

Legal(ly) Linked Data

Over de noodzaak tot betere standaardisatie van juridisch bronmateriaal

Computerrecht 2018/51

De juridische informatie die wordt opgeslagen is heel divers en vooral weinig gestructureerd. Die input moet verbeteren met meer structurering als we juridische informatie willen laten verwerken door computers. Pas als dit voorwerk niet meer nodig is en alle juridische informatie op een eenduidige en samenhangende wijze wordt opgeslagen en verwerkt, kunnen we gaan denken aan echt geavanceerde computertoepassingen in het juridische bedrijf.

1. Inleiding

Al geruime tijd gaat het recht de computer in: vrijwel alle basisinformatie van de democratische rechtsstaat (wetten, uitspraken, parlementaire geschiedenis) is inmiddels gedigitaliseerd; hetzelfde geldt voor rechtswetenschappelijke literatuur, registers en andere secundaire bronnen. Tegelijkertijd komt recht in toenemende mate ook uit de computer: overeenkomsten rollen na opgave van een paar variabelen uit de printer, expertsystemen produceren steeds complexere beschikkingen en juridische zoekmachines kunnen de gebruiker steeds beter van bruikbare resultaten voorzien.

Er zit evenwel nog een grote kloof tussen de huidige stand van *Legal Tech* en de beloften die ons worden voorgespiegeld,² bijvoorbeeld op het gebied van *RegTech* (technologie voor het monitoren van nakoming van financiële regelgeving),³ *smart contracts* in blockchains,⁴ voor de verbetering van de 'relevantiebepaling' in zoekmachines,⁵ advocaatkeuze⁶ of voor de vergezichten van 'Legal Big Data'.⁷ Deels heeft de huidige *underperformance* te maken met de eigenaardigheden van het juridisch domein, maar voor een belangrijk deel ook met de mismatch tussen de kwaliteit van het 'in-

put-recht' en de eisen die daaraan kunnen worden gesteld, bezien vanuit het 'output-recht'.

Garbage in, garbage out is een welbekend informaticaprincipe. En omdat in de wereld van internet en open data de makers van Legal-Tech-applicaties geen controle hebben over de *input-garbage*, proberen zij de kwaliteit ervan te verbeteren. De verschillende componenten die daarvoor worden gebruikt, duiden we hier aan met de verzamelterm 'rechtzeters'. Deze rechtzeters brengen bijvoorbeeld structuur aan in ongestructureerde documenten,⁸ herkennen concepten,⁹ clusteren aan de hand van trefwoorden, zien onderlinge verbanden en argumentatiepatronen, optimaliseren teksten voor zoekmachines, etc.

Dankzij allerlei technologieën, te scharen onder het containerbegrip 'artificiële intelligentie', worden dergelijke rechtzeters steeds beter. Desalniettemin kleven er aanzienlijke nadelen aan een te grote rol voor deze rechtzeters: ze moeten het input-recht interpreteren, ze maken fouten, generaliseren en de werking ervan is niet inzichtelijk voor eindgebruikers. Bovendien zijn rechtzeters onderhoudsgevoelig en maatschappelijk bezien duur: iedere producent van output-recht-applicaties moet zijn eigen rechtzeters bouwen, die onderling niet of nauwelijks interoperabel zijn.

Uit een oogpunt van rechtszekerheid, maatschappelijke efficiency en ter stimulering van innovatie zou het daarom de voorkeur verdienen om de kwaliteit van het input-recht zodanig te verbeteren dat rechtzeters uiteindelijk niet meer nodig zijn. Hiertoe moet het juridisch bronmateriaal op zodanige manier worden ge(her)structureerd dat het door computers beter kan worden gelezen, verwerkt en begrepen. Deze drie begrippen vormen een oplopende schaal, zowel qua bruikbaarheid voor output-recht-applicaties, als qua complexiteit om ze binnen het input-recht te realiseren. 'Leesbaarheid' ziet alleen op tekstuele leesbaarheid, terwijl 'verwerkbaarheid' ook de aanwezigheid van documentstructuren, identificatiecodes en metadata impliceert. De informatie is dan weliswaar geordend, maar voor de computer nog steeds zonder betekenis. Daartoe moet ook (juridische) betekenis worden toegevoegd, bijvoorbeeld in de vorm van gestructureerde terminologie en geduide relaties. Dit terrein is volop in beweging; voor een volledig overzicht van de stand van zaken ontbreekt in dit artikel helaas de ruimte. Daarom zal de problematiek casuïstisch

1 Mr. dr. M. van Opijnen is adviseur rechtsinformatica bij het Kennis- en Exploitatiecentrum Officiële Overheidspublicaties (UBR|KOOP) en functioneel architect van de in dit artikel besproken LinkeXtractor.

