Handreiking voor gedragsregels autonome systemen

ECP NL
Platform voor eNederland

Handreiking voor gedragsregels autonome systemen

Juridische aandachtspunten bij de bouw en het gebruik van autonome systemen

Een productie van:

Colofon

Dit is een uitgave van ECP.NL, Platform voor eNederland.

Teksten
mr. Bart W. Schermer

Ontwerp omslag en binnenwerk:
ECP.NL / Efficiënta Offsetdrukkerij BV

Druk
Efficiënta Offsetdrukkerij BV

ISBN
ISBN-10: 90-76957-20-7
ISBN-13: 978-90-76957-20-3

Voorwoord

Kunstmatige intelligentie is een veelbelovende technologie die een belangrijke invloed kan hebben op onze economie en maatschappij. Op steeds meer plaatsen zien wij dat met behulp van kunstmatige intelligentie systemen kunnen handelen zonder directe tussenkomt van een mens. Dergelijke 'autonome systemen' dragen bij aan onze welvaart, veiligheid en kwaliteit van leven.

ECP.NL, Platform voor eNederland, is van mening dat het van groot belang is dat de brede introductie van autonome systemen in Nederland op een gecoördineerde en maatschappelijk verantwoorde manier plaatsvindt. Alleen zo kan Nederland maximaal profiteren van de voordelen van autonome systemen.

Het is op dit moment nog onduidelijk hoe autonome systemen zich in nabije de toekomst zullen gaan ontwikkelen en welke plaats zij in gaan nemen in onze maatschappij. Het lijkt daarom voorbarig om nu reeds wetgeving op te stellen aangaande het gebruik van autonome systemen. Toch is enige coördinatie in dit stadium van de ontwikkeling wel wenselijk. Door een aantal juridische aandachtspunten aan te stippen en bijbehorende gedragsregels te formuleren kunnen bouwers en gebruikers van autonome systemen juridische vragen rondom het gebruik van autonome systemen vermijden.

ECP.NL heeft daarom besloten om een handreiking te doen voor het opstellen van dergelijke gedragsregels. ECP.NL hoopt hiermee bij te dragen aan een goed klimaat voor het gebruik van autonome systemen en het vertrouwen in deze technologie te stimuleren.

mr. A.J.M. van Bellen
Directeur ECP.NL, Platform voor eNederland

Inhoudsopgave

1 Inleiding

De *Handreiking gedragsregels autonome systemen* is een zelfregulerend initiatief dat is opgestart door ECP.NL, Platform voor eNederland, in samenwerking met marktpartijen, beleidsmakers en wetenschappers. Doel van dit initiatief is te komen tot een maatschappelijk verantwoord gebruik van autonome systemen en het vertrouwen in deze veelbelovende technologie te vergroten.

In 2004 is door ECP.NL een eerste stap gezet tot het verhelderen van het juridisch kader voor autonome systemen door de juridische aspecten van autonome systemen in kaart te brengen. Uit de door ECP.NL verrichte studie bleek dat er in het kader van autonome systemen een aantal juridische vraagstukken zijn die aan de orde kunnen komen, nu of in de toekomst. Verder bleek uit de studie dat het op dit moment invoeren van specifieke wetgeving voor autonome systemen vooralsnog niet wenselijk of noodzakelijk is.

Omdat wetgeving in dit stadium van de ontwikkeling niet het juiste instrument is om het gebruik van autonome systemen te reguleren, lijkt zelfregulering een waardevol alternatief om het gebruik van autonome systemen in goede banen te leiden en het vertrouwen erin te vergroten. Om deze reden heeft ECP.NL een aantal juridische aandachtspunten geïdentificeerd die in het kader van autonome systemen van belang zijn of van belang gaan worden in de (nabije) toekomst. Wanneer aanbieders en gebruikers rekening houden met deze aandachtspunten vermindert de kans op juridische onzekerheden bij het gebruik van autonome systemen.

Belangrijk:
In deze rapportage wordt een bepaalde basiskennis over de technische aspecten van kunstmatige intelligentie en autonome systemen verondersteld. Het gaat dan met name over de verschillende soorten autonome systemen en hun toepassingen. Het gaat de scope van de gedragsregels te buiten om verschillende autonome systemen uitgebreid te bespreken. Voor meer informatie over autonome systemen kunt u de ECP.NL rapportage *Juridische Aspecten van Autonome Systemen* (2005) raadplegen. U kunt deze publicatie gratis downloaden van de site van ECP.NL (www.ecp.nl)

1.1 Vertrouwen is de sleutel

Kunstmatige-intelligentie-toepassingen worden reeds op grote schaal in onze maatschappij gebruikt en de verwachting is dat steeds meer en geavanceerdere toepassingen het levenslicht zullen zien in de komende jaren. In de nabije toekomst zal kunstmatige-intelligentie technologie dan ook een substantiële invloed hebben op de Nederlandse economie en maatschappij. Bij de toepassing van nagenoeg alle technologieën die grote gevolgen hebben voor de maatschappij ontstaan er vragen omtrent mogelijke risico's en of negatieve effecten. Autonome systemen vormen hierop geen uitzondering. Acceptatie van een nieuwe technologie valt en staat met het vertrouwen in de technologie. Het is daarom zaak angst en onzekerheid weg te nemen.

