

Gerechtelijke uitspraken: een statistisch onderzoek aan de hand van text mining

Wolter Hassink, Kim van der Kraats, Remme Verkerk (allen hoogleraar aan de Universiteit Utrecht)

26 september 2022

In een aantal lopende statistisch onderzoeken naar civiele procedures van Diogo Leitao Requena, Remme Verkerk, Frans van Dijk, en Wolter Hassink, worden administratieve gegevens van het LDCR geanalyseerd. De informatie omvat alle civiele procedures in Nederland over de periode 2001 - 2021.

Eén van deze onderzoeken heeft als werkhypothese dat rechtspersonen die regelmatig eiser of gedaagde zijn (ook wel aangeduid als “repeat players”), specifieke kennis kunnen opbouwen, waardoor ze de uitkomsten van gerechtelijke procedures in hun voordeel kunnen beïnvloeden. Eén van de uitkomstmaten van een gerechtelijke procedure is daarbij de duur vanaf de start van de procedure tot aan de gerechtelijke uitspraak. “Repeat players” als eisende partij hebben relatief kortere gerechtelijke procedures, terwijl ze als gedaagde partij juist relatief langere procedures hebben.

Een andere maatstaf zou de uitspraak van de rechter kunnen zijn: de uitspraak is in het voordeel van de gedaagde dan wel ten gunste van de verweerder. Hierbij doet zich het probleem voor dat deze informatie niet in administratieve bestanden ruim beschikbaar is. Een gedegen dossieronderzoek is tijdrovend, zodat slechts een beperkt gegevensbestand kan worden opgebouwd. Dit staat niet in verhouding tot de omvangrijke administratieve data van het LDCR.

Doel van de beoogde analyse is om via een statistische text mining analyse uitspraken van rechters te analyseren. Aan de hand van teksten van uitspraken van rechters worden uitkomstmaten verkregen.. Het gaat daarbij om de volgende kenmerken (dit is in dit stadium niet uitputtend):

- Dossiernummer
- Namen van de eisende en gedaagde partijen (alleen bij rechtspersonen).
- Kenmerken partij (wel of niet natuurlijk persoon)
- Gemachtigde/ geen/ advocaat/ incassobureau/ deurwaarder
- Financieel belang
- Toegewezen bedrag
- Uitspraak van de rechter: de mate waarin de partij die zaak aanspant wordt door de rechter in het gelijk gesteld.

Vanzelfsprekend dient de te verzamelen informatie AVG-proof te zijn.

Voor zover we kunnen nagaan zijn teksten van de uitspraken in drie vormen beschikbaar.

1. Publicaties van gerechtelijke uitspraak. Dit is echter een beperkte, selectieve groep van uitspraken
2. Gerechtelijke uitspraken die door rechtbanken beschikbaar zijn gesteld in een zgn. e-archief.
3. Een uitgebreide verzameling van dossiers die door de rechtbank zelf digitaal wordt gearhiveerd.

Geen van deze mogelijke informatiebronnen is volledig met betrekking tot de gerapporteerde rechtszaken. Sommige rechtszaken worden niet opgenomen in deze archieven. Het is een selectieve groep, waarbij de rechtbanken beslissen welke teksten beschikbaar te stellen. Bijkomend probleem is dat de gegevensbronnen 2 en 3 niet voor onderzoekers openbaar beschikbaar zijn. Onze indruk is

echter dat het e-archief van deze drie opties het meest veelbelovend is. Op dit moment zijn we vooral op zoek naar de mate van vernieuwing van deze onderzoeksvraag (is er al eerder onderzoek gedaan) en naar de haalbaarheid.

We hebben de volgende twee vragen:

- 1) Zijn uitspraken van de rechter via het e-archief eerder ontsloten voor wetenschappelijk onderzoek?
- 2) Heeft de Rvdr ervaring met het e-archief? Is er inzicht in de representativiteit van de zaken?

Graag zouden we hierover een gesprek met de Rdvr willen aangaan.