2 Zie bijvoorbeeld N. Netten e.a., 'The Rise of Smart Justice: on the Role of AI in the Future of Legal Logistics', in: *Proceedings AI4J - Artificial Intelligence for Justice, Workshop at the 22nd European Conference on Artificial Intelligence*, (Den Haag 30 augustus 2016). Online publiek.

3 UK Government Chief Scientific Adviser, *Distributed Ledger Technology: beyond block chain*. Online publiek.

4 H. Schuringa, 'Enkele civielrechtelijke aspecten van blockchain', in: *Computerrecht* 2017/254.

5 M. van Opijnen & C. Santos, 'On the Concept of Relevance in Legal Information Retrieval', in: *Artificial Intelligence and Law* 2017, nr. 25, p. 65-87. Online publiek.

6 Kritisch: G. van Dijk, 'Neem kiezen advocaat per computer met korreltje zout', in: *Het Financieel Dagblad* 3 oktober 2016, p. 9.

7 B. Custers & F. Leeuw, 'Legal big data. Toepassingen voor de rechtspraktijk en juridisch onderzoek', in: *NJB* 2017/1854.

8 Zie bijvoorbeeld M. Trompper & R. Winkels, 'Automatic Assignment of Section Structure to Texts of Dutch Court Judgments', in: F. Bex & S. Villata (red.), *Legal Knowledge and Information Systems - JURIX 2016: The Twenty-Ninth Annual Conference*, (Nice: IOS Press, 2016), p. 167-172.

9 Zie bijvoorbeeld: R. Nanda e.a., 'Concept Recognition in European and National Law', in: A. Wyner & G. Casini (red.), *Legal Knowledge and Information Systems - JURIX 2017: The 30th International Conference* (Luxembourg: IOS Press, 2017), p. 193-198.

worden geïllustreerd aan de hand van een rechtzetter die onlangs in het publieke domein in gebruik is genomen: de LinkeXtractor.

2. Linkextracten

Iemand die onderzoek wil doen naar de toepassing van artikel 138 Sr in de rechtspraak is wel even aan het zoeken. Het veld 'wetsverwijzingen' op Rechtspraak.nl is meestal niet gevuld en kan ook niet worden doorzocht. Men moet dus zoeken in de tekst zelf. Maar zo'n verwijzing kan zijn geschreven als 'Artikel 138 Wetboek van strafrecht', 'Art. 138 Sr' of 'art 138 WvSr'. De zoekmachine weet niet dat dat allemaal hetzelfde is, en dus moet de gebruiker eerst al die mogelijke varianten zelf bedenken, om er vervolgens op te zoeken. Maar dan nog zal je de uitspraak missen waarin de rechter aan het begin van de uitspraak schrijft: "*Wetboek van Strafrecht (hierna: de Wet)*", en drie alinea's verder: "*Art. 138 van de Wet*". Ingewikkelder nog wordt het bij verwijzingen naar Europese regelgeving ('Executieverordening' = 'EEX-Vo' = 'Verordening (EG) nr. 44/2001' = 'Verord. (EEG) 44-2001', etc.) of jurisprudentie ('NJ 2006/606' = 'JA 2005, 34' = 'LJN AR5213' = 'Wrongful life' = 'Baby Kelly').

Deze zoekfrustraties zijn sinds kort verleden tijd. Op Linked-data.overheid.nl (hierna ook: 'LiDO-website') kan namelijk gezocht worden op alle bovengenoemde citatievarianten. De zoekmachine begrijpt de zoekvraag, retourneert de correcte titel van het bedoelde document en geeft direct ook een overzicht van (citatie)relaties. Zo wordt bij de Executieverordening getoond dat er 1.567 documenten bekend zijn die ernaar verwijzen, waaronder 1.561 rechterlijke uitspraken – ongeacht de schrijfwijze die de rechter heeft gehanteerd om de verwijzing te maken. Vervolgens kan men eindeloos door het hele netwerk van relaties klikken, bijvoorbeeld om te onderzoeken hoe een arrest van de Hoge Raad, het Hof van Justitie of het EHRM doorwerkt in latere jurisprudentie. Vanaf iedere uitspraak op Rechtspraak.nl kan men nu – door te klikken op de link 'Verrijkte uitspraak' – de van hyperlinks voorziene versie op de LiDO-website oproepen. Ook bij een wetsartikel op Wetten.overheid.nl kan men rechtstreeks de gerelateerde jurisprudentie opvragen.¹⁰

Al deze hyperlinks worden aangebracht door de 'LinkeXtractor',¹¹ een rechtzetter die met grote precisie juridische verwijzingen herkent en herleidt naar gestandaardiseerde *Uniform Resource Identifiers* (URI's). URI's zijn codes waarmee informatieobjecten op een gestandaardiseerde, unieke en persistente manier worden geïdentificeerd. Als dat op internet gebeurt spreken we van een 'http-URI'. Zo'n http-URI lijkt op een gewoon webadres (een URL: *Uniform*

Resource Locator) maar is conceptueel fundamenteel anders: URL's kunnen veranderen en verdwijnen, URI's zijn door een bevoegde autoriteit 'gemunt', exact beschreven en persistent. Ze zijn niet bedoeld om een bepaald document te tonen (al gebeurt dat vaak wel), maar zorgen ervoor dat alle computers zonder enig misverstand over hetzelfde brokje informatie kunnen praten.

