

Vrije Universiteit Brussel

From the Selected Works of Mireille Hildebrandt

2012

Oordeelsvorming door mens en machine: heuristieken, algoritmes en legitimatie

Mireille Hildebrandt



Available at: https://works.bepress.com/mireille_hildebrandt/38/

**OORDEELSVORMING DOOR MENS EN MACHINE:
HEURISTIEKEN, *ALGORITMES* EN LEGITIMATIE**

Mireille Hildebrandt

**This is an extended version of a paper presented at the
Symposium Juridische Argumentatie
24th June 2011**

A shorter version will be published in the Conference Proceedings (forthcoming)

Oordeelsvorming door mens en machine: heuristieken, *algoritmes* en legitimatie

Mireille Hildebrandt

Samenvatting

Gezien de ontwikkelingen op het terrein van ‘data mining’, ‘machine learning’, en neurale netwerken is het niet onwaarschijnlijk dat geautomatiseerde ‘rechtsvinding’ op middellange termijn een grote vlucht zal nemen. In deze bijdrage zal ik onderzoeken hoe heuristiek en legitimatie van het juridische oordeel zich verhouden tot juridische expertsystemen. In navolging van Moretti kunnen we het gebruik van dit soort expertsystemen kwalificeren als ‘lezen op afstand’. Om te voorkomen dat juridische argumentatie zich gedachteloos gaat baseren op de aannamen die in de gebruikte computertechnieken schuilen, zullen juristen moeten leren om niet alleen juridische teksten maar ook de bevindingen van expertsystemen te interpreteren. De kwaliteit van juridische argumentatie zal in de toekomst mede afhangen van de mate waarin wordt voorzien dat juridische teksten op verschillende wijze ‘op afstand gelezen’ kunnen worden, afhankelijk van de gebruikte algoritmes.

1	Rechtsinformatica	3
1.1	<i>Geautomatiseerde toepassing van juridische kennis</i>	3
1.2	<i>Het genereren van nieuwe juridische kennis</i>	4
2	Drie toepassingen van rechtsinformatica	6
2.1	<i>Kunnen computers rechtspreken?</i>	7
2.2	<i>Geautomatiseerde wetstoepassing</i>	8
2.3	<i>Ondersteuning van de rechterlijke macht</i>	10
3	De algoritmiek en de heuristiek van <i>het juridische oordeel</i>	11
4	Van ‘close reading’ naar ‘distant reading’ – en terug	12
4.1	<i>Lezen-op-afstand</i>	12
4.2	<i>Naar een hermeneutiek van machinaal leren</i>	13
4.2.1	Vooronderstellingen	13
4.2.2	Aanbevelingen	15
5	Afsluiting	16

1 Rechtsinformatica

Rechtsinformatica ziet op automatisering van de opslag en toegang tot juridische kennis en op de toepassing van kunstmatige intelligentie op rechtsvragen.¹ In alle gevallen gaat het om computationele technieken, dat wil zeggen technieken waarbij juridische kennis wordt omgezet in voor machines leesbare discrete data zodat manipulatie (computatie) van deze data mogelijk is ten einde kennis (1) terug te vinden, (2) toe te passen of (3) te genereren.

De meest eenvoudige computationele technieken betreffen de opslag en indexering van grote hoeveelheden juridische teksten (verdragen, wetgeving, jurisprudentie en rechtsgeleerde boeken en artikelen) en het doorzoeken van deze teksten met behulp van trefwoorden. Hierdoor kunnen enorme tekstbestanden worden ontsloten op een manier die voorheen onmogelijk was: bepaalde woorden, zinsdelen of wetsartikelen kunnen in zeer korte tijd worden teruggevonden zonder dat de onderzoeker een heel boek moet doorlezen om alle vindplaatsen te achterhalen. Deze computationele technieken (*knowledge management* en *information retrieval*) noemen we geen expertsystemen omdat hier uitsluitend sprake is van het terugvinden van kennis die als zodanig in het systeem is opgeslagen. Het gaat hier niet om het toepassen van kennis op nieuwe zaken en al helemaal niet om het genereren van nieuwe kennis.

Ik zal hieronder kort een aantal technieken bespreken die worden gebruikt om juridische kennis toe te passen (productieregels en case-based reasoning). Vervolgens zal ik nader ingaan op technieken die juridische kennis genereren (met name machinaal leren).

1.1 Geautomatiseerde toepassing van juridische kennis

De eerste juridische expertsystemen beperkten zich tot de opslag van productieregels (als-dan regels), gebaseerd op relevante regelgeving, waarbij geen aandacht werd besteed aan jurisprudentie.² De redeneerpatronen in dit type expertsystemen zijn deductief van aard. Inmiddels zijn productieregels ontwikkeld op basis van alternatieve logica's, met name fuzzy productieregels, deontische logica en systemen gebaseerd op niet-monotone logica.³ Het gebruik van fuzzy logica kan een aantal problemen oplossen die te maken hebben met de ambiguïteit die inherent is aan natuurlijke taal; deontologische logica staat toe om normatieve verbanden 'door te rekenen' en niet-monotone logica houdt rekening met het feit dat nieuwe informatie tot nadere kwalificatie van een hoofdregel kan leiden.

Een volgende generatie juridische expertsystemen legt zich toe op 'case-based reasoning' (CBR), een term die verwijst naar computertechnieken die beogen te bepalen of sprake is van vergelijkbare gevallen.⁴ Door een situatie te ontleden in een

¹ (Aikenhead, 1995; Leenes, 1998).

² Zie overigens (Prakken, 2005) die precedents integreert in argumentatieschema's.

³ (Hollatz, 1999). (L. Royakkers, 2000; L. L. Royakkers, 1998). (Sartor, 1993).

⁴ (Rissland, Ashley, & Branting, 2005; Weber, Martins, & Barcia, 1998). (Ashley & Brüninghaus, 2009)

aantal factoren kan het programma 'meten' in hoeverre een nieuw geval overeenkomt met eerdere gevallen en dezelfde regel toepassen op zaken die voldoende 'eender' zijn. Deze techniek lijkt zeer toepasselijk op juridische problemen omdat de regel wordt afgeleid uit beslissingen in voorafgaande zaken. Daarmee lijkt het een goede imitatie van wat in de Common Law precedentwerking wordt genoemd, waarbij in plaats van een deductieve redeneertrant een analogieredenering wordt gevolgd.⁵ Nieuwenhuis noemt casusvergelijking één van de belangrijkste heuristieken, dat wil zeggen manieren 'om argumentatiemateriaal te vinden'.⁶ Voor zover de factoren die gebruikt worden om de gelijkenis meetbaar te maken vooraf zijn gedefinieerd is sprake van een gesloten systeem, dat bijvoorbeeld niet in staat zal zijn om hard cases te herkennen;⁷ het systeem kan niet buiten de vooraf geformuleerde parameters denken.⁸ Daardoor kunnen dit soort systemen goede diensten bewijzen bij het mechanisch bewerken van grote hoeveelheden 'voorgesorteerde' tekst, maar zullen ze falen als instrument van rechtsontwikkeling; ze zijn per definitie niet creatief.⁹ Een voordeel is dat bij deze systemen volstrekt transparant is waar de analyse op is gebaseerd, waardoor ze een interessant reservoir aan argumenten kunnen leveren en op die manier bij kunnen dragen aan juridische oordeelsvorming.

