BUYER GROUP

AI BEELDHERKENNING MET DRONES

MARKTCONSULTATIE – RAPPORTAGE

JULI 2022

INHOUDSOPGAVE

1 INLEIDING

Kunstmatige intelligentie (AI) gebruiken om beeldmateriaal van drones te verwerken, biedt steeds meer mogelijkheden om maatschappelijke uitdagingen het hoofd te bieden. Slimme beeldherkenningstechnologie heeft potentie bij allerlei vraagstukken en het verbeteren van publieke taken. De Buyer Group AI beeldherkenning met drones beoogt betrokken organisaties en experts samen te brengen met als doel om kennis over AI beeldherkenning, zoals met drones, te vergroten en het biedt een gezamenlijk vertrekpunt om ambities bij toekomstige aanbestedingen te bepalen. In de Buyer Group bekijken overheden welke huidige en toekomstige wensen er spelen, brengen die behoeftes in kaart, met daarbij aandacht voor randvoorwaarden en mogelijkheden uit de praktijk.

Deze Buyer Group is een initiatief van de Nederlandse AI Coalitie in samenwerking met PIANOo Expertisecentrum Aanbesteden[1]. De groep bestaat uit een breed consortium van deelnemers waaronder; ProRail, Nederlandse Voedsel- en Warenautoriteit, Gemeente Amsterdam, Rijkswaterstaat, Politie, Het Waterschapshuis, en Waterschap Drents Overijsselse Delta.

Binnen de Buyer Group werken opdrachtgevers aan een gezamenlijke visie voor de markt. De deelnemers van de Buyer Group bepalen de koers, geven invulling aan de visie, en implementeren die visie en inkoopstrategie bij nieuwe projecten. Dit doen ze op basis van eigen use-cases en ervaringen die ze met elkaar uitwisselen.

1.1 Interactie met de markt

De Buyer Group versterkt publieke organisaties door gezamenlijk naar thema's te kijken op het gebied van AI beeldherkenning, zoals met drones, om daarmee een vertaling te maken naar een gezamenlijke marktvisie en strategie. Bij het ontwikkelen van de gezamenlijke visie betrekt de Buyer Group marktpartijen via marktconsultaties. Tijdens deze marktconsultaties wordt een dialoog aangegaan met marktpartijen. Voor de Buyer Group is deze consultatie een uitgelezen kans om vroegtijdig in contact te komen met belangstellenden voor mogelijke aanbesteding in de toekomst. De marktconsultatie richt zich daarbij op het verkrijgen van inzicht in de volgende thema gebieden:

- Cyber security / informatiebeveiliging
- Delen van data, AI en informatie
- Borgen van kwaliteit
- Ethiek en publieke acceptatie
- Toekomstige ontwikkelingen bij drone technologie

Op basis van deze thema's zijn er specifieke vragen geformuleerd en door middel van het marktconsultatiedocument[2] (22 maart 2022) en bijeenkomsten (19 en 21 april 2022) met diverse marktpartijen behandeld.

1 Wilt u meer weten over alle andere Buyer Groups die er bestaan en benieuwd wat de deelnemers van Buyer Groups afgelopen jaar bereikt hebben? U kunt erover lezen via: Public Impact Report 2021.

2 https://www.tenderned.nl/aankondigingen/overzicht/256534/details.

1.2 Open en eerlijk

Deze rapportage van de marktconsultatie is gebaseerd op de schriftelijke en mondelinge bijdragen van marktpartijen. De Buyer Group stelt afsluitend dit verslag op, waarin op hoofdlijnen de bevindingen zijn samengevat. Daarin zal geen concurrentiegevoelige informatie gedeeld worden. Het verslag wordt, zoals eerder beschreven in het marktconsultatiedocument van maart 2022, na afloop met alle deelnemers gedeeld en het staat de Buyer Group vrij dit op elk ander gewenst moment ook met anderen te delen. De bevindingen zullen daar waar de Buyer Group het mogelijk en zinvol acht worden geïmplementeerd in de gezamenlijke marktvisie.

1.3 Leeswijzer

Deze rapportage bevat een hoofdtekst met drie hoofdstukken en een bijlage (het marktconsultatiedocument, beschikbaar via TenderNed). Zoals beschreven in het marktconsultatiedocument maakt het geen deel uit van een specifieke aanbestedingsprocedure. Het is bedoeld om de markt en experts te betrekken, zodat zij waardevolle input kunnen leveren voor de marktvisie van de Buyer Group.

Afbeelding 1: Keten beeldherkenning met drones

2 UITWERKING MARKTCONSULTATIE

Ter overzichtelijkheid zijn de vragen hieronder opgenomen per paragraaf, de volledige vraagstelling is in het marktconsultatiedocument (d.d. 21 maart 2022) te lezen. De onderwerpen/vragen zijn niet in volgorde van belangrijkheid opgenomen.