1.2 Het stimuleren van vertouwen

Vertrouwen is een abstract begrip dat zich moeilijk laat definiëren. Een definitie van vertrouwen zou kunnen zijn: *het geloof van partijen binnen economische en sociale relaties dat de wederpartij de beste bedoelingen jegens hen heeft, eerlijk is, en in voldoende mate competent is om invulling te geven aan de relatie en de doelen die daarbinnen worden nagestreefd.*

Vertrouwen strekt zich ook uit tot technologie. In zijn algemeenheid kan gesteld worden dat een technologie enkel gebruikt wordt wanneer gebruikers een voldoende mate van zekerheid hebben omtrent de betrouwbaarheid ervan. Wanneer we kijken naar IT-systemen zien we dat met name de volgende elementen van belang zijn bij het stimuleren van vertrouwen:

Integriteit
Integriteit wordt in de informatiebeveiliging veelal gelijk gesteld aan het begrip betrouwbaarheid. Integriteit heeft betrekking op de correcte werking van het systeem en de veiligheid van de daarin opgeslagen informatie. In het kader van autonome systemen kunnen wij een onderscheid maken tussen de betrouwbaarheid/integriteit van het autonoom systeem zelf (het redeneersysteem en de werking daarvan) en de in het autonoom systeem opgeslagen informatie.

Identificatie/authenticatie/autorisatie
Identificatie- en authenticatiemethoden worden gebruikt om de identiteit van een entiteit (bijvoorbeeld een ander autonoom systeem of van een gebruiker) vast te stellen. Identificatie kan nodig zijn voor de technische werking van het systeem, maar in het kader van vertrouwen is identificatie met name van belang voor het toewijzen van rechten, plichten en verantwoordelijkheden.

Vertrouwelijkheid
Met vertrouwelijkheid wordt bedoeld dat gegevens opgeslagen in een systeem alleen te benaderen zijn door iemand die gerechtigd is deze te benaderen.

Beschikbaarheid
Beschikbaarheid heeft betrekking op de toegankelijkheid van een systeem voor een (geautoriseerde) gebruiker. Beschikbaarheid wordt gedefinieerd als het percentage tijd dat een systeem storingsvrij is.

Continuïteit
Continuïteit hangt nauw samen met beschikbaarheid en heeft

betrekking op de tijd dat een systeem onafgebroken beschikbaar is.

Transparantie
Transparantie heeft betrekking op de inzichtelijkheid van de handelingen die het systeem uitvoert. Transparantie bevordert de voorspelbaarheid van het systeem.

Toetsbaarheid
Een laatste element dat bij kan dragen aan het vertrouwen in systemen is de toetsbaarheid van de handelingen. Een dergelijke toetsing vind meestal achteraf plaats en vaak slechts wanneer daar een noodzaak toe bestaat.

1.3 Het creëren van vertrouwen

Er bestaan verschillende instrumenten om het vertrouwen in een bepaalde technologie te vergroten bij zowel bedrijven als burgers. Deze instrumenten dienen complementair aangewend te worden. Om het vertrouwen in autonome systemen te vergroten moeten bedrijven en burgers allereerst weten wat autonome systemen zijn en wat zij betekenen voor de samenleving. Hiervoor is heldere, onpartijdige voorlichting noodzakelijk. Daarnaast moeten bedrijven en burgers voldoende juridische waarborgen worden geboden bij het gebruik van en de omgang met autonome systemen. Een derde instrument om het vertrouwen in autonome systemen te vergroten is een adequaat niveau van beveiliging.

Voorlichting
Bedrijven en burgers hebben op dit moment onvoldoende inzicht in de (technische) werking van autonome systemen waardoor er een incompleet of zelfs verkeerd beeld ontstaat over de technologie en haar toepassingen hetgeen negatieve consequenties kan hebben. Goede voorlichting vormt daarom een essentieel onderdeel bij het stimuleren van vertrouwen in autonome systemen.

Regulering
Een gedegen regelgevend kader neemt onzekerheden weg.

Daarom speelt regulering een belangrijke rol bij het creëren van vertrouwen in autonome systemen. Zoals reeds aangegeven is het op dit moment onverstandig (en waarschijnlijk ook onmogelijk) om het gebruik van autonome systemen via wetgeving te reguleren. Op dit moment is het identificeren van juridische aandachtspunten en het geven van handvatten voor bouw en gebruik in de vorm van gedragsregels het hoogst haalbare.

Uiteindelijk kan zelfregulering zo nodig in samenhang met wetgeving het gebruik van autonome systemen reguleren. Partijen die autonome systemen gebruiken en zich binden aan een gedragscode, kunnen door de wederpartij op de naleving daarvan worden aangesproken. Niet uitgesloten is dat de normstelling die van een gedragscode uitgaat een zodanig algemene aanvaarding krijgt, dat deze ook in rechte afdwingbaar wordt.

Beveiliging
Beveiliging speelt in zijn algemeenheid een belangrijke rol bij informatie- en communicatietechnologie. Dit geldt in het bijzonder voor autonome systemen. De correcte werking van autonome systemen is namelijk in belangrijke mate afhankelijk van de integriteit van het autonome systeem en het platform waar deze op draait.

1.4 Het gebruik van de aandachtspunten

Voor een verantwoord gebruik van autonome systemen is het noodzakelijk dat rekening wordt gehouden met juridische aandachtspunten. Deze aandachtspunten dienen meegenomen te worden bij zowel de bouw als het gebruik van autonome systemen.