Naast systemen die afleiden of voorspellen hoe nieuwe casus op grond van bestaande regels of op grond van rechterlijke uitspraken (moeten) worden beslist, worden ook systemen ontwikkeld die uitspraken samenvatten,¹⁰ en/of systemen die inzicht geven in de argumentatiestructuur van rechterlijke uitspraken,¹¹ alsmede combinaties van verschillende systemen. De tot nog toe besproken systemen lijken zowel voorspelbaar als narekenbaar omdat zij volledig bepaald worden door de algoritmes (stappenplannen) die door de programmeur in de software zijn vastgelegd.

1.2 Het genereren van nieuwe juridische kennis

De hiervoor beschreven vormen van geautomatiseerde toepassing zijn deterministisch van aard. Zij kunnen alleen tevoren geprogrammeerde stappen zetten, die zijn voorzien met het oog op een eindig aantal vooraf uitgewerkte mogelijke probleemsituaties. In geval van productieregels is de rechtsregel (de major) met het oog op mechanische toepassing herschreven en vervolgens op een minor toegepast, die daartoe in hapklare brokken (*bites*) in de database moet worden ingevoerd. Juridische argumentatie kenmerkt zich echter door wederzijdse afstemming van regel

⁵ Ook in Civil Law jurisdicties speelt precedentwerking in het kader van de rechtszekerheid een grote rol.

⁶ (Nieuwenhuis, 1976). p. 495.

⁷ Zie daarover (Leenes, 1998; Van der Linden-Smith, 2000).

⁸ Over het belang van parameters ook (Groothuis, 2004), die erop wijst dat gevallen die zich niet gemakkelijk zonder verlies van betekenis laten vertalen in parameters niet geschikt zijn voor automatisering. Parameters (geslacht, leeftijd, temperatuur) moeten immers met discrete waarden worden ingevuld.

⁹ (Bench-Capon & Prakken) werken aan uitbreiding van CBR analyses zodat zij hypothetische en waardegerelateerde redeneringen kunnen incorporeren. Net als bij reguliere CBR beperkt de creatieve inbreng zich hierbij tot die van de programmeurs.

¹⁰ (Grover, Hachey, & Korycinski, 2003), in combinatie met case-based reasoning. (Merkl & Schweighofer, 1997), met gebruik van neurale netwerken.

¹¹ (Prakken, 2005).

en geval op een wijze die niet volledig vooraf is bepaald. In die zin is in geval van productieregels eigenlijk geen sprake van juridische argumentatie. Dat gaat ook op voor CBR voor zover de parameters aan de hand waarvan zaken in het systeem worden ingevoerd van te voren vaststaan. Gedetermineerde expertsystemen kunnen uit de aard der zaak niet omgaan met omgevingen die vragen om het voortdurend bijstellen van parameters. Het zijn geen lerende systemen.

Een alternatieve manier om tot oplossing te komen op basis van de uitkomsten van eerdere casus is om kunstmatige neurale netwerken in te zetten. Ook deze werken met juridische casus in plaats van juridische regels. Een neurale netwerk is een computermodel dat opereert overeenkomstig ons zenuwstelsel (connectionistisch); het gaat daarbij vooral om patroonherkenning. Zodra verschillende casus dezelfde soort patronen vertonen worden ze als hetzelfde type zaken geclassificeerd. Deze classificatie is niet van tevoren gegeven, maar wordt aan de hand van voorbeelden 'geleerd' (gecontroleerd leren). Ook hier lijkt een analogieredenering aan de orde. Anders dan bij CBR is een neurale netwerk een zwarte doos in de zin dat zelfs de programmeur niet kan achterhalen op grond waarvan het netwerk bepaalde patronen heeft herkend. Daardoor is er geen sprake van transparantie en levert het netwerk in eerste instantie geen argumenten op waarmee de oplossing van een zaak beredeneerd kan worden. Een voordeel is dat neurale netwerken niet werken met vooraf ingegeven parameters, maar juist in staat zijn niet voorziene patronen te herkennen. Het gebrek aan creativiteit en vernieuwing dat eigen is aan CBR is hier dus niet het probleem, maar het neurale net kan geen redenen aanleveren voor de keuzes die gemaakt worden.

Neurale netwerken zijn een voorbeeld van computersystemen die op grond van voorbeelden en/of feedback in staat zijn om te leren.¹² Het betreft hier de sub-discipline van machinaal leren.¹³ Kort gezegd gaat het bij machinaal leren niet meer om de vraag hoe we computers het beste kunnen programmeren om een bepaalde taak uit te voeren, maar hoe we ze het beste kunnen programmeren zodat ze zichzelf gaan programmeren met het oog op die taak. Machinaal leren betekent dat computersystemen automatisch verbeteren op grond van hun ervaring: '[W]e say that a machine learns with respect to a particular task T, performance metric P, and type of experience E, if the system reliably improves its performance P at task T, following experience E.'¹⁴ Ik bespreek eerst gecontroleerd- en semi gecontroleerd leren en vervolgens ongecontroleerd leren.

In geval van gecontroleerd lerende expertsystemen (*supervised learning*) gaat het om computationele technieken waarbij het systeem wordt getraind aan de hand van vooraf aangedragen model-oplossingen, ook wel 'trainingsdata' genoemd. Bij semi gecontroleerd leren gaat het om technieken waarbij het systeem door middel van kritiek achteraf wordt bijgestuurd. In het eerste geval analyseert het systeem een set van goede oplossingen (de trainingsset), waarna het de kans krijgt om de verworven kennis toe te passen op nieuwe data. Die toepassing kan worden 'nagekeken' (met behulp van een testset), en het proces kan worden herhaald als blijkt dat er te veel foute of ongewenste oplossingen worden aangedragen. Bij semi gecontroleerd leren

¹² (Merkl & Schweighofer, 1997). Zie ook (Van der Linden-Smith, 2000), p. 72-3.

¹³ (Mitchell, 2006). Zie (Maat, Krabben, & Winkels, 2010) over hoe machine learning (ML) op dit moment presteert bij de classificatie van juridische teksten.

¹⁴ (Mitchell, 2006), p. 1.

krijgt het systeem geen modeloplossingen, maar worden achteraf ongewenste of foute oplossingen verworpen, waardoor het de oplossingsrichting kan bijstellen (*reinforcement learning*). De data analyse die door deze systemen wordt uitgevoerd kan gebeuren door bijvoorbeeld gebruik te maken van beslisbomen, regressie analyse, regel inductie of kunstmatige neurale netwerken. Een cruciaal probleem van neurale netwerken is dat zij functioneren als een zwarte doos: het is ook voor de programmeur niet zichtbaar op grond waarvan het systeem tot bepaalde oplossingen komt.

Mitchell merkt op dat je machinaal leren in beginsel zo kunt ontwerpen dat het systeem zelf ‘trainingsdata’ gaat zoeken en op die manier zichzelf permanent kan bijscholen.¹⁵ Dat betekent dat je een machinaal lerend systeem in beginsel als een kunstmatige agent op pad kunt sturen in een dynamisch databestand met rechterlijke uitspraken; terwijl het bestand dagelijks wordt aangevuld met nieuwe zaken, verzamelt de agent de nieuwe ‘oplossingen’ en optimaliseert aldus voortdurend de eigen algoritmes.¹⁶

In geval van ongecontroleerd lerende expertsystemen (*unsupervised learning*) gaat het om computationele technieken waarbij door middel van bijvoorbeeld clusteren nieuwe patronen worden gevonden die niet van tevoren zijn aangedragen en ook niet achteraf door een programmeur zijn bijgestuurd. Het gaat hier vooral om exploratief onderzoek van een gegevensbestand en om het ontdekken van nog onbekende correlaties. Deze vorm van leren lijkt in veel sterkere mate dan (semi) gecontroleerd leren tot werkelijk nieuwe en onverwachte oplossingen te kunnen leiden omdat het systeem niet afhankelijk is van vooraf bedachte correlaties. Het is goed mogelijk dat een analyse op grond van ongecontroleerd machinaal leren curieuze verbanden toont tussen het ontbijt van de rechter en het type uitspraken dat zij doet (legal realism come true?). Willekeur en inconsistentie kunnen zichtbaar worden, en allerlei correlaties tussen factoren enerzijds en uitspraken anderzijds die niets van doen hebben met de argumenten ter legitimatie van de uitspraak worden aangevoerd.

