

# Van juridische naar computertaal

Radboud Winkels\*

De meeste juridische beslissingen zijn simpel en kunnen beter en efficiënter machinaal uitgevoerd worden; in de praktijk gebeurt dit ook steeds meer. Hiervoor moet juridische ‘kennis’ die in rechtsbronnen zit in computersystemen worden ondergebracht. Dat is een tijdrovend en foutgevoelig proces. Is ook dat te automatiseren?

## 1 Inleiding

Bij het Leibniz Center for Law houden wij ons bezig met het toepassen van methoden en technieken uit de informatica op het recht, zowel om de rechtspraktijk te ondersteunen als om beter te begrijpen hoe het recht in elkaar steekt en *of en hoe* machines juridisch kunnen redeneren. Gottfried Wilhelm Leibniz (1646-1716), de filosoof, wiskundige en jurist naar wie wij ons centrum hebben genoemd, droomde al van een ‘calculator’ om alle menselijke problemen, waaronder juridische, op te lossen:

[...] the plan I have had for a long time to reduce all human thinking to a calculation, such as we know it in algebra or the *ars combinatoria*... so that many arguments could be solved, the certain could be distinguished from the uncertain and even grades of probability could be measured. Then if two were arguing, they could say to each other “Let us calculate”.<sup>1</sup>

Drie eeuwen later voorspelde collega Van den Herik dat de computer de menselijke wereldkampioen schaken zou verslaan en kreeg daarin met 10 jaar verschil gelijk. Vervolgens voorspelde hij in 1991 dat de computer op termijn ook beter dan mensen zou kunnen rechtspreken.<sup>2</sup> Is dat zo en zo ja, wat is daarvoor nodig en hoe ver zijn we nu?

## 2 Stand van zaken

Bij ‘het recht’ denken de meeste mensen aan rechtszaken, maar dat is slechts het topje van de ijsberg. De meeste rechtshandelingen die

mensen elke dag verrichten zijn relatief eenvoudig en standaard en leiden totaal niet tot geschillen. Daar waar dat wel gebeurt, wordt het meestal onderling opgelost. Slechts een klein aantal van de geschillen komt voor de rechter of een andere beroepsinstantie en ook dan wordt er nog vaak onderling geschikt.

**De meeste rechtshandelingen die mensen elke dag verrichten zijn relatief eenvoudig en standaard en leiden totaal niet tot geschillen**

De meeste van deze (eenvoudige) beschikkingen en beslissingen zouden heel goed door computers kunnen worden afgehandeld en dat gebeurt ook op steeds grotere schaal, met name in de uitvoering door de overheid. Zo worden vrijwel alle particuliere belastingaangiftes automatisch afgehandeld, net als beperkte snelheidsovertredingen en sinds kort vele verblijfsaanvragen bij de IND.

Software kan eenvoudige juridische beslissingen *sneller* en *consistenter* nemen dan mensen. Doorlooptijden van zaken kunnen dramatisch verkort worden (van maanden naar dagen, van dagen naar minuten). Een andere manier om deze systemen in te zetten, is ze aan burgers en bedrijven beschikbaar te

\* Dr. R.G.F. Winkels is als universitair hoofddocent werkzaam bij het Leibniz Center for Law van de Faculteit Rechtsgeleerdheid van de Universiteit van Amsterdam. Veel van het hier te beschrijven onderzoek heb ik gedaan als dagelijks begeleider van E. de Maat die bij ons promoveerde op het proefschrift *Making Sense of Legal Texts* (diss. Amsterdam UvA), Ridderkerk: Ridderprint 2012.

1 Brief aan Jacob Spener, juli 1687, geciteerd door R. Meyer, *Leibniz und die europäische Ordnungskrise*, Hamburg: Hansischer Gildenverlag 1948, p. 233.

2 H.J. van den Herik, *Kunnen computers recht spreken?* (oratie Leiden), Arnhem: Gouda Quint 1991.

stellen om hun kansen beter in te schatten bij bijvoorbeeld de aanvraag van een vergunning of het aangaan van een rechtszaak.

## Software kan eenvoudige juridische beslissingen sneller en consistenten nemen dan mensen

### 2.1 Mogelijke bezwaren

Deze ontwikkelingen, die zoals gezegd al in gang zijn gezet, kunnen een aantal bezwaren oproepen. Ten eerste is er de ervaring dat mensen nogal snel geneigd zijn automatische systemen te vertrouwen. Zijn ze in het begin misschien een beetje argwanend, als het systeem een paar keer goed gepresteerd heeft, wordt men steeds makkelijker en neemt men beslissingen klakkeloos over.<sup>3</sup> Andere mensen werpen op dat het principieel ongewenst is dat machines morele, ethische of juridische beslissingen nemen. Machines hebben immers geen moraal en geen empathisch vermogen.