1.4.1 Functie van de handreiking
Het doel van de *Handreiking gedragsregels autonome systemen* is het onder de aandacht brengen van de mogelijke juridische aandachtspunten die spelen bij het gebruik van autonome systemen. Zowel aanbieders als gebruikers van autonome systemen dienen rekening te houden met de mogelijke juridische consequenties die het gebruik van autonome systemen met zich mee kan brengen. Door het vroegtijdig signaleren en benoemen

van juridische vraagstukken die kunnen spelen bij het gebruik van autonome systemen wordt onzekerheid weggenomen.

Het doel van deze handreiking is het bewust maken van aanbieders en gebruikers van autonome systemen van de juridische aspecten van autonome systemen en zorgen dat bij de bouw en het gebruik van autonome systemen deze juridische aspecten in ogenschouw worden genomen. De handreiking kan dan ook op verschillende manieren worden gebruikt.

Deze handreiking dient allereerst als een *richtlijn voor bouwers van autonome systemen*. Een belangrijk gegeven is dat de architectuur van een autonoom systeem de mogelijkheden ervan definieert. Wanneer bij de bouw van een autonoom systeem geen rekening wordt gehouden met de in deze handreiking genoemde aandachtspunten dan is het risico op juridische kwesties groter. Wanneer echter al bij de bouw rekening wordt gehouden met de genoemde aandachtspunten, dan kunnen risico's en juridische onduidelijkheden tot een minimum beperkt worden.

De handreiking kan verder dienen als *checklist voor potentiële gebruikers*. Zij kunnen aan de hand van de in de handreiking genoemde aandachtspunten inventariseren of de voor hen gebouwde autonome systemen wel op een verantwoorde manier kunnen worden ingezet of dat additionele waarborgen noodzakelijk zijn.

Uiteindelijk kunnen in een later stadium, wanneer de markt wat volwassener is en dit noodzakelijk blijkt, de aandachtspunten en de daaraan gekoppelde gedragsregels als basis dienen voor een *gedragscode*. Deze mogelijkheid wordt in hoofdstuk 3 nader toegelicht.

Tot slot is een vanuit technisch oogpunt interessante toepassing het verwerken van de aandachtspunten uit deze handreiking in de *'algemene voorwaarden'* van agentplatformen.[1] Software

[1] Voor meer informatie over agentplatformen zie de ECP.NL rapportage Juridische Aspecten van Autonome Systemen

agenten die gebruik maken van het agentplatform moeten zich automatisch confirmeren aan de eisen die in deze algemene voorwaarden gesteld zijn. Wanneer de gedragsregels uit deze handreiking worden verwerkt in de algemene voorwaarden van agentplatformen, dan voldoen de op het agentplatform draaiende agenten automatisch aan de gedragsregels.

1.4.2 Toepasbaarheid

De toepasbaarheid van de regels uit deze handreiking hangt nauw samen met de ontwikkeling van autonome systemen. Zoals beschreven in het rapport *Juridische Aspecten van Autonome Systemen* bevindt de ontwikkeling van autonome systemen zich nog in een relatief vroege fase. De autonome systemen die nu worden ingezet hebben nog niet een bijzonder hoge graad van autonomie (waardoor ze zelfstandig nog weinig juridisch relevante handelingen kunnen uitvoeren) en opereren voornamelijk in gesloten omgevingen die goed te controleren zijn. Voor de korte en middellange termijn zullen de juridische vraagstukken waarop een deel van deze gedragsregels betrekking hebben daarom in de praktijk nog geen voorname rol spelen. Wanneer autonome systemen op de lange termijn echter meer autonomie en verantwoordelijkheid krijgen en hun taken in semi-open tot volledig open omgevingen gaan uitvoeren, worden de gedragsregels steeds relevanter.

1.5 Opbouw

Aan de in hoofdstuk 2 genoemde aandachtspunten zijn bijbehorende gedragsregels gekoppeld. Deze gedragsregels moeten niet als bindend worden gezien, maar veeleer als een leidraad voor het verantwoord gebruik van autonome systemen. Als zodanig vormen de aandachtspunten uit deze handreiking in hun onderlinge samenhang en in hun formulering een set van mogelijke gedragsregels, waaraan een partij die autonome systemen gaat bouwen en/of gebruiken zich tenminste zou moeten binden om bij wederpartijen voldoende vertrouwen te wekken.

2 Juridische aandachtspunten autonome systemen

In dit hoofdstuk worden juridische aandachtspunten bij de bouw en het gebruik van autonome systemen uiteengezet. In kaders zijn (waar mogelijk) gedragsregels opgesteld die als basis kunnen dienen voor een gedragscode.

2.1 Algemeen

Voordat aandacht zal worden besteed aan de specifieke juridische aandachtspunten die spelen bij het gebruik van autonome systemen dient nog een aantal algemene juridische aandachtspunten te worden geadresseerd.

2.1.1 De rol van partijen

Het eerste algemene aandachtspunt betreft de rollen die partijen vervullen bij het gebruik van autonome systemen. Het zijn deze rollen die de juridische verhouding tussen partijen bepalen. Verschillende rollen die we kunnen onderscheiden zijn: de bouwer van een autonoom systeem, de aanbieder van een autonoom systeem, de eigenaar van het platform waarop een autonoom systeem draait en de (eind)gebruiker van een autonoom systeem. Bij de eindgebruiker is het voorts van belang onderscheid te maken tussen professionele gebruikers (bedrijven) en particuliere gebruikers (consumenten). Deze laatste groep geniet vanuit het recht een hogere mate van bescherming dan de eerste categorie.