2 Drie toepassingen van rechtsinformatica

Qua toepassing kunnen drie verschillende domeinen worden onderscheiden bij het gebruik van expertsystemen. Ten eerste kennen we de *automatisering van wetstoepassing*, met name in een bestuursrechtelijke context, waar bijvoorbeeld grote hoeveelheden beschikkingen op geautomatiseerde wijze worden geproduceerd (Wet Studiefinanciering, Wet administratiefrechtelijke handhaving verkeersvoorschriften, Vreemdelingenwet).¹⁷ Ten tweede kan de rechtelijke macht zich laten ondersteunen door *juridische expertsystemen*, die bijvoorbeeld zorgen voor samenvattingen van juridische bronnen, argumentatieschema’s binnen juridische uitspraken, analyse van een databank met rechterlijke uitspraken, of vergelijkingen van straftoematingsbeslissingen. Ten derde zou de taak van de rechter – volgens

¹⁵ (Mitchell, 2006), p. 5.

¹⁶ De term agent verwijst hier naar een AI software programma dat autonoom bepaalde taken uitvoert binnen één of meer computersystemen. Voor een vergelijking met het juridische agent-begrip zie (Koops, Hildebrandt, & Jacquet-Chiffelle, 2010).

¹⁷ (Kordelaar, Teeseling, & Hoogland; Van der Linden-Smith, 2000).

sommigen – op den duur deels of geheel door *juridische beslissystemen* kunnen worden overgenomen. In deze tekst gaat het vooral om de vraag hoe de kennis die binnen deze drie toepassingsdomeinen wordt gebruikt zich verhoudt tot de heuristiek en legitimatie van het juridische oordeel.

In een rechtsstaat heeft de rechter een bijzondere taak en een specifieke discretie die samenhangt met het feit dat zij in zekere zin het laatste woord heeft bij de interpretatie van rechtsnormen. Degene die de norm vaststelt (de wetgever) is niet degene die gezaghebbend oordeelt over de juiste toepassing (de rechter).¹⁸ Wanneer ik hieronder over het *juridische* oordeel spreek gaat het mij om de anticipatie op het rechterlijk oordeel, geheel in lijn met het bekende adagium van Oliver W. Holmes: ‘The prophecies of what the courts will do in fact, and nothing more pretentious, are what I mean by the law’. Interessante vraag is of een computationele anticipatie op het rechterlijk oordeel gelijk mag worden gesteld met een *juridisch* oordeel.

2.1 Kunnen computers rechtspreken?

Het derde toepassingsdomein is geënt op een computationele theorie van de geest.¹⁹ Cognitieve wetenschappers die deze theorie aanhangen zullen waarschijnlijk hoge verwachtingen koesteren van rechtsprekende machines. De computationele theorie, die samenhangt met de representationele theorie van de geest, gaat ervan uit dat de menselijke geest opereert door middel van het stapsgewijs manipuleren van symbolen of andere representaties van objecten. In de meest extreme versie van deze theorie komt semantiek voort uit syntaxis en heeft geen eigenstandig belang (behalve wellicht als illusie). Het ligt voor de hand dat op basis van zulke aannames het ontwikkelen van goede oordelen afhangt van de volledigheid van de input en van de juistheid van de te volgen algoritmes. Wie het computationele model van het menselijke brein volgen zullen hoge verwachtingen hebben van algoritmische rechtstoepassing, en mogelijk zelfs dromen van betere dan wel perfecte rechtspraak: Dworkin’s Hercules is eindelijk in zicht!²⁰ Dit toepassingsdomein houdt direct verband met de vraag of artificiële intelligentie (AI) kan leiden tot agenten die net zo verstandig *handelen als* of zelfs nog verstandiger *beslissen dan* mensen.²¹

In Nederland stelde Van den Herik in zijn Leidse oratie van 1991 dat computers niet alleen in staat zijn om rechterlijke oordeelsvorming te ondersteunen maar op den duur in bepaalde gevallen de rechter ook zullen kunnen vervangen.²² Zijn positie lijkt

¹⁸ *Iudex non rex est lex loqui*, zie (Schoenfeld, 2008; Witteveen, 1991).

¹⁹ (Horst, 2009).

²⁰ (Dworkin, 1991, first printed 1986), waarmee ik niet wil suggereren dat Dworkin een aanhanger zou zijn van de computationele theorie van de geest. Zie overigens (Gigerenzer, 2008) over de mate waarin dagelijkse handelingen worden bepaald door logische deducties, statistische inferenties en heuristieken. Heuristieken zijn vuistregels die het mogelijk maken om snel te beslissen in geval van onzekerheid; zij functioneren dankzij selectieve waarneming en productief vergeten. Hun succes wordt gemeten aan de hand van de mate waarin zij een adequate ‘fit’ met de omgeving bevorderen, in plaats van de mate waarin zij overeenkomen met de wetten van logica of statistiek.

²¹ Dit roept de vraag op of deze rechtsprekende machines rechtspersoonlijkheid moet worden toegekend. Klassieke teksten over de vraag wanneer niet-mensen als rechtspersoon erkend kunnen worden zijn (French, 1979; Solum, 1992; Teubner, 2007). Zie ook (Hildebrandt, forthcoming).

²² (Herik, 1991). Maar zie ook, met wat meer relativisering (Franken & Van den Herik, 2002).

eerder pragmatisch dan metafysisch: ‘Na drie maanden feilloos advies is de computer de rechter, wat zij er ook beiden van vinden.’ Hij vervolgt: ‘De rechtsprekende computer is synoniem met rechtszekerheid en rechtsgelijkheid’. Het eerste lijkt mij het gevolg van een begrijpelijke cognitieve bias bij de ondersteunde rechter.²³ Geautomatiseerde rechtspraak die steeds weer met kwalitatief goede beslissingen komt die ook nog eens narekenbaar zijn in termen van argumentatie zal ertoe leiden dat de aandacht naar andere taken verschuift (een gezonde vorm van *selective perception bias*). Daarnaast zal het steeds moeilijker worden om nog te achterhalen wat de rechter zou hebben besloten als zij niet door een juridisch expertsysteem was geadviseerd, bijvoorbeeld omdat de manier waarop het systeem de zaken presenteert *default* wordt (een zorgelijk *framing effect*).²⁴ Hoewel we bij het gebruik van een wasautomaat niet treuren om het verlies van de vaardigheid van het wassen met de hand, denk ik dat dit bij rechtspraak anders ligt. Dat hangt samen met het feit dat de tweede stelling veel vragen oproept. Is er wel sprake van gelijke behandeling van gelijke gevallen als computers werken met vooraf ingegeven parameters? Gaat het dan niet om het mechanisch kwalificeren van bepaalde probleemsituaties als gelijke gevallen, die bij nader inzien ook heel anders gekwalificeerd kunnen worden? Kan de rechtzoekende daar nog achter komen of moet die zich neerleggen bij de manier waarop de computer de casus ‘leest’.²⁵ In hoeverre wordt de rechtszekerheid gediend als computers die werken met machinaal leren zelf ‘verzinnen’ welke gevallen als gelijk worden gezien zonder dat we kunnen achterhalen waarop die gelijksoortigheid is gebaseerd?