Tegen beide punten kan worden ingebracht dat (a) mensen *ook*, en wel *meer* fouten maken en (b) bezwaar en beroep tegen (machinale) beslissingen altijd mogelijk is en men uiteindelijk bij een menselijke beslisser uitkomt. Aan de andere kant kan men zeggen dat machines (in principe) voorspelbaar zijn en feitelijk gelijke gevallen ook gelijk zullen behandelen. Maar wanneer zijn twee gevallen gelijk? Men kan het ene geval *altijd* onderscheiden van het andere als men wil; het is de vraag of het onderscheid van belang is voor een juiste of rechtvaardige beslissing. Ten slotte is het recht natuurlijk nooit 'af'; de wereld verandert en ook onze ideeën over wat goed of rechtvaardig is veranderen.<sup>4</sup>

Met andere woorden, automatisch beslissen dient sommige rechtsprincipes beter (rechtszekerheid, rechtsgelijkheid), andere moeilijker of slecht (recht doen aan individuele geval, subsidiariteitsbeginsel).

Automatisch beslissen dient sommige rechtsprincipes beter (rechtszekerheid, rechtsgelijkheid), andere moeilijker of slecht (recht doen aan individuele geval, subsidiariteitsbeginsel)

### 3 Wat is er nodig?

Voor juridische toepassingen moeten de informele rechtsbronnen vertaald worden naar een formele representatie waar machines iets mee kunnen. Wetten en regelingen zijn *door* mensen geschreven *voor* mensen en bevatten vele onduidelijkheden, impliciete aannames en inconsistenties. Voor een deel is dit bewust en noodzakelijk. De wetgever kan en wil niet alle situaties van tevoren voorzien en door bepaalde zaken impliciet of 'vaag' te houden kunnen uitvoeringsinstanties of rechters er in de praktijk meer mee. Voor een ander deel zijn fouten of ambiguïteiten in regelingen onbedoeld en te voorkomen; ik kom daar nog op terug.

Wetten en regelingen zijn door mensen geschreven voor mensen en bevatten vele onduidelijkheden, impliciete aannames en inconsistenties. Voor een deel is dit bewust en noodzakelijk

Computers kunnen slecht met vaagheden en inconsistenties overweg, zeker als ze ook nog moeten uitleggen en rechtvaardigen waarom ze een bepaalde beslissing hebben genomen. Gelukkig moeten uitvoeringsinstanties dat ook; in de praktijk worden vage begrippen of omstandigheden dan ook veelal weer nader uitgewerkt in jurisprudentie en uitvoeringsbesluiten. Daar kunnen we voor onze intelligente programma's goed gebruik van maken.

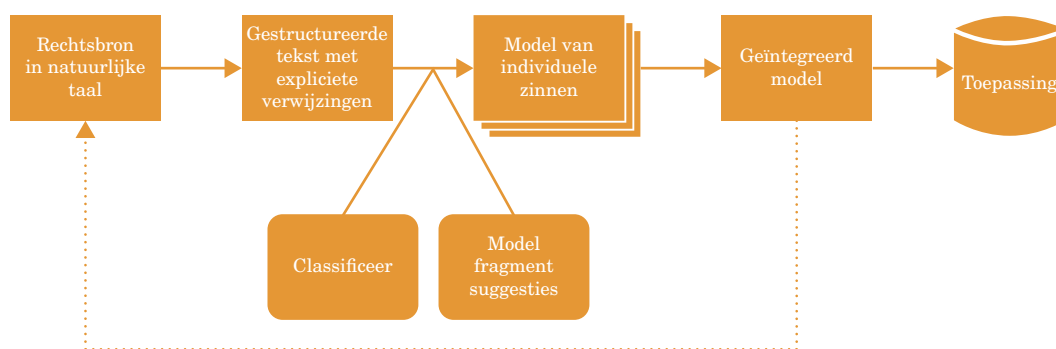
Het vertalen van de informele en incomplete rechtsbronnen naar uitvoerbare instructies voor de praktijk is een tijdrovende, foutgevoelige en kostbare zaak; ook voor uitvoeringsinstanties. De kans dat twee mensen voor dezelfde tekst tot een ander model komen is aanzienlijk. Ons onderzoek is voor een groot deel gewijd aan hoe we dat proces kunnen versnellen, verbeteren en automatiseren.

### 4 Verdeel en heers op zoek naar patronen

Het proces van informele juridische taal in rechtsbronnen naar formele computertaal is ingewikkeld: één van de standaardmethoden uit de informatica is het opdelen van het proces in kleinere stappen. Een andere is het zoeken naar terugkerende patronen. Beide passen wij toe in ons onderzoek.

<sup>3</sup> Zie bijvoorbeeld: M. Groothuis, *Beschikken en digitaliseren*, Den Haag: Sdu Uitgevers 2005.