2.1.2 Specifieke regulering

De in deze handreiking aangedragen aandachtspunten hebben betrekking op autonome systemen in het algemeen. Er wordt in deze handreiking geen rekening gehouden met specifieke juridische eisen die binnen een bepaalde branche/sector aanvullend aan autonome systemen kunnen worden gesteld. Naast branchespecifieke eisen kunnen aanvullende juridische eisen ook voortvloeien uit de verhouding tussen partijen. Zo bestaan in de verhouding bedrijf-consument (b2c) andere juridische eisen dan

in de verhouding bedrijf-bedrijf (b2b). Bij het gebruik van autonome systemen moet daarom altijd rekening worden gehouden met specifieke, aanvullende (zelf)regulering.

2.2 Identificatie

In het maatschappelijk verkeer is identificatie belangrijk. Omdat steeds meer communicatie over afstand plaatsvindt kan het onduidelijk zijn met wie partijen precies te maken hebben. Met name in het handelsverkeer is het vaak wenselijk om de wederpartij te kunnen identificeren. Dit wil overigens niet zeggen dat identificatie ten alle tijden noodzakelijk of verplicht is. Wanneer voldoende waarborgen voor het vaststellen van de verantwoordelijkheid voor een bepaalde handeling aanwezig zijn, kunnen transacties ook anoniem (of pseudoniem) plaatsvinden.

Bij de identificatie van autonome systemen moeten we onderscheid maken tussen twee verschillende vormen van identificatie: 1) de identificatie van het autonoom systeem zelf, en 2) de identificatie van de achterliggende gebruiker van het autonome systeem. Vanuit technisch oogpunt is de identificatie van een autonoom systeem zelf vaak een vereiste. Identificatie van een autonoom systeem vindt plaats aan de hand van een unieke identifier zoals een identificatienummer. De identiteit van een autonoom systeem dient bestendig te zijn, met andere woorden, de identiteit mag niet te wijzigen zijn en blijft gedurende de levenscyclus van een autonoom systeem gehandhaafd.

In veel gevallen is het ook noodzakelijk om vast te kunnen stellen wie de aanbieder of eindgebruiker is van een autonoom systeem. Het is immers niet mogelijk om computerprogramma's of robots die geen rechtssubject zijn verantwoordelijk te houden voor een bepaalde handeling. Er moet daarom een koppeling zijn tussen het autonoom systeem en de achterliggende aanbieder en/of gebruiker, opdat aan deze de verantwoordelijkheid voor het handelen van het autonoom systeem kan worden toegerekend. Een dergelijke koppeling kan tot stand worden gebracht met behulp van onder andere cryptografische technieken zoals de elektronische handtekening.

In een aantal gevallen bestaat er ook een wettelijke plicht tot identificatie. In het civiele recht is dit bijvoorbeeld het geval wanneer een overeenkomst alleen in schriftelijke vorm geldig of onaantastbaar tot stand kan komen. Volgens de huidige wet is het niet langer noodzakelijk dat een overeenkomst met een schriftelijkheidsvereiste daadwerkelijk 'op papier' wordt vastgelegd, met andere woorden, een dergelijke overeenkomst kan ook langs elektronische weg tot stand komen. Maar in dat geval moet de identiteit van partijen wel met voldoende zekerheid vastgesteld kunnen worden (6:227a BW).

Een plicht tot identificatie kan ook voor bedrijven gelden. Zo moeten bedrijven in het kader van de Wet Elektronische Handel (3:15d lid 1 sub a BW) en de Wet Koop op Afstand (7:46c lid 1 sub a BW) via hun website bepaalde informatie over zichzelf aan afnemers en consumenten verstrekken.

Buiten het civiele recht zijn er uiteraard ook situaties waarin men zich moet identificeren. Met name in relatie tot de overheid is identificatie vaak noodzakelijk. Voor het aanvragen van een vergunning bij de gemeente is bijvoorbeeld een geldig legitimatiebewijs nodig.

Directe identificeerbaarheid (een koppeling tussen een autonoom systeem en een achterliggende persoon of organisatie) staat wel de anonimiteit van de gebruiker in de weg. Het kan zijn dat een partij liever anoniem wenst te blijven. Dat is legitiem, want het Nederlandse recht verplicht een partij niet zich bij iedere transactie te identificeren. Waar passend dient de mogelijkheid tot het anoniem gebruik van autonome systemen te worden geboden. Hierbij moet ook rekening worden gehouden met branchespecifieke eisen.

Van anonimiteit kan pseudonimiteit worden onderscheiden. In de Wet Elektronische Handtekeningen wordt de mogelijkheid geopend om in een digitaal certificaat niét de identiteit te vermelden van de persoon aan wie het certificaat toebehoort. We spreken dan van pseudonieme digitale certificaten. Omdat het certificaat en daarmee het autonoom systeem niet rechtstreeks

aan een persoon gekoppeld kan worden, blijft de anonimiteit en de privacy van de gebruiker beschermd.

Een laatste aandachtspunt dat speelt in het kader van identificatie is de omgang met persoonsgegevens. Uitgangspunt bij het gebruik van autonome systemen dient te zijn dat wanneer er geen noodzaak of (verifieerbare) wettelijke verplichting toe bestaat, autonome systemen geen persoonsgegevens mogen verzamelen of prijsgeven.