2.2 Geautomatiseerde wetstoepassing

Het eerste toepassingsdomein (geautomatiseerde wetstoepassing bij het nemen van beschikkingen) betreft beslissystemen die relatief eenvoudige administratiefrechtelijke regels toepassen.²⁶ Groothuis merkt op dat overheidsbesluiten die door ICT systemen genomen beslissingen aan een aantal eisen moeten voldoen: rechtvaardiging (de logica op grond waarvan de automatische beslissing werd genomen moet worden geëxpliciteerd); toegankelijkheid (het systeem moet toegankelijk zijn voor belanghebbenden; integriteit (het moet op basis van de juiste input tot de juiste beslissingen komen); transparantie (de manier waarop het systeem functioneert moet zichtbaar zijn); tegenspraak (degene ten aanzien van wie een geautomatiseerde beslissing wordt genomen heeft de mogelijkheid hiertegen in beroep te gaan).²⁷ Van der Linden-Smith merkt op dat mechanische wetstoepassing op

²³ Over de rol van ‘cognitive bias’ in de cognitieve wetenschappen, kunstmatige intelligentie, behavioral economics: (Gigerenzer, 1991).

²⁴ (Van der Linden-Smith, 2000), 18 wijst op drie mogelijke gevolgen van geautomatiseerde beschikkingen, die wellicht ook hier gelden: ‘immunisering van de bureaucratie, door juridisch “dichtgetimmerde” beschikkingen; zelfreferentiële juridische sfeer, beschikkingen die niet noodzakelijkerwijs recht doen aan de specifieke omstandigheden van het individuele geval; deprofessionalisering van de ambtenaar: mogelijk verminderde kennis van regels bij de ambtenaar en de vaardigheid om ze toe te passen, of om beschikkingen “met de hand” op te stellen’.

²⁵ Een mooie verwijzing naar de zorgelijke hilariteit van computerondersteund beslissen: http://www.youtube.com/watch?v=0ZAoMv_QnAU (Computer says no).

²⁶ (Citron, 2007; Groothuis, 2004; Leenes, 1998; Van der Linden-Smith, 2000).

²⁷ Zie ook art. 15 Richtlijn gegevensbescherming, die een recht toekent aan een ieder ‘om niet onderworpen te worden aan een besluit waaraan voor hem rechtsgevolgen zijn verbonden of dat hem in

gespannen voet staat met de discretionaire ruimte die nodig is om algemene regels toe te passen op individuele gevallen. Het gaat haar om door de wetgever toegekende beleids- en beoordelingsvrijheid, de interpretatievrijheid die inherent is aan onbepaalde of vage begrippen, conflicterende regels en discrepantie tussen juridische regels en maatschappelijke verwachtingen. Van der Linden-Smith is vooral geïnteresseerd in de manier waarop geautomatiseerde systemen van invloed zijn op de omgang met hard cases, die zij definieert als zaken waarbij de feiten, en/of de regel, en/of de match tussen beide onduidelijk zijn, of waarbij de uitkomst onaanvaardbaar is. Haar conclusie is dat geautomatiseerde wetstoepassing in theorie een aantal voordelen heeft, zoals 'snellere en beter gemotiveerde beschikkingen, verbetering van rechtsgelijkheid, en efficiency-winst voor de betrokken uitvoeringsorganisaties'.²⁸ In de uitvoeringspraktijk liggen de zaken echter anders, met name omdat geautomatiseerde systemen de hard cases niet als zodanig herkennen. Daardoor ontstaan efficiency verlies (als het systeem een duidelijk geval ten onrechte als hard case ziet en die doorstuurt voor handmatige afhandeling) of onjuiste belissingen (als het systeem een hard case verwerkt alsof het een duidelijk geval is). Om dit te vermijden zouden alleen regels moeten worden toegepast die bijvoorbeeld zo min mogelijk open en vage termen en zo veel mogelijk categorieën bevatten, en zouden de systemen zowel foutloos (sic!) en transparant moeten zijn en de mogelijkheid moeten bevatten tot afwijking. Deze conclusies komen in vergaande mate overeen met die van Citron, in haar spraakmakende artikel over *Technological Due Process*. Citron maakt hierin onderscheid tussen regels en standaarden, waarbij zij regels definieert als voorschriften (ex ante) en standaarden als instrument voor interpretatie en weging (ex post). Standaarden veronderstellen altijd discretie. Automatisering is volgens Citron gebaseerd op regels en dat betekent dat zaken die beter via standaarden kunnen worden afgedaan niet moeten worden geautomatiseerd. De vraag is in hoeverre haar betoog ook opgaat voor case-based-reasoning en machinaal leren nu beiden zich richten op het verbinden van voorafgaande casus met toekomstige gevallen. Waarschijnlijk worden deze methoden van 'rechtsvinding' weinig of niet bij geautomatiseerde beschikkingen gebruikt.²⁹ Citron bepleit herinrichting van procedures in de geautomatiseerde staat. *Due process* betekent hier allereerst dat de totstandkoming van de beslissing altijd narekenbaar moet zijn: de systemen moeten audit trails produceren die de feiten en regels opslaan die ten grondslag liggen aan de beslissingen; alle stappen die het systeem zet moet gedocumenteerd zijn en voorkomen moet worden dat de beslissingen als onfeilbaar of objectief worden voorgesteld. Ten tweede moet worden voorzien in hoor en wederhoor; beslissingen moeten aanvechtbaar zijn in een procedure waar de belanghebbende het eigen tegenverhaal kan vertellen. Ten derde moeten nieuwe procedures worden ontworpen voor democratische regelgeving: open source software, notice and comment, testprocedures van de software en participatie van het publiek aan de manier waarop interpretatie- en beleidsregels in de software worden geïmplementeerd.

aanmerkelijke mate treft en dat louter wordt genomen op grond van een geautomatiseerde gegevensverwerking die bestemd is om bepaalde aspecten van zijn persoonlijkheid, zoals beroepsprestatie, kredietwaardigheid, betrouwbaarheid, gedrag, enz. te evalueren'. Uitzonderingen zijn onder meer mogelijk indien wettelijk vastgesteld en voorzien van de nodige bescherming voor de belanghebbende.

²⁸ (Van der Linden-Smith, 2000), p. 75.

²⁹ Zie (Lodder, 2006) over het voorstel van Hamburg voor een computermodel voor het ondersteunen van euthanasiebeslissingen. Lodder meent dat rechtszekerheid beter gediend is met regelgebaseerde computermodellen.

2.3 Ondersteuning van de rechterlijke macht

Het tweede toepassingsdomein (geautomatiseerde kennissystemen ter ondersteuning van de rechtspraak) betreft expertsystemen die een rechtspraak analyse kunnen uitvoeren via case-based reasoning of op basis van machinaal leren. In geval van cased-based reasoning en beslisbomen geeft het systeem inzicht in de totstandkoming van de oplossing van een nieuwe casus; bij neurale netwerken en bij clustering is dat niet of minder duidelijk. Anders dan bij het derde toepassingdomein, waar het expertsysteem zelf als rechter zou functioneren, gaat het hier om de voorbereiding van rechterlijke beslissingen. Anders dan bij het tweede toepassingsdomein, waar het expertsysteem zich beperkt tot eenvoudige zaken en vaak semi-automatisch functioneert, gaat het hier om de productie van juridische kennis die dienstbaar is aan de uitoefening een taak die de grondslag vormt van de rechtsstaat.