<sup>4</sup> En worden idealiter steeds beter.



Figuur 1: Het proces van rechtsbronnen naar formeel model en een computertoepassing

5 Het formaat op *wetten.nl* is hier nog niet geschikt voor. Afzonderlijke zinnen zijn bijvoorbeeld niet aangegeven. Wij zetten de wetten daarom om naar CEN MetaLex-formaat en houden een schaduwadministratie bij. Zie: R. Hoekstra, 'The MetaLex Document Server – Legal Documents as Versioned Linked Data', in L. Aroyo e.a. (red.), *Proceedings of the International Semantic Web Conference*, Berlijn: Springer 2011, p. 128-143.

6 Dit lukt met 95%-99% correctheid en vrijwel geen onterecht gevonden verwijzingen voor wetgeving. Zie: E. de Maat, R. Winkels & T. van Engers, 'Automated detection of reference structures in law', in: T. van Engers (red.), *JURIX 2006, volume 152 of Frontiers in Artificial Intelligence and Applications*, Amsterdam: IOS Press 2006, p. 41-50. Voor jurisprudentie ligt het nog wat lager op 85%-90%; zie R. Winkels e.a., 'Towards a Legal Recommender System', in: R. Hoekstra (red.), *JURIX 2014, Volume 271 of Frontiers in Artificial Intelligence and Applications*, Amsterdam: IOS Press 2014, p. 169-178.

7 R. Winkels, J. de Ruyter & H. Kroese, 'Determining Authority of Dutch Case Law', in: K.M. Atkinson (red.), *JURIX 2011, Volume 235 of Frontiers in Artificial Intelligence and Applications*, Amsterdam: IOS Press 2011, p. 103-112; en M. van Opijnen, *Op en in het web – hoe de toegankelijkheid van rechterlijke uitspraken kan worden verbeterd* (diss. Amsterdam UvA), 2014.

8 R. Winkels e.a. 2014, p. 169-178.

Om te beginnen richten wij ons op de officiële, geschreven bronnen: wet- en regelgeving en jurisprudentie. Bij twijfel over interpretatie raadplegen wij experts. Dan volgt een proces als weergegeven in Figuur 1. Soms leidt het formaliseringsproces zelfs tot aanpassing van de wet of regeling (de gestippelde pijl in Figuur 1).

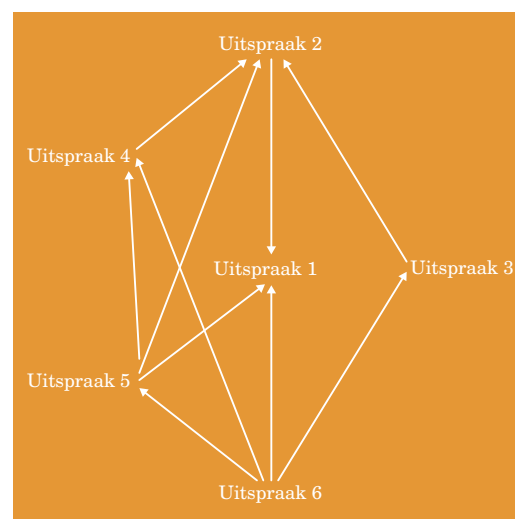
Als de tekst nog niet gestructureerd is in hoofdstukken, paragrafen, zinnen, enzovoorts dan doen we dat automatisch. Vervolgens maken we alle verwijzingen in de tekst expliciet en voor machines begrijpelijk. Het gaat hierbij om zowel *interne* (binnen het document zelf, bijvoorbeeld een verwijzing in de Wet Inkomstenbelasting naar een artikel verder in die wet) als *externe* verwijzingen (een verwijzing naar een ander document, bijvoorbeeld een andere wet). Bovendien kunnen we *directe* verwijzingen ('artikel 31 van deze wet') onderscheiden van *indirecte* ('het vorige artikel'). Wil de verwijzing voor de machine te volgen zijn, dan moet ondubbelzinnig kunnen worden vastgesteld en aangegeven *waar* de verwijzing naar is, dus bij Nederlandse wetgeving naar het nummer van de wet in het Basiswettenbestand en daarbinnen het precieze artikel en eventuele lid of zinsdeel.<sup>5</sup>

Dit proces kan automatisch gedaan worden doordat men over het algemeen met vaste patronen naar wetgeving of uitspraken verwijst en alle namen van wetten en regelingen bekend zijn.<sup>6</sup>