Partijen dragen zorg voor de identificeerbaarheid van autonome systemen.

Waar noodzakelijk dient de aanbieder/en of eindgebruiker van een autonoom systeem identificeerbaar te zijn. Deze identiteit moet aan het autonoom systeem gekoppeld kunnen worden.

Waar mogelijk moet ruimte worden geboden voor anoniem of pseudoniem gebruik van autonome systemen.

Autonome systemen dienen zorgvuldig en terughoudend te zijn bij de verwerking van persoonsgegevens.

2.3 Transparantie

Autonome systemen voeren hun taken (deels) zonder directe tussenkomst van mensen uit. Hoewel dit ontegenzeggelijk grote voordelen biedt, heeft het vanuit juridisch oogpunt het nadeel dat het onduidelijk kan zijn hoe een autonoom systeem tot een bepaalde handeling is gekomen.

Inzicht in de werking van een systeem is noodzakelijk voor het opbouwen van vertrouwen in dit systeem. Hoewel een eindgebruiker geen volledig inzicht hoeft te hebben in de wijze waarop een autonoom systeem werkt, moet het gedrag van een autonoom systeem wel tot op zekere hoogte verklaarbaar en

voorspelbaar zijn. Met andere woorden, een gebruiker moet zich een redelijk beeld kunnen vormen van hoe het autonoom systeem omgaat met de input die het krijgt. Gebeurt dit niet dan kan het voorkomen dat het autonoom systeem onverwachte en wellicht ongewenste handelingen verricht voor de gebruiker. Wanneer de bouwer, aanbieder en de eindgebruiker een overeenstemmend beeld hebben van de globale werking, mogelijkheden en beperkingen van het gebruikte autonoom systeem (een functioneel model), dan kunnen dergelijke situaties tot een minimum worden beperkt. Bij voldoende duidelijkheid over hoe een autonoom systeem tot een bepaalde handeling komt, kan een beter antwoord worden geformuleerd op de vraag wie de verantwoordelijkheid danwel aansprakelijkheid draagt voor het handelen van een autonoom systeem.

Wanneer een eindgebruiker globaal weet hoe een autonoom systeem op basis van door de hem of haar ingevoerde criteria tot een handeling komt, dan kan de gebruiker zich ook een redelijk beeld vormen van de consequenties die het gebruik van een autonoom systeem met zich meebrengt. Een voorbeeld kan dit illustreren.

Stel, een gebruiker wil een autonoom systeem een boek laten kopen. Op basis van de verwachtingen die de gebruiker heeft omtrent de werking van het autonoom systeem en de mogelijkheden die het systeem biedt voor input zal de gebruiker bepaalde criteria invoeren (genre, schrijver, maximale prijs). De gebruiker verwacht vervolgens een resultaat dat in lijn is met de door hem opgegeven criteria en de ideeën die hij heeft rondom de werking van het systeem. Wanneer er vervolgens een voor de gebruiker totaal onverwachte en/of onvoorspelbare handeling tot stand wordt gebracht (de maximale prijs wordt bijvoorbeeld overschreden), dan is het de vraag of de verantwoordelijkheid voor dit handelen bij de gebruiker moet liggen. Het kan namelijk zijn dat er een fout zit in het autonoom systeem of het platform waar deze op draait, waardoor de onvoorspelbare handeling tot stand is gekomen. Andersom geldt hetzelfde: wanneer op basis van de input van de gebruiker een handeling tot stand wordt gebracht met een voor de gebruiker voorspelbare uit-

komst, dan ligt het juist weer voor de hand de aansprakelijkheid voor deze handeling bij de eindgebruiker neer te leggen.

Het is daarom bovenal zaak dat alle betrokken partijen een duidelijk en overeenstemmend beeld hebben over de werking en de mogelijkheden van een autonoom systeem.

Om bij te dragen aan de transparantie van autonome systemen kunnen bouwers en aanbieders van autonome systemen in overweging nemen om in de door hen gebouwde of aangeboden autonome systemen een voor gebruikers toegankelijke 'geschiedenis' in te bouwen. Een dergelijke functie kan de gebruiker inzicht geven in de handelingen die het autonome systeem heeft uitgevoerd. Dit draagt daarmee bij aan de transparantie van het autonome systeem.

Bouwers en aanbieders van autonome systemen geven indien mogelijk duidelijk inzicht in de werking van de door hen gebouwde of aangeboden autonome systemen.

Partijen moeten er rekening mee houden dat zij een overeenstemmend beeld hebben van de mogelijkheden en onmogelijkheden van het gebruikte autonoom systeem.

Bouwers en aanbieders verzorgen voor de gebruiker indien mogelijk inzicht in de handelingsgeschiedenis van de door hen gebouwde of aangeboden autonome systemen.

2.4 Integriteit

De integriteit van een autonoom systeem, de daarin opgeslagen informatie en de overdracht daarvan, vormen noodzakelijke voorwaarden voor een gerechtvaardigd vertrouwen in autonome systemen. Wanneer de integriteit van autonome systemen, de daarin opgeslagen informatie of de overdracht van informatie ontregeld worden, dan kan de correcte werking van een autonoom systeem niet meer gegarandeerd worden. Daarom dienen

partijen (bouwers, aanbieders en gebruikers) zorg te dragen voor de veiligheid en integriteit van hun autonome systemen. Indien een autonoom systeem een bepaald platform nodig heeft om op te draaien (bijvoorbeeld een agentplatform) dan dient de eigenaar of de beheerder van dit platform ook de integriteit van het platform te waarborgen.