De exponentieel toenemende hoeveelheid juridische teksten die binnen een jurisdictie – inmiddels vaak ook online – beschikbaar komen, vraagt om slimme zoekstrategieën ten einde door de bomen nog iets van een bos te ontwaren. Daar komt bij dat de internationalisering in veel gevallen een redelijke kennis van internationaal en transnationaal recht vereist, alsmede het vermogen om in voorkomende gevallen recht uit een andere jurisdictie toe te passen. De aloude tactiek van het systematiseren van het recht via codificaties (continentale rechtstelsels) en *stare decisis* (common law) schieten hopeloos te kort als het gaat om de aanwassende stroom wetten, uitspraken, doctrinaire handboeken en artikelen, verdragen en beginselen in kaart te brengen en te reconstrueren tot een non-contradictoir geheel van regels en beginselen. Hoewel tijdschriftredacties van gerenommeerde tijdschriften nog steeds een belangrijke rol spelen bij het vormen van een gezaghebbend oordeel is de hoeveelheid tijdschriften met een grote reputatie op internationaal niveau dusdanig toegenomen dat verschillende officiële rankings ontstaan om kwaliteit te oormerken.

Het is inmiddels volstrekt ondenkbaar dat enig jurist nog in staat is zelfs maar het eigen vakgebied in detail te overzien. Op enig moment zal het voor de rechter ondoenlijk zijn om door het zelf lezen van arresten, nieuwe regelgeving en wetenschappelijke tractaten op de hoogte te blijven van wat er omgaat binnen de eigen rechtsorde, inclusief alle raakvlakken en overlappingsen met andere jurisdicties. Expertsystemen zullen nodig zijn om kennis te genereren uit de brei van mogelijk relevante teksten. Die expertsystemen zullen waarschijnlijk handzame overzichten (samenvattingen) genereren, afgestemd op specifieke zoekvragen (bv: geef mij de stand van de jurisprudentie inzake rechtsverwerking bij sportwedstrijden, of inzake noodweer bij mishandelde vrouwen). Dan kan het gaan om classificaties (op basis van gecontroleerd machinaal leren), gevolgd door vergelijkingen tussen de relevante zaken (clustering), gevolgd door argumentatie-analyses (CBR). Interessanter is de mogelijkheid om op basis van gecontroleerd en ongecontroleerd machinaal leren oplossingen te genereren van nieuwe zaken en die als oplossingsrepertoire voor te leggen aan de rechter. Deze oplossingen en de eventueel bijgeleverde redeneringen (beslisbomen, regressie-analyse) kunnen door de rechter worden in gezet als heuristieken bij het bepalen van het rechterlijk oordeel.

3 De algoritmiek en de heuristiek van *het juridische oordeel*

Voor zover het gaat om geautomatiseerde wetstoepassing die volgens vooraf bepaalde stappen plaatsvindt lijkt mij dat evident geen sprake is van een juridisch oordeel. Dit hangt samen met het feit dat een juridisch oordeel steeds tegelijk een kwalificatie van de feiten en een nadere uitleg van het juridische normenstelsel is; beiden worden als het ware op elkaar afgestemd.³⁰ De buigzaamheid die hiervoor nodig is wordt bepaald door de ambiguïteit van de natuurlijke taal en de performativiteit van het menselijk oordeel. De onbuigzaamheid van de algoritmische ‘rechtsvinding’ maakt deze vorm van redeneren zeer geschikt als instrument om snel en tegen lage kosten grote hoeveelheden beslissingen te produceren. Voor zover het voor burgers duidelijk is hoe hun handelen door het systeem zal worden opgeslagen en verwerkt, kan het langs die weg zelfs de rechtszekerheid dienen. Maar omdat het noodgedwongen werkt met vooraf vastgestelde kwalificaties kan het geen rechtsgelijkheid bieden; het systeem kan niet zien wanneer het ongelijke gevallen onterecht over één kam scheert. Het systeem kan wel redeneren in de zeer technische zin van algoritmische stappenplannen, maar niet argumenteren in de zin dat het de eigen configuratie op grond van nieuwe informatie kan herzien.

Bij machinaal lerende systemen lijkt dit wel het geval te zijn. Kenmerk van deze systemen is nu juist dat ze in staat zijn hun eigen programma te herschrijven om met betere resultaten te komen. In tegenstelling tot productieregels en CBRs reageren machinaal lerende systemen op hun omgeving door patronen te herkennen in functie van het doel waarvoor ze ontworpen zijn (althans bij semi-gecontroleerd machinaal leren). Zij gaan dus niet alleen af op de manier waarop de programmeur naar de werkelijkheid kijkt, maar zijn in staat om zelf nieuwe kennis te genereren. Deze kennis is niet gebaseerd op inzicht in oorzaken of redenen, maar puur op correlaties tussen verschillende data. Voor zover deze correlaties voorspellende waarde hebben, lijken ze een nieuw type kennis te genereren die probabilistisch van aard is.³¹ In zekere zin zou je die kennis objectief kunnen noemen, nu het niet gaat om tevoren door de onderzoeker of gebruiker bedachte hypothesen maar om door het systeem ‘ontdekte’ patronen. In paragraaf 2.2.1 zal ik aangeven waarom ook deze ‘objectiviteit’ biased is, en in paragraaf 2.2.2 aangeven hoe juristen kunnen voorkomen dat zij deze *framing bias* over het hoofd zien.

Het lijkt erop dat machinaal lerende systemen – in tegenstelling tot gedetermineerde systemen – heuristieken ontwikkelen die veel weg hebben van het soort heuristiek dat eigen is aan menselijke cognitie.³² In plaats van perfecte en volledig doorgerkende antwoorden kiest het systeem indien mogelijk voor snelle, effectieve vuistregels die optimale oplossingen blijken te bieden. Grondslag voor die oplossing is niet een correcte redenering of een causaal verband maar de ervaring van het systeem. Het zou interessant zijn om te onderzoeken hoe deze vorm van machinaal leren zich verhoudt tot het leerstuk van de juridische *topoi* (vindplaatsen, heuristieken). Evident is dat voor zover computationele heuristieken verwantschap vertonen met menselijke cognitie, de gelijkenis betrekking heeft op onbewuste processen. Hoewel de vorming van het juridische oordeel voor een groot deel is ingebed in onbewuste –

³⁰ (Glastra van Loon, 1987/1956).

³¹ (Anderson, 2008; Ayres, 2007).

³² (Gigerenzer, 2007).

geautomatiseerde – processen, wordt de ontwikkeling van juridisch relevante heuristieken echter voortdurend bijgestuurd door de bewust gearticuleerde argumentatieve praktijk waarbinnen het rechterlijk oordeel tot stand komt. Het meest cruciale aspect van die praktijk is de mogelijkheid van tegenspraak en de daarmee samenhangende eis dat partijen ingaan op relevante tegenargumenten. Het meest on-juridische aspect van machinaal lerende systemen zou kunnen zijn dat het systeem buiten de orde van argument, tegenargument en rechtsoordeel tot een ‘juist’ oordeel kan komen, op basis van computationele heuristieken die ‘correct’ inschatten welk oordeel door een jurist geveld zou worden – zonder daar evenwel redenen voor te kunnen geven. Het probleem is dat een dergelijke aanspraak op correcte oordeelsvorming ten onrechte abstraheert van de bias die inherent is aan de lerende machine, zoals ik hieronder nader zal toelichten.