Het resultaat van deze stap is een complex netwerk van rechtsbronnen dat op zich al nuttige toepassingen biedt. Zo kun je bijvoorbeeld op basis van een analyse van het netwerk van verwijzingen in jurisprudentie voorspellingen doen over het belang van uitspraken (zie Figuur 2 voor een voorbeeld). Een simpele variant is dat een uitspraak waarnaar veel verwezen wordt gezaghebbender is dan een waar niet naar verwezen wordt. Een stap verder is dat een uitspraak waarnaar verwe-

zen wordt door een gezaghebbende uitspraak weer interessanter is dan een waarnaar verwezen wordt door een minder gezaghebbende uitspraak (net als het PageRank-algoritme van Google). Er zijn meer varianten denkbaar. Zeker binnen de Angelsaksische traditie is dit bruikbaar, maar ook binnen Nederland is aangetoond dat dit werkt.<sup>7</sup> Op eenzelfde wijze kunnen we gebruikers van bijvoorbeeld de Nederlandse portal voor wetgeving *wetten.nl* die een bepaald wetsartikel aanklikken, suggesties doen voor andere relevante rechtsbronnen.<sup>8</sup>

De volgende stap is het trachten te begrijpen van de inhoud van de rechtsbronnen. We beginnen met individuele zinnen in wet- en regelgeving. We proberen eerst te achterhalen met welk type zin we te maken hebben. Wat is het doel van de zin? Geeft deze bijvoorbeeld een definitie van een begrip of een beschrijving van een verplichting, of verandert deze iets in een andere wet? In totaal onderscheiden we vijftien types in Nederlandse wets teksten, elk weer herkenbaar aan stereotiep woordgebruik, grotendeels terug te voeren



Figuur 2: Voorbeeld van een netwerk van uitspraken

Type	Voorbeelden zinspatronen	%
Definitie	... onder ... wordt (mede) verstaan ...	3
Fictie	... wordt geacht ...	1
Delegatie	... kan regels stellen ...	3
Norm – Recht	... kan/kunnen ... ; ... is bevoegd ...	11
Norm – Stellend	... (geen patroon)	36
Van toepassingsverklaring	... is/zijn (niet) van toepassing ...	8
Straf	... wordt gestraft met ...	1
Waardetoekenning	... is/zijn ... vermeerderd/verminderd met ...; ... bedraagt/bedragen ...	0
Wijziging – 5 types	... komt te vervallen... ; ... wordt gewijzigd in ...; ... wordt ingevoegd ...	33
Inwerkingstreding	Deze wet treedt in werking ...	3
Citeertitel	Deze wet wordt aangehaald als ...	1
	<b>Totaal</b>	<b>100</b>

Tabel 1: Type zinnen in Nederlandse wetgeving met herkenningpatronen en percentage voorkomen

op de *Aanwijzingen voor de regelgeving* en wetgevingstraditie.<sup>9</sup> Tabel 1 geeft een samenvatting.

Definities kunnen dus automatisch herkend worden aan het gebruik van woorden als ‘wordt verstaan onder’, bijvoorbeeld in de Tabakswet, artikel 1, aanhef en onderdeel a:

In deze wet en de daarop berustende bepalingen wordt verstaan onder:

- a. tabaksproducten: producten die voor roken, snuiven, zuigen of pruimen bestemd zijn en die [...] uit tabak bestaan;

Een fictie is te herkennen aan het gebruik van de term: ‘wordt geacht’, bijvoorbeeld in de Wet Inkomstenbelasting 2001, artikel 2.2 tweede lid:

Een Nederlander die in dienstbetrekking staat tot de Staat der Nederlanden, wordt steeds geacht in Nederland te wonen indien hij is uitgezonden [...]

Normen zijn lastiger te herkennen, vooral omdat de *Aanwijzingen voor de regelgeving* het gebruik van de normatieve termen ‘moeten’ en ‘dienen’ afraden.<sup>10</sup> De meeste normen zijn dan ook ‘stellend’ geformuleerd, als presentatie van een feit. Bijvoorbeeld artikel 46, eerste lid van de Wet op de lijkbezorging:

Op een gesloten begraafplaats worden geen lijken begraven.

Aangezien alle andere types wel duidelijk herkenbaar zijn, ondervangen we dit probleem door als laatste alle zinnen die niet geassocieerd zijn toe te wijzen aan de klasse ‘Norm – Stellend’.

Uit enkele experimenten blijkt dat ook dit heel goed automatisch te doen is. Ongeveer 95% van alle zinnen en 90% van opsommingen wordt correct geassocieerd.<sup>11</sup> Normen vormen het grootste deel (bijna de helft) van de zinnen en daarna nog eens een derde van de wijzigingen. Waardetoekenningen waren in

deze experimenten te verwaarlozen, maar komen relatief vaker voor in belastingwetgeving en sociale zekerheid.

## 5 Het genereren van modelfragmenten

Voor sommige type zinnen, zoals wijziging, inwerkingstreding en citeertitel, is de classificatie al bijna genoeg voor de computer om te begrijpen wat het betekent. Interpretatie van de wijzigingen kan gebruikt worden om wetgeving automatisch te consolideren.<sup>12</sup> Dat wil zeggen dat we machinaal van de huidige versie van een wet de nieuwe kunnen maken op basis van interpretatie van de wijzigingswet.