Wanneer wordt gesproken over integriteit dan moet een onderscheid worden gemaakt tussen de integriteit van het autonoom systeem zelf (met name het redeneersysteem dat het autonoom systeem hanteert) en dat van de data die het autonoom systeem gebruikt. Dit onderscheid is met name van belang bij het bepalen van de aansprakelijkheid voor eventuele schade die ontstaat door ongewenst handelen van een autonoom systeem. Wanneer de ongewenste handeling bijvoorbeeld is te wijten aan een fout in het redeneersysteem, dan ligt het niet voor de hand om de eindgebruiker aansprakelijk te stellen wanneer deze fout niet het gevolg is van zijn handelen. Wanneer de ongewenste handeling het gevolg is van corrupte data, dan moet de aansprakelijkheid eerder worden gezocht bij degene die de verantwoordelijkheid voor de data draagt.

Een ander punt waar partijen rekening mee moeten houden bij het gebruik van autonome systemen betreft het opsporen en afhandelen van incidenten. Wanneer de integriteit van een autonoom systeem of platform wordt geschonden, dan moeten passende maatregelen voorhanden zijn om deze schendingen te detecteren en te verhelpen.

Partijen dragen zorg voor de integriteit van het autonoom systeem, de daarin opgeslagen informatie en de overdracht daarvan.

Partijen nemen passende maatregelen om schendingen van de integriteit van een autonoom systeem te kunnen detecteren en maken afspraken over de acties die ondernomen dienen te worden wanneer een schending geconstateerd is.

De eigenaar en/of beheerder van een platform waarop autonome systemen draaien, draagt zorg voor de integriteit van dit platform.

2.5 Vertrouwelijkheid

Het is mogelijk dat autonome systemen vertrouwelijke informatie bevatten zoals persoonsgegevens of bedrijfsgeheimen. Het is in dergelijke gevallen noodzakelijk dat de aanbieders en gebruikers van autonome systemen afdoende veiligheidsmaatregelen nemen om de vertrouwelijkheid van deze informatie te garanderen. Informatiebeveiliging en passende identificatie, authenticatie- en autorisatiemechanismen zijn hiervoor noodzakelijk.

De vertrouwelijkheid van de in het autonoom systeem opgeslagen informatie, en in sommige gevallen het handelen van het autonoom systeem zelf, dient gewaarborgd te zijn. Het onderwerp vertrouwelijkheid hangt daarom nauw samen met het onderwerp integriteit. Om de vertrouwelijkheid te kunnen waarborgen is de veiligheid en de integriteit van het autonome systeem (en het platform waar deze eventueel op draait) een belangrijke randvoorwaarde. Partijen (bouwers, aanbieders en gebruikers) dienen zorg te dragen voor adequate beveiliging om de vertrouwelijkheid te kunnen waarborgen.

Omdat een autonoom systeem dat draait op een platform van een derde zich in principe buiten de controle van zijn eigenaar

of gebruiker bevindt, is het van belang dat eigenaar of gebruiker afdoende waarborgen worden geboden waaraan het vertrouwen kan worden ontleend dat de vertrouwelijkheid gegarandeerd is. Deze eis heeft primair betrekking op het gebruik van autonome systemen in gesloten omgevingen. Momenteel bevinden wij ons in een fase van gesloten omgevingen voor autonome systemen (voorbeelden zijn robots in een autofabriek en software agenten in gesloten agent-systemen). De verwachting is dat autonome systemen in de toekomst steeds vaker zullen opereren in open omgevingen zoals het internet.[2] In dergelijke open omgevingen zal de vertrouwelijkheid van informatie opgeslagen in autonome systemen moeilijker te handhaven en te verifiëren zijn.

Bij de vertrouwelijkheid van autonome systemen speelt het opsporen en afhandelen van incidenten wederom een belangrijke rol. Wanneer de vertrouwelijkheid wordt geschonden, dan moeten passende maatregelen voorhanden zijn om deze schendingen te detecteren en te verhelpen.

Partijen dragen zorg voor de vertrouwelijkheid van de in door hen gebouwde of gebruikte autonome systemen opgeslagen informatie.

Partijen nemen passende maatregelen om onrechtmatige openbaringen van vertrouwelijke informatie te kunnen detecteren en maken afspraken over de acties die ondernomen dienen te worden wanneer een onrechtmatige openbaring geconstateerd is.

2.6 Beschikbaarheid en continuïteit

De beschikbaarheid en continuïteit van autonome systemen vormen een belangrijk onderdeel van de mate van vertrouwen die

[2] Zie het rapport *Juridische Aspecten van Autonome Systemen* (ECP.NL 2005) voor een uitgebreidere beschrijving van deze ontwikkelingen.

het publiek heeft in deze systemen. Gebruikers en derden moeten erop kunnen vertrouwen dat de autonome systemen die zij gebruiken of waar zij mee in aanraking komen blijven functioneren. Het is van belang dat aanbieders, maar ook gebruikers van autonome systemen zorg dragen voor de beschikbaarheid en continuïteit van hun autonome systemen. Het gaat met name om de beschikbaarheid en de continuïteit van het redeneersysteem, de in het autonoom systeem opgeslagen informatie en de identiteit van het autonoom systeem.