De LinkeXtractor herkent thans verwijzingen naar officiële bekendmakingen (zoals Staatsblad en Staatscourant), parlementaire documenten, Europese regelgeving, Nederlandse regelgeving (zoals beschikbaar op Wetten.overheid.nl) en jurisprudentie van alle Nederlandse en Europese instanties. Op dit moment zijn alle uitspraken van Rechtspraak.nl 'ge-linkextract', het verwerken van andere documenten (zoals Kamerstukken) is in voorbereiding. De miljoenen linkjes van de LinkeXtractor bieden tal van mogelijkheden: naast de beschreven functionaliteit op de LiDO-website kunnen ze dienen als grondstof voor academisch en beleidsmatig onderzoek,¹² onderwijskundige toepassingen¹³ en statistische analyses.¹⁴

3. Tekortkomingen

Ook al is de herkenningsgraad hoog, een gebruiker van de LiDO-website zal af en toe de wenkbrauwen fronsen, want sommige verwijzingen worden foutief of helemaal niet herkend. Dit kan verschillende oorzaken hebben.

Allereerst zijn nog niet alle brongegevens beschikbaar. Zo worden verwijzingen naar Kamerstukken vanaf 1995 wel herkend, maar oudere (nog) niet. Ook komt het voor dat de teksten door de auteur onvoldoende zijn gecontroleerd. Zo kan de rechter bijvoorbeeld door een tikfout verwijzen naar een verordening die niet bestaat of kunnen verwijzingen incompleet zijn ('artikel 126 BW'). Ook kunnen citeertitels en afkortingen ambigu zijn. Zo bestaan er talloze regelingen met de naam 'Aanwijzing opsporingsambtenaren', kan 'Rv' opeens naar twee varianten van dezelfde wet verwijzen¹⁵ en

10 Klik daartoe bij het desbetreffende wetsartikel aan de rechterzijde op het knopje 'Opties' en selecteer 'Relaties'.

11 Zie voor een technische beschrijving: M. van Opijnen, N. Verwer & J. Meijer, 'Beyond the Experiment: the eXtensible Legal Link eXtractor', in: *Proceedings of Workshop on Automated Detection, Extraction and Analysis of Semantic Information in Legal Texts, held in conjunction with the 2015 International Conference on Artificial Intelligence and Law (ICAIL, San Diego, 2015)*. Online publiek.

12 Het onderzoek van jurisprudentiële netwerkrelaties wint aan belangstelling. Zie bijvoorbeeld J. Jacobs & M. Vols, 'Juristen als rekenmeesters: Over de kwantitatieve analyse van jurisprudentie', in: G. van den Berg (red.), *In dienst van het recht: Opstellen aangeboden aan prof. mr. J.G. Brouwer ter gelegenheid van zijn afscheid als hoogleraar Algemene Rechtswetenschap aan de Rijksuniversiteit Groningen*, Den Haag: Boom Juridisch 2017, p. 89-104; T. Agnoloni & U. Pagallo, 'The Power Laws of the Italian Constitutional Court, and Their Relevance for Legal Scholars', in: A. Rotolo (red.) *Legal Knowledge and Information Systems – JURIX 2015: The Twenty-Eighth Annual Conference* (Braga), p. 1-10.

13 Zie bijvoorbeeld D. van Kuppevelt & G. van Dijck, 'Answering Legal Research Questions About Dutch Case Law with Network Analysis and Visualization', in: A. Wyner & G. Casini (red.), *Legal Knowledge and Information Systems – JURIX 2017: The 30th International Conference* (Luxembourg), p. 95-100.

14 Zoals het 'Model for Automated Rating of Case law', een computeralgoritme waarmee het juridisch belang van rechterlijke uitspraken kan worden berekend: M. van Opijnen, *Op en in het web. Hoe de toegankelijkheid van rechterlijke uitspraken kan worden verbeterd* (diss. Amsterdam UvA, 2014). Online publiek.