4 Van ‘close reading’ naar ‘distant reading’ – en terug

4.1 Lezen-op-afstand

Op dit punt aangekomen denk ik dat juristen kunnen leren van de digitale revolutie binnen de geesteswetenschappen.³³ Juist omdat het in beide gevallen gaat om het computationeel analyseren van grote tekstbestanden, is het interessant om de relevantie van de theorievorming hieromtrent voor de computationele ondersteuning van het juridisch oordeel te onderzoeken.

In hun ‘How not to read a million books’ stellen Clement et al. een strategie voor van niet-lezen ofwel lezen-op-afstand.³⁴ Daarmee doelen zij enerzijds op computertechnieken die teksten geautomatiseerd samenvatten, zodat de lezer de verkorte inhoud kan bekijken zonder de tekst zelf te lezen. Anderzijds richten zij zich op datamijn technieken (niet-lezen) als ‘provocatie tot lezen’. Hun MONK project (*metadata offer new knowledge*) biedt geesteswetenschappers een digitale omgeving waarin zij patronen kunnen ontdekken en analyseren in grote hoeveelheden literaire teksten.³⁵ Patronen kunnen worden gezocht in een enkele tekst (Gertrude Stein’s *The Making of Americans*), in een deelverzameling (Shakespeare) of in de hele verzameling (evolutionaire ontwikkelingen in de literatuur). Het belang van dit type projecten komt voort uit de simultane reductie en abstractie die mogelijk wordt door het lezen-op-afstand. Moretti schrijft daarover:

“Distant reading”, I have once called this approach; where distance is however not an obstacle, but a specific form of knowledge: fewer elements, hence a sharper sense of their overall interconnection. Shapes, relations, structures. Forms. Models.³⁶

Zijn doel is het ontwikkelen van inzicht in de evolutie van literaire genres en met instemming citeert hij Fernand Braudel:

³³ Zie bv het e-humanities programma van de KNAW
<http://www.knaw.nl/Pages/DEF/29/284.bGFuZz1FTkc.html>.

³⁴ (Clement, Steger, Unsworth, & Uszkalo, 2008).

³⁵ Zie <http://www.monkproject.org/>, er zijn totaal 2585 werken en 151.500 woorden opgeslagen.

³⁶ (Moretti, 2005), p. 1.

An incredible number of dice, always rolling, dominate and determine each individual existence: uncertainty, then, in the realm of individual history; but in that of collective history simplicity and consistency.³⁷

Hoewel Moretti zich lijkt te verliezen in de honger naar patronen op bovenindividueel niveau, die een ‘more rational literary history’ mogelijk zouden maken, vragen Clement et al. juist ook aandacht voor de manier waarop deze patronen nieuwe inzichten genereren in de individuele teksten. Niet-lezen is in feite een lezen-op-afstand dat nieuwe manieren van *close-reading* mogelijk maakt; enerzijds doordat niet gelezen teksten ineens interessant blijken, anderzijds omdat reeds gelezen teksten vanuit een ander perspectief nieuwe betekenis krijgen. Een kwantitatieve benadering van literaire teksten luidt dan niet het einde in van het lezen zelf, maar biedt de lezer een nieuw begin. Dit vraagt dan wel om een nieuwe hermeneutiek, die zich richt op het ‘lezen’ – in de zin van interpreteren - van de uitkomsten van machinaal leren.

4.2 Naar een hermeneutiek van machinaal leren

Voor deze nieuwe hermeneutiek moeten we niet bij Moretti zijn, nu hij lijkt te geloven dat kwantitatieve analyse tot objectieve uitkomsten leidt. Tegenover zijn aantrekkelijke maar wat naïeve enthousiasme voor computationele kennisproductie stellen Sculley en Pasanek een hier wel bijzonder relevant scepticisme. Zij spreken van de oppervlakkige objectiviteit van computationele methoden en beargumenteren dat in de digitale geesteswetenschappen nood is aan een rigoureuze normering van machinaal leren.³⁸ Zij onderscheiden vier veronderstellingen die deel uitmaken van de praktijk van machinaal leren. Voorzover aan deze veronderstellingen geen aandacht wordt besteed zal de kennis die wordt geproduceerd gemakkelijk incompleet, eenzijdig, onjuist of onzinnig zijn. Deze veronderstellingen zijn vervolgens aanleiding tot het formuleren van een vijftal aanbevelingen. Ik bespreek hier eerst kort de veronderstellingen en vervolgens de aanbevelingen en verwijs daarbij direct naar de relevantie voor het juridische domein. Voor zover mogelijk vermijd ik daarbij technisch jargon uit de discipline van machinaal leren, maar voor een goed begrip van de implicaties van het gebruik van dit type technologie is het noodzakelijk dat juristen zich enigszins verdiepen in de wijze waarop dit soort kennis tot stand komt.

4.2.1 Vooronderstellingen

Om te beginnen veronderstelt machinaal leren dat de dataset waarbinnen wordt gemijnd gekenmerkt wordt door *een constante verdeling van waarschijnlijkheden*. Deze veronderstelling is nodig om de idee van *generalisering te formaliseren*, die het mogelijk maakt om de opgedane kennis toe te passen op nieuwe teksten. Dit stelt allereerst eisen aan het samenstellen van de dataset, ten einde atypische steekproeven te vermijden. Zowel bij literaire teksten als bij juridische teksten komt daar bij dat

³⁷ Idem, p. 4, verwijzing naar Fernand Braudel, ‘L’histoire, mesure du monde’, in *Les écrits de Fernand Braudel*, vol. 11, Paris 1997.

³⁸ (Sculley & Pasanek, 2008).

juist de singulariteit, de ambiguïteit en de instabiliteit van de relevante teksten van het allergrootste belang zijn. Bij literaire teksten kan het dan nog gaan om ‘the pleasure of the text’ die schuilt in de ontregeling van voorspelbare patronen.³⁹ Bij juridische teksten gaat het er bovendien om dat recht wordt gedaan aan het individuele geval, in het licht van de regel waaronder het wordt gebracht. Een bepaalde oplossing kan een *outlier* zijn in termen van machinaal leren, maar bij nader inzien het beginpunt zijn van een nieuwe lijn in de jurisprudentie. Het performatieve karakter van de kwalificatie van de feiten maakt het lastig deze ‘sprong van Scholten’ in te bouwen in het proces van machinaal leren.

De tweede veronderstelling houdt in dat *de hypotheseruimte beperkt is*, dat wil zeggen dat de verzameling van mogelijke uitkomsten eindig is en dus moet worden onderworpen aan specifieke restricties. Deze eis hangt samen met het feit dat het programma *moet leren om gevonden patronen toe te passen op nieuwe gevallen*, in plaats van simpelweg te onthouden welke uitkomst bij welke input hoort. De programmeur zal dus vooraf moeten gissen of zij bijvoorbeeld op zoek is naar lineaire scheidbaarheid in een hoogdimensionale ruimte, dan wel er van uit kan gaan dat de datapunten het resultaat zijn van een duidelijk zichtbaar statistisch proces, ofwel dat de hypothese het beste gerepresenteerd kan worden als een verzameling van logische disjuncties. Deze keuzes bepalen of de gegevensanalyse bijvoorbeeld plaatsvindt aan de hand van regressie analyse, support vector machines, Bayesiaanse netwerken, verborgen Markov modellen of beslisbomen. Het is niet ondenkbaar dat verschillende technieken tot verschillende uitkomsten leiden, zonder dat op voorhand kan worden vastgesteld welke uitkomsten de juiste zijn. De vraag is sowieso of er een archimedisch punt is van waaruit – los van de gebruikte technieken – kan worden vastgesteld welke uitkomsten de juiste zijn. Bij semi-gecontroleerd leren is dat punt ‘gewoon’ de expert die aangeeft of de uitkomsten naar haar oordeel kloppen. Dat impliceert dat dit type technieken niet zozeer objectief zijn, maar vooral consistent; als een bepaalde techniek eenmaal gekozen is berekent de computer in beginsel de daarbij horende uitkomsten. Die zijn daardoor nog niet beter dan die van een andere techniek. In geval van ongecontroleerd machinaal leren is sowieso niet duidelijk in welk opzicht de uitkomsten objectief zouden zijn.