Voor de andere zinnen, die echt over het te regelen domein gaan, is verdere analyse noodzakelijk. Redekundige (en in mindere mate taalkundige) ontleding van de zinnen brengt uitkomst. Hiervoor maken we gebruik van de Nederlandse Alpino-parser.<sup>13</sup> We zien de normatieve zinnen als beschrijvingen van een ongewenste of gewenste situatie. Het hoofdwerkwoord van de zin is de actie die al dan niet geoorloofd is en de andere elementen kwalificeren en specificeren de actie. We maken onderscheid tussen de *agent*, de *patient* en de *recipient* van de actie. In een actieve zin gaan we ervan uit dat het onderwerp van de zin de *agent* is, het lijdend voorwerp de *patient* en het meewerkend voorwerp de *recipient*. Bijvoorbeeld in de zin: ‘Onze Minister vaardigt een dwangbevel uit aan de nalatige’ is ‘uitvaardigen’ het hoofdwerkwoord en dus de actie. De *agent* is ‘Onze Minister’, de *patient* het ‘dwangbevel’ en de *recipient*

Recht/Verplichting	
<i>Actie</i>	Uitvaardigen
<i>Agent</i>	Onze Minister
<i>Patient</i>	Dwangbevel
<i>Recipient</i>	Nalatige

de ‘nalatige’. De hele zin drukt een recht of verplichting uit en is weer stellend geformuleerd (zie hiernaast).

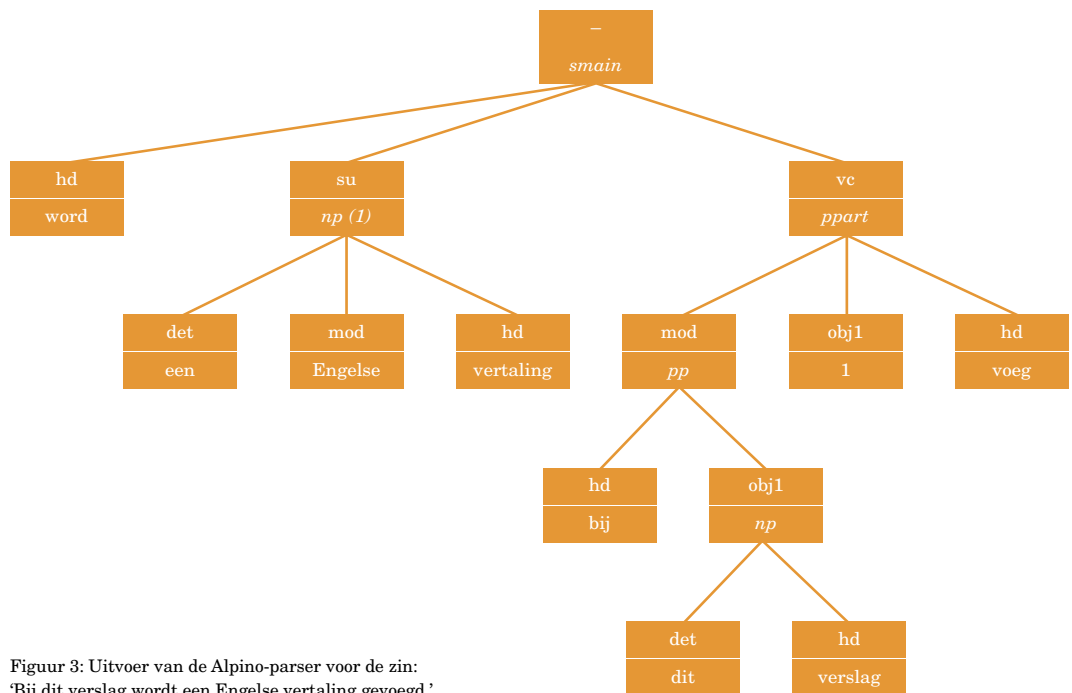
<sup>9</sup> *Aanwijzingen voor de regelgeving*. Circulaire van de Minister-President. Origineel in *Stcrt.* 1992, 230. Zie: <http://wetten.overheid.nl/BWBR0005730>.

<sup>10</sup> Aanwijzing, p. 53: ‘Voorschriften worden, tenzij dit onvermijdelijk is, niet met behulp van de werkwoorden “moeten” of “dienen” geformuleerd.’

<sup>11</sup> Zie: E. de Maat & R. Winkels, ‘Automatic Classification of Sentences in Dutch Laws’, in: E. Francesconi, G. Sartor & D. Tiscornia (red.), *Jurix 2008, volume 189 of Frontiers in Artificial Intelligence and Applications*, Amsterdam IOS Press 2008, p. 207-216.