15 *Wetboek van Burgerlijke Rechtsvordering* (geldt in geval van digitaal procederen) (<https://wetten.nl/BWBR001827>) of *Wetboek van Burgerlijke Rechtsvordering* (geldt in geval van niet-digitaal procederen) (<https://wetten.nl/BWBR0039872>).

kan 'Verordening 2015/75' zowel slaan op Verordening 75 uit het jaar 2015 als op Verordening 2015 uit het jaar 1975.¹⁶ Belangrijker evenwel is dat redactionele citeerinstructies, zoals de Aanwijzingen voor de regelgeving¹⁷ en de Leidraad voor juridische auteurs,¹⁸ nog uit het papieren tijdperk stammen. Deze voorschriften veronderstellen dat de lezer met behulp van de gegevens in de verwijzing de bedoelde informatie kan opzoeken, maar dergelijke citaties leveren geen constructen op die direct en foutloos naar een digitaal object leiden, terwijl de geciteerde bronnen inmiddels wel als digitale objecten bestaan. Een voorbeeld van een citatiemethode die wel aan deze eis voldoet is de *European Case Law Identifier* (ECLI), een identificatiecode voor rechterlijke uitspraken die in 2010 door de Raad van de EU is vastgesteld¹⁹ en thans in zestien lidstaten en bij drie Europese gerechten is geïmplementeerd.²⁰ In 2013 heeft in Nederland de ECLI het Landelijk Jurisprudentienummer (LJN) vervangen, is ECLI de geadviseerde citatiemethode geworden in de Leidraad voor juridische auteurs, alsmede opgenomen op de pas-toe-of-leg-uit-lijst van het Forum Standaardisatie.²¹ Per 1 januari 2018 is ECLI ook voorgeschreven in de Aanwijzingen voor de Regelgeving.²² Het opzoeken van een ECLI op Rechtspraak.nl, de ECLI-zoekmachine van het Europese e-Justice portaal²³ of Google leidt eigenlijk altijd naar de bedoelde uitspraak.²⁴

Ook de LinkeXtractor herkent natuurlijk ECLI's; bovendien wordt de ECLI gebruikt als de standaard-URI voor rechterlijke uitspraken: met behulp van complexe logica worden archaïsche verwijzingen (zoals zaaknummers en vindplaatsen) omgezet naar ECLI's. Voor andere documentsoorten worden andere URI's gehanteerd: voor Europese regelgeving wordt het op EUR-Lex gehanteerde CELEX-nummer gebruikt,²⁵ voor Kamerstukken en officiële bekendmakingen de (technische) codering van Officielebekendmakingen.nl²⁶ en voor Nederlandse wetgeving de BWB-id-standaard.²⁷ In tegenstelling tot de ECLI worden al deze URI's echter niet gebruikt door citerende juristen, waardoor de kans op fou-

tieve, ambigue, onvolledige of onduidelijke verwijzingen nog steeds bestaat.

Om de kwaliteit van de verwijzingen in de bronteksten te verbeteren zou, als eerste stap, de concipiënt van een uitspraak (of ander juridisch document) reeds bij het schrijven de LinkeXtractor – of vergelijkbare software – kunnen gebruiken om de redactionele en juridische kwaliteit van de uitspraak te controleren. Dat impliceert overigens niet dat de aangebrachte hyperlinks onderdeel moeten gaan uitmaken van de rechterlijke uitspraak zelf. Naast de authentieke uitspraaktekst zonder hyperlinks, zou een versie met de geverifieerde links beschikbaar kunnen worden gesteld op Rechtspraak.nl, op vergelijkbare wijze als waarop de versie met niet-geverifieerde hyperlinks thans op de LiDO-web-site is te zien.

4. Wetsverwijzingen

Door de LinkeXtractor in te zetten in het tekstproductieproces zou ook een ander lastig probleem kunnen worden opgelost: het aanduiden van de versie van de geciteerde wet(sartikel). Als de rechter refereert aan 'art. 138 Sr', dan bedoelt hij daarmee (impliciet) te verwijzen naar dit wetsartikel in zijn algemeenheid (als uitdrukking van een bepaald concept of beginsel) of naar de versie zoals geldig op de dag van uitspraak. Alleen als een andere versie wordt bedoeld of als verwarring op de loer ligt, wordt de geldigheidsdatum nader gespecificeerd. Daarvoor worden echter vaak uitdrukkingen gebruikt die voor een computer moeilijk zijn te interpreteren, zoals: 'pleegdatum delict', 'ten tijde van het aangaan van de overeenkomst' of 'zoals thans ahangig'. Van de LinkeXtractor vergt dit eerst (uitputtende) herkenning van dit soort constructen,²⁸ en vervolgens het herleiden ervan naar de bedoelde datum. Een geschoold juridisch oog kan de 'pleegdatum delict' snel vinden, maar voor een computer is dit een helse opgave. De LinkeXtractor zet daarom – gemakshalve – wetsverwijzingen op de versie zoals geldig op de uitspraakdatum omdat dit in de meeste gevallen correct is.

