication](#)

*Citation for published version (APA):*

van Opijnen, M. (2014). Op en in het web: Hoe de toegankelijkheid van rechterlijke uitspraken kan worden verbeterd

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <http://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

# HOOFDSTUK 5

---

## Selectie en relevantie

In § 2.4.5 hebben we ‘hanteerbaarheid’ aangemerkt als één van de aspecten die een rol spelen bij de toegankelijkheid van jurisprudentiedatabanken. Daarbij hebben we zes deelaspecten onderscheiden: begrijpelijkheid, contextualiteit, centralisatie van aanbod, metadatering, (taal)grensoverschrijdende bruikbaarheid en doorzoekbaarheid. Bij dat laatste deelaspect gaat het erom dat de gebruiker in staat moet zijn om makkelijk en snel de relevante uitspraken te vinden, aangenomen tenminste dat die in de databank aanwezig zijn. In deze laatste zin zijn twee cruciale vraagstukken inzake jurisprudentiedatabanken samengevat. Ten eerste, hoe weet je welke uitspraken je voor een databank moet selecteren, zeker als je niet weet wat de gebruiker uiteindelijk relevant zal vinden? En ten tweede, wat is dat eigenlijk precies: ‘relevant’?

Deze twee thema’s, selectie en relevantie, staan in dit hoofdstuk centraal. In § 5.1 onderzoeken we eerst het selectievraagstuk: wat is het doel van uitsprakeselectie, hoe hebben selectiecriteria zich ontwikkeld en hoe werkt selectie in de praktijk? En natuurlijk leggen we een relatie naar het relevantievraagstuk, dat we in § 5.2 behandelen. Uit die bespreking zal blijken dat ‘relevantie’ tal van aspecten kent die aandacht behoeven als we de hanteerbaarheid, en in het bijzonder de doorzoekbaarheid van uitsprakendatabanken willen optimaliseren. We zullen ons vervolgens in het bijzonder richten op één specifiek deelaspect van relevantie: de ‘domeinrelevantie’, zijnde de mate waarin de juridische gemeenschap een algemeen, zaaksoverstijgend belang toekent aan een rechterlijke uitspraak. In § 5.3 ten slotte trekken we enkele conclusies en blikken we vooruit.

## 5.1 Het selectievraagstuk

De selectie van uitspraken ter opname in jurisprudentieverzamelingen is reeds op meerdere plaatsen in dit boek aan de orde gekomen. Bij het juridisch kader in hoofdstuk 2 bespraken we Aanbeveling R(95)11 van het Comité van Ministers van de Raad van Europa<sup>1674</sup> en de (oude en nieuwe) selectiecriteria van Rechtspraak.nl.<sup>1675</sup> Selectie kwam ook reeds ter sprake bij het rechtsvergelijkend onderzoek,<sup>1676</sup> de bespreking van rechtswetenschappelijke kritiek<sup>1677</sup> en de feitelijke publicatie op Rechtspraak.nl.<sup>1678</sup>

Bezien vanuit de primaire doelstellingen van uitsprakenpublicatie – publieke controle en rechtsvorming<sup>1679</sup> – is selectie een fundamenteel vraagstuk. Continu moet een balans worden gezocht tussen ‘te weinig’ – waardoor publieke controle niet effectief kan worden uitgeoefend en het rechtsvormende werk van de rechter onvoldoende zichtbaar is – en ‘te veel’ – waardoor gebruikers te veel tijd kwijt zijn met het scheiden van kaf en koren.

1674 *Vide supra*: § 2.3.2.2.4.2.1.

1675 *Vide supra*: § 2.3.2.2.3.

1676 *Vide supra*: § 2.6.2.1.

1677 *Vide supra*: § 2.7.1.

1678 *Vide supra*: § 3.4.2.

1679 *Vide supra*: § 1.1.

Bij het doel van de selectie staan we nader stil in § 5.1.1. Vervolgens bespreken we hoe het denken over selectiecriteria zich de laatste decennia heeft ontwikkeld (§ 5.1.2) en hoe dit in de praktijk uitwerkt (§ 5.1.3). In § 5.1.4 ten slotte stellen we een andere benadering van het selectievraagstuk voor.

### 5.1.1 Het doel van selectie

Het selectievraagstuk bestond reeds toen databanken er nog niet waren. Ook toen jurisprudentie nog uitsluitend in papieren tijdschriften werd gepubliceerd, speelde reeds de vraag welke beslissingen op te nemen en welke niet. Het publiceren van alle uitspraken was niet alleen uit commercieel oogpunt onaantrekkelijk, het had ook geen toegevoegde waarde voor de lezer, die immers om zou komen in het informatieaanbod.

Om de bruikbaarheid voor de ‘consument’ te onderstrepen komen we daarom bij het substantief ‘selectie’ regelmatig het adjectief ‘representatief’ tegen. Zo refereert Moltmaker-II aan de kritische geluiden die er in de pers waren over “(D)e representativiteit van de geselecteerde uitspraken”<sup>1680</sup> en rept de Toegankelijkheidsnota van “(E)en representatief beeld van rechterlijke uitspraken”<sup>1681</sup> dat van overheidswege zou moeten worden verstrekt.

Ook in de doctrine wordt regelmatig gesproken over (het gebrek aan) een ‘representatieve’ selectie.<sup>1682</sup>

Vaak wordt echter nauwelijks uitgelegd wat met ‘representatief’ precies wordt bedoeld. Van Dale definieert ‘representatief’ als: “In zijn kenmerken een groep of geheel vertegenwoordigend, er een goede indruk, een beeld van gevend.” In het dagelijks spraakgebruik wordt het woord vaak geassocieerd met de ‘representatieve’ steekproef. In de statistiek wordt daarmee bedoeld dat de steekproef een getrouwe afspiegeling vormt van de gehele populatie. Om dat te bereiken dient men er voor te zorgen dat diverse strata (subpopulaties) naar evenredigheid in de steekproef zijn vertegenwoordigd.<sup>1683</sup> Voor de selectie van publicatiewaardige uitspraken noemt Aanbeveling R(95)11 in art. I.2 van Appendix II als strata: rechterlijke hiërarchie, geografische verdeling, rechtsgebieden en onderwerpen. Art. II begint vervolgens met de woorden:

*If a selection takes place, it should be carried out in such a manner that objectivity and representativity of the database are ensured.*

*The selection should ensure, on the one hand, broad and comprehensive access to information on court decisions and, on the other hand, that the accumulation of useless information is avoided.*

1680 [Werkgroep Moltmaker 1994, par. 6.2.1].

1681 Kamerstukken II 1996-1997, 20 644, nr. 30, p. 7.

1682 Zo is het eerste hoofdstuk van [Kottenhagen en Kaptein 1989] geheel gewijd aan ‘De representativiteit van gepubliceerde jurisprudentie in Nederland’. De conclusie luidt dat “(D)e gepubliceerde rechtspraak zeker niet altijd een representatief beeld geeft van alle rechtspraak op een bepaald rechtsgebied” (p. 15). Vgl. ook [van Dunné 1994] en [Kottenhagen 1994].

1683 [Nijdam 2004, p. 55].

Uit de tweede zin mogen we afleiden dat een zuiver statistische uitleg van het begrip ‘representativiteit’ dus niet de bedoeling is. Anders zouden routine-uitspraken een evenredig deel van de gepubliceerde verzameling uitmaken, en dat dient als ‘*accumulation of useless information*’ te worden vermeden. Een zuiver statistische toepassing zou er bovendien toe leiden dat juridisch belangwekkende uitspraken regelmatig buiten de boot vallen, en dat is natuurlijk ook niet de bedoeling.<sup>1684</sup>

In het vervolg van het geciteerde art. II wordt nadere invulling gegeven aan het begrip ‘selectie’. Alhoewel ook deze nadere uitwerking voor tal van problemen zorgt – waarover later – maakt dit artikel duidelijk dat het woord ‘representatief’ betekent dat zowel belangrijke rechtsvormende uitspraken moeten worden opgenomen, als ook (zie art. II.4) feitelijke uitspraken en beslissingen die staande jurisprudentie bevestigen.

Op het belang van de publicatie van ogenschijnlijk oninteressante gevallen werd reeds gewezen door Snijders,<sup>1685</sup> die onderzoek deed naar door de rechter vastgestelde omgangsregelingen. Hij concludeerde dat uit de ongepubliceerde uitspraken een geheel ander beeld naar voren komt dan uit de gepubliceerde uitspraken. In later werk stelt Snijders vervolgens:

*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’.*<sup>1686</sup>

Alhoewel Snijders niet bepleit om de omvang van het gepubliceerde corpus sterk uit te breiden, legt hij hiermee het dilemma met betrekking tot de ‘representatieve selectie’ op adequate wijze bloot.<sup>1687</sup> Dit dilemma heeft niet alleen betrekking op de criteria die worden gehanteerd om tot een publicatiewaardige selectie te komen, maar ook op de verschillende verwachtingen die er leven met betrekking tot die selectie. Iedere jurist heeft, vanuit zijn specifieke praktijk of wetenschappelijke interesse, belangstelling voor een bepaald segment van gedane uitspraken. Omdat selecteurs, los van persoonlijke voorkeuren,<sup>1688</sup> altijd selecteren voor de grootste gemene deler van de doelgroep die zij met hun periodiek of databank willen bereiken, krijgt niemand maatwerk en zal iedereen ontevreden zijn. Openlijk beleden commentaar is daarom vrijwel altijd negatief.<sup>1689</sup> Pogingen om aan deze onvrede tegemoet

1684 Vgl. Kamerstukken II 2005-2006, 29 279, nr. 35, p. 6, waarin ook de minister twee dingen door elkaar haalt: “De omvang van de uitsprakendatabank groeit snel en daarmee ook de representativiteit ervan.”

1685 [Snijders 1977].

1686 [Snijders 1978, p. 208].

1687 Vergelijkbare geluiden zijn te vinden bij onder meer [Giard 2010a], [Bergwerf en Houweling 2008] en [Mommers, Zwenne en Schermer 2010].

1688 *Vide infra*: § 5.1.2.

1689 Ten aanzien van Rechtspraak.nl *vide supra*: § 2.7, ten aanzien van jurisprudentieperiodieken bijv. [Malsch, Ijpelaar en Nijboer 2007, p. 112] en [Kottenhagen en Kaptein 1989, hfd. 2].

te komen waren in eerste instantie vooral gericht op het verhelderen en verbeteren van de selectiecriteria. Hierop zullen we in de volgende paragraaf nader ingaan.

### 5.1.2 Ontwikkeling van selectiecriteria in Nederland

Veel Nederlandse jurisprudentietijdschriften stellen het zonder expliciete selectiecriteria. Het fingerspitzengefühl van de redacteurs is een belangrijke leidraad. De redacteur van de Nieuwsbrief Strafrecht verwoordde het aldus:

*Ik lees alle uitspraken en beoordeel of ik het zelf de moeite waard of zinvol vind. Dat kan om verschillende redenen zijn. Ofwel de uitspraak is juridisch nieuw, of juist niet nieuw maar bevestigend. De uitspraak is aardig, gek of actueel. Ik beoordeel dat met mijn eigen verstand. De criteria zijn niet vastgesteld en staan niet op papier. Het gaat om wat ik er zelf van vind.*<sup>1690</sup>

Het vermoeden dat de selectie van uitspraken sterk wordt gekleurd door persoonlijke voorkeuren, wordt door uitspraken van andere redacteurs bevestigd.<sup>1691</sup>

Voor zover er wel expliciete selectiecriteria zijn geformuleerd, zijn deze zelden gepubliceerd. Die van de NJ<sup>1692</sup> zijn bijvoorbeeld bekend.

1. *juridisch belang van de zaak;*
2. *actualiteitswaarde;*
3. *vindt rechtsvorming plaats;*
4. *worden 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.*<sup>1693</sup>

Deze criteria zijn bijzonder summier; er schijnen ook nog enkele meer specifieke criteria te bestaan,<sup>1694</sup> maar deze zijn niet publiekelijk bekend.<sup>1695</sup>

Van andere tijdschriften komt af en toe iets over de selectiecriteria naar buiten. Zo bood enkele jaren terug het Tijdschrift voor Consumentenrecht en Handelszaken in een redactioneel commentaar inzicht in de gehanteerde criteria:

1. Evenwichtige spreiding over de verschillende onderdelen van het consumentenrecht;

<sup>1690</sup> Geciteerd in: [Malsch, Ijpelaar en Nijboer 2007, p. 108].

<sup>1691</sup> Zie bijvoorbeeld de opmerkingen van [Hesseling 1982] over het verschil dat hij zal maken in het selectiebeleid van de Praktijkids, in vergelijking met zijn voorganger.

<sup>1692</sup> De criteria van de NJ worden eveneens door NJFS gebruikt ([Malsch, Ijpelaar en Nijboer 2007, p. 110]).

<sup>1693</sup> [Redactie NJ 1982], geciteerd in: [MvJ 1992, p. 13] en [Kottenhagen 1994, p. 43].

<sup>1694</sup> [Kottenhagen 1994, p. 43-44].