2.1 Begrijpelijk maken van het gebruik

Vraag 1

Met betrekking tot het communiceren/informeren over het gebruik van AI beeldherkenning technologie

a) *Tot op welke hoogte kunt u de werking van algoritmen van AI beeldherkenningstechnologie duidelijk en begrijpelijk maken? Welke andere manieren zijn er in de markt en/of wat is de ervaring ermee in de markt?*

Met betrekking tot het communiceren/informeren over het gebruik van drones

b) *Kunt u het doelgebruik van inzet van drones duidelijk en begrijpelijk maken voor een breder publiek? Graag een toelichting*

Vraag 1a

Het communiceren/informeren over het gebruik van AI beeldherkenning technologie is een aspect dat valt binnen één van de thema's van de Buyer Group, gericht op ethiek en publieke acceptatie. Het doorlopen van het proces op hoofdlijnen en de uitkomsten daarvan kunnen inzichtelijk worden gemaakt:

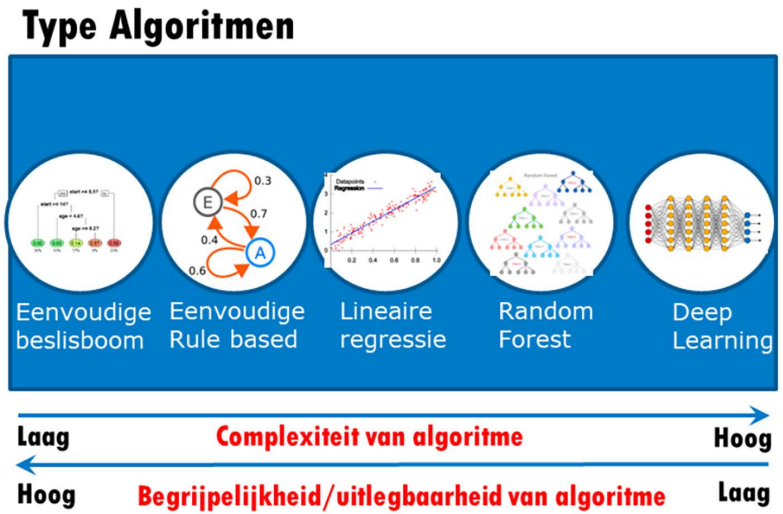
- Inzichtelijk maken welke data in het model gaat.
- Inzicht in keuze van het type model met een balans tussen complexiteit en verklaarbaarheid.
- Het optimaliseren van het algoritme, hoe het model wordt getraind
- Transparantie in het gebruik van de uitkomst van het model

Deelnemende partijen lieten weten dat over het algemeen het mogelijk is om de werking van algoritmen van AI beeldherkennings-technologie begrijpelijk te maken. Transparantie is aan te brengen qua (a) gegevensinvoer, (b) in AI-oplossing, (c) in het gebruik van de uitkomst. Het inzicht geven kan op verschillende manieren worden gedaan, zoals over hoe het is opgebouwd, hoe het is getraind, hoe de parameters zijn gekozen, en waarom + hoe de uitkomst wordt gebruikt.

Partijen beantwoorden tevens de vraag welke manieren er in de markt zijn om inzicht te kunnen geven in de werking van de AI technologie. In dat opzicht werd bij random forest naast het uitleggen van de beslisbomen ook feature importance genoemd. Verder is een decision forest moeilijker te begrijpen maar de werking is soortgelijk, dus als de eerste stap uit te leggen is, dan is het model vervolgens inzichtelijker te maken. Ook werden andere methodes voor het uitleggen van complexe modellen werden benoemd, zoals TCAV, SHAP of LIME. Daarnaast is door middel van een loss-functie de optimalisatie van algoritmen uit te leggen. De werking van AI technologie (en brondata) begrijpelijk maken voor publiek kan verder door visualisatie, geannoteerde beelden, of (simpele) praktijkcasus te beschrijven. Partijen gaven aan dat het communiceren afhangt van de doelgroep.

Ook werd de publicatie[3] van 'Richtlijnen voor het toepassen van algoritmen door overheden en publieksvoorlichting over data-analyses' uit 2021 benoemd. Aandachtsgebieden die daarin zijn uitgewerkt:

- Bewustzijn risico's
- Transparantie & Uitlegbaarheid
- Gegevensherkenning
- Auditeerbaarheid
- Verantwoording
- Validatie
- Toetsbaarheid



Afbeelding 2: Type algoritmen

Uitleggen *hoe* een model tot een uitkomst is gekomen is lastiger. Dit geldt al helemaal voor ingewikkeldere modellen, zoals neurale netwerken. Daarnaast werd door één van de partijen naar voren gebracht dat ingewikkelde modellen weliswaar kunnen zorgen voor nauwkeurigere resultaten, alleen dat het per casus afhangt of die nauwkeurigheid ook per se nodig is. In praktijk zal het zoeken zijn naar de balans tussen verklaarbaarheid en complexiteit. Advies van een partij daarbij was om modellen in eerste instantie klein en relatief eenvoudig te houden, zodat het te overzien is

Tot slot onderstreepten meerdere partijen het belang van brondata beschikbaar houden. En daarbij aan te geven welke data ervan is gebruikt om tot het model te komen, alsook welke codering van het model uit andere modellen komt (die op andere data is getraind).