Ter verbetering van de kwaliteit zou ook voor dit soort situaties de concipiënt van de tekst de regie moeten krijgen. Dat zou niet eens arbeidsintensief hoeven te zijn. Zo zouden bij het bovenomschreven controleren van de bronverwijzingen alle wetsverwijzingen standaard de geldigheidsdatum 'vandaag' kunnen krijgen, tenzij de concipiënt expliciet opdracht geeft om bij bepaalde verwijzingen een andere datum te gebruiken. De software kan daarbij natuurlijk een lijstje produceren van de mogelijke geldigheidsdata.²⁹

Er zijn overigens meer problemen met het identificeren van regelgeving(selementen). Iemand die jurisprudentie zoekt over artikel 101 VWEU, wil waarschijnlijk ook jurispruden-

16 Veroorzaakt door de veranderde nummering van (onder meer) verordeningen per 1 januari 2015. Harmonising the numbering of EU legal acts, EUR-Lex.

17 Aanwijzingen voor de regelgeving, Overheid.nl (online publiek).

18 M.H. Bastiaans e.a., *Leidraad voor juridische auteurs 2016*, (Deventer: Wolters Kluwer 2016).

19 *PbEU* 2011, C 127/01.

20 Zie het implementatieoverzicht op <https://bo-ecli.eu> alsmede: M. van Opijnen, 'Gaining Momentum. How ECLI Boosts Accessibility of Case Law in Europe', in: *Journal of Open Access to Law* 2017, 5 (online publiek).

21 ECLI, Forum Standaardisatie (online publiek).

22 Aanwijzing 3.44.

23 ECLI search engine, e-Justice Portal (online publiek). Zie ook: M. van Opijnen & B. Veenman, 'Jurisprudentie zoeken op Europese schaal', in: *Nederlands tijdschrift voor Europees Recht* 2016, nr. 4, p. 137-140.

24 De ECLI is te beschouwen als URI, maar is geen http-URI. Een voorstel daartoe is inmiddels wel geformuleerd: M. van Opijnene.a., 'Towards ECLI 2.0', in: *Proceedings CeDEM17 International Conference for E-Democracy and Open Government 2017* (Krems).

25 What is a CELEX number?, EUR-Lex (online publiek).

26 Voor zover mij bekend is de identificatiemethode van Officielebekendmakingen.nl de enige met een beschreven URI-structuur, te vinden op de website Juriconnect.nl, onder 'Implementatie'.

27 Ook deze standaard staat op de pas-toe-of-leg-uit-lijst van Forum Standaardisatie. BWB, Forum Standaardisatie (online publiek).

28 Voor wetgeving is dit wel gedaan, zie bijvoorbeeld E. de Maat, *Making Sense of Legal Texts*, (diss. Amsterdam UvA 2012), maar in vergelijking met de citatiepraktijk in de rechtspraak is het aantal verwijzingsconstructen in wetsteksten beperkt te noemen.

29 Vergelijk het lijstje met mogelijke geldigheidsdata dat gebruikt kan worden bij de versievergelijker op Wetten.nl.

tie over (het gelijklopende) artikel 81 EG-verdrag.³⁰ Maar wanneer is nog sprake van 'gelijkheid' tussen twee versies, al dan niet na vernummering? En wie bepaalt dat? Weerbarstig is ook het probleem van regelgeving die met terugwerkende kracht in werking treedt. Daarbij speelt niet alleen de inwerkingtredingsdatum een rol, maar ook de zichtdatum: kon je op het moment dat je handelde naar de wettekst al weten dat deze – met terugwerkende kracht – was gewijzigd?

De hiervoor reeds genoemde BWB-id-standaard houdt rekening met het verschijnsel van terugwerkende kracht, maar kent wel enkele andere fundamentele problemen, die voor een groot deel zijn terug te voeren op de (onvolkomen) manier waarop de wetgeving is gestructureerd. Zo hebben soms meerdere artikelen binnen één regeling hetzelfde nummer, of ontbreekt een aanduiding bij regelingsonderdelen. Nederland is natuurlijk niet het enige land dat worstelt met dit soort vraagstukken, daarom zijn in internationaal verband verschillende initiatieven ontplooid om te komen tot een gestandaardiseerde informatiearchitectuur voor wetgevingsdocumenten. In CEN-Metalex³¹ zijn belangrijke theoretische fundamenten gelegd, URN:LEX³² is een uitgebreide standaard voor identificatie en verwijzing. Beide vinden in de praktijk echter weinig toepassing. De meest uitgewerkte standaard is Akoma Ntoso (AKN), sinds 2017 een OASIS-standaard.³³ Deze standaard heeft naast een *naming convention* (afspraken over de URI-structuur)³⁴ ook een uitgebreid XML-schema voor het beschrijven van documentstructuren.³⁵ In steeds meer jurisdicties worden de toepassingsmogelijkheden voor Akoma Ntoso onderzocht, onder meer in het Verenigd Koninkrijk³⁶ en de Europese Unie.³⁷ Onlangs is UBR|KOOP³⁸ een onderzoek gestart naar de mogelijkheden voor Akoma Ntoso in Nederland. Daarbij wordt tevens onderzocht of de *European Legislation Identifier* (ELI) kan worden geïmplementeerd.³⁹ ELI is primair ontworpen voor bekendmakingen en richt zich alleen op identificatie, niet op documentstructuren. De standaard kent enkele beperkingen, maar inmiddels hebben verschillende lidstaten, alsmede de EU zelf (delen van) de standaard geïmplementeerd.⁴⁰