<sup>1695</sup> [Crommelin 1988, p. 754] noemt bijv. kracht van gewijsde.

2. Gepubliceerde uitspraken moeten representatief zijn voor de betrokken instantie;
3. De aard van het tijdschrift (is het een consumentenzaak of niet);
4. Het leveren van een bijdrage aan de rechtsvorming;
5. Is er voldoende ruimte in het tijdschrift.<sup>1696</sup>

Ook deze criteria maken duidelijk dat er bij individuele uitspraken regelmatig debat zal ontstaan over de vraag of ze publicatiewaardig zijn.

Een meer wetenschappelijk en beleidsinhoudelijk debat over selectiecriteria is vooral gevoerd met betrekking tot (plannen voor) publieke dataverzamelingen. Kottenhagen en Kaptein<sup>1697</sup> stellen, in navolging van Voskuil,<sup>1698</sup> een tweelagensysteem voor. Ieder gerecht zou een eigen databank op moeten bouwen waarin een eerste selectie van uitspraken wordt opgenomen. Deze selectie zou moeten worden gemaakt door een per instantie aan te wijzen documentalist. Uit deze lokale databanken zou vervolgens een tweede selectie van publicatiewaardige uitspraken moeten worden gemaakt door een landelijke dienst, bestaande uit:

*(E)en college van wetenschappers, leden van de rechterlijke macht, vertegenwoordigers van de uitgevers en vertegenwoordigers van het ministerie.*<sup>1699</sup>

Voor die tweede selectie zijn de volgende criteria voorgesteld:

*Een uitspraak wordt gepubliceerd wanneer:*

1. *een nieuwe rechtsregel wordt geformuleerd, een bestaande rechtsregel wordt gewijzigd of gemodificeerd of de aandacht gevestigd wordt op een bestaande rechtsregel die tot dan toe nog niet is toegepast;*
2. *een bestaande rechtsregel toegepast wordt op een feitenconstellatie die significant afwijkt van eerdere toepassingen;*
3. *de geschiedenis, toepassing of verwerking van wettenrecht of rechtersrecht nader uitegelegd of gekritiseerd wordt;*
4. *tegenstrijdige rechtsregels gecreëerd worden dan wel tot een oplossing worden gebracht;*
5. *het een rechtsregel en/of feitenconstellatie betreft die van algemeen belang is;*
6. *een beslissing in beroep of cassatie vernietigd wordt, tenzij deze vernietiging het gevolg is van een tussentijdse verandering van een wet of van de jurisprudentie dan wel van het feitenmateriaal en overigens niet aan genoemde criteria wordt voldaan;*
7. *in de uitspraak het rechtsprekend college de publicatie van de beslissing bij wijze van straf heeft bevolen;*

---

1696 [Hondius 2005].

1697 [Kottenhagen en Kaptein 1989, par. 6.3].

1698 [Voskuil, de Jong en van Essen-Kruseman 1976-1980]. *Vide supra*: § 3.1.1.

1699 [Kottenhagen en Kaptein 1989, p. 105].

8. *de publicatie van de uitspraak van een rechtscollege terug gaat op een wettelijke regeling;*
9. *de beslissing vergezeld gaat van een concurring of een dissenting opinion.*<sup>1700</sup>

De uitspraken die niet door de tweede selectie komen, zouden in de vorm van een rechtspraakregister toch toegankelijk moeten zijn voor wetenschap, rechterlijke macht of advocatuur.<sup>1701</sup>

Moltmaker-I<sup>1702</sup> volgt, zonder daarnaar enige verwijzing te maken,<sup>1703</sup> in beginsel dit tweelagenmodel, maar de te gebruiken selectiecriteria worden niet geconcretiseerd. Dit gebeurt wel door Franken,<sup>1704</sup> die het organisatiemodel van Moltmaker-I volgt. Zijn voorstellen sluiten ook nauw aan bij de structuur van Aanbeveling R(95)11, in het bijzonder het gebruik van negatieve selecties (voor de hoogste gerechten) en positieve selecties (voor de lagere gerechten).<sup>1705</sup> Anders dan Kottenhagen/Kaptein doet Franken een concreet voorstel voor criteria voor de eerste selectie. Deze criteria zijn objectief van aard.<sup>1706</sup> Uitspraken worden opgenomen indien ze:

- *volgens de wettelijke voorschriften zijn uitgewerkt, dat wil zeggen voldoen aan de (grond)wettelijke motiveringseis;*
- *op tegenspraak zijn gewezen of, wanneer het om verstekzaken gaat, een beslissing met betrekking tot een ontvankelijkheidsvraag inhouden;*
- *een rechtsbeslissing bevatten.*<sup>1707</sup>

Voor selectie in de centrale databank geldt vervolgens voor de uitspraken van de hoogste en Europese gerechten het negatieve criterium: alle uitspraken worden opgenomen, tenzij ze een pro-formakarakter hebben of slechts een standaardmotivering bevatten. Uitspraken van de appelcolleges en rechtbanken worden geselecteerd volgens zowel de negatieve als de positieve selectiecriteria. Uitspraken worden niet opgenomen indien ze:

- a) *zijn gemotiveerd met behulp van een standaardformulering c.q. clause de style: zogenaamde bouwstenen dan wel tekstverwerkingsmodulen maken deze standaardformulering herkenbaar;*
- b) *bewijsvragen betreffen, welke overeenstemmen met de gevestigde jurisprudentie. Hier gaat het om feitelijke oordelen, die niets toevoegen aan de kennis betreffende rechtsvinding en rechtsvorming.*<sup>1708</sup>

1700 Ibid., p. 114.

1701 Vraagstukken omtrent bijv. de regeling van de feitelijke toegang tot dit register en de anonimisering van uitspraken blijven overigens grotendeels onbesproken.

1702 [Commissie Moltmaker 1993]. *Vide supra*: § 3.1.1.

1703 De vraag lijkt gerechtvaardigd of de Commissie-Moltmaker überhaupt wel kennis heeft genomen van [Voskuil, de Jong en van Essen-Kruseman 1976-1980] en/of [Kottenhagen en Kaptein 1989].

1704 [Franken 1995].

1705 *Vide supra*: § 2.3.2.2.4.2.1.

1706 Door Franken 'formele criteria' genoemd.

1707 [Franken 1995, p. 56].

1708 Ibid., p. 57.



En ze worden wel opgenomen indien:

- 1) *de uitleg van een begrip of een wettelijke term wordt gegeven, dat wil zeggen een rechtsregel wordt geformuleerd of gewijzigd;*
- 2) *een methode van interpretatie wordt gehanteerd, zodat een bestaande rechtsregel wordt toegepast op een feitenconstellatie die afwijkt van eerdere toepassingen;*
- 3) *een methode van argumentatie wordt geconstateerd, die afwijkt van eerdere toepassingen;*
- 4) *vragen met betrekking tot de bevoegdheid van het gerecht worden beslist;*
- 5) *de motivering aangeeft, dat er dubia bestaan. die – zonder afbreuk te doen aan de onderhavige beslissing – voorwerp voor nader onderzoek opleveren, met andere woorden de beslissing van het college is niet onomstreden;*
- 6) *het een rechtsregel en/of feitenconstellatie betreft die van algemeen belang is (signaleringsfunctie voor de wetgever of de wetenschap).*<sup>1709</sup>

Enkele objectieve criteria van Kottenhagen/Kaptein zijn door Franken niet opgenomen, zoals publicatie bij wijze van straf,<sup>1710</sup> of vanwege een wettelijke verplichting.<sup>1711</sup> Ook in de subjectieve criteria zijn er aanmerkelijke verschillen in formulering. Beide sets van criteria zijn reeds specifiekere dan de in het begin van deze paragraaf geciteerde criteria van de NJ, maar laten nog steeds veel ruimte voor interpretatie.

Een voorbeeld van een verdere uitwerking van subjectieve criteria is te vinden bij de Judok-databank van de Raad van State.<sup>1712</sup> Een uitspraak zou daarin moeten worden opgenomen indien aan een van de volgende voorwaarden is voldaan:

1. *De Afdeling geeft een uitleg/interpretatie van een (pseudo-) wettelijk voorschrift.*
2. *De Afdeling behandelt een juridisch leerstuk.*
3. *De Afdeling geeft een verduidelijking van een juridisch probleem.*
4. *De Afdeling “gaat om”.*
5. *De Afdeling handhaaft in uitdrukkelijke bewoordingen haar eerdere jurisprudentie.*<sup>1713</sup>
6. *De Afdeling handhaaft stilzwijgend haar jurisprudentie, nadat deze in bureau-nota, raadkamer of vakliteratuur ter discussie is gesteld.*
7. *De Afdeling geeft een verfijning op haar eerdere jurisprudentie.*
8. *De Afdeling wijzigt de formulering van haar eerdere jurisprudentie.*
9. *De Afdeling dwingt het overheidsorgaan tot een beleidswijziging.*
10. *Het geschil is onderwerp van beraad geweest in het Voorzitters, het chef-juristen- of het kolom-/kameroverleg.*

---

1709 Ibid., p. 58.

1710 *Vide supra*: § 2.3.2.1.1.3.

1711 *Vide supra*: § 2.3.2.1.

1712 *Vide supra*: § 3.1.1.

1713 Zie ook [Lindo 1993], waarin de NJ-redactie erop wijst dat het nuttig kan zijn om uitspraken die eerdere jurisprudentie bevestigen te publiceren teneinde zinloze beroepen te voorkomen.

11. *De uitspraak is het vervolg op een eerdere, in JUDOK opgenomen zaak.*  
 12. *De Afdeling spreekt haar bedenkingen uit over bepaald overheidsbeleid zonder dat dit tot vernietiging behoeft te leiden.*<sup>1714</sup>

Sommige van deze criteria zijn bijzonder objectief, zoals de criteria 10 en 11, maar deze criteria veronderstellen kennis die niet blijkt uit de uitspraaktekst. Andere (subjectieve) criteria zullen in bepaalde gevallen duidelijk tot publicatie leiden, maar ook op basis van deze criteria zullen er voldoende twijfelgevallen overblijven.

De Selectiecriteria Rechtspraak.nl 1999<sup>1715</sup> wijken van de hierboven besproken criteria af omdat ze niet alleen op juridische, maar ook op maatschappelijke relevantie zien.<sup>1716</sup> Opvallend in deze selectiecriteria is dat er, ondanks de publicaties die er inmiddels over het onderwerp waren verschenen – inclusief Aanbeveling R(95)11 – geen onderscheid wordt gemaakt tussen hogere en lagere instanties. Wel is er een (onbenoemd) onderscheid in positieve en negatieve criteria, maar deze worden niet gerelateerd aan het soort gerecht. Objectieve maatstaven komen alleen voor bij de negatieve criteria; voor de positieve criteria bestaan alleen subjectieve toetsen. Laatstgenoemde laten veel ruimte voor interpretatie, sommige zijn zelfs onbegrijpelijk geformuleerd.<sup>1717</sup>

De Selectiecriteria Rechtspraak.nl 2012<sup>1718</sup> daarentegen hebben een veel helderder structuur, die in de toelichting bovendien verder wordt geëxpliciteerd. In art. 3 wordt een objectief positief criterium gebruikt met betrekking tot (specifieke kamers van) bepaalde gerechten. Van deze afdelingen moeten alle uitspraken worden gepubliceerd, tenzij ze voldoen aan een objectief negatief criterium. Art. 4 lid 1 formuleert een aantal objectieve positieve criteria met betrekking tot de grondslag van de procedure, eventueel in combinatie met de expliciete beslissing van de rechter betreffende openbaarmaking (sub b en c). In art. 4 lid 2 staan enkele andersoortige objectieve positieve criteria. Ze hebben betrekking op de aard van het gevoerde verweer en de reactie van de rechter daarop (sub a), voorgaande publicatie van een uitspraak in dezelfde rechtsgang (sub b), de aard van het ten laste gelegde delict (sub c) of opgelegde straf (sub d). De bepaling (sub e) dat alle uitspraken in wrakingsprocedures moeten worden gepubliceerd, is ook een objectief positief criterium maar zou qua structuur beter in lid 1 hebben gepast.

Art. 5 ziet op de subjectieve positieve criteria. Deze hebben betrekking op zowel maatschappelijke als juridische relevantie. De eerste twee criteria gebruiken externe belangstelling als maatstaf (respectievelijk van publieke media en juridische vakpers). Uit het in de

1714 Geciteerd in [Mv] 1992, p. 10].

1715 *Vide supra*: § 2.3.2.2.3. Voor de tekst zij verwezen naar Bijlage 3.

1716 Bij uitzondering wordt het NJ-criterium 'actualiteitswaarde' ook gebruikt om zaken met maatschappelijke relevantie te selecteren. Door de NJ-redactie is als voorbeeld de zaak-Menten genoemd ([Redactie NJ 1988], geciteerd in [Kottenhagen en Kaptein 1989, p. 34]).

1717 Zoals het laatste criterium van de tweede sectie: "(H)et antwoord op een rechtsvraag op andere wijze wordt geformuleerd of gemotiveerd en omgekeerd bijzonderheden voorkomen in verband met de aard van de procedure, hetzij een bijzonder dictum heeft."