Vraag 1b

Met betrekking tot het communiceren/informeren over duidelijk en begrijpelijk maken van het doelgebruik van drones voor een breder publiek kwamen de meeste partijen met soortgelijke inzichten. Goede communicatie en heldere doelstellingen bij de inzet van drones vergroot de publieke acceptatie. Dit kan ook door use-cases duidelijk te maken en daarbij aan te geven waarom een drone wordt gebruikt (en waarom het niet op een andere manier kan). Met name bij use-cases om de veiligheid te verbeteren of voor andere duidelijk maatschappelijke doelstellingen worden dan aanvaardbaar.

Het is belangrijk dat duidelijk is wie de operator is, niet enkel door de kleur van een drone, maar ook valt bijvoorbeeld te denken aan een informatiepunt zoals een website of een open toegankelijke app[4]. Verder is het belangrijk om transparant te zijn welke data verworven wordt en hoe het verwerkt wordt. Hoe kleiner de hoeveelheid of

3 https://www.rijksoverheid.nl/documenten/richtlijnen/2021/09/24/richtlijnen-voor-het-toepassen-van-algoritmen-door-overheden-en-publieksvoorlichting-over-data-analyses.
4 Europese Commissie werkt aan U-Space, dat hier mogelijk in kan voorzien.

hoe specifieker de ingewonnen data, zorgt voor een betere publieke acceptatie.

Eén partij benadrukte daarbij dat in het ideale geval de hele keten transparant wordt, inclusief alle data die opgeslagen wordt. Want als het alleen op een begrijpelijke manier (voor een breed publiek) wordt gecommuniceerd, dan is het vaak te algemeen en worden details weggelaten, waardoor op een later moment experts niet goed kunnen achterhalen wat er exact gedaan is. Dus beter een volledige beschrijving van de hele keten: op deze locatie(s) zijn opnames gemaakt, de opnames gaan in dit model, dat herkent deze beelden, dit wordt vervolgens zo vastgelegd, om uiteindelijk tot deze uitkomst te komen.

2.2 Levenscyclus

Vraag 2

c) Hoe maakt u de levenscyclus van een algoritme inzichtelijk voor overheden, zodat er vooraf en tijdens afname van uw product/dienst de kwaliteit wordt geborgd?

d) Hoe voorkomt u onbewuste vooringenomenheid (bias) in de uitkomst van algoritmes, en hoe is dit te signaleren?

Vraag 2a

Het markconsultatiedocument bevatte een fasering die bij veel partijen herkenbaar is bij het inzichtelijk maken van de levenscyclus van een algoritme voor overheden, om vooraf en tijdens afname de kwaliteit te borgen. Waarbij er een feedbacklus is om de kwaliteit te verhogen. De fases:

1e Fase – onderzoeksanalyse / doelbepaling / acceptatiecriteria
2e Fase – ontwerp/ontwikkeling/training en testen van het algoritme
3e Fase – in productie

Algemeen wordt beaamd dat de beste manier om inzicht te creëren in een algoritme voor gebruikers/overheden is door ze tijdens het gehele proces actief betrokken te houden. Er zijn een aantal manieren waarop de levenscyclus inzichtelijk gemaakt kan worden zoals Explainable AI (XAI) technieken die kunnen helpen om inzicht te geven in welke informatie door het model wordt gebruikt om voorspellingen te doen. Verder is een goed CI/CD (Continuous Integration / Continuous Deploymen) programma een belangrijk element, zoals dat ook bij andere software trajecten is. Maar het principe ‘Concept drift’ ligt altijd op de loer omdat er met dynamische data wordt gewerkt waar een model mee werkt.

Vraag 2b

Onbewuste vooringenomenheid in uitkomst van algoritme (een bias) kan weliswaar ontstaan in alle fasen van het ontwikkelproces, maar voornamelijk tijdens de trainingsfase. Door tijdens alle fasen een kritische houding aan te nemen kan het worden gesignaleerd. De focus ligt hierbij op de dataset, de architectuur en de aannames die hierbij zijn gemaakt. Drie zaken zijn van belang:

1) de juiste datavoorbereiding: alle objecten/patronen in relatief gelijke hoeveelheden worden weergegeven en er is een diversiteit van voorwaarden (opnamen vanuit verschillende hoeken, lichtomstandigheden, etc.)
2) domeinkennis: voor het beoordelen van data en parameters
3) een juiste validatie: zoals testen t.o.v. een gold data set

Meestal treedt bias op tijdens de trainingsfase. Daarom is extra aandacht nodig bij het samenstellen van een trainingsdataset. Een zo gevarieerd mogelijke dataset is van belang. Ook is het een manier om juist gebruik te maken van technieken waarbij gewerkt wordt met ongebalanceerde datasets, datasets met een mogelijke bias erin. 'Concept drift' zorgt ervoor dat het onontkoombaar is dat AI-modellen tijdens de levensduur periodiek bijgetraind en bijgewerkt worden. Hierbij hoort periodieke validatie, op basis van nieuw materiaal.