De derde veronderstelling houdt in *dat de gegevens goed zijn gerepresenteerd*. Nu gebeurtenissen of entiteiten moeten worden ingevoerd op een manier die door het betreffende computersysteem wordt ‘verstaan’, zal per definitie sprake zijn van een reductie of transformatie. Wil machinaal leren tot betrouwbare uitkomsten leiden dan *moeten de relevante kenmerken van de gebeurtenissen of entiteiten zijn opgenomen*. Bij tekst analyse kunnen we de tekst bijvoorbeeld opslaan als een ‘zak met woorden’, eventueel aangevuld met bepaalde relaties tussen woorden, relaties met woordvolgorde, of we kiezen andere eenheden (zinsdelen, hele zinnen of andere tekstblokken). Wie een gedicht wil onderzoeken zal manieren moeten vinden om lettergrepen, maat, alliteratie, eindrijm enzovoort en zo verder te representeren. Wat een goede representatie is hangt dus af van het doel waarvoor het databestand zal worden gebruikt, maar wat bijvoorbeeld de relevante kenmerken van een bepaald type casus in het strafrecht zijn is niet objectief vast te stellen. Ook hier is onvermijdelijke sprake van normatieve *bias*.

³⁹ (Sculley & Pasanek, 2008).

De vierde veronderstelling betreft het *no free lunch* theorema van Wolpert en Macready.⁴⁰ Het zou mooi zijn als we een meta-lerend systeem konden ontwerpen dat alle hypothese ruimte aftast en alle typen representatie van gegevens uitprobeert. Op die manier zouden we wellicht het optimale ontwerp kunnen localiseren. Wolpert en Macready toonden aan dat dit niet mogelijk is, omdat er geen algoritme is dat in alle gevallen de juiste uitkomst genereert. Zoals Sculley en Pasanek opmerken gaat het bij Wolpert en Macready om natuurwetenschappelijke experimenten, waar in ieder geval nog een extern perspectief lijkt te zijn dat kan bevestigen of de hypothese bevestigt dan wel gefalsificeerd wordt. Bij het juridische oordeel ligt dat anders, omdat het externe perspectief hier ligt bij de juridische expert die op basis van haar juridische kennis en ervaring toetst of het systeem de juiste uitkomsten genereert. Die kennis is echter fundamenteel betwistbaar; de rechtszekerheid van het moderne recht ligt immers niet (meer) in het natuurrecht (de moraal) maar in de precare verhouding van de rechtszekerheid (vertrouwensbeginsel) met de rechtvaardigheid (rechtsgelijkheid) en de doelmatigheid (doelgerichtheid en effectiviteit) van het recht. In dat opzicht wordt de rechtszekerheid bepaald door het complexe spel van rechtsstatelijke beslissingsbevoegdheden van de democratische wetgever en de controlerende rechter.

4.2.2 Aanbevelingen

Sculley en Pasanek geven niet alleen aan welke vooronderstellingen de uitkomsten van machinaal leren bepalen. Zij presenteren ook een vijftal aanbevelingen om de omgang met dit type kennisproductie te normeren. In zekere zin bieden zij de aanzet tot een hermeneutiek van de geautomatiseerde, algoritmische rechtsvinding. De vraag is immers hoe wij de resultaten van datamijnen gaan interpreteren. Wanneer we eenmaal zijn genezen van de idee dat hier sprake is van optimale of objectieve kennis, die geen theorie meer nodig heeft, is het zaak nader in kaart te brengen welke rol dit type kennis kan en mag spelen in de argumentatie ter onderbouwing van het juridisch oordeel.

De eerste aanbeveling is *het expliciteren van vooronderstellingen eigen aan de betrokken wetenschappelijke disciplines*. Het gaat niet om bovengenoemde vooronderstellingen maar betreft de impliciete uitgangspunten die experts uit verschillende domeinen meebrengen wanneer zij samen een expertsysteem ontwerpen. Computerwetenschappers en juristen spreken in allerlei opzichten een andere taal, hetgeen niet alleen tot wederzijds onbegrip en irritatie kan leiden maar ook kan resulteren in systemen die rechtsregels vertalen op een manier die niet strookt met de rol die het recht in een constitutionele democratie zou moeten spelen. Zo kan het reduceren van het juridische begrip privacy tot het computerwetenschappelijke begrip *data confidentiality* of tot een probabilistische conceptie van anonimiteit leiden tot een vloed aan technische oplossingen van een probleem dat niet zo veel te maken heeft met het vage begrip privacy, dat zich niet restloos om laat zetten in machinetaal.

De tweede aanbeveling is dat *teksten op verschillende manieren worden gerepresenteerd en dat ze met behulp van verschillende hypotheseruimtes worden*

⁴⁰ (Wolpert & Macready, 1997).

geanalyseerd. Daardoor ontstaat een pluraliteit aan mogelijke oplossingen die om te beginnen duidelijk maken dat er geen sprake kan zijn van objectief juiste antwoorden en daarnaast zichtbaar maken dat de uitkomsten pas betekenis krijgen als ze geïnterpreteerd worden. De grafieken, tabellen, diagrammen, boomstructuren, clusters en associaties die voortkomen uit machinaal leren spreken niet voor zichzelf. Ze moeten in het kader een voorliggend geval worden gezien in het licht van andere grafieken en tabellen, die dezelfde dataset beschrijven. Alleen op die manier kan duidelijk worden wat voor betekenis een bepaalde toepassing van machinaal leren heeft voor de casus die ‘aan de hand is’.

De derde aanbeveling is dat alle experimenten bij het ontwerp van machinaal leren worden gerapporteerd, ook wanneer zij geen interessante of relevante resultaten opleveren. In het kader van *evidence based medicine* is dit inmiddels een belangrijke normatieve standaard; het verzwijgen van experimenten die geen of negatieve effecten tonen kan immers leiden tot levensbedreigende situaties. Maar ook in geval van juridische expertsystemen kan het van belang zijn om transparant te zijn over mislukte experimenten, die bijvoorbeeld aanleiding kunnen zijn om algoritmes bij te stellen en alsnog tot interessante resultaten te komen.

De vierde aanbeveling is om *zowel alle gegevens als de gebruikte methoden beschikbaar te maken zodat de uitkomsten door onafhankelijke onderzoekers kunnen worden gefalsificeerd*. In het geval van juridische expertsystemen is dat zeker niet vanzelfsprekend omdat juridische kennis een belangrijk deel van het vermogen kan vormen van bijvoorbeeld commerciële advocatenkantoren. Het is echter van groot belang dat de kennis aanspraken die in de resultaten verpakt zitten in rechte kunnen worden aangevochten. Op rechtsstatelijke gronden ligt een pleidooi voor *open source* code voor de hand.