1718 *Vide supra*: § 2.3.2.2.3. Voor de tekst zij verwezen naar Bijlage 4.

toelichting gegeven voorbeeld van een journalist die een volledige politierechterzitting bijwoont, blijkt dat dit criterium niet zo objectief is als op het eerste gezicht lijkt.

De criteria sub c en d vragen een (subjectieve) beoordeling omtrent het mogelijke belang dat de uitspraak voor anderen zou kunnen hebben. Het laatste criterium van art. 5 betreft het juridisch belang van een uitspraak. In vergelijking met de ruime uitwerking die dit criterium heeft gehad in enkele van de hierboven besproken selectiecriteria, is de korte redactie opmerkelijk:

*(Een uitspraak wordt altijd gepubliceerd indien ... D)e uitspraak een jurisprudentievormend karakter heeft, bijvoorbeeld doordat een eerdere jurisprudentielijn wordt gewijzigd, genuanceerd, ingeperkt of uitgebreid, of doordat een bepaalde casuïstiek of wettelijke bepaling voor het eerst aan het oordeel van de rechter is onderworpen.*

Bij inhoudelijke vergelijking blijkt hierin evenwel vrijwel hetzelfde te staan als in de gehele sectie 2 van de Selectiecriteria Rechtspraak.nl 1999. De toelichting merkt hierbij op dat een ruime interpretatie geëigend is:

*(T)wijfel over het juridisch belang van een uitspraak is op zichzelf eigenlijk reeds voldoende om tot publicatie te besluiten.*

Art. 6 lid 1 ten slotte ziet op alle uitspraken die niet krachtens de andere artikelen moeten worden gepubliceerd. Met behulp van een objectief negatief criterium wordt eerst bepaald welke uitspraken daarvan niet voor publicatie in aanmerking komen – “(V)oorzover daarin (...) uitsluitend standaardformuleringen worden gebruikt” – en vervolgens wordt gesteld dat de resterende uitspraken “(Z)oveel als mogelijk” worden gepubliceerd. In het licht van het zelfbindende karakter van de Selectiecriteria Rechtspraak.nl 2012<sup>1719</sup> kan dit worden gelezen als een inspanningsverplichting, waarbij grenzen kunnen worden gevonden in technische mogelijkheden of financiële middelen.<sup>1720</sup>

### 5.1.3. Selecteren in de praktijk

Alhoewel er per gerecht ook interne databanken bestaan (huisdatabanken en later E-archief),<sup>1721</sup> is er thans geen sprake van een tweelagenmodel, zoals dat voorzien was door Voskuil,<sup>1722</sup> Kottenhagen/Kaptein,<sup>1723</sup> Moltmaker-I<sup>1724</sup> en Franken:<sup>1725</sup> er wordt geen keuze gemaakt uit de interne verzamelingen die per gerecht worden opgeslagen, maar per uitspraak wordt direct besloten of deze voor de interne en/of de externe databank in aan-

1719 *Vide supra*: § 2.3.2.2.3.

1720 Vgl. art. 48a van het Oostenrijkse *Gerichtsorganisationsgesetz*, *supra*: § 2.6.2.1.1.

1721 *Vide supra*: § 3.3.

1722 [Voskuil, de Jong en van Essen-Kruseman 1976-1980].

1723 [Kottenhagen en Kaptein 1989].

1724 [Commissie Moltmaker 1993].

1725 [Franken 1995].

merking komt.<sup>1726</sup> De selectiebeslissingen voor Rechtspraak.nl worden door de gerechten zelf genomen,<sup>1727</sup> en niet door een landelijke commissie. Een afwijkende situatie is er bij het bestuursrecht, waar het Landelijk stafbureau bestuursrecht ten behoeve van een specifieke collectie op Porta Iuris<sup>1728</sup> en de Jurisprudentienieuwsbrief Bestuursrecht<sup>1729</sup> een tweede selectie maakt uit lokaal geselecteerde uitspraken. Ook bij (de meeste) gerechten is er geen tweelagenmodel; rechters bepalen zelf of een uitspraak voor publicatie op Rechtspraak.nl in aanmerking komt;<sup>1730</sup> de regels voor opname in huisdatabank/E-archief zijn diffuus.

Op een iets abstracter niveau beschouwd is in het hele selectieproces het opstellen van selectiecriteria slechts een eerste stap. Ondanks de verschillende nuances die daarin kunnen worden aangebracht – zie de verschillende modellen in de vorige paragraaf – is dit de eenvoudigste stap. Veel lastiger is het vervolgens om te beoordelen of een individuele uitspraak aan de selectiecriteria voldoet, vooral als het om subjectieve criteria gaat.<sup>1731</sup> In een tweelagenmodel wordt de – lastigste – tweede selectie gedaan door zeer deskundige mensen. Alleen bij tijdschriftredacties en de hierboven genoemde bestuursrechtelijke collectie wordt een dergelijk model gehanteerd, voor Rechtspraak.nl geschiedt de selectie in één stap.<sup>1732</sup> Tijdgebrek, zelfoverschatting of juist zelfonderschatting, en onvoldoende zicht op wat de buitenwereld belangrijk vindt, kunnen daarom tot een weinig consistente toepassing van overigens heldere selectiecriteria leiden.

Een bijzonder probleem wordt daarnaast gevormd door de perceptie van de gebruiker. Met betrekking tot strafrechtelijke uitspraken is recent onderzoek gedaan door Malsch c.s.<sup>1733</sup> Hieruit komt naar voren dat juristen de (in tijdschriften en op Rechtspraak.nl) gepubliceerde aantallen schromelijk overschatten; ze kennen de ‘gemiddelde’ uitspraak derhalve een te hoge jurisprudentiële waarde toe.<sup>1734</sup> Lezers van periodieken blijken bovendien in de veronderstelling te verkeren dat de redacties welomschreven criteria gebruiken, al verschillen de omschrijvingen<sup>1735</sup> van deze veronderstelde criteria nauwelijks van de *gutfeeling* waarop veel redacties vertrouwen. Dat geldt eveneens voor de criteria die vermeend een rol spelen bij de selectie door de gerechten zelf (voor publicatie op Rechtspraak.nl of voor toezending

---

1726 *Vide supra*: § 3.3.

1727 *Vide supra*: § 3.4.2.

1728 *Vide supra*: § 3.3.1.3.

1729 Deze wordt zowel op Porta Iuris als op Rechtspraak.nl gepubliceerd.

1730 *Vide supra*: § 3.4.2.9.1.

1731 Vgl. [Hartlief 1990] in zijn bespreking van [Kottenhagen en Kaptein 1989].

1732 *Vide supra*: § 3.4.2.9.1.

1733 [Malsch, Ijpelaar en Nijboer 2007].

1734 Er is een opvallende gelijkheid tussen deze gebruikersbeoordeling van door mensen vastgestelde relevantie en de door [Blair en Maron 1985, p. 293] geconstateerde overschatting van computerprestaties bij het vinden van relevante documenten. In laatstgenoemd onderzoek dachten juristen dat het systeem 75% van de (volgens het onderzoeksprotocol) relevante documenten had gevonden, terwijl dat in werkelijkheid slechts 20% was. Andere onderzoeken tonen aan dat hoe groter de documentcollectie en de omvang van de individuele documenten, des te groter de misconceptie van de gebruiker omtrent de relevantieprestaties van het systeem ([Turtle 1995, p. 24]). *Vide infra*: § 5.2.1.1.

1735 [Malsch, Ijpelaar en Nijboer 2007, p. 109].

naar een tijdschrift); hierbij valt echter wel op dat gebruikers vaak denken dat ‘ijdelheid van de rechter’ een selectiecriteria is.<sup>1736</sup> Malsch c.s. concluderen uiteindelijk:

*Met de instelling van Rechtspraak.nl is er veel gepubliceerde rechtspraak bijgekomen, maar uit de opmerkingen van geïnterviewden en uit de resultaten van ander onderzoek blijkt dat gebruikers nogal wat kritiek hebben op de samenstelling van het aanbod op Rechtspraak.nl.*<sup>1737</sup>

Of deze gebruikers – waarvan in ieder geval met betrekking tot de (vermeende) omvang van het gepubliceerde corpus is vastgesteld dat ze geen goed beeld hebben van de werkelijkheid – gelijk hebben in hun kritiek, is onbekend:

*De gepubliceerde jurisprudentie is niet inhoudelijk door ons onderzocht. We kunnen dus niet vaststellen of het aanbod inderdaad onevenwichtig is.*<sup>1738</sup>

De volgende conclusie, dat:

*Striktere en helderder richtlijnen en procedures kunnen helpen om een evenwichtige publicatie van jurisprudentie te krijgen.*<sup>1739</sup>

verbaast dan ook enigszins. Het is immers niet aangetoond dat gebruikersperceptie verandert door striktere selectieregels.<sup>1740</sup>

#### 5.1.4 Een andere benadering

De discussie over selectiecriteria werd reeds gevoerd voordat digitale databanken hun intrede deden; ook toen er nog slechts papieren jurisprudentieperiodieken waren, bestond er reeds vrees voor informatieovervloed. Zo stelde Crommelin in 1988 nog dat een beleid (van de NJ) om alle beslissingen van de Hoge Raad op te nemen:

*(G)een enkele bijdrage (zou) leveren aan de toegankelijkheid van de rechtspraak. De lezer is veel meer gediend met de strengst mogelijke selectie waarbij hem uitsluitend voorgeschoteld wordt wat volgens vastgestelde criteria voor de praktijk van belang is.*<sup>1741</sup>

Maar met de groei van de publicatiemogelijkheden wil de gebruiker steeds meer zelf kunnen uitmaken wat voor zijn eigen praktijk van belang is. De keerzijde wordt natuurlijk gevormd door informatiestress<sup>1742</sup> en de vrees voor nog verdere toename daarvan.<sup>1743</sup>

1736 Ibid., p. 111. Vgl. [van de Pol 1988].

1737 [Malsch, Ijpelaar en Nijboer 2007, p. 113].

1738 Ibid.

1739 Ibid.

1740 Voor meer onderzoek naar gebruikerspercepties, *vide infra*: § 5.2.2.

1741 [Crommelin 1988, p. 753].

1742 Bijv. [van Woensel 2002], [van Hoek 2002].

1743 Bijv. [NVvR 2007], [Buruma 2012].

Het selectieprobleem wordt daarmee steeds meer verlegd van de input-zijde naar de output-zijde. Vroeger was er alleen het tijdschrift, en – ook al publiceerde de redactie misschien te weinig – men kon er redelijk op vertrouwen dat (vrijwel) alle uitspraken lezenswaardig waren. Dit selectiewerk, dat voorheen alleen door de redacties werd gedaan, moet tegenwoordig door de gebruiker zelf worden verricht. Naarmate de technische mogelijkheden tot publicatie verbeteren zal – op basis van art. 6 lid 1 Selectiecriteria Rechtspraak.nl 2012 – een steeds verdere verschuiving plaatsvinden van toepassing van positieve selectiecriteria naar negatieve selectiecriteria. De vraag zal dan niet meer zijn: ‘Is er reden deze uitspraak te publiceren?’, maar: ‘Is er reden deze uitspraak niet te publiceren?’

En ook al worden de gezamenlijke gebruikers dan tegemoet gekomen in hun wens van een ruimere selectie, voor de individuele gebruiker zal een steeds kleiner percentage van het gepubliceerde corpus de moeite waard zijn. Met andere woorden: het vinden van de naald wordt lastiger naarmate de hooiberg groeit.

Inherent aan de erkenning dat selectie meer als een gebruikersprobleem en minder als een redactioneel probleem moet worden benaderd, is het loslaten van het dichotoom axioma waarop de selectiediscussie is gebaseerd. Voor een redactie is de vraag ‘publiceren of niet’, maar voor een gebruiker is de vraag ‘meer of minder relevant’? Dat is geen booleaanse keuze, maar een glijdende schaal,<sup>1744</sup> die bovendien voor iedere gebruiker anders kan liggen.

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?<sup>1745</sup> 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.

## 5.2 Relevantie: een theoretisch kader

Het begrip ‘relevantie’ wordt door Van Dale beschreven als ‘het van gewicht of betekenis zijn.’ Het begrip zal in eerste instantie weinig vraagtekens oproepen, intuïtief wordt het door iedereen begrepen.<sup>1746</sup> Desalniettemin is er rond dit begrip een hele wetenschap ontstaan: ‘information retrieval’ (hierna ook wel aangeduid als ‘IR’ of ‘retrieval’<sup>1747</sup>).

1744 Vgl. [van den Berg 1994, p. 962]. Anders: [Bing 1991, p. 28-30], die de relevantie van een bron als een binaire waarde beschouwt. Aan de juridische argumenten die aan de bron kunnen worden ontleend, kent hij wel een ordinale waarde toe.