Daarnaast is het belang van documentatie door meerdere partijen benoemd, zoals hoe data verzameld wordt en hoe geannoteerd. Om zo de samenstelling en kwaliteit van de data te controleren die wordt gebruikt om de modellen te trainen, en ook de uitkomsten van de modellen voor verscheidene subgroepen.

Tot slot merkt een partij op dat er ook technologie is die in staat stelt om te beoordelen of een bias in het algoritme zit, en er zijn partijen in de markt die met kennisverspreiding en/of (open source) toolkits het verantwoorde gebruik van AI proberen te bevorderen[5].

2.3 Verantwoording over het resultaat van het algoritme

Vraag 3

Wat zijn (uw) methodes om processtappen in de ontwikkelfase van het algoritme te valideren? Waar ligt volgens u de verantwoordelijkheid bij de opdrachtnemer en hoe waarborgt u dit juridisch?

Over methodes om processtappen in de ontwikkelfase van het algoritme te valideren, laten meerdere partijen weten dat training-, test- en validatiedata daarbij essentieel zijn. Validatie wordt grotendeels gewaarborgd door voor en tijdens het ontwikkelproces af te stemmen tussen opdrachtgever en opdrachtnemer wat het gewenste resultaat is en wat de ondergrens van de nauwkeurigheid van het algoritme mag zijn. Heldere communicatie over de gebruikte methodes en mogelijke aandachtspunten worden hierbij extra belicht en tijdens het ontwikkelproces worden de resultaten op testdata getoetst. Met aan de voorkant van het proces duidelijk specificeren welke functionaliteiten geborgd moeten zijn en wat daarvan de geëiste voorwaarden in betrouwbaarheid en snelheid moeten zijn.
Meerdere partijen lieten daarbij op hoofdlijnen volgende stappen weten:
De eerste stap is de controle van het algoritme op basis van de trainingsdataset

- De tweede stap is de controle van het algoritme op basis van een testdataset (dat kan een subset zijn van de trainingsdataset, maar die niet gebruikt is in de training)
- De derde stap is de controle van het algoritme d.m.v. een validatiedataset

Het (juridisch) borgen van de kwaliteit kan bijvoorbeeld door KPIs (het gewenste resultaat is en wat de ondergrens van de nauwkeurigheid van het algoritme mag zijn) vast te leggen waaraan het detectie algoritme dient te presteren. En de afwegingen die in het ontwikkelproces worden gemaakt, deze ook goed vastleggen in een logboek. Daarbij werd wel door meerderen de kanttekening geplaatst dat een (AI) model een benadering is van de werkelijkheid, en deze zal nooit 100% accuraat, volledig of nauwkeurig zijn.

5 *15 Open Source Responsible AI Toolkits and Projects to Use Today | by ODSC - Open Data Science | Medium.*

Waar verder volgens deelnemende partijen de verantwoordelijkheid ligt, is dat opdrachtgever de uiteindelijke norm stelt waaraan moet worden voldaan, en ook de uiteindelijke beslissing neemt of uitkomsten voldoende zijn om mee te handelen. Tegelijkertijd is de opdrachtnemer daarbij verantwoordelijk om duidelijk over te brengen wat haalbaar is en te informeren over de mogelijke afwijkingen. Partijen laten weten dat het idealiter een samenspel is, zodat opdrachtnemer samen met opdrachtgever bekijkt welke data nodig is voor een goede validatie door een aantal representatieve use cases te definiëren. Ook ten aanzien van ethiek wordt het gezien als een gedeelde verantwoordelijkheid tussen opdrachtgever en opdrachtnemer, waarbij opdrachtgever in principe de leidende partij is die weet of de opdracht die verstrekt wordt ook op dat vlak toelaatbaar is en de opdrachtnemer daarbij tegelijkertijd ook vooraf en tijdens de uitvoering de verantwoordelijkheid heeft om aan te geven als daar iets op aan te merken is.

2.4 Informatiebeveiliging

Vraag 4: Veranderende ontwikkelingen

In de praktijk worden normen gevraagd in aanbestedingen. In de jaren na aanbesteding vindt er contractuitvoering plaats. Bij eisen rondom informatiebeveiliging gelden veelal periodieke actualisering, en het bijhouden van marktontwikkelingen en kennis voor het onderhoud.

a) *Hoe gaat u in deze context om met periodieke actualiseringen?*

b) *Wat verwacht u tijdens contractuitvoering van opdrachtgevers die bij aanbesteding normen hebben vastgesteld, om (gezamenlijk) grip te houden op veranderende ontwikkelingen?*

Vraag 4a

Bij informatiebeveiliging is actualisering en bijhouden van ontwikkelingen van belang. Deelnemende partijen geven aan dat bij start van een project het nodig is om af te spreken welke controls vanuit ISO27001 en/of BIO effectief zullen zijn, of dat er bijvoorbeeld een periodiek overleg tussen opdrachtgever en opdrachtnemer plaatsvindt om te valideren en te bepalen of er aanvullende maatregelen nodig zijn (binnen de reikwijdte van de aanbesteding). En dat er op BIO-niveau gekeken wordt welke maatregelen daadwerkelijk van belang zijn binnen de scope van het project, alsook dat mogelijk aanvulling nodig is, zoals handvatten die bijvoorbeeld uit de Cybersecurity implementatie richtlijn (CSIR[6]) te halen zijn, op basis van weerstandsniveaus.