De beschikbaarheid en continuïteit kan door verschillende oorzaken in het gedrang komen, bijvoorbeeld door een programmeerfout in het autonoom systeem, of een fout in het platform waarop deze draait. Ongeacht de oorzaak van een eventueel probleem dat de werking van een autonoom systeem kan hinderen, is het zaak dat dergelijke fouten niet tot gevolg hebben dat het autonoom systeem of het platform waar deze op draait aangetast wordt of volledig verloren gaat.

Partijen dragen zorg voor de continuïteit en beschikbaarheid van de door hen aangeboden of gebruikte autonome systemen.

Partijen nemen passende maatregelen om de gevolgen van fouten in een autonoom systeem of het platform waarop deze draait, welke kunnen leiden tot aantasting van het systeem of het platform, tot een minimum te beperken.

2.7 Toetsbaarheid en traceerbaarheid

Mocht het handelen van een autonoom systeem leiden tot een juridische kwestie, dan is het van belang te kunnen achterhalen hoe de handelingen tot stand zijn gekomen en wie bij de handeling betrokken zijn geweest. Het handelen van een autonoom systeem moet daarom toetsbaar en traceerbaar zijn. Het bijhouden van logs lijkt hiervoor de meest geschikte methode.

Wanneer gebruik wordt gemaakt van logging, dan is het zaak dat de integriteit van deze logs gegarandeerd is en dat de inhoud niet door de betrokken partijen betwist wordt.

Een aandachtspunt dat wel speelt rondom het bijhouden van logs is de vertrouwelijkheid van deze logs. Niet iedereen mag zomaar kennis nemen van deze logs. Om deze reden is het zaak om ook de gegenereerde logs te beveiligen, zodat hun vertrouwelijkheid gegarandeerd is.

Partijen dragen zorg voor een de traceerbaarheid en toetsbaarheid van de door een autonoom systeem verrichte handeling.

Partijen dragen zorg voor de integriteit van de door autonome systemen gegenereerde logs.

Partijen dragen zorg voor de vertrouwelijkheid van gegenereerde logs.

2.8 Intellectuele eigendom

Bij het gebruik van autonome systemen dient rekening te worden gehouden met vraagstukken omtrent het intellectuele eigendom. Het intellectuele eigendom op een autonoom systeem ligt meestal bij de bouwer van een autonoom systeem, maar dit is sterk afhankelijk van de juridische verhouding tussen partijen en de gemaakte afspraken. De situatie is minder duidelijk wanneer een autonoom systeem nieuwe vaardigheden leert. Ligt het intellectuele eigendom van de nieuwe vaardigheden bij de oorspronkelijke bouwer of aanbieder van een autonoom systeem, of ligt het intellectuele eigendom in dit geval bij de gebruiker? Om dergelijke discussies omtrent het intellectuele eigendom te vermijden moeten dus duidelijke afspraken worden gemaakt tussen de bouwer, aanbieder en/of gebruiker van een autonoom systeem

Een tweede aandachtspunt betreft het verwerken van informatie welke is beschermd door het intellectueel eigendomsrecht. Het gaat dan met name om auteursrechten. Wanneer een autonoom systeem auteursrechtelijk beschermde werken zonder de toestemming van de rechthebbende verveelvoudigt of openbaar maakt, dan handelt het autonoom systeem in strijd met het auteursrecht. Het is aan de bouwers, aanbieders en eindgebruikers om te voorkomen dat een autonoom systeem inbreuk maakt op het auteursrecht. Maar wanneer auteursrechtelijk beschermde werken niet als zodanig door een autonoom systeem te herkennen zijn, rijst de vraag in hoeverre de auteursrechthebbenden een verantwoordelijkheid hebben om kenbaar te maken dat het gaat om auteursrechtelijk beschermde werken.

Een derde aandachtspunt betreft werken die worden voortgebracht door een autonoom systeem zélf. Het is zeer goed mogelijk dat autonome systemen zelf werken voortbrengen (bijvoorbeeld afbeeldingen, gedichten of verhalen). De vraag is dan wie rechthebbende is op deze werken. Er bestaat nog veel onduidelijkheid over het antwoord op deze vraag en het is daarom zinvol dat partijen daar onderling afspraken over maken.

Partijen maken afspraken omtrent het intellectuele eigendom bij het gebruik van autonome systemen.

Partijen dragen, waar mogelijk, zorg voor de correcte omgang met auteursrechtelijk beschermde werken van derden.

Partijen maken afspraken omtrent het intellectuele eigendom op door een autonoom systeem voortgebrachte werken.

3 Van gedragsregels naar gedragscode

Wanneer partijen zich expliciet willen conformeren aan de gedragsregels zoals vastgelegd in dit document dan kunnen partijen overwegen zich aan de gedragsregels te binden door middel van een gedragscode. In deze paragraaf wordt een aantal algemene aandachtspunten bij het opstellen en hanteren van de gedragscode uiteengezet.

3.1 Bekendmaking

Om vertrouwen te kunnen genereren dienen partijen bekend te maken dat zij zich zullen gedragen overeenkomstig de gedragsregels. Partijen kunnen dit doen door de gedragscode expliciet te onderschrijven.