1745 Vgl. o.m. [Smits 1996, p. 188].

1746 Vgl. [Saracevic 1975, p. 324].

1747 Van Dale definieert het uit het Engels overgenomen ‘retrieval’ als: “*Het opzoeken en zichtbaar maken van in een databank opgeslagen informatie.*” Volgens deze definitie is het Nederlandse ‘retrieval’ een synoniem voor het Engelse ‘*information retrieval*’ en zou ‘informatieretrieval’ derhalve een pleonasme zijn.

*The processes in IR were geared toward relevance as their reason d'être because of the desire to provide effective approaches to the problem of dealing with the 'bewildering array of knowledge'.<sup>1748</sup>*

In deze paragraaf onderzoeken we eerst de verschillende betekenissen van het begrip 'relevantie' (§ 5.2.1), zowel binnen de IR in het algemeen, als bij jurisprudentiedatabanken in het bijzonder. Vervolgens gaan we in op de vraag hoe relevantie kan worden gemeten (§ 5.2.2).

## 5.2.1 Het begrip 'relevantie'

In deze paragraaf proberen we inhoud te geven aan het begrip 'relevantie'. Daarbij gaan we eerst in op de betekenissen die het begrip kan hebben in de IR (§ 5.2.1.1). Vervolgens onderzoeken we hoe we het begrip moeten opvatten in relatie tot jurisprudentiedatabanken (§ 5.2.1.2).

### 5.2.1.1 Generieke theorie

Binnen de informatica en vooral binnen het subdomein van de retrieval wordt, vanuit verschillende invalshoeken, al lang over het relevantiebepgrip gefilosofeerd.<sup>1749</sup> Er is geen alom geaccepteerde en alomvattende theorie, maar om een beeld te krijgen van het relevantievraagstuk volgen we in grote lijnen Saracevic, die over een periode van 30 jaar<sup>1750</sup> veel aan de theorievorming heeft bijgedragen.

Saracevic onderscheidt om te beginnen vier thema's: aard (wat is relevantie?), manifestatie (wat voor verschijningsvormen heeft relevantie?), gedrag (relevantie binnen het zoekgedrag van gebruikers) en effecten (hoe kan relevantie effectief worden geoperationaliseerd?).<sup>1751</sup>

Met betrekking tot het eerste thema, de aard van relevantie, onderkent Saracevic een aantal essentiële attributen, die samenkomen in zijn definitie:

*As a cognitive notion relevance involves an interactive, dynamic establishment of a relation by inference, with intentions toward a context.<sup>1752</sup>*

Relevantie kan zich vervolgens op verschillende manieren manifesteren. Deze manifestaties – we kunnen ook spreken over 'relevantiesoorten' – komen in grote lijnen overeen

1748 [Saracevic 1996, p. 201]. Vgl. ook [Mizzaro 1997, p. 810] alsmede de daar aangehaalde auteurs.

1749 Zie voor een historisch overzicht onder meer: [Mizzaro 1997], [Saracevic en Kantor 1997], [Saracevic 2007b], [Saracevic 2007a].

1750 In het bijzonder [Saracevic 1975], [Saracevic 1996], [Saracevic en Kantor 1997], [Saracevic 2007a], [Saracevic 2007b].

1751 [Saracevic 1996, p. 202].

1752 Ibid., p. 205.

met de invalshoeken vanwaaruit relevantie in de loop der jaren is onderzocht.<sup>1753</sup> Saracevic onderscheidt vijf manifestaties op een glijdende schaal. Deze (hierna te bespreken) schaal loopt van het zuivere IT-perspectief naar het menselijk perspectief.

In deze typologie blijven *interface design* en *user navigation* overigens buiten beschouwing. Deze disciplines voegen zelf niets toe aan het bepalen van de relevantie,<sup>1754</sup> alhoewel ze grote invloed kunnen hebben op hoe de gebruiker de relevantiebepaling kan ervaren, zowel positief als negatief:

*In general, users like features that make retrieval systems look smart or that keep them from looking stupid, although it's hard to be very specific about what 'smart' or 'stupid' means in this context. These kinds of features can cause problems if they convince users that the system is more intelligent than it is. Effective retrieval will probably always require that users be actively engaged in the process.*<sup>1755</sup>

De eerste relevantiemanifestatie van Saracevic is de systeemrelevantie of algoritmische relevantie. Deze relevantie ziet op de relatie tussen een query<sup>1756</sup> en documenten in het computersysteem. Binnen deze manifestatie kan een document bijvoorbeeld relevant zijn als het gezochte woord erin voorkomt.<sup>1757</sup> Maar dat zegt niets over het onderwerp van het document, terwijl dat voor de relevantie van het document voor een gebruiker waarschijnlijk van groot belang is.

Daartoe onderscheiden we de tweede manifestatie: onderwerpsrelevantie. Hierbij wordt een relatie gelegd tussen het onderwerp waarop de gebruiker wil zoeken en het onderwerp van een document.

De derde manifestatie, de cognitieve relevantie of toepasselijkheid (*pertinence*) ziet op de vraag in hoeverre de gevonden informatie daadwerkelijk aansluit bij de informatiebehoefte van de gebruiker. Een document dat relevant is op grond van de eerste twee manifestaties

1753 Mizzaro gebruikte in zijn metastudie ([Mizzaro 1997]) een iets ander model, waarin we echter dezelfde grootheden herkennen die Saracevic als attributen en manifestaties benoemt. Relevantie kan volgens Mizzaro worden opgevat als een relatie tussen twee entiteiten uit twee verschillende groepen. In de eerste groep vinden we: document, surrogaat (een representatie van het document) en (in een document besloten liggende) informatie. In de tweede groep bevinden zich: het probleem van de gebruiker dat om een oplossing vraagt; de informatiebehoefte die de gebruiker uit het probleem heeft gedistilleerd; de informatievraag ('request' – een representatie van de informatievraag in menselijke taal); en ten slotte een query (*vide infra*, noot 1756). Ieder van de zeven entiteiten kan worden onderverdeeld in drie componenten, waarvan er één of meer van toepassing kunnen zijn: het onderwerp ('topic'); de taak waarop de activiteit van de gebruiker betrekking heeft; en de context. De vierde dimensie die Mizzaro toevoegt is tijd. We kunnen aldus bijvoorbeeld spreken over de relevantie van een document voor een query betreffende een specifiek onderwerp op een bepaald tijdstip.

1754 Al spelen ze natuurlijk wel een belangrijke rol bij bijv. systemen waarbij de gebruiker een relevantiebepalend systeem kan trainen door aan te geven of bepaalde documenten relevant zijn; [Turtle 1995, par. 6].

1755 *Ibid.*, p. 15.

1756 Van Dale: "informatieverzoek aan een databank". [Turtle 1995, p. 20-21] wijst op twee opvattingen omtrent het wezen van de query. Ten eerste kan een query worden gezien als een klein document, waarbij het doel is de overeenkomst tussen dit documentje en de documenten in de te doorzoeken collectie vast te stellen. Daarnaast kan een query worden opgevat als een regel of instructie voor het vinden van relevante documenten.

1757 De hiermee samenhangende meetproblemen bespreken we in § 5.2.2.



kan cognitief volstrekt irrelevant zijn, bijvoorbeeld omdat het te oud is, omdat de gebruiker het al kent, of omdat de gebruiker zijn zoekvraag verkeerd heeft geformuleerd.

Maar ook als het document wel blijkt aan te sluiten bij de informatiebehoefte, wil dat nog niet zeggen dat de gebruiker er iets aan heeft. De vierde manifestatie, de situationele relevantie of bruikbaarheid, breidt het perspectief van de gebruiker daarom uit naar de context van zijn zoekactie: helpt de gevonden informatie om het achterliggende probleem op te lossen?

Ten slotte benoemt Saracevic een vijfde manifestatie, die hij affectieve relevantie (*'motivational or affective relevance'*) noemt en omschrijft als:

*(The) relation between the intents, goals, and motivations of a user, and texts retrieved by a system or in the file of a system, or even in existence. Satisfaction, success, accomplishment, and the like are criteria for inferring motivational relevance.*<sup>1758</sup>

Cosijn en Ingwersen<sup>1759</sup> bouwden kritisch voort op deze typologie en stelden de vraag of de laatste manifestatie niet eerder als een attribuut van relevantie zou moeten worden beschouwd, en op de glijdende schaal van relevantiemanifestaties zou moeten worden vervangen door socio-cognitieve relevantie, ook wel contextuele- of domeinrelevantie te noemen:

*Socio-cognitive relevance is measured in terms of the relation between the situation, work task or problem at hand in a given socio-cultural context and the information objects, as perceived by one or several cognitive agents. It encompasses the system, a group of individual users or agents, and the socio-organisational environment.*<sup>1760</sup>

De analyse van citaties van wetenschappelijke publicaties wordt door Cosijn en Ingwersen aangehaald als voorbeeld van socio-cognitieve relevantie, die een domeinafhankelijke kwaliteitsmaat kan bieden.<sup>1761 1762</sup> We sluiten ons aan bij deze nuancering van Cosijn en Ingwersen op het model van Saracevic en beschouwen de domeinrelevantie als vijfde relevantiemanifestatie. We bakenen de domeinrelevantie zelfs nog iets scherper af en abstraheren geheel van de relatie met de zoekvraag, de informatiebehoefte of het gebruikersprobleem. Voor ons ziet domeinrelevantie dan uitsluitend op de mate waarin een document door domeinspecialisten belangrijk wordt gevonden binnen het desbetreffende kennisdomein.

### 5.2.1.2 Toepassing op jurisprudentiedatabanken

In de voorgaande paragraaf hebben we het begrip 'relevantie' in zijn algemeenheid onderzocht. In deze paragraaf passen we de bevindingen toe op jurisprudentiedatabanken, waarbij 'juridische domeinrelevantie' onze bijzondere aandacht heeft. Allereerst bespreken

1758 [Saracevic 1996, par. 4], zie ook [Saracevic 2007a].

1759 [Cosijn en Ingwersen 2000].

1760 Ibid., p. 541.

1761 Ibid., p. 548.

1762 De link met bibliometrie komen we ook bij anderen tegen. Zie bijvoorbeeld [Hjørland 2002].

we in § 5.2.1.2.1 hoe de verschillende relevantiemanifestaties in jurisprudentiedatabanken kunnen worden geoperationaliseerd. Om iets meer beeld te krijgen van hoe juridische domeinrelevantie in de praktijk kan worden gebruikt, bekijken we in § 5.2.1.2.2 een jurisprudentiedatabank die een importantie-indicator heeft geïmplementeerd.

### 5.2.1.2.1 Relevantiemanifestaties

We kunnen de theorievorming uit § 5.2.1.1 vertalen naar hulpmiddelen waarmee uitspraken-databanken kwantitatief hanteerbaar kunnen worden gemaakt. Een samenvattende matrix met de relevantiemanifestaties is opgenomen in Figuur 5-1.

De vijf verschillende relevantiemanifestaties kunnen worden vertaald naar juridische retrieval.<sup>1763</sup> De hulpmiddelen die we vinden op het niveau van de systeemrelevantie zijn weinig specifiek voor het juridisch domein, al kan finetuning voor het juridisch domein tot betere resultaten leiden: zo moeten lijsten die voor stemming<sup>1764</sup> worden gebruikt juridische terminologie bevatten.

Voor onderwerpsrelevantie kan het gebruik van een taxonomie van grote waarde zijn. Daarbij kan een spanningsveld ontstaan tussen gewenste detaillering enerzijds en bruikbaarheid anderzijds: hoe meer lemmata zich in de taxonomie bevinden, hoe groter de kans om relevante documenten te vinden, maar tegelijkertijd wordt de classificatie van documenten moeilijker. De taxonomie moet vanzelfsprekend juridisch van aard zijn, maar niet iedere willekeurige juridische taxonomie is ook voor classificatie van rechterlijke uitspraken de meest geschikte. Indexering van documenten gebeurt traditioneel door redacteurs, maar daaraan kleven de nodige problemen:

*Manual indexing is only as good as the ability of the indexer to anticipate questions to which the indexed document might be found relevant. It is limited by the quality of its thesaurus. It is necessarily precoordinated and is thus also limited in its depth. Finally, like any human enterprise, it is not always done as well as it might be.*<sup>1765</sup>

In toenemende mate gebeurt het indexereren ook door computers, maar ook dat leidt niet altijd tot bevredigende resultaten.<sup>1766</sup>

1763 Al zijn natuurlijk andere indelingen mogelijk, zie bijvoorbeeld [Sutton 1994, p. 187], die voor juridische retrieval drie manifestaties hanteert, die grofweg overeenkomen met onderwerpsrelevantie, situationele relevantie en domeinrelevantie. Sutton wijst daarnaast op de specifieke denkwijze van juristen, waarmee tot nog toe onvoldoende rekening wordt gehouden in juridische retrievalssystemen. Kritiek op de gebrekkige aansluiting van jurisprudentiedatabanken op de *mindset* van de jurist is al zo oud als de databanken zelf, vgl. [Dabney 1986] en [Smith 2007].

1764 *Vide infra*: § 5.2.2.1.1.

1765 [Dabney 1986, p. 14].