De meeste partijen laten weten dat bij een project waarvoor informatiebeveiliging vereist is, dat het de verantwoordelijkheid van opdrachtnemer is om eventuele veranderingen in de gaten te houden. Ze merken op goed op de hoogte te zijn van ontwikkelingen rondom certificeringen en vinden dat opdrachtnemers de opdrachtgevers tijdig horen te informeren over mogelijke gevolgen voor hun bedrijfsvoering als gevolg van deze ontwikkelingen. Het omgaan met periodieke updates van informatiebeveiliging zal over het algemeen de principes van het information security management system (ISMS) volgen. In die zin verschilt het volgens enkele partijen niet

6 https://cip-overheid.nl/media/1706/csir-34-definitief-concept-20210914.pdf.

veel van andersoortige IT software aanbestedingen. Het contract moet de vereisten bevatten om een actief ISMS te hebben. Er wordt bijvoorbeeld jaarlijks als onderdeel van het managementsysteem gekeken of er een verandering is in dreigingsbeeld in de regelgeving en in de ontwikkelingen op een AI.

Vraag 4b

Voor effectieve risicoanalyse van functionele, technische en organisatorische eisen laten deelnemende partijen weten dat ze het essentieel vinden dat de opdrachtgever meedenkt en meewerkt. Het inschatten van risico's zijn een belangrijk onderdeel van informatiebeveiliging.

Dit zijn niet alleen risico's van de leverancier, maar vooral ook de risico's vanuit het gezichtspunt van de opdrachtgever, en waar de opdrachtgever zicht op heeft. Deelnemende partijen geven aan dat ze verwachten van opdrachtgevers een inventarisatie hebben gedaan voor de aanbesteding over welk niveau van informatiebeveiliging van toepassing is, Daarbij is van belang dat opdrachtgever bij risicoanalyse de materiedeskundige personen betrekt.

Een proactieve rol van opdrachtgever en opdrachtnemer wordt verwacht. Zo is het van belang dat er een open cultuur wordt gecreëerd waarin iedereen zich kan uitspreken en waarin niemand zich beperkt voelt. Alsook waarbij er bijvoorbeeld periodieke (contract)management overleggen worden belegd waarin informatiebeveiliging een vast thema is op de agenda waarin gewenste veranderingen op het gebied van informatiebeveiligings-normen besproken kunnen worden (vanuit Nederlandse en/of Europese regelgeving of instanties als de autoriteit persoonsgegevens op het gebied van gegevensverwerkingen).

Vraag 5: Afspraken nakomen en aantoonbaarheid

Gedurende de looptijd van een contract vindt er regelmatig validatie van afspraken plaats, audits, en opvolgingen ervan. Zo is het ook nodig om na te gaan of afnemer(s) de afspraken nakomen voor het waarborgen van informatiebeveiliging. Zoals het updaten van beveiligingsplannen. Daarbij speelt het aantoonbaar kunnen maken van het nakomen van afspraken.

Wat voor afspraken zijn volgens u met elkaar te maken over de aantoonbaarheid van nakoming van afspraken, om zo als opdrachtgever-opdrachtnemer tijdens contractuitvoering verder aan te kunnen werken?

Deelnemende partijen laten weten dat onderdelen uit ISO27001 een graadmeter kunnen zijn, alsook dat rapportage en logging een belangrijke rol spelen. In bepaalde gevallen (met gegronde redenen) kunnen audits van opdrachtgever ook een mogelijkheid zijn als hierover afspraken zijn gemaakt, waarbij de derde partij in gezamenlijkheid wordt gekozen.

Enkele deelnemende partijen geven aan dat de midden- en kleinbedrijven regelmatig geen volledige ISO27001 certificering hebben of verkrijgen, en merken op dat certificering ook geen garantie is. Dus dat vandaar ook advies is om in contracten te zorgen dat met transparantie en het leveren van bijvoorbeeld periodieke rapportages voldoende zekerheid te krijgen is, en vooraf duidelijk is welke controles en het aanleveren van informatie het belangrijkste in zijn.

Net als bij vraag 4b wordt ook hier het op een open wijze communiceren en de periodieke (contract)managementoverleggen benoemd. Om zo ten aanzien van informatiebeveiliging te bespreken of er nog aspecten zijn veranderd, of type data is veranderd, of er nog steeds aan afgesproken procedures worden gehouden, enzovoort.