5. Parlementaire documenten en het wetgevingsproces

Op OfficiëleBekendmakingen.nl zijn ook de parlementaire stukken doorzoekbaar en gekoppeld aan Wetten.nl, de LinkeXtractor detecteert verwijzingen naar deze Kamerstukken en ook op de websites van Eerste en Tweede Kamer zijn alle documenten aanwezig. Desondanks kan de vraag: "Geef mij een overzicht van de parlementaire geschiedenis van artikel 12 van de Wet op de rechtsbijstand" nog niet met een enkele muisklik worden beantwoord. Integendeel, je bent er al snel een dag mee bezig. Het linkextracten van de Kamerstukken zal daarvoor niet de (volledige) oplossing bieden.

Veel meer dan bij rechterlijke uitspraken het geval is, is de kans namelijk groot dat de wetsverwijzingen die door de LinkeXtractor in de Kamerstukken worden gevonden, niet correct zijn. Dit vindt zijn grond in de gewoonte om wetsvoorstellen gedurende de parlementaire behandeling te hernummeren, waarbij een additionele hernummering kan plaatsvinden bij definitieve vaststelling. De transponeringstabellen⁴¹ waarin deze veranderingen vaak worden bijgehouden zijn wel computerleesbaar maar niet computerverwerkbaar.

Een tweede probleem is dat bij de Handelingen en veel andere Kamerstukken niet is opgenomen op welke wetsartikelen ze betrekking hebben, al zijn deze stukken wel gekoppeld aan parlementaire dossiers. Ook in de documenten zelf is geen betekenisvolle en computerverwerkbare informatie opgenomen over de inhoud ervan. Er kan wel worden verwezen naar een memorie van toelichting, maar het is niet mogelijk te verwijzen naar een specifieke paragraaf binnen de artikelsgewijze toelichting. Op dit moment mist een memorie van toelichting de daartoe benodigde betekenisvolle structuur; bovendien is er geen standaard om dergelijke verwijzingen te construeren. Die standaard is er overigens ook niet voor identificatie van een parlementair document als geheel; Overheid.nl, Rijksoverheid.nl, Eerste Kamer en Tweede Kamer gebruiken allemaal eigen coderingen voor digitale identificatie van dezelfde documenten, maar deze coderingen zijn niet interoperabel en zelden gedocumenteerd.⁴² Voor hergebruikers is het daarom niet of nauwelijks mogelijk om informatie van verschillende sites te integreren, zonder daarvoor rechtzetters te ontwikkelen. Net als parlementaire documenten hebben ook de officiële elektronische bekendmakingen – zoals de publicaties in Staatsblad, Tractatenblad en Staatscourant – geen betekenisvolle documentstructuren.

Hiermee raken we aan een veel wezenlijker probleem: het wetgevingsproces is nog steeds gebaseerd op negentiende-eeuwse concepten, en voldoet niet aan de eisen die je in de eenentwintigste eeuw zou mogen stellen. Terwijl er vroeger een beperkt wettelijk corpus was, met relatief

30 Aanwijzing 6.17 van de Aanwijzingen voor de regelgeving raadt artikelvernummeringen af, maar laat tegelijkertijd uitzonderingen toe; ook in Nederland gebeurt het regelmatig. Vernummeringen van leden binnen één artikel is een nog veel frequenter verschijnsel.

31 De Metalex-standaard is gegoten in de vorm van een *CEN Workshop Agreement*. Zie: www.metalex.eu.

32 URN:LEX is een *draft proposal* binnen de *Internet Engineering Task Force*. Zie A Uniform Resource Name (URN) Namespace for Sources of Law (LEX), IETF Datatracker (online publiek).

33 Akoma Ntoso Version 1.0. Part 1: XML Vocabulary, OASIS (online publiek).

34 Akoma Ntoso Naming Convention Version 1.0, OASIS (online publiek).

35 Akoma Ntoso Version 1.0. Part 2: Specifications, OASIS (online publiek).

36 Legislative drafting, amending and publishing tools, Legislation.gov.uk.

37 Het Europees Parlement gebruikt Akoma Ntoso binnen 'AT4AM', software voor het redigeren en routeren van amendementen: www.at4am.org/overview/.

38 <https://koopoverheid.nl>.

39 *PbEU* 2012, C 325/02.