1766 De Amerikaanse commerciële systemen LexisNexis en WestLaw hebben hun content op vele duizenden (hiërarchisch gestructureerde) onderwerpen georganiseerd. Bij WestLaw delen redacteurs uitspraken handmatig in, LexisNexis laat de classificatie door de computer doen. Een vergelijking van de relevantiescores laat zien dat de mens het tot nu toe beter doet, maar dat een gebruiker altijd aanvullende zoekwegen moet bewandelen om alle relevante informatie te verkrijgen ([Mart 2010]).

Relevantiemanifestatie	Mogelijke middelen
Systeemrelevantie	Standaard retrievaltechnieken.
Onderwerpsrelevantie	Taxonomie
Cognitieve relevantie / toepasselijkheid	Thesaurus. Vertaling naar lekenterminologie / matching met casuïstiek
Situationele relevantie / bruikbaarheid	Citatie-indexen, vertalingen, toelichtingen, inhoudsindicaties
Domeinrelevantie	Importantie-indicator

**Figuur 5-1.** Matrix voor mogelijke hulpmiddelen in jurisprudentiedatabanken voor de verschillende relevantiemanifestaties.

Voor de cognitieve relevantie kan het inzetten van een thesaurus – waarin gerelateerde begrippen en synoniemen zijn opgenomen – een hulpmiddel zijn. Hiermee kan een gebruiker zoekvragen meer toesnijden op de specifieke problematiek die de aanleiding vormt voor zijn informatievraag. Voor niet-juridische gebruikers kan er meerwaarde zitten in een koppeling tussen lekentermen en juridische begrippen.<sup>1767</sup> De grens tussen cognitieve relevantie en situationele relevantie is vrij diffuus, vooral omdat de situatie van de gebruiker zich aan het oog van het systeem onttrekt. Maar de kans dat de informatie als bruikbaar wordt ervaren kan worden vergroot door het aanbieden van metadata als samenvattingen, toelichtingen, formele en materiële relaties, dat laatste bijvoorbeeld in de vorm van citatie-indexen.

De domeinrelevantie ten slotte vraagt naar de jurisprudentiële waarde van een uitspraak: wat is de juridische importantie van deze uitspraak in het algemeen? Deze domeinrelevantie staat los van de systeemrelevantie die de uitspraak heeft voor een bepaalde query, de onderwerpsrelevantie van de uitspraak met betrekking tot een bepaald thema of rechtsgebied, en de cognitieve relevantie die de uitspraak heeft voor de informatievraag van de gebruiker. Er kan wel een relatie zijn tussen situationele relevantie en domeinrelevantie: er zijn veel situaties denkbaar waarin de bruikbaarheid van een gevonden uitspraak niet alleen wordt bepaald door het probleem dat de gebruiker tot de zoekactie aanzette, maar ook door de jurisprudentiële waarde ervan. Er zijn zelfs queries denkbaar waarin situationele relevantie en domeinrelevantie volledig samenvallen, bijvoorbeeld in het geval de gebruiker alle juridisch belangrijke uitspraken wil zien die de afgelopen week door een bepaalde rechtbank zijn gedaan. Beide relevantiemanifestaties moeten echter scherp van elkaar worden onderscheiden: bij het bepalen van de situationele relevantie is altijd een individuele gebruiker betrokken, bij het vaststellen van de domeinrelevantie nooit.

<sup>1767</sup> Zoals bijv. gerealiseerd in het BEST-project: [Hoekstra 2009].

### 5.2.1.2.2 'Importance levels' in HUDOC

Wat wordt bedoeld met domeinrelevantie bij jurisprudentiedatabanken kan worden geïllustreerd met de HUDOC-databank van het EHRM, waar een specifiek zoekfilter voor deze relevantiemaniëfestatie beschikbaar is onder de naam 'importance level'.<sup>1768</sup> De gebruiker kan zijn zoekvraag beperken tot één of meer van vier niveaus; tot aan de herziening van de website in juni 2012<sup>1769</sup> waren er drie niveaus. De benamingen en omschrijvingen van de huidige categorieën luiden:<sup>1770</sup>

- **Case Reports**

*Judgments, decisions and advisory opinions delivered since the inception of the new Court in 1998 which have been published or selected for publication in the Court's official Reports of Judgments and Decisions. The selection from 20007 onwards had been made by the Bureau of the Court following a proposal by the Jurisconsult.*

*Judgments of the former Court (published in Series A and Reports) and cases published in the former Commission's series Decisions and Reports have not been included in the Case Reports category and are therefore classified by levels 1, 2 and 3 only.*

- **1 = High importance**

*All judgments, decisions and advisory opinions not included in the Case Reports which make a significant contribution to the development, clarification or modification of its case-law, either generally or in relation to a particular State.*

- **2 = Medium importance**

*Other judgments, decisions and advisory opinions which, while not making a significant contribution to the case-law, nevertheless go beyond applying existing case-law.*

- **3 = Low importance**

*Judgments, decisions and advisory opinions of little legal interest, namely judgments and decisions that simply apply existing case-law, friendly settlements and strike outs (unless raising a particular point of interest).*

Vóór juni 2012 bestonden alleen de categorieën 1, 2 en 3; *Case Reports* was geen afzonderlijke categorie. Alhoewel een indeling in vier categorieën fijnmaziger is dan in drie categorieën, heeft het nieuwe systeem enkele nadelen. Ten eerste is het verwarrend dat er één categorie is met een afwijkende benaming. Voor de niet-goed-ingevoerde gebruiker is het even puzzelen om te begrijpen hoe *Case Reports* zich verhoudt tot de drie genummerde categorieën.<sup>1771</sup>

1768 *Vide supra*: § 2.6.1.1.

1769 *Vide supra*: § 2.6.1.1.

1770 Naast de toevoeging in 2012 van de *Case-Reports*-categorie zijn ook de beschrijvingen van de andere categorieën gewijzigd, maar deze wijzigingen zijn louter redactioneel van aard.

1771 Waarbij nog opgemerkt kan worden dat de nummering contra-intuïtief is; op veel websites worden waarderungen gebruikt die worden gevisualiseerd met sterren. Daarbij geldt: hoe meer sterren, hoe beter. De nummering van HUDOC werkt precies andersom.

Ten tweede is de *Case-Reports*-categorie weinig flexibel. De toelichting vermeldt dat een uitspraak ten aanzien waarvan nog geen publicatiebesluit is genomen, voorlopig in een van de andere drie categorieën wordt opgenomen. Maar indien het besluit tot publicatie eenmaal is genomen, dan blijft een zaak tot aan het einde der tijden in de *Case-Reports*-categorie, terwijl uitspraken in de andere drie categorieën wel van de ene naar de andere categorie kunnen verhuizen als gevolg van een veranderende perceptie inzake hun juridisch belang. Daar komt nog bij dat oude uitspraken (bijvoorbeeld uit *Series A*) nooit hoger komen dan categorie 1. De toevoeging van de *Case-Reports*-categorie in het nieuwe HUDOC heeft daarom de zuiverheid van het classificatiesysteem enigszins aangetast.

In Figuur 5-2 is te zien hoe de uitspraken op drie verschillende tijdstippen over de categorieën waren verdeeld. Hieruit blijkt niet alleen dat het overgrote deel van de uitspraken in de laagste categorie valt, maar ook dat uitspraken regelmatig van de ene naar de andere categorie verhuizen. Bij de creatie van de *Case-Reports*-categorie in juni 2012 lijken er veel uitspraken uit met name de *High-Importance*-categorie naar deze nieuwe categorie te zijn verplaatst. Veel uitspraken uit de andere categorieën maakten daardoor ook promotie. Het volledige overzicht, inclusief een onderscheid naar *judgments* en *decisions*, is opgenomen in Bijlage 10.

## 5.2.2 Het meten van relevantie

Nu we hebben vastgesteld welke relevantiemaniëfestaties er zijn, kunnen we onderzoeken hoe relevantie in de praktijk kan worden gemeten. Ook daarbij maken we een onderscheid tussen generieke theorie (§ 5.2.2.1) en de toepassing op jurisprudentiedatabanken (§ 5.2.2.2).

### 5.2.2.1 Generieke theorie

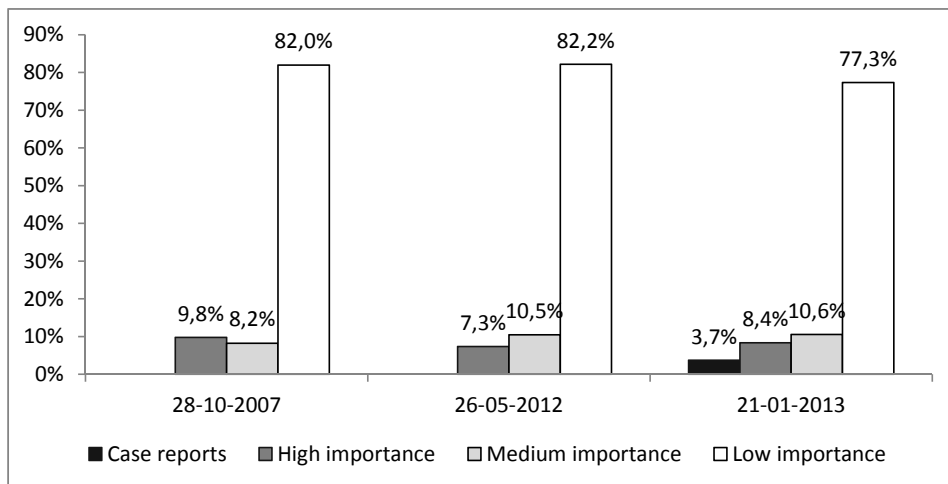
In de voorgaande paragrafen hebben we gezien dat relevantie een meervoudig begrip is. En wanneer we in staat zijn om het begrip ‘relevantie’ binnen een bepaalde context te definiëren, stuiten we direct op de vraag hoe we die relevantie kunnen meten:

*(L)ike so many other things in life, relevance is relative. Some things are more relevant than others, some things are not relevant any more, and some have a chance of becoming relevant.*<sup>1772</sup>

In deze paragraaf zullen we ingaan op de problemen die samenhangen met het kwantificeren van relevantie. In § 5.2.2.1.1 behandelen we eerst enkele algemene vraagstukken met betrekking tot relevantiemeting. Ook de vraag hoe meetsystemen beoordeeld moeten worden komt daarbij aan de orde. Omdat we in het vervolg van dit onderzoek naast klassieke statistiek ook gebruik zullen maken van sociaalnetwerkanalyse, introduceren we in § 5.2.2.1.2 hieruit enkele basale begrippen.

---

1772 [Saracevic 1975, p. 325].



**Figuur 5-2.** Verdeling van EHRM-uitspraken over de *importance levels*, in 2007, 2012 en 2013.

### 5.2.2.1.1 Het meten van een relatief begrip

In § 5.2.1.1 hebben we – in navolging van Saracevic – een onderverdeling gemaakt in vier thema's met betrekking tot relevantie, waarvan we er twee in de genoemde paragraaf hebben behandeld: aard en manifestatie. In deze paragraaf behandelen we de twee andere thema's: gedrag en effecten. 'Gedrag' heeft betrekking op hoe mensen relevantie beoordelen: welke factoren vinden zij belangrijk? 'Effecten' ziet op de vraag waardoor het menselijk oordeel over relevantie (en dus het gedrag) wordt beïnvloed. Beide thema's zijn nauw gerelateerd, zoals reeds bleek bij een studie van Gull in 1956 waarbij de prestaties van twee verschillende zoeksystemen werden getest.<sup>1773</sup> De onbedoelde en onverwachte uitkomst was niet zozeer dat de zoeksystemen verschilden, maar dat de beoordelaars het niet eens konden worden over de vraag welke documenten als relevant aangemerkt zouden moeten worden. In onze bespreking zullen we dan ook niet te strikt vasthouden aan het onderscheid tussen gedrag en effecten, maar ons richten op de vraag hoe prestaties van retrievalssystemen voor de verschillende relevantiemaniëfestaties kunnen worden gemeten.

We beginnen met de systeemrelevantie. Hiervoor zijn in de loop der jaren verschillende maten ontwikkeld. Eerder<sup>1774</sup> kwamen we al de begrippen *recall* en *precision* tegen, klassieke maten voor de prestaties van retrievalssystemen.<sup>1775</sup> *Recall* is het percentage documenten dat correct wordt gevonden binnen het totaal aantal documenten dat had moeten worden gevonden. *Precision* is het percentage documenten in een antwoordset dat zich daarin te-

<sup>1773</sup> [Gull 1956], ook besproken in [Saracevic 2007b, p. 2134], [Saracevic 2008, p. 773].

<sup>1774</sup> *Vide supra*: § 4.3.3.

<sup>1775</sup> Zie o.m. [Baeza-Yates en Ribeiro-Neto 1999, p. 75].

recht bevindt. Op het eerste gezicht lijken dergelijke maten objectief, maar de exacte criteria moeten heel nauwkeurig worden omschreven: als we zoeken naar documenten over ‘fiet-sen’, moeten dan ook documenten worden gevonden waarin ‘fietsen’ slechts een substring vormt (racefietsen, bakfietsen, snorfietsen, waterfietsen, luchtfietsen)? Moet er ‘gestemd’<sup>1776</sup> worden gezocht, zodat we ook documenten vinden waarin de woorden ‘fietst’ of ‘gefietst’ voorkomen? En moet er ook worden gezocht naar synoniemen (rijwielen), bredere termen (tweewielers) of deelverzamelingen (mountainbikes)?