2.5 Bewustzijn ten behoeve van informatiebeveiliging

Vraag 6

Kent u best practices[7] op welke manier u en/of anderen hierin zijn geprikkeld?

Graag een toelichting

Op het gebied bewustwording zijn er een heel aantal technieken. Over het algemeen geven de meeste partijen aan dat bedrijfsbrede communicatie belangrijke factor is. Andere suggesties waren onder meer neppe phishing mails, mystery guests, ethische hackers, e-learnings en gamefication, en ransomware-simulaties (waarbij een gebruiker echt het gevoel/schrik heeft dat er op dat moment gehackt wordt). Alsook in specifieke projecten en/of bij bepaalde (risicovolle) eindgebruikers kennis en ervaringen delen, of door gebruikersessies te organiseren rond het thema informatiebeveiliging. Tot slot is het belangrijk dat organisaties lid zijn van de gebruikersgroepen op de diverse fora, zodat er kennis is van nieuwe bedreigingen.

2.6 Transparantie ten behoeve van informatiebeveiliging

Vraag 7

a) *Is een dergelijke methode zoals een SBOM/HBOM werkbaar in het werkveld AI Beeldherkenning en drones? Welke kansen en risico's ziet u?*

b) *Kent u vergelijkbare manieren om informatieveiligheid te borgen? Zijn er voor andere aspecten nog andere methoden bruikbaar? Graag een toelichting*

Vraag 7a

De meeste deelnemende partijen gaven aan dat een Software Bill of Materials (SBOM) een goed werkbare manier kan zijn om risico's te beperken. Volgens hen kan een opdrachtgever – wanneer het gaat om het samenstellen en ontwikkelen van (AI) softwareproduct – dit als een onderdeel meenemen in aanbestedingen. In zo'n SBOM staan dan de exacte versies van alle software- en hardwarecomponenten die toegepast worden, zodat er gelijk een mededeling opkomt voor potentiële kwetsbaarheden bij een controle ten opzichte van Common Vulnerabilities and Exposures (CVE) lijst.
Kanttekening die enkele deelnemers aangeven is dat een SBOM lijst niet een statische lijst is, dus het zal updates vergen wanneer er nieuwe (versies van) componenten worden gebruikt. De SBOM is niet een waterdichte oplossing, aangezien kwetsbaarheden die niet-publiek bekend zijn (zoals 'zero-day' kwetsbaarheden) door het toepassen van SBOMs niet geïdentificeerd zullen worden.

7 Een effectieve techniek, werkmethode of activiteit.

Daarnaast zou controle op kwetsbaarheden vanzelfsprekend het beste plaatsvinden tijdens het ontwikkelproces.

Een ander punt om rekening mee te houden is de leesbaarheid van een SBOM. Hiervoor is het van belang om te weten voor welke doelgroep (expert of leek) en met welke doelstelling de SBOM geschreven wordt. Zo zal een expert meer gebaat zijn bij de inzichten van metadata en dergelijke om bijvoorbeeld een toets te kunnen doen op AVG compliancy. Terwijl een persoon die minder deskundig is van dit specifieke werkgebied meer gebaat is bij inzicht in brondata, model en trainingsdata. Iets wat meer inzicht en begrip geeft over de informatieuitkomsten en de herkomst van data.

Vraag 7b

Deelnemende partijen geven aan dat het van belang is dat leveranciers transparant durven te zijn over kwetsbaarheden. Voor leveranciers kan het verleidelijk zijn om informatie intern te houden, terwijl dat juist effectieve informatiebeveiliging belemmert (ook voor andere afnemers van dezelfde componenten). Een open samenwerking gericht op verbetering is nodig. Verder merken een aantal ook het gebruik van open source software op en het open source maken van ontwikkelde broncode.

Andere manier om informatieveiligheid te borgen is bijvoorbeeld door bij oplevering een beveiligingstoets te laten uitvoeren op de ingerichte IT-voorziening door een gespecialiseerd bedrijf. Alsook periodieke beveiligingsonderzoeken op de gebruikte systemen en hardware, zoals door middel van

- Secure codereview,
- Configuratiereviews en Attack & Penetration testen,
- Assessments op (naleven van) procedures die opdrachtgever heeft om de informatiebeveiliging van de gebruikte componenten te waarborgen.

Met dergelijke periodieke scans kunnen (ernstige) kwetsbaarheden gemeld worden en direct opgepakt worden, indien nodig.

2.7 Open data standaarden

Vraag 8

Welke open data standaarden ziet u als leidend, en kunt u het toelichten met voor- en nadelen?