3.2 Definities

Door gehanteerde begrippen die in de gedragscode meermalen voorkomen te definiëren bevorderen partijen de duidelijkheid en leesbaarheid van de gedragscode. Een belangrijk aandachtspunt in een gedragscode betreft de gehanteerde definities. Met name in het kader van autonome systemen, dat een veelvoud aan technologieën en toepassingsvormen omvat, is een heldere definiëring van begrippen relevant. Dit bevordert de duidelijkheid en leesbaarheid van de gedragscode en voorkomt interpretatieverschillen.

3.3 Reikwijdte

Partijen moeten duidelijk het werkingsgebied (juridische reikwijdte, verbindendheid bij gebruik) van hun gedragscode aangeven. Hierbij dient ook rekening te worden gehouden met de vraag aan wie de gedragscode gericht is (werknemers, consumenten, bezoekers et cetera)

3.4 Openheid en transparantie

Aanbieders en gebruikers van technologie van autonome systemen dienen een open en transparante houding richting derden aan te nemen. Een algemene regel die in een gedragscode vastgelegd kan worden is de verplichting om een open en transparante houding richting de buitenwereld aan te nemen met betrekking tot het gebruik van autonome systemen.

4 Samenvatting gedragsregels

Hieronder vindt u nogmaals alle gedragsregels opgesomd.

Identificatie
Partijen dragen zorg voor de identificeerbaarheid van autonome systemen.

Waar noodzakelijk dient de aanbieder/en of eindgebruiker van een autonoom systeem identificeerbaar te zijn. Deze identiteit moet aan het autonoom systeem gekoppeld kunnen worden.

Waar mogelijk moet ruimte worden geboden voor het anoniem of pseudoniem gebruik van autonome systemen.

Autonome systemen dienen zorgvuldig en terughoudend te zijn bij de verwerking van persoonsgegevens.

Transparantie

Bouwers en aanbieders van autonome systemen geven indien mogelijk duidelijk inzicht in de werking van de door hen gebouwde of aangeboden autonome systemen.

Partijen moeten er rekening mee houden dat zij een overeenstemmend beeld hebben van de mogelijkheden en onmogelijkheden van het gebruikte autonoom systeem.

Bouwers en aanbieders verzorgen voor de gebruiker indien mogelijk inzicht in de handelingsgeschiedenis van de door hen gebouwde of aangeboden autonome systemen.

Integriteit
Partijen dragen zorg voor de integriteit van het autonoom systeem, de daarin opgeslagen informatie en de overdracht daarvan.

Partijen nemen passende maatregelen om schendingen van de integriteit van een autonoom systeem te kunnen detecteren en maken afspraken over de acties die ondernemen dienen te worden wanneer een schending geconstateerd is.

De eigenaar en/of beheerder van een platform waarop autonome systemen draaien, draagt zorg voor de integriteit van dit platform.

Vertrouwelijkheid

Partijen dragen zorg voor de vertrouwelijkheid van de in door hen gebouwde of gebruikte autonome systemen opgeslagen informatie.

Partijen nemen passende maatregelen om onrechtmatige openbaringen van vertrouwelijke informatie te kunnen detecteren en maken afspraken over de acties die ondernomen dienen te worden wanneer een onrechtmatige openbaring geconstateerd is.

Continuïteit

Partijen dragen zorg voor de continuïteit van de door hen aangeboden of gebruikte autonome systemen.

Partijen nemen passende maatregelen om te voorkomen dat een fout in een autonoom systeem of het platform waarop deze draait, leidt tot het volledig verloren gaan van een autonoom systeem.

Toetsbaarheid en traceerbaarheid

Partijen dragen zorg voor een de traceerbaarheid en toetsbaarheid van de door een autonoom systeem verrichte handeling.

Partijen dragen zorg voor de integriteit van de door autonome systemen gegenereerde logs.

Partijen dragen zorg voor de vertrouwelijkheid van gegenereerde logs.

Intellectuele eigendom

Partijen maken afspraken omtrent het intellectuele eigendom bij het gebruik van autonome systemen.

Partijen dragen, waar mogelijk, zorg voor de correcte omgang met auteursrechtelijk beschermde werken van derden.

Partijen maken afspraken omtrent het intellectuele eigendom op door een autonoom systeem voortgebrachte werken.

5 Appendix: Deelnemers Werkgroep

De handreiking voor gedragsregels is samengesteld door de werkgroep Autonome Systemen van ECP.NL. De volgende personen hebben op persoonlijke titel zitting in de werkgroep:

C.	Stuurman	Van Doorne / Universiteit Tilburg (voorzitter)
J.J.F.M.	Borking	Borking Consultancy
F.M.T.	Brazier	Vrije Universiteit Amsterdam / Stichting NLnet
H.J.	Geurts	AKD Prinsen van Wijmen
N.	Hagemans	Ministerie van Justitie
H.J.	van den Herik	Universiteit Maastricht
F.A.M.	van der Klaauw-Koops	Universiteit Leiden
R.P.J.	Megens	Tryllian Solutions BV
A.	Oskamp	Vrije Universiteit Amsterdam
M.	den Uyl	Parabots / Sentient
J.P.G.M.	Verbeek	Universiteit Maastricht
M.B.	Voulon	Duthler Associates
N.J.E.	Wijngaards	Decis Labs
M.	Durinck	ECP.NL
B.W.	Schermer	ECP.NL (secretaris)