Hierin zien we reeds een nauwe verwevenheid met de onderwerpsrelevantie, waarbij wordt getoetst of een document relevant is voor het onderwerp waarop is gezocht. Voor de beoordeling van de resultaten spelen verschillende factoren een rol, zoals de uitgebreidheid van de taxonomie die is gehanteerd, en de vraag of ook documenten uit een moeder- of dochtercategorie moeten worden gevonden. Voor het testen van ontwikkelde zoekalgoritmen hebben we bovendien niet alleen te maken met de beoordeling door de eindgebruiker, maar ook met de classificatie van de documenten (door mens of computer).

Bij de cognitieve relevantie en de situationele relevantie wordt de subjectieve component nog groter. Onderzoeken<sup>1777</sup> laten dan ook zien dat die menselijke beoordelaar bij het ontwikkelen van zoekalgoritmen een onzekere en wispelturige factor is.<sup>1778</sup> Deze wispelturigheid komt op twee momenten tot uitdrukking. Allereerst bij het vaststellen van de gouden standaard: het benoemen van de set documenten die op een bepaalde zoekvraag gevonden zou moeten worden. Saracevic<sup>1779</sup> benoemt vier (groepen van) personen die vaak worden ingezet om te komen tot een gouden standaard: eindgebruiker, domeinspecialist, informatiespecialist, buitenstaander. Vervolgens komt de instabiliteit van de relevantietoets ook tot uiting bij het beoordelen van de testset tegen de gekozen gouden standaard. Uit verschillende onderzoeken komen als algemene conclusies naar voren:

1. Hoe groter de expertise van de beoordelaars, hoe groter de overlap en consistentie in het oordeel;
2. Hoe meer beoordelaars, hoe kleiner de overlap en de consistentie.<sup>1780</sup>

Zowel bij het vaststellen van de gouden standaard, als bij het vergelijken van een concrete resultaatset met die gouden standaard, wordt een belangrijk obstakel dus gevormd door de beperkingen van de steekproef.

Het afzien van steekproeven heeft pas recent de aandacht getrokken, onder meer omdat de technische mogelijkheden hiertoe inmiddels aanwezig zijn.<sup>1781</sup> ‘Crowdsourcing’<sup>1782</sup> kan veel meer kennis genereren dan welke steekproef dan ook, en lijkt succesvol voor een aantal

1776 Bij *stemming* moeten ook alle woorden worden gevonden die gebaseerd zijn op dezelfde woordstam.

1777 Een uitgebreid overzicht van verschillende onderzoeken is opgenomen in [Saracevic 2007b].

1778 [Saracevic 2008].

1779 Ibid., p. 769.

1780 [Saracevic 2007b, p. 2137].

1781 Door Saracevic wordt aan dit thema vreemd genoeg vrijwel geen aandacht besteed.

1782 Door Van Dale vrij beperkt gedefinieerd als: “*Vorm van productverbetering waarbij producenten de kennis en vaardigheden van hun klanten gebruiken om hun producten of diensten te verbeteren.*”

toepassingen.<sup>1783</sup> Crowdsourcing lijkt effectief om de onderwerpsrelevantie van informatie-objekten beter te bepalen, omdat het steekproefprobleem kan worden geëlimineerd. Als bijvoorbeeld iedereen kan meehelpen met het taggen<sup>1784</sup> van foto's op internet kan meer (gespecialiseerde) kennis worden ingeschakeld dan wanneer dit werk zou zijn voorbehouden aan een kleine groep.

Crowdsourcing kan ook worden ingezet door mensen niet expliciet om een bijdrage te vragen, maar door gebruik te maken van opgeslagen gegevens omtrent hun gedrag. Ten behoeve van cognitieve relevantie kan bijvoorbeeld worden onderzocht welke documenten gebruikers van een website vanuit een resultaatpagina hebben geopend nadat zij op een bepaald woord hebben gezocht. Op eenzelfde manier kan crowdsourcing worden gebruikt voor het bepalen van situationele relevantie, bijvoorbeeld door de geografische locatie van een gebruiker te gebruiken om te onderzoeken wat relevant is voor het geografisch nabije crowdsegment.

Voor domeinrelevantie heeft crowdsourcing een bijzondere betekenis. We hebben domeinrelevantie gedefinieerd als de relevantie die een informatieobject heeft binnen een specifiek kennisdomein, zoals gepercipieerd door mensen die goed zijn ingevoerd in dat domein, en onafhankelijk van een specifieke gebruikersvraag. Als voorbeeld noemden we reeds<sup>1785</sup> citatieanalyse, die onder meer wordt ingezet om de importantie van auteurs en tijdschriften te meten. Hierbij wordt geen testpanel ingezet, maar wordt gebruik gemaakt van alle citaties die in het corpus beschikbaar zijn. De reden hiervoor is dat het vooralsnog onmogelijk is gebleken om de computer te laten voorspellen hoe populair een wetenschappelijke publicatie zal zijn.

Ook de hieronder<sup>1786</sup> nog te bespreken algoritmen die door internetzoekmachines worden gebruikt, zijn voor een belangrijk deel gebaseerd op citatie, en daarmee ook op de *wisdom of the crowd*. Als veel pagina's naar website X verwijzen, en als website X veel wordt bezocht, dan zal deze ook wel belangrijk zijn.<sup>1787</sup> Deze zeer uitgebreide wereldwijde domeinkennis wordt voor specifieke queries gecombineerd met algoritmen uit andere relevantiemaniëfestaties. De *clues* daarvoor komen uit de metadata die op een pagina of website aanwezig zijn, of de inhoud van documenten en perifere gegevens, zoals gegevens over de individuele gebruiker.<sup>1788</sup>

1783 [Surowiecki 2005]. Mede met betrekking tot het overheidsdomein: [TNO 2008].

1784 Van Dale: "Labels hechten aan iets, bv. aan woorden in een corpus."

1785 *Vide supra*: § 5.2.1.1.

1786 *Vide infra*: § 5.2.2.1.2.

1787 Juist omdat er geen gebruik wordt gemaakt van steekproeven, is het aantrekkelijk en soms relatief eenvoudig om zoekalgoritmen te misleiden. Voor citatie-indexen is dat bijvoorbeeld overvloedige zelfcitatie, voor zoekmachines de 'Google bomb'.

1788 De precieze implementatie van de rankingalgoritmen van de zoekgiganten behoort tot de bedrijfsgeheimen.



### 5.2.2.1.2 Traditionele statistiek versus netwerkanalyse

Voor het berekenen van de relevantie van een document (om het even in welke relevantie-manifestatie) wordt gebruik gemaakt van allerlei mathematische en statistische modellen. Daarbij wordt vaak een vergelijking gemaakt tussen de zoekvraag enerzijds en een aantal *clues* in de geïndexeerde documenten anderzijds. Die zoekvraag kan als query zijn geformuleerd, maar ook in termen van een probleem waar een gebruiker mee worstelt. Vervolgens wordt gekeken in hoeverre de documenten zich individueel verhouden tot de zoekvraag. Goffman was een van de eersten die erop wees dat dit niet volstaat; ook de relatie tussen de objecten onderling kan van invloed zijn op het uiteindelijke resultaat.<sup>1789</sup> Door deze bevinding werd het noodzakelijk om op een andere manier te gaan kijken naar dataverzamelingen. Sociaalnetwerkanalyse kreeg hierin een belangrijke plaats. Omdat deze analysemethode een belangrijke rol speelt in ons verdere onderzoek gaan we in deze paragraaf kort in op de basisprincipes<sup>1790</sup> en de belangrijkste conceptuele verschillen met de traditionele statistiek.

In deze traditionele statistiek worden overeenkomsten en verschillen tussen objecten (bijvoorbeeld mensen of documenten) onderzocht door hun kenmerken (attributen) met elkaar te vergelijken. In netwerkanalyse echter staan naast de objecten niet de attributen centraal, maar de relaties ('edges' of 'links') die er tussen de objecten ('nodes') bestaan. In de traditionele statistiek kunnen we tot uitdrukking brengen dat A en B zijn getrouwd, maar niet dat ze juist met elkaar in de echt zijn verbonden. Kennis van dergelijke relaties is echter essentieel als men wil analyseren hoe sociale netwerken in elkaar zitten. De attributen uit de traditionele statistiek worden in de sociaalnetwerkanalyse beschouwd als relaties. Als bijvoorbeeld A en B als attribuut hebben dat ze van het geslacht 'vrouw' zijn, hebben ze in sociaalnetwerkanalyse een relatie met elkaar met betrekking tot het kenmerk 'vrouw'. Door deze constructie zijn binnen de sociaalnetwerkanalyse in beginsel ook alle methodieken uit de traditionele statistiek toepasbaar.

Een tweede belangrijk punt waarop sociaalnetwerkanalyse zich onderscheidt van de traditionele statistiek is de omvang van de geanalyseerde dataverzameling. In de statistiek neemt de steekproef een belangrijke plaats in, terwijl men in de netwerkanalyse met de volledige verzameling wil werken.<sup>1791</sup> Immers, als men in netwerkanalyse een object voor een steekproef zou willen selecteren, dan zou men, om een goed beeld te krijgen van dit object, ook alle objecten moeten selecteren waarmee het geselecteerde object een relatie heeft. En omdat de eigenschappen van die objecten worden beïnvloed door weer andere objecten, moeten ook die in de steekproef worden meegenomen, et cetera. Indien technisch ook maar enigszins mogelijk wordt daarom de voorkeur gegeven aan analyse van het volledige netwerk.<sup>1792</sup>

1789 [Goffman 1970].

1790 Zie voor een uitgebreider inleiding o.m. [Hanneman en Riddle 2005].

1791 Vgl. o.m. [de Nooy, Mrvar en Batagelj 2011, p. 6].

1792 Er bestaan wel methoden om in netwerkanalyse steekproeven te trekken. Daarbij wordt begonnen met een aantal objecten en worden slechts gerelateerde objecten tot in de n-de graad in de steekproef betrokken. Een belangrijk nadeel van deze methoden is dat *isolates* (objecten zonder connectie) buiten de steekproef blijven.

Binnen sociaalnetwerkanalyse kunnen relaties tussen objecten gericht (*directed*) of ongericht (*undirected*) zijn. Een gerichte relatie is (in beginsel) asymmetrisch, zoals 'houden van' of 'citeren'; een ongerichte relatie is altijd symmetrisch, zoals 'getrouwd zijn met'. Netwerkrelaties kunnen worden weergegeven in matrices, maar voor een beter (visueel) begrip worden ze ook vaak weergegeven in een graaf.

Van oudsher werden netwerken vooral beschouwd als *random networks*, waarin de toevoeging van nodes en links willekeurig geschiedt. De verdeling van het aantal links heeft daardoor een Poisson-distributie. Het World Wide Web (WWW) blijkt echter een '*scale free network*'<sup>1793</sup> te zijn, waarvan de verdeling voldoet aan een *Power law*: de distributie van links is in hoge mate ongelijk, er zijn enkele objecten met een groot aantal inkomende links en daarnaast heel veel objecten die heel weinig (of zelfs geen enkele) inkomende link hebben. *Scale free networks* kenmerken zich door een dynamische groei van het aantal nodes en *preferential attachment*:<sup>1794</sup> het verschijnsel dat links niet naar willekeurige andere nodes worden gelegd, maar naar die nodes die reeds een bovengemiddeld aantal links hebben. Het vermogen van een node om een bovengemiddeld aantal inkomende links te verwerven zou onder meer afhankelijk zijn van leeftijd en kwaliteit.<sup>1795</sup>

Netwerken kunnen op talloze manieren worden geanalyseerd. De aandacht kan daarbij worden gericht op het gehele netwerk (bijvoorbeeld de dichtheid van de relaties binnen het netwerk), op het bestaan van subnetwerken of op individuele nodes. Bij dat laatste gaat het met name om de vraag in hoeverre een node een centrale positie heeft in een netwerk. Om dat te berekenen zijn een aantal *centrality measures* ontwikkeld, waarvan we de belangrijkste hier bespreken.

Om te beginnen kunnen we tellen hoeveel relaties een node heeft, hetgeen we uitdrukken met de term '*degree*'. In een gericht netwerk kan daarbij een onderscheid worden gemaakt tussen de inkomende relaties ('*in-degree*') en uitgaande relaties ('*out-degree*').<sup>1796</sup> Met '*betweenness*' drukken we uit in hoeverre een node een cruciale positie inneemt op het pad dat moet worden gevolgd door andere nodes om met elkaar in contact te komen.<sup>1797</sup> Een derde maat voor *centrality* is '*closeness*', waarmee we kunnen bepalen hoe ver nodes van elkaar zijn verwijderd. We kunnen dat berekenen door het aantal relaties te tellen dat moet worden gepasseerd om van de ene bij de andere node te komen. Ook zijn er modellen, zoals *power-centrality*<sup>1798</sup> en *proximity prestige*,<sup>1799</sup> waarin *degree* en *closeness* worden gecombineerd. In deze modellen wordt niet alleen rekening gehouden directe relaties, maar ook met indirecte relaties.