Deelnemende partijen geven als reactie aan dat het lastig is een generiek antwoord mogelijk te geven, aangezien het afhangt van het domein en de functie. Tegelijkertijd zijn een heel aantal formaten als standaarden door partijen benoemd:

- OGC standaarden van het Open geospatial consortium o.a.;
  - WCS
  - GeoTIFF
  - netCDF
  - LAS/LAZ
  - HDF5
  - GeoJSON,
  - WMS
  - WMTS
  - WFS
  - CityGM
- Esri shape (uitwisselingsformaat voor geografische informatie)
- Voor de publieke data sets van de overheid veelal het ISO 19139 XML schema gebruikt voor non-spatial data (of het meer recentere ISO/TS 19139 format).
- GEMINI 2.3 voor spatial metadata (gebaseerd op ISO 19139)
- Protobuf (Protocol Buffers)
- ONNX, (Open Neural Network Exchange)

Het overgrote deel van de deelnemende partijen laat weten dat ze voorstander zijn van open formaat standaarden. Aangezien data erdoor makkelijker uitwisselbaar zijn en niet gebonden aan één software omgeving. Het verkleint het risico van een vendor lock-in. Het gebruikmaken of het vereisen van open formaat standaarden wordt gezien als een goede stap richting het integraler benutten van data over verschillende domeinen. Daarnaast kan het 'data-waste' (verzamelde data die ongebruikt blijft) verminderen doordat hergebruik makkelijker wordt.

Tegelijkertijd merken enkele partijen op dat met name bij domeinen waar een bepaalde volwassenheid optreedt de open formaat standaarden het meest effectief zijn. Bij innovatieve ontwikkelingen kunnen de voordelen minder sterk zijn als er nog geen consensus is over de inhoud en interpretatie van de data. Door dan open formaat standaarden voor te schrijven, kan er mogelijk data verloren gaan, om het in standaard voorgeschreven formaat in te passen.

Overigens zegt open formaat standaard niets over het gebruik van de data. Het is in ieder geval essentieel om het formaat van de uitwisseling goed te documenteren, bijvoorbeeld met een OpenAPI specificatie (OAS of AsyncAPI). Anders gezegd, wat je opslaat in het formaat – de metadata – is belangrijk bij het aanleveren van data, zodat partijen weten wat ze krijgen aangeleverd.

2.8 Exitfase, overdraagbaarheid en leerprocessen van het algoritme

Vraag 9

a) Kunt u aangeven welke mogelijkheden u ziet op het vlak van overdraagbaarheid van algoritmes en/of benutting van leerprocessen?

b) En wat is voor u een acceptabele exit strategie bij oplevering?

Vraag 9a

Een meerderheid van de deelnemende partijen is van mening dat overdraagbaarheid in de praktijkmogelijk is. Om werkelijk algoritmes (codes) over te dragen, zullen er wel van tevoren afspraken nodig zijn, onder andere over intellectueel eigendom.
Wanneer het niet wenselijk is voor een leverancier om het algoritme over te dragen, is er mogelijk nog te kiezen om bepaalde kennis uit het algoritme te laten delen zoals door middel van de set gewichten, of dat de data is aan te bieden. Door brondata over te dragen kunnen andere partijen op basis daarvan hun eigen AI model ontwikkelen. Het beschrijven van de aard van brondata is belangrijk voor bruikbaarheid.

Als bij een project een algoritme nog ontwikkeld wordt – op basis van samenwerking tussen opdrachtnemer en opdrachtgever – dan is het werk eventueel over te nemen door een andere partij of andere professionals binnen een organisatie wanneer op een ordelijke manier te werk is gegaan.

Een AI-project bestaat uit de volgende kernonderdelen:
- gelabelde trainings- en validatiedata;
- het ontwerpen en het methodisch verder ontwikkelen van de AI-modellen;
- software om de modellen te trainen;
- productiesoftware.

Documentatie is bij het ontwerpen, ontwikkelen en monitoren van AI belangrijk. Dat bevordert de uitlegbaarheid, controleerbaarheid, en het begrijpen van onderliggende behoeften van de opdrachtgever. Daarnaast geven deelnemende partijen aan dat ook de gebruikte trainingsdata, backbone model en modeltype waardevol is. Alsook de ontwerpkeuzes, constructie van broncode, beperkingen, selectie van variabelen, parameters.

Als het gaat om het werkelijk overdragen van algoritmes (codes) dan is het vastleggen van de volgende zaken waardevol:
- documentatie in de code (opmerkingen erbij);
- documentatie bij/voor de code voor het beschrijven van de afhankelijkheden en een 'how-to' om de programmatuur aan te roepen;
- gebruikersvoorbeelden in code;
- demonstraties (bij oplevering); video (screen recording); duidelijk beschrijving van input, output en mogelijke foutmeldingen.

Vraag 9b

Bij aanvang van een project zijn duidelijke afspraken nodig, en dat verschilt per project/use case/diensten. Een aantal deelnemende partijen geven aan dat er weinig verschil is met andersoortige software projecten, waaronder een CI/CD programma ook als element. Een open cultuur en vertrouwen is van belang bij projectwerk, en samen een acceptabele exit strategie te bepalen met wederzijdse afspraken over verschillende rechten en plichten (acceptatie, garantie, intellectueel eigendom, beheer en gebruik, alsook bijvoorbeeld het mogen delen van codes in communities indien gewenst).