40 ELI implementation status, EUR-Lex (online publiek).

41 Zie bijvoorbeeld in <https://zoek.officiëlebekendmakingen.nl/kst-34506-A.html>.

42 Voor zover mij bekend is de identificatiemethode van Officiëlebekendmakingen.nl de enige met een beschreven URI-structuur, te vinden op de website Juriconnect.nl, onder 'Implementatie'.

weinig onderlinge samenhang en een overzichtelijk aantal wijzigingsmomenten, kennen we tegenwoordig een interdependente kluwen met regelingen van verschillende staatkundige niveaus, talloze actoren en een complexe wetgevingskalender. Het meest fundamentele probleem waar het output-recht mee wordt geconfronteerd is dat het positief recht, voor zover vastgelegd in de wet, als zodanig niet digitaal kenbaar is: de afgekondigde wet is niet computerverwerkbaar en de wel computerverwerkbare geconsolideerde versie heeft geen officiële status.

Mochten we dit proces vandaag de dag (her)ontwerpen, dan zou het voor de hand liggen om daarvoor een structuur te kiezen die veel weg heeft van *Git*,⁴³ een gedistribueerd versiebeheersysteem dat wereldwijd door softwareontwikkelaars wordt gebruikt. In dit systeem kunnen tal van mensen tegelijkertijd samenwerken aan dezelfde (gedeelde) broncode. Dit leidt uiteindelijk tot een nieuwe versie van het geheel – softwareprogramma dan wel geconsolideerde wet – waarbij op ieder moment nog keuzes kunnen worden gemaakt aangaande welke ontwikkelde componenten wel of niet in een bepaalde release meegenomen moeten worden. In de huidige wetgevingssystematiek daarentegen worden wijzigingswetten in natuurlijke taal geformuleerd, omschrijvende wat er aangepast moet worden in een wet die wellicht reeds talloze malen door andere wijzigingswetten is veranderd. Bij gebreke aan een computerverwerkbaar structuur van de wijzigingswetten, is ook de consolidatie vervolgens mensenwerk en dus imperfect,⁴⁴ al zijn er zorgvuldige processen om de gevolgen daarvan te minimaliseren.⁴⁵ Weliswaar zijn stappen gezet om tot digitalisering van het proces te komen,⁴⁶ maar het Draaiboek voor de regelgeving⁴⁷ leert dat het in de praktijk toch nog vooral een papieren proces is, ook al zijn de documenten digitaal.

Om de kwaliteit van het output-recht te verbeteren, maar ook ter verbetering van het wetgevingsproces zelf, zal het uiteindelijk onvermijdelijk zijn als in de hele wetgevingsketen dezelfde computerverwerkbare – en bij voorkeur ook computerbegrijpbare taal wordt gesproken. Eén architectuur voor documentstructuren, één architectuur voor identificeren en verwijzen, één vocabulaire voor het beschrijven van de processtappen. De hierboven genoemde Akoma-Ntoso-standaard biedt voor het gehele wetgevingsproces een uitgebreide gereedschapskist; onderzoek naar bruikbaarheid zou opportuun zijn. Aan een andere, nationale, standaard wordt gewerkt binnen het programma 'Aan de slag met de omgevingswet': deze 'Standaard Officiële Publicaties' (STOP) voorziet onder meer in computerbegrijpbare 'annotaties' en automatisch consolideren.⁴⁸

43 <https://git-scm.com/>.

44 Een aardige illustratie van wat er mis kan gaan kan men krijgen door op Wetten.overheid.nl op de pagina 'uitgebreid zoeken' te zoeken op het woord 'abusievelijk' in het veld 'Wetstechnische informatie'.

45 L. Vester & M. van der Loop, 'Reactie op 'The World according to www.overheid.nl'', in: *NJB* 2014/2002.

46 *Kamerstukken II* 2012/13, 33605 VI, 1, sectie B-3-IV.

47 Draaiboek voor de regelgeving, Kenniscentrum Wetgeving en Juridische zaken (online publiek).

48 Vanaf 19-2-2017 is deze standaard in publieke consultatie. Consultatie Standaarden Omgevingsdocumenten, Geonovum (online publiek).

6. Secundaire Literatuur

Minstens zo problematisch als het citeren van regelgeving is het citeren van rechtswetenschappelijke literatuur.⁴⁹ De Leidraad voor juridische auteurs vult vele pagina's met voorschriften, maar ook deze zijn gebaseerd op het idee dat de lezer moet worden voorzien van een set zoekvariabelen waarmee hij zelf een zoektocht moet ondernemen. Maar als je de door de Leidraad voorgeschreven citaties letterlijk als zoekopdracht invoert in de diverse zoekmachines van de juridische uitgevers krijg je zelden het document dat je hoopte te vinden.