<sup>12</sup> Zie bijvoorbeeld E. de Maat e.a., ‘Automated Handling of Amending Documents and Resulting Consolidations’, in: G. Governatori (red.), *Jurix 2009, volume 205 of Frontiers of Artificial Intelligence*, Amsterdam: IOS Press 2009, p. 116-125.

<sup>13</sup> G. Bouma, G. van Noord & R. Malouf, ‘Alpino: Wide Coverage Computational Analysis of Dutch’, in: W. Daelemans e.a. (red.), *Computational Linguistics in the Netherlands CLIN 2000*, Amsterdam: Rodopi 2001, p. 45-59.



Figuur 3: Uitvoer van de Alpino-parser voor de zin: 'Bij dit verslag wordt een Engelse vertaling gevoegd.'

Bij passieve zinnen, waar rechtsbronnen ook vol mee zitten, werkt dit niet direct zo; daar is immers het lijdend voorwerp het onderwerp geworden. Neem de zin: 'Bij dit verslag wordt een Engelse vertaling gevoegd'.<sup>14</sup> In de uitvoer van de Alpino-parser hierboven (Figuur 3) is te zien dat het werkwoordelijk deel van de zin (*verb clause* of 'vc' in de figuur) de zin in actieve vorm bevat, waarbij het onderwerp ('su') nu lijdend voorwerp ('obj1') is geworden.<sup>15</sup> Door alleen het werkwoordelijk deel te gebruiken krijgen we het correcte model zonder hulpwerkwoord. In dit geval is het onderwerp (*agent*) impliciet en leeg.

Een ander probleem is de 'negatie'. Neem de eerder genoemde zin: 'Op een gesloten begraafplaats worden geen lijken begraven'. Dit is een verplichting en de *patient* van de zin is 'geen lijken'. Dat zou tot de volgende representatie leiden:

Verplichting	
Actie	Begraven
Patient	geen lijken
Voorwaarde	op een gesloten begraafplaats

14 Art. 33, derde lid, tweede volzin van de Wet bescherming Antarctica.

15 'hd' is het werkwoordelijk hoofddeel van de zin; 'su' (*subject*) is het onderwerp; 'vc' het verbaal complement; 'det' (*determiner*) een lidwoord; 'mod' een bijwoordelijke bepaling; 'obj1' is lijdend voorwerp.

Dit lijkt te impliceren dat men verplicht is op een gesloten begraafplaats iets te begraven dat 'geen lijken' is. De bedoelde betekenis is natuurlijk dat het verboden is op een gesloten begraafplaats lijken te begraven:

Verplichting	
Niet	
Actie	Begraven
Patient	geen lijken
Voorwaarde	op een gesloten begraafplaats

Als we een negatie tegenkomen in een zin, dan leidt dat tot de negatie van de hele representatie zoals in het voorbeeld hierboven de 'niet' bovenaan, en verdwijnt het woord ('geen' in dit geval) uit de representatie zelf.

### 5.1 Definities

Definities bestaan uit drie delen:

- 1 De term die gedefinieerd wordt, het *definiendum*;
- 2 De definitie of *definiens*; en
- 3 Optioneel een *scope*-bepaling die aangeeft voor welke rechtsbronnen de definitie geldt.

Bijvoorbeeld deze zin uit de Geneesmiddelenwet, artikel 1, aanhef en onderdeel c: 'In deze wet en de daarop berustende bepalingen wordt verstaan onder immunologisch geneesmiddel: een vaccin, toxine, serum of allergeen' kan eenvoudig worden omgezet in de representatie:

Definitie	
Scope	deze wet en de daarop berustende bepalingen
Definiendum	immunologisch geneesmiddel
Definiens	een vaccin, toxine, serum of allergeen

Eventueel kan de uitvoer van de parser nog gebruikt worden om de *definiens* op te splitsen in de afzonderlijke delen. Bij het besluiten dat wel of niet te doen is nog menselijke interventie nodig. Zo wordt in de Drank- en Horecawet in artikel 1 ‘sterke drank’ gedefinieerd als: ‘de drank, die bij een temperatuur van twintig graden Celsius voor vijftien of meer volumeprocenten uit alcohol bestaat, met uitzondering van wijn’. Voor de meeste toepassingen is het volledig representeren van deze definitie vermoedelijk overbodig.