Voor de manier waarop informatie op het WWW met elkaar is verbonden, en wat daaruit is af te leiden omtrent het belang van websites, zijn afzonderlijke, op eigenvectoren

1793 [Barabási en Bonabeau 2003].

1794 Ibid., p. 54.

1795 [Bianconi en Barabási 2001].

1796 [Freeman 1979].

1797 [Freeman 1977].

1798 [Bonacich 1987].

1799 [Liu 2011, p. 274].

gebaseerde modellen ontwikkeld. Kleinberg introduceerde het ‘*Hyperlink-Induced Topic Search*’-algoritme (HITS),<sup>1800</sup> waarin een onderscheid wordt gemaakt tussen ‘*hubs*’ (pagina’s die vooral een verwijzingsfunctie hebben) en ‘*authorities*’ (pagina’s waarnaar veel wordt verwezen).<sup>1801</sup> Voor beide eigenschappen krijgt een website een waardering, vergelijkbaar met *in-degree* en *out-degree*. De gedachte daarbij is dat een *hub* sterk is als deze verwijst naar veel sterke *authorities*, en dat een *authority* sterk is als er door veel sterke *hubs* naar wordt verwezen. HITS is query-specifiek: voor elke nieuwe vraag wordt er een berekening uitgevoerd.<sup>1802</sup>

Dit model is doorontwikkeld door Brin en Page<sup>1803</sup> in het *PageRank*-algoritme, dat de basis vormt voor het ranking-algoritme van Google. Anders dan HITS is *PageRank* niet query-specifiek, en levert het geen waarden op voor *hubs*, maar alleen voor *authorities*. Beide maten houden geen rekening met het onderwerp van de webpagina’s. Sommige pagina’s gaan over tal van onderwerpen, en voor de gebruiker kan het relevant zijn dat de link tussen twee pagina’s betrekking heeft op het specifieke onderwerp van zijn interesse. Daarom wordt inmiddels geëxperimenteerd met modellen waarin een semantische laag over HITS en *PageRank* heen wordt gelegd.<sup>1804</sup>

In een groot netwerk kan de positie van *nodes* soms worden verhelderd door het netwerk eerst op te delen in kleinere, coherente subgroepen (‘*cores*’). Met algoritmen zoals ‘*generalized core*’<sup>1805</sup> kan de optimale opdeling van een netwerk worden berekend.

Een andere specifieke toepassing van *centrality* vinden we in de bibliometrie, en meer in het bijzonder bij citatieanalyse. De *impact factor* werd ontwikkeld om de invloed van wetenschappelijk werk te meten aan de hand van het aantal inkomende citaties. Het oorspronkelijke idee dat in 1955 werd gelanceerd had betrekking op artikelen,<sup>1806</sup> later werd de methode als ‘*journal impact factor*’ (JIF) doorontwikkeld voor volledige tijdschriften.<sup>1807</sup> In de praktijk leiden deze maten regelmatig tot problemen, bijvoorbeeld omdat in sommige wetenschappelijke domeinen citatie meer gebruikelijk is dan in andere.<sup>1808</sup> Bovendien kan manipulatie worden uitgelokt, bijvoorbeeld omdat wetenschappers willen publiceren in goedscorende tijdschriften.<sup>1809</sup> Als reactie daarop zijn meer verfijnde algoritmen ontwikkeld. Zo maakt de ‘*eigenfactor*’<sup>1810</sup> een weging van tijdschriften door rekening te houden met het totaal aantal citaties dat in het citerende tijdschrift verschijnt. De ‘*article influence*’ berekent

1800 [Kleinberg 1999].

1801 Alhoewel het WWW voor Kleinberg de inspiratie was het ontwikkelen van dit algoritme, wijst hij erop dat het ook op andere zaken dan internetpagina’s kan worden toegepast.

1802 Vgl. [Signorini 2005].

1803 [Brin en Page 1998].

1804 Bijv. [Nie, Davison en Qi 2006].

1805 [Batagelj en Zaveršnik 2010].

1806 [Garfield 1955].

1807 Zie voor een historisch overzicht [Garfield 2005].

1808 Ibid., p. 5.

1809 Zie bijv. [Schutte en Svec 2007].

1810 [Bergstrom 2007].

een vergelijkbare maat voor artikelen.<sup>1811</sup> Als alternatieven voor het berekenen van de impact van tijdschriften zijn ook HITS en PageRank voorgesteld.<sup>1812</sup> Met de ‘*h-index*’ (Hirsch-index)<sup>1813</sup> kan de productiviteit en impact van individuele wetenschappers worden gemeten.

### 5.2.2.2 Toepassing op jurisprudentiedatabanken

De bevindingen van specifiek op het juridisch domein gericht onderzoek naar relevantiebepaling en relevantietoetsing wijken niet af van het meer generieke onderzoek dat daarvoor is gedaan.<sup>1814</sup> menselijke beoordelaars van relevantie(maatstaven) zijn vaak verdeeld en inconsistent, en objectivering blijkt bijzonder lastig. Blair en Maron<sup>1815</sup> constateerden bijvoorbeeld dat dezelfde juristen op verschillende momenten verschillende indexterminen aan hetzelfde document toekennen. Meer gericht op domeinrelevantie was het onderzoek van Eldridge,<sup>1816</sup> die tot de conclusie kwam dat een meer dan binaire maatstaf de onenigheid vergroot: hij verzocht juristen om rechterlijke uitspraken met een vierpuntsschaal op juridische relevantie te beoordelen. Bij het oordeel ‘irrelevant’ waren alle vier het vaak eens, maar bij alle uitspraken waarover ten minste één van hen een ander oordeel had, kwamen ze slechts in 4,3% van de gevallen tot een eensluidend oordeel.

Inzet van computers bij relevantiebepaling vormt een aantrekkelijk perspectief. Niet alleen omdat computers minder last hebben van subjectiviteit, maar ook omdat bij grotere collecties handmatige relevantiebepaling van documenten een prijzige aangelegenheid wordt. Bij de HUDOC-databank geschiedt de classificatie naar *importance level*<sup>1817</sup> handmatig. De bescheiden jaarlijkse productie van het EHRM, het beperkte rechtsterrein waarop wordt geopereerd en een goed onderlegd stafbureau maken dit mogelijk, maar voor omvangrijker collecties wordt zo iets al snel onbetaalbaar.

Hieronder geven we een kort overzicht van onderzoek dat de afgelopen jaren is gedaan naar de mogelijkheden om juridische relevantie op geautomatiseerde wijze te laten bepalen. Veel van dit onderzoek is gebaseerd op citatieanalyse.

Voordat citaties werden gezien als een bron voor evaluatief wetenschapsonderzoek, was het gebruik van juridische citatie-indexen, met name in *common-law*-landen, reeds wijdverspreid.<sup>1818</sup> In eerste instantie bestonden deze systemen vooral op papier, maar reeds in 1946 voorzag de Amerikaanse jurist Kelso de ‘*Lawdex*’: een geautomatiseerd systeem waarin

---

1811 Bergstrom (ibid., p. 315) doet dit door de eigenfactor te delen door het aantal artikelen dat in het tijdschrift verschijnt.

1812 [Cheng, YunTao, JunPeng et al. 2009].

1813 [Hirsch 2005].

1814 *Vide supra*: § 5.2.2.1.1.

1815 [Blair en Maron 1985, p. 290], alhoewel deze vraag niet betrekking had op het belang van een document, maar op het toekennen van indexterminen aan dat document.

1816 [Eldridge 1968].

1817 *Vide supra*: § 5.2.1.2.2.

1818 Zie voor een historisch overzicht [Shapiro 1992].

alle juridische informatie zou zijn opgeslagen.<sup>1819</sup> Een deels door mensen en deels door een machine toegekende importantie-indicator zou daarvan onderdeel moeten uitmaken:

(S)o that the lawyer can, if he prefers, review only four-star data (...).<sup>1820</sup>

Het duurde lang voordat dit idee verder werd uitgewerkt. In 1970 constateerde Marx<sup>1821</sup> dat 'shepardization' (het volgen van ingaande en uitgaande citaties)<sup>1822</sup> weliswaar voor juristen een van de belangrijkste onderzoeksmethoden was, maar dat juridische informatiesystemen daarentegen vooral waren gebaseerd op (omvangrijke) taxonomieën. Hij pleitte daarom voor meer gebruik van de citaties in retrievalsysteem, en toonde in een onderzoek van bescheiden omvang aan dat dit tot goede resultaten zou kunnen leiden. Eenzelfde pleidooi werd in 1981 gehouden door Tapper,<sup>1823</sup> die naast uitspraakcitaties ook oog had voor het belang van andere variabelen.<sup>1824</sup> Het (vector)model dat Tapper (voor een experimentele setting en met een beperkte dataset) ontwikkelde, gebruikte hij vooral om de cognitieve en situationele relevantieberekening te verbeteren. Domeinrelevantie speelde daarin een rol, maar was geen doel op zichzelf.

Ander onderzoek is meer exploratief van aard, en richt zich op de aard van jurisprudentie-citatie-netwerken<sup>1825</sup> en de betekenis ervan voor het wezen van het recht zelf. Zo bouwen Post en Eisen<sup>1826</sup> voort op de theorie van Balkin<sup>1827</sup> dat het recht kan worden beschouwd als een fractalstructuur. Hun hypothese dat de omvang van deze structuur kan worden gemeten in inkomende en uitgaande citaties, wordt door hun onderzoek slechts in zeer beperkte mate ondersteund (zij het wel statistisch significant).

Smith<sup>1828</sup> stelt in een vergelijkbare maar uitgebreidere<sup>1829</sup> studie dat het netwerk van het recht (*web of law*) vergelijkbaar is met een *scale free network*,<sup>1830</sup> en daarom ook aan de daarvoor geldende wetten voldoet. Hij toont aan dat de scheefheid in citatieverdeling grote overeenkomst vertoont met die van citatie-indexen van wetenschappelijke artikelen: enkele uitspraken worden heel vaak geciteerd,<sup>1831</sup> terwijl het merendeel van de uitspraken slechts één keer of helemaal nooit wordt aangehaald. Dit strookt met onze eerdere constatering

---

1819 [Kelso 1946].

1820 Ibid., p. 390.

1821 [Marx 1970].

1822 Deze Amerikaanse term is afgeleid van het Shepard-systeem (zie o.m. [Spriggs en Hansford 2000]).

1823 [Tapper 1982].

1824 Waaronder diverse variabelen die ook door ons zijn onderzocht. *Vide infra*: hfd. 7.

1825 Onderzoek naar de netwerkeigenschappen van andere juridische bronnen is zeldzaam. Zie voor een onderzoek naar citaties van juridische literatuur [Shapiro 1985], [Shapiro 1991], en voor een (kleinschalig) onderzoek met betrekking tot wetgeving [Mazzega, Bourcier en Boulet 2009].

1826 [Post en Eisen 2000].

1827 [Balkin 1986].

1828 [Smith 2007].

1829 Onder meer door de grotere dataverzameling, die alle door LexisNexis opgenomen zaken omvatte.

1830 *Vide supra*: § 5.2.2.1.2.

1831 0,025% van het aantal uitspraken krijgt circa 80% van het aantal inkomende citaties; [Smith 2007, p. 346].

dat de meeste uitspraken niets bijdragen aan de rechtsontwikkeling, maar alleen een titel-verstreckende of notariële functie hebben.<sup>1832</sup>

Fowler c.s.<sup>1833</sup> onderzochten in 2007<sup>1834</sup> bijna 27.000 meervoudige-kamerbeslissingen van het *US Supreme Court*, met de vraag:

*(H)ow can we identify the most legally central cases at the U.S. Supreme Court at a given point in time in an empirically rigorous, reliable, and valid manner?*<sup>1835</sup>

Met behulp van het HITS-algoritme<sup>1836</sup> tonen ze aan dat hiermee betere resultaten kunnen worden bereikt dan met een standaard *degree*-telling.<sup>1837</sup> Vervolgens stellen ze vast dat met behulp van de *inward relevance* en *outward relevance* de kans op citatie door het *US Supreme Court* zelf en door lagere gerechten beter kan worden voorspeld dan met andere maten, zoals de publicatie in juridische standaardwerken<sup>1838</sup> of landelijke nieuwsbladen.<sup>1839</sup> De reden van citatie – bijvoorbeeld *overruling*, bevestiging of illustratie – wordt door Fowler c.s. niet meegewogen, en evenmin wordt rekening gehouden met het ‘gewicht’ van de citatie: het zou verschil kunnen uitmaken of de geciteerde uitspraak één of twintig keer wordt aangehaald door de citerende uitspraak.<sup>1840</sup> Alhoewel ze het algoritme in drie netwerken van verschillende omvang testten, gebruikten Fowler c.s. HITS zonder dat daaraan een specifieke query ten grondslag lag.