Om dit probleem op te lossen is binnen het Juriconnect-platform⁵⁰ de Juriconnect Doctrine Identifier (JCDI) ontwikkeld: een korte betekenisloze identificatiecode (URI) waarmee werken voor rechtswetenschap en rechtspraak uniek en persistent kunnen worden geïdentificeerd. Door deze code toe te voegen aan een tekstuele verwijzing, kan deze altijd worden teruggevonden in het Juridische Doctrine Register (JuDoReg).⁵¹ Dit register is thans in opbouw; het aantal deelnemers is nog beperkt, maar diverse uitgevers bereiden participatie voor. En als iedereen de JCDI vervolgens toevoegt aan zijn literatuurverwijzingen, dan is de vraag: "Waar is mijn artikel geciteerd?" straks ook met een enkele muisklik te beantwoorden. Ook uit wetenschappelijk oogpunt biedt een dergelijke citatie-index interessante mogelijkheden.⁵²

7. Tot slot

In voorgaande paragrafen is vooral ingegaan op de noodzaak om te komen tot een betere computerverwerkbaarheid van het input-recht, in het bijzonder het juridisch bronmateriaal van de overheid. Eenduidige identificatie en fijnmazige documentstructuren zijn daarvan essentiële onderdelen. Hoe nuttig een rechtzetter als de LinkeXtractor ook moge zijn, het creëert *Legal Linked Data* zonder enige autoriteit. Om geavanceerde Legal Tech-applicaties te kunnen ontwikkelen is het echter noodzakelijk dat er *Legally Linked Data* worden geproduceerd: computerverwerkbare data waarvan identificatie en structuur bij de totstandkoming zijn geborgd en dus juridisch betrouwbaar zijn.

Eigenlijk kunnen we pas dan serieus aan de gang met de volgende stap: het computerbegrijpbaar maken van juridische informatie. De LinkeXtractor kan nu bijvoorbeeld wel vaststellen dat een uitspraak verwijst naar een wetsartikel, maar we weten dan nog niet waarom die verwijzing is gemaakt.⁵³ Om dit soort vraagstukken op te lossen moet juridische semantiek in kaart worden gebracht en gestan-

49 Voor een analyse van deze problematiek: M. van Opijnen, 'The European Legal Doctrine Identifier – a Missing Link?', in: G. Peruginelli & S. Faro (red.), *La Dottrina Giuridica e la sua Diffusione* (Turijn: G. Giappichelli Editore, 2017). Online publiek.

50 <https://juriconnect.nl>.

51 <https://judoreg.nl>.

52 Zie ook W. Soetenhorst, 'Een juridische citatie-index. Het proof of concept is voorhanden', in: *NJB* 2017/915.

53 Zie voor een experiment bijvoorbeeld: R. Winkels e.a., 'Towards a Legal Recommender System', in: *27th International Conference on Legal Knowledge and Information Systems (JURIX 2014, Krakow)*. IOS Press 2014, p. 169-178.

daardiseerd. Op dit gebied wordt academisch veel onderzoek gedaan,⁵⁴ maar ontwikkeling in de praktijk is complex en weerbarstig. Ook hier is standaardisatie echter onontbeerlijk, maar gelukkig worden ook op dit vlak inmiddels stappen gezet. Als voorbeelden kunnen worden genoemd: de Overheid Web Metadata Standaard (OWMS),⁵⁵ de Stelselcatalogus,⁵⁶ de Justitiethesaurus⁵⁷ en de Nederlandse Fiscale Leerstukken Taxonomie (NFLT).⁵⁸

Een stap verder gaat het nog als de computer niet alleen de juridische terminologie kan begrijpen maar ook juridisch kan redeneren en mogelijk zelfs beslissen. Ongeacht of men dit een schrikbeeld⁵⁹ of een onvermijdelijkheid⁶⁰ vindt, ook hier is standaardisatie noodzakelijk.⁶¹

Artificiële intelligentie schetst ook voor het rechtsbedrijf intrigerende stippen op de horizon; innovatieprojecten die zich richten op recht uit de computer zijn hip en sexy. Het recht in de computer krijgen daarentegen is een minder populair onderwerp: het heeft een saai imago en noodzaak en complexiteit worden stelselmatig onderschat. Gezien de toenemende dataficering van het recht, het bestuurlijk belang dat wordt gehecht aan digitale innovatie en de economische voordelen die – zowel door publieke als private partijen – kunnen worden behaald, verdient de kwaliteit van het input-recht echter een hogere prioriteit op de agenda.

54 Zie voor een recent overzicht: P. Casanovas e.a., 'Special Issue on the Semantic Web for the Legal Domain Guest Editors' Editorial: The Next Step', in: *Semantic Web Journal* 2016. Online publiek.

55 <https://standaarden.overheid.nl/owms>.

56 www.stelselcatalogus.nl/.

57 <https://data.overheid.nl/data/dataset?q=justitiethesaurus>.

58 <https://fiscaletaxonomie.pleio.nl/>.

59 M. Februari, 'Robotrechters produceren McDonald's-Rechtspraak', in: *NRC Handelsblad* 25 augustus 2015, p. 16.

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

61 Zo is er een OASIS-werkgroep (LegalRuleML TC, oasis.open.org) en is er in Nederland het initiatief Wendbare Wetsuitvoering: www.wendbarewetsuitvoering.nl.