Zelfs als we niet van plan zijn een compleet automatisch beslissysteem te ontwerpen, dan kan het opsporen en in kaart brengen van begrippen en definities al helpen in de uitvoeringspraktijk. Bij de Belastingdienst en de Migratie- en Naturalisatiedienst (IND) gebeurt dat nu nog met de hand, om opleidings- en informatiemateriaal te maken of om de impact van nieuwe regelgeving op de bedrijfssystemen te bepalen. In een haalbaarheidsstudie hebben wij aangetoond dat automatische ondersteuning mogelijk is.<sup>16</sup> De computer vindt wel veel te veel concepten: bijna 7.000 in zes belastingwetten rond nalatenschap die in totaal zo’n 23.000 keer voorkwamen. Dat komt onder andere doordat we grote naamwoordzinnen en al hun onderdelen ook als concept opnamen. Zo ontstonden er concepten bestaande uit 21 woorden als: ‘kader van een regeling voor onderling overleg op grond van het verdrag ter afschaffing van dubbele belasting in geval van winstcorrecties’. Zowel deze hele lange zin als concepten bestaande uit één woord zijn niet zo interessant voor de professionals bij de Belastingdienst. Als we de meest extreme lengtes aan beide kanten schrappen, houden we een lijst over die door de experts bekeken en aangepast kan worden. Dit is in ieder geval efficiënter dan alles met de hand doen. Verder onderzoek zal moeten uitwijzen waar de optimale lengte ligt voor een volledig automatisch proces.

## 6 Integreer van modelfragmenten

Als we alle afzonderlijke modellen hebben van alle zinnen, moeten deze nog geïntegreerd worden. Allereerst kunnen gebruikte termen aan hun definities gebonden worden. In principe komt het neer op het matchen van deze termen met de *definienda*. Hierbij kunnen drie kleine problemen optreden. Ten eerste kan de term net anders geformuleerd zijn; enkelvoud kan meervoud zijn geworden, een werkwoord kan in een andere tijd gebruikt worden. De standaardoplossing is het terugbrengen van woorden tot hun stam of lemma. De Alpino-

parser doet dat al voor ons. Ten tweede kan de betreffende definitie in een ander document staan. In lagere regelgeving is dat vaak het geval, getuige de *scope*-bepaling hierboven. Deze lagere regelgeving verwijst op haar beurt weer naar de wet waaraan zij haar bevoegdheid ontleent, zodat de definities te achterhalen zijn. Ten slotte kunnen er nog meerdere definities van een term zijn, allemaal met een verschillend bereik. De meest nauwe definitie waarbinnen de term valt moet dan gekozen worden. Zo stelt artikel 3.6, eerste lid, aanhef en onderdeel a van de Wet arbeid en zorg: ‘Voor de toepassing van deze paragraaf wordt verstaan onder werknemer: de werknemer, bedoeld in artikel 1:1, onderdeel b, met uitzondering van degene die op grond van de Eerste Afdeling, Paragraaf 2, van de Ziektewet geen werknemer in de zin van die wet is’. Dit artikel geeft een inperking van de *scope* van de definitie van werknemer in artikel 1.1 (‘deze wet’) voor deze paragraaf en de term ‘werknemer’ hierbinnen moet aan deze tweede definitie gelinkt worden.

Een andere manier waarop modelfragmenten verbonden kunnen worden is als dezelfde term gebruikt wordt zonder dat hij gedefinieerd wordt. Dat kan soms in dezelfde zin zijn, bijvoorbeeld in de hoofdzin en een bijzin. Soms wordt dat expliciet aangegeven als bijvoorbeeld in ‘een verzoek als bedoeld in lid 7’, soms met aanwijzende voornaamwoorden als ‘die’ en ‘deze’. Zo bijvoorbeeld in de Mijnbouwwet, artikel 21, vijfde lid, aanhef en onderdeel a: ‘De vergunning vervalt van rechtswege: a. als de houder een natuurlijke persoon is, met ingang van de dag na die waarop die persoon is overleden’, waar ‘die persoon’ verwijst naar de ‘natuurlijke persoon’. In andere gevallen gebeurt het door eerst een onbepaald lidwoord te gebruiken (‘een beslissing’) en daarna te verwijzen met een bepaald lidwoord (‘de beslissing’). Deze verwijzingen zijn automatisch oplosbaar, maar dat is niet altijd het geval. De Successiewet 1956, artikel 2, derde lid stelt bijvoorbeeld: ‘In geval van schenking door een rechtspersoon wordt de plaats, waar de schenker is gevestigd, als zijn woonplaats aangemerkt’. ‘De schenker’ is niet eerder geïntroduceerd, maar een ‘schenking’ impliceert een ‘schenker’ en in dit geval is dat een rechtspersoon. Dit kan de machine niet afleiden zonder extra informatie. Integratie kan dan ook niet zonder menselijke tussenkomst gebeuren.