De vijfde aanbeveling is dat de gebruikte methodologie aan peer review wordt onderworpen. Juristen zijn geen experts in machinaal leren. Juist omdat de gebruikte methoden de bandbreedte van de juridische interpretatie in vergaande mate zullen bepalen is het van groot belang dat samenwerkingsverbanden tussen juristen en data mijners elkaar over en weer controleren, zodat degenen die expertsystemen gebruiken enige zekerheid hebben dat het om betrouwbare systemen gaat.

5 Afsluiting

Tot voor kort konden juristen het zich nog wel permitteren geen acht te slaan op juridische expertsystemen. Inmiddels is het doorzoeken van grote hoeveelheden juridische teksten aan de orde van de dag (met behulp van min of meer slimme zoekmachines), evenals geautomatiseerde wetstoepassing (met behulp van productieregels, CBRs of combinaties daarvan), terwijl de rechter en het OM zich in toenemende mate laten ‘leiden’ door informatiesystemen die de opslag en afstemming van relevante feiten, regels en dossiers voor hun rekening nemen (kennismanagement systemen). Hoewel juridische expertsystemen tot nog toe veelal gebaseerd zijn op het geautomatiseerd – algoritmisch – verwerken van eenvoudige gevallen, zou ik durven voorspellen dat de invloed van expertsystemen die creatiever omgaan met de input

niet lang op zich zal laten wachten. In deze bijdrage heb ik het gebruik van productieregels en CBRs gecontrasteerd met lerende expertsystemen, die op basis van zelf gegenereerde heuristieken in staat zijn te anticiperen op het rechterlijk oordeel. Op zoek naar een antwoord op de vraag hoe computationele heuristieken zich verhouden tot de argumentatieve praktijk die constitutief is voor het juridisch oordeel heb ik theorievorming bekeken die is ontwikkeld in reactie op de digitale revolutie binnen het geesteswetenschappelijk onderzoek. Daaruit concludeer ik dat het van groot belang is dat juristen een hermeneutiek ontwikkelen van de computationele anticipatie van het rechterlijk oordeel. Het werk van met name Sculley en Pasanek kan daartoe een eerste aanzet vormen.

Bibliografie

- Aikenhead, M. (1995). Legal Knowledge-based systems: some observations on the future. *Web Journal of Current Legal Issues*, <http://webjcli.ncl.ac.uk/articles2/aiken2.html>.
- Anderson, C. (2008). The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. *Wired Magazine*, 16(7).
- Ashley, K., & Brüninghaus, S. (2009). Automatically classifying case texts and predicting outcomes. *Artificial Intelligence and Law*, 17(2), 125-165.
- Ayres, I. (2007). *Super crunchers : why thinking-by-numbers is the new way to be smart*. New York: Bantam Books.
- Bench-Capon, T., & Prakken, H. Using argument schemes for hypothetical reasoning in law. *Artificial Intelligence and Law*, 18(2), 153-174.
- Clement, T., Steger, S., Unsworth, J., & Uszkalo, K. (2008). How Not to Read a Million Books. *Harvard University, Cambridge, MA*, from <http://www3.isrl.illinois.edu/~unsworth/hownot2read.html>
- Gigerenzer, G. (2007). *Gut feelings : the intelligence of the unconscious*. New York: Viking.
- Glastra van Loon, J. F. (1987/1956). *Norm en Handeling. Bijdrage tot een kentheoretische fundering van de sociale wetenschappen*. Groningen: Wolters-Noordhoff.
- Groothuis, M. M. (2004). *Beschikken en digitaliseren. Over normering van de elektronische overheid*. Den Haag: SDU.
- Grover, C., Hachey, B., & Korycinski, C. (2003). Summarising legal texts: sentential tense and argumentative roles, *Proceedings of the HLT-NAACL 03 on Text summarization workshop - Volume 5*: Association for Computational Linguistics.
- Hollatz, J. (1999). Analogy making in legal reasoning with neural networks and fuzzy logic. *Artificial Intelligence and Law*, 7(2), 289-301.
- Koops, B. J., Hildebrandt, M., & Jacquet-Chiffelle, D.-O. (2010). Bridging the Accountability Gap: Rights for New Entities in the Information Society? *Minnesota Journal of Law Science & Technology*, 11(2), 497-561.
- Leenes, R. (1998). *Hercules of Karneades. Hard cases in recht en rechtsinformatica*. Enschede: Twente University Press.
- Maat, E. d., Krabben, K., & Winkels, R. G. F. (2010). Machine Learning Versus Knowledge Based Classification of Legal Texts. In R. G. F. Winkels (Ed.),

- Legal Knowledge and Information Systems. JURIX 2010* (pp. 87-96). Amsterdam: IOS Press.
- Merkel, D., & Schweighofer, E. (1997). En Route to Data Mining in Legal Text Corpora: Clustering, Neural Computation, and International Treaties. In IEEE (Ed.), *Proceedings of the Eighth International Workshop on Database and Expert Systems Applications* (pp. 465-470).
- Mitchell, T. M. (2006). *The Discipline of Machine Learning*: Carnegie Mellon University, School of Computer Science, available at <http://www-cgi.cs.cmu.edu/~tom/pubs/MachineLearningTR.pdf>.
- Moretti, F. (2005). *Graphs, Maps, Trees. Abstract Models for a Literary History*. London: Verso.
- Nieuwenhuis, J. H. (1976). Legitimatie en heuristiek van het rechterlijke oordeel. *Rechtsgeleerd Magazijn Themis*, 494-515.
- Prakken, H. (2005). AI & Law, logic and argument schemes. *Argumentation*, 19, 303-320.
- Rissland, E. L., Ashley, K. D., & Branting, L. K. (2005). Case-based reasoning and law. *Knowl. Eng. Rev.*, 20(3), 293-298.
- Royakkers, L. (2000). Combining deontic and action logics for collective agency. In J. Breuker, R. Leenes & R. Winkels (Eds.), *Legal Knowledge and Information Systems. Jurix 2000: The Thirteenth Annual Conference* (pp. 135-146). Amsterdam: IOS Press.
- Royakkers, L. L. (1998). *Extending Deontic Logic for the Formalisation of Legal Rules*. Dordrecht: Springer.
- Sartor, G. (1993). A simple computational model for nonmonotonic and adversarial legal reasoning, *Proceedings of the 4th international conference on Artificial intelligence and law*. Amsterdam, The Netherlands: ACM.
- Schoenfeld, K. M. (2008). Rex, Lex et Judex: Montesquieu and la bouche de la loi revisited. *European Constitutional Law Review*, 4, 274-301.
- Sculley, D., & Pasanek, B. M. (2008). Meaning and Mining: The Impact of Implicit Assumptions in Data Mining for the Humanities. *Literary and Linguistic Computing*, 23(4), 409-424.
- Van der Linden-Smith, T. (2000). *Een duidelijk geval: geautomatiseerde afhandeling*.
- Weber, R., Martins, A., & Barcia, R. M. (1998). On Legal Texts and Cases. In *Textual Case-Based Reasoning: Papers from 1998 Workshop, Technical Report WS-98-12* (pp. 40-50): AAAI Press.
- Witteveen, W. J. (1991). *Evenwicht van machten*. Zwolle: W.E.J. Tjeenk Willink.
- Wolpert, D. H., & Macready, W. G. (1997). No Free Lunch Theorems for Optimization. *IEEE Transactions on Evolutionary Computation*, 1(1).