Omdat uitspraken betrekking kunnen hebben op meerdere juridische vraagstukken, kan het voor eindgebruikers relevant zijn om te weten of uitspraakcitaties betrekking hebben op een specifiek onderwerp. Om hiermee rekening te houden bouwden Zhang en Koppaka<sup>1841</sup> voort op het werk van Nie c.s.<sup>1842</sup> en ontwikkelden een model dat linkanalyse combineert met semantiek.

Ook voor andere doeleinden is citatieanalyse inmiddels ingezet. Hellyer<sup>1843</sup> gebruikte het om te onderzoeken of de digitale beschikbaarheid van grote hoeveelheden actuele uitspraken van invloed is op het citatiegedrag van rechters; een hypothese<sup>1844</sup> waarvoor hij geen bewijs vond. Whalen<sup>1845</sup> onderzocht of citatiefrequentie kan voorspellen of een uitspraak van het Amerikaanse *Supreme Court overruled* gaat worden.

1832 *Vide supra*: § 3.4.2.9.2.3. Vgl. [Smith 2007, p. 340].

1833 [Fowler, Johnson, Spriggs II et al. 2007].

1834 In 2008 aangevuld met een rechtshistorische analyse van de bevindingen: [Fowler en Jeon 2008].

1835 [Fowler, Johnson, Spriggs II et al. 2007, p. 325].

1836 *Vide supra*: noot 1800.

1837 [Fowler, Johnson, Spriggs II et al. 2007, par. 3].

1838 [Fowler en Jeon 2008, p. 22-24].

1839 [Fowler, Johnson, Spriggs II et al. 2007, p. 339].

1840 De keuze om deze ‘multipliciteit van de citatie’ (*vide infra*: § 7.2.3.1) buiten beschouwing te laten vloeide voort uit het gebruik van het Shepard-systeem (*vide supra*: noot 1822), waarin deze informatie niet aanwezig is.

1841 [Zhang en Koppaka 2007].

1842 [Nie, Davison en Qi 2006], zie ook *supra*: § 5.2.2.1.2.

1843 [Hellyer 2005].

1844 O.m. geuit in [Oskamp 2009].

1845 [Whalen 2013].

Naast exploratief en theoretisch onderzoek zijn ook praktische toepassingen ontwikkeld. Citatie-indexen ('*citators*') zijn vooral in *common-law*-jurisdicties onmisbare hulpmiddelen in de juridische praktijk. De meeste systemen zijn commercieel, zoals Shepard's<sup>1846</sup> (inmiddels overgenomen door LexisNexis) en het concurrerende KeyCite van WestLaw. Ook deze systemen hebben veel handmatige input, en kunnen dan ook aanzienlijk verschillen in de (relevantie van de) uitkomsten.<sup>1847</sup> Er zijn inmiddels ook *free-access*-initiatieven, zoals LawCite.<sup>1848</sup> In dit initiatief van AustLII wordt (onder meer)<sup>1849</sup> een importantie-indicator toegekend aan uitspraken op basis van citatieanalyse. Daarbij wordt rekening gehouden met citatie in andere uitspraken en (sommige) tijdschriften. Alhoewel er over het project is gepubliceerd,<sup>1850</sup> zijn de onderliggende algoritmen onbekend.<sup>1851 1852</sup>

Een van de weinige recente<sup>1853</sup> onderzoeken buiten *common-law*-jurisdicties is verricht door Geist<sup>1854</sup> die een analyse maakte van de citaties in de redactionele samenvattingen van uitspraken van het Oostenrijkse *Oberster Gerichtshof*, aanwezig in het publieke *Rechtssystem*.<sup>1855</sup> Hij vond een *Power law* distributie, waarbij 19% van de uitspraken geen enkele maal werd geciteerd, en 68% minder dan drie keer. Als gouden standaard gebruikte hij publicatie in een jurisprudentietijdschrift. Gebaseerd op een eenvoudige *in-degree* telling, trof hij inderdaad een verband aan tussen publicatie in het tijdschrift en citatie in latere uitspraken.<sup>1856</sup>

Vergelijkbaar onderzoek is gedaan door Malmgren,<sup>1857</sup> die het corpus van HvJ EU met *in-degree*, HITS-*authority* en *PageRank* analyseerde.<sup>1858</sup> Hij beperkte het onderzoek tot uitspraken over een kleine set artikelen van het VWEU en gebruikte als gouden standaard de uitspraken die in een standaardwerk over EU-recht het meest uitgebreid worden besproken.<sup>1859</sup> Malmgren concludeert dat de netwerkalgoritmen betere resultaten opleveren dan een standaard zoekalgoritme, maar dat geen van de drie algoritmen duidelijk beter presteert, al

1846 *Vide supra*: noot 1822.

1847 Zie voor een vergelijking van de twee genoemde *citators* [Mart 2010, p. 244-249].

1848 [Redactie LawCite 2010].

1849 Daarnaast worden ook 'parallele citaties' getoond, een functie vergelijkbaar met de Nederlandse LJN-index (*vide supra*: § 3.2.1.3).

1850 [Mowbray, Chung en Greenleaf 2009].

1851 LawCite verkeert nog steeds in de 'Alpha'-fase. D.d. 20-02-2013 zegt de site: "We are aiming to publish a fairly finalised 'beta' version of LawCite before the end of 2010, with a 'production' version in 2011" <[www.austlii.edu.au/LawCite/doc/overview.html](http://www.austlii.edu.au/LawCite/doc/overview.html)>.

1852 Er zijn meer initiatieven, nochtans zonder concrete resultaten. Zie bijvoorbeeld [Leiter 2011b].

1853 Van oudere datum is bijvoorbeeld het statistisch-exploratief onderzoek van [Wagner-Döbler 1994] naar frequentieverdelingen van citaties in Duitse jurisprudentie.

1854 [Geist 2009].

1855 *Vide supra*: § 2.6.2.1.1.

1856 Alhoewel niet uitgedrukt in een statistische maat of anderszins gespecificeerd.

1857 [Malmgren 2011].

1858 Hij gebruikte hiervoor de publieke data van EUR-Lex. In de metadata hiervan zijn de citaties, net als in het Shepard-systeem (*vide supra*, noot 1840) ongewogen opgenomen.

1859 [Malmgren 2011, par. 5.3.3].

doet *PageRank* het licht beter dan HITS.<sup>1860</sup> Een vergelijkbaar resultaat werd gevonden bij een kleinschalige netwerkanalyse van Nederlandse uitspraken.<sup>1861</sup>

Vermeldenswaard ten slotte zijn de voorzichtige stappen die worden gezet om netwerkanalyse in te zetten bij het onderzoeken van inhoudelijke juridische vraagstukken. Dibadj gebruikte netwerkanalyse om rechterlijk beleid met betrekking tot enkele civielrechtelijke leerstukken in het Amerikaanse recht in kaart te brengen,<sup>1862</sup> Siems zette netwerkanalyse in bij rechtsvergelijkend onderzoek naar de bescherming van crediteuren en aandeelhouders,<sup>1863</sup> en Gelter en Siems onderzochten kruisverbanden tussen de uitspraakcorpora van de hoogste gerechten van tien Europese landen.<sup>1864</sup>

### 5.3 Conclusies en vooruitblik

We zijn dit hoofdstuk begonnen vanuit het toegankelijkheidsaspect ‘hanteerbaarheid’, in het bijzonder vanuit het deelaspect ‘doorzoekbaarheid: hoe kan in steeds grotere jurisprudentiedatabanken informatie snel en doeltreffend worden gevonden. We hebben daarbij eerst de aandacht gericht op de input. Lange tijd is getracht om kritiek op de samenstelling van uitspraakcollecties voor te zijn door goede selectiecriteria te formuleren. Objectieve en bruikbare selectiecriteria blijken echter lastig te formuleren, en tijd en capaciteit om individuele uitspraken zorgvuldig aan deze criteria te toetsen ontbreekt bij publieke databanken. Het is echter vooral de algemene roep om substantieel meer gepubliceerde uitspraken die er voor zorgt dat het voor publicatie geselecteerd zijn steeds minder zegt over de juridische relevantie van een uitspraak. Daardoor wordt hanteerbaarheid een probleem dat aan de outputzijde van een databank moet worden opgelost, en daarmee wordt het een relevantievraagstuk uit de *information retrieval*: hoe kan een gebruiker ‘relevante’ informatie vinden.

Voor het begrip ‘relevantie’ hebben we vijf betekenissen (‘manifestaties’) onderkend: systeemrelevantie, onderwerpsrelevantie, cognitieve relevantie, situationele relevantie en domeinrelevantie. Al deze relevantiemanifestaties zijn van belang voor de toegankelijkheid van jurisprudentiedatabanken, en idealiter worden ze bij het ontwikkelen van praktische toepassingen in samenhang beschouwd.

Omwille van de noodzakelijke beperking richten wij ons in het vervolg van het onderzoek evenwel op slechts één enkele manifestatie: de domeinrelevantie. We zeggen daarmee niet dat de andere manifestaties minder belangrijk zijn, of dat jurisprudentiedatabanken goed doorzoekbaar zouden zijn als de domeinrelevantie maar inzichtelijk zou worden gemaakt; integendeel, een ideale zoekoplossing zal met alle relevantiemanifestaties rekening moeten houden.

---

1860 Ibid., par. 6.1.

1861 [Winkels, de Ruyter en Kroese 2011].

1862 [Dibadj 2008], [Dibadj 2009].

1863 [Siems 2010].

1864 [Gelter en Siems 2012].



Dat het meten en kwantificeren van domeinrelevantie lastig is, blijkt onder meer uit de problematiek die is gesignaleerd bij de selectie van uitspraken aan de inputzijde. Juridische relevantie is een relatief begrip, dat intuïtief wel wordt begrepen, maar vooralsnog moeilijk objectiveerbaar lijkt.

Onderzoek naar de manier waarop juridische relevantie kan worden gemeten is relatief schaars, en richt zich met name op citatieanalyse. Opvallend is dat de meeste onderzoeken zich beperken tot de hoogste instanties, zoals het *US Supreme Court*, het HvJ EU en het Oostenrijkse hooggerechtshof. Citatieanalyse kan worden verondersteld in zekere mate effectief te zijn voor dergelijke corpora, omdat rechtsontwikkeling en rechtseenheid voor deze gerechten een belangrijke bestaansreden vormen. Bij dergelijke uitspraken zullen de aantallen uitgaande citaties – die een uitspraak *well-founded in law* maken – en inkomende citaties – die een uitspraak *authoritative* maken – relatief hoog zijn. Gezien de meer op primaire geschilbeslechting gerichte functie van de lagere gerechten, is het aantal inkomende en uitgaande citaties bij uitspraken van lagere gerechten veel kleiner.<sup>1865</sup> Naarmate het aantal citaties afneemt, neemt waarschijnlijk ook de bruikbaarheid ervan als importantie-indicator af. Daarbij moet bedacht worden dat veel van het bovengenoemde onderzoek heeft plaatsgevonden in de Verenigde Staten, waar, dankzij het leerstuk van *stare decisis*,<sup>1866</sup> uitspraken die geen precedentes citeren relatief zeldzaam zijn.<sup>1867</sup>

Ook de invloed van andere variabelen zou daarom nadere analyse waard zijn, maar alleen<sup>1868</sup> Fowler<sup>1869</sup> onderzoekt naast citaties ook andere variabelen – die hij overigens alleen gebruikt om de gouden standaard vast te stellen. Een andere kanttekening die bij de tot nu gedane onderzoeken kan worden gemaakt, betreft de bruikbaarheid ervan in praktische toepassingen. Het besproken werk is vooral exploratief van aard, zonder dat duidelijk is hoe eindgebruikers hiervan zouden kunnen profiteren.

Ons onderzoek naar juridische relevantie bouwt voort op eerder onderzoek, maar breidt dit ook in aanzienlijke mate uit. We doen onderzoek op grote datasets, waarbij uitspraken van lagere gerechten in ruime mate vertegenwoordigd zijn. Omdat citaties daardoor waarschijnlijk minder belangrijk worden, zullen we ook op zoek gaan naar andere indicatoren voor juridisch belang. Ons doel is om uiteindelijk te komen tot een maat voor contextonafhankelijke juridische relevantie. Voordat we in hoofdstuk 7 daartoe het *Model for Automated Rating of Case law* ontwikkelen, bespreken we eerst in hoofdstuk 6 het verzamelen van de onderzoeksgegevens: de samenstelling van de datacollecties en de wijze waarop allerlei voor het model belangrijke gegevens aan rechterlijke en wetenschappelijke teksten zijn onttrokken. Daarbij gaan we tevens invulling geven aan het deelaspect ‘contextualiteit’ van het toegankelijkheidsaspect ‘hanteerbaarheid’.

1865 Zie voor de Verenigde Staten [Smith 2007] (*vide supra*: noot 1831), voor Nederland komt dit in § 7.2.3 aan de orde.

1866 *Vide infra*: § 7.2.5.

1867 [Tapper 1982, p. 137].

1868 In beperkte mate [Smith 2007], die ook kijkt naar citaties in rechtswetenschappelijke literatuur.

1869 [Fowler en Jeon 2008].