Als er een geïntegreerd model van de relevante rechtsbronnen is gemaakt, is het omzetten naar een werkende toepassing nog wel een technische uitdaging, maar juridische kennis is dan niet meer nodig. Natuurlijk moet

16 R. Winkels & R. Hoekstra, ‘Automatic Extraction of Legal Concepts and Definitions’, in: B. Schäfer (red.), *JURIX 2012. Volume 250 of Frontiers in Artificial Intelligence and Applications*, Amsterdam: IOS Press 2012, p. 156-165.



het model onderhouden worden als het recht of de maatschappij verandert, maar dat zal dan makkelijker en sneller kunnen gebeuren. De impact van wijzigingen is nu meteen duidelijk, omdat alle afhankelijkheden tussen rechtsbronnen expliciet zijn gemaakt. Sterker nog, ons onderzoek heeft aangetoond dat het formaliseren van wetsontwerpen kan leiden tot het vinden en repareren van fouten en onbedoelde vaagheden.<sup>17</sup>

**Ons onderzoek heeft aangetoond dat het formaliseren van wetsontwerpen kan leiden tot het vinden en repareren van fouten en onbedoelde vaagheden**

### 7 Conclusies en vooruitblik

Ik heb betoogd dat juridische toepassingen kunnen helpen bij het beter en efficiënter uitvoeren van veel juridische taken. Om dat te kunnen doen, moet onder andere de juridische ‘kennis’ die in rechtsbronnen zit in computersystemen worden ondergebracht. Dat is een tijdrovend en foutgevoelig proces. Automatisering van dat proces zou kunnen helpen. Ik heb laten zien hoe we de grote stap van bronnen in natuurlijke, juridische taal naar computerprogramma’s kunnen opdelen in stukken en door gebruik te maken van terugkerende patronen en taaltechniek een heel eind kunnen komen. Zo lukt het goed verwijzingen in en tussen juridische documenten expliciet te maken en op basis hiervan nieuwe informatie over deze rechtsbronnen te genereren, zogenaamde metadata. Het automatisch vinden van concepten in rechtsbronnen kan professionals bij uitvoeringsinstanties helpen bij het beheren van hun informatie- en lesmateriaal en het inschatten van de impact van wetswijzigingen op hun uitvoeringstaak. Ook is het mogelijk op basis van machinale interpretatie van wijzigingswetten tot automatische consolidatie van wetgeving te komen.

Er blijven echter nog genoeg problemen over en menselijke interventie is onvermijdelijk. Te denken valt aan het volgende.

De Alpino-parser is een mooi stukje techniek, maar hij werkt niet perfect voor juridische teksten. De uitvoer bestaat uit verschillende mogelijke ontledingen van een zin en hoewel de correcte er altijd bij zit, is het lang niet altijd de eerste. Alpino is getraind op een

krantencorpus en juridische teksten zijn significant anders; ze worden gekenmerkt door lange en vaak ingewikkelde zinnen, zoals we hierboven al even zagen bij een concept van 21 woorden. Een mogelijke oplossing is Alpino te hertrainen op een juridisch corpus.

**Ik heb betoogd dat juridische toepassingen kunnen helpen bij het beter en efficiënter uitvoeren van veel juridische taken**

Ten tweede kan gewezen worden op het feit dat rechtsbronnen worden geschreven door mensen, voor mensen. Wij gebruiken een heleboel kennis van de wereld bij het schrijven en begrijpen van zulke teksten. Maar zolang de machine deze kennis niet heeft, kan correcte interpretatie een probleem blijven; ik heb daar hierboven al enkele voorbeelden van gegeven.

Een oplossing voor dit probleem zou zijn om aan de bron te beginnen. Als wetgevingsjuristen eenduidiger zouden aangeven wat de bedoeling is, zou dat ons een hoop werk besparen. Ze zouden dan samen met uitvoeringsinstanties automatische beslis-systemen kunnen produceren in plaats van stapels papier. Voor bijvoorbeeld belasting- en administratief recht is dat nu al goed mogelijk. Voor andere terreinen zou het al schelen als wetgevingsjuristen automatische ondersteuning zouden krijgen bij het creëren en (vooral) veranderen van wetgeving, zodat ze niet nodeloos nieuwe termen introduceren, maar wel consistent verwijzen en de impact van veranderingen beter kunnen inschatten.

**De kwaliteit van wet- en regelgeving kan aan de bron al aanmerkelijk verbeterd worden**

Dat grote automatiseringsprojecten bij de IND of het Uitvoeringsinstituut Werknemersverzekeringen mislukken of in ieder geval in aanzienlijke mate tijd en budgetten overschrijden, heeft onder andere te maken met de grote afstand tussen geschreven rechtsbronnen en uitvoerbare machinecode. Onderzoek als hier beschreven is dan uitermate nuttig, maar de kwaliteit van wet- en regelgeving kan aan de bron al aanmerkelijk verbeterd worden.

<sup>17</sup> Zie bijvoorbeeld: T. van Engers e.a., ‘POWER: Using UML/OCL for Modelling Legislation – an application report’, *Proceedings of the 8th International Conference on Artificial Intelligence and Law*, New York: ACM Press 2001, p. 157-167.