Diverse deelnemende partijen laten weten dat wanneer een algoritme op basis van een specifieke use case of in de context van opdrachtgever wordt ontwikkeld, er ook te bespreken is dat opdrachtgever eigenaar van broncode en brondata wordt en alles over te dragen is. Op die manier kan opdrachtgever bijvoorbeeld ook na realisatie van het algoritme vervolgens het beheer en onderhoud op zich nemen (indien opdrachtgever die wens heeft en over de expertise beschikt). Zoals in vraag 9a naar voren kwam, is daarbij dan ook nodig dat onderliggende documentatie volledig is. Verder kan kennisoverdracht te regelen zijn door met elkaar als opdrachtgever en opdrachtnemer tijdens het project bij elke stap samen te werken, zodat gemeenschappelijke kennis wordt opgebouwd. Daarnaast valt ook nog te denken aan bijvoorbeeld training, zodat opdrachtgever (of derde partijen) in staat is om vervolgens de algoritmes te gebruiken en/of door te ontwikkelen, of mogelijke ondersteuningsrol van leverancier bij opkomende vragen naderhand.

Ook is de suggestie gedaan om bij exit strategie op te nemen dat het succesvol hertrainen van algoritme met vergelijkbare resultaten een voorwaarde is om een overdracht te kunnen doen. Op die manier kunnen opdrachtgevers een bepaalde zekerheid inbedden. En dit eveneens contractueel vast te leggen bij de opvolgende opdrachtnemer, zodat ook bij het einde van die contractperiode de doorontwikkeling wordt geborgd.

2.9 Datadelen en gefedereerde data

Vraag 10

Welke oplossingsrichtingen kent u / heeft u op het vlak van AI beeldherkenning en drones om data te delen in een gefedereerde data infrastructuur met verschillende overheidsdiensten en gebruikers? Graag een toelichting indien mogelijk

Enkele deelnemende partijen merken op dat ze een oplossing/portaal/platform/hub kunnen aanbieden, voor het gefedereerd delen van data. Waarbij de tussenpartij als neutrale tussenpartij kan fungeren voor het beheer. Zodat met de oplossing die wordt gefaciliteerd de gefedereerde data toegankelijk is en een centrale toegang ertoe is. Een belangrijke factor voor datadeling is dat data en de toegang tot de data ondubbelzinnig beschreven wordt en dat machines/software die beschrijving kunnen begrijpen. Zo kunnen de in 2016 geïntroduceerde internationale FAIR-principes voor datamanagement hierin van toepassing zijn (Findable, Accessible, Interoperable, and Reusable)[8]. Deze principes zijn relevant voor zowel wetenschappelijke als niet wetenschappelijke data. Of anderzijds dat data in leveranciersonafhankelijk format wordt gebruikt, zodat diverse applicaties erop aan te sluiten zijn. Zodat data niet vast zit in specifieke applicaties, met risico tot vendor lock-in.

Eén van de partijen bracht naar voren dat vanuit governance perspectief er nieuwe besturingsmodellen nodig zijn bij het delen van data met gefedereerde entiteiten. Het datadelen gaat niet alleen om het technisch faciliteren van data-overdracht door middel van digitale infrastructuren. Het gaat ook om het begrijpen van de afhankelijkheden met data aanbieders, softwareleveranciers, gebruikers en een nieuwe verdeling van verantwoordelijkheden en eigenaarschap van data op orde houden. En daarbij afspraken te maken over onder andere het borgen van maatschappelijke waarden als transparantie, privacy en data soevereiniteit.

Afsluitend zijn nog enkele richtingen benoemd om datadeling te bewerkstelligen. Zo is er vanuit de Rijksoverheid het Programma Regie op Gegevens[9], dat onder andere uitgebreide toelichting geeft op verschillende vormen van datadeling aan de hand van een voorstudie[10]. Daarnaast is er een Werkgroep Data Delen[11] bij de Nederlandse AI Coalitie, die eind 2021 een gids[12] presenteerde. Een tweede richting zou sectorspecifieke initiatieven kunnen zijn, zoals reeds gebeurd in het zorg- of hypotheekdomein. Tot slot is als derde richting een publiek-private coalitie of open samenwerkingscommunity mogelijk. Op die manier is transformatie te versnellen en datadeling te organiseren, met gezamenlijke sturing op bijvoorbeeld interoperabiliteit en de borging van veiligheid, juistheid en rechtmatigheid van datadeling. Zo is er bijvoorbeeld in de watersector het open innovatieplatform Digishape[13] van ondernemers, kennisinstellingen en overheden die samen de potentie van digitalisering willen benutten.

8 https://nl.wikipedia.org/wiki/FAIR-principes.
9 https://www.digitaleoverheid.nl/overzicht-van-alle-onderwerpen/regie-op-gegevens/.
10 Voorstudie naar een kader voor regie op gegevens · Regie op Gegevens (pleio.nl).
11 Data Delen - Nederlandse AI Coalitie (nlaic.com).
12 https://nlaic.com/nieuws/nl-aic-presenteert-gids-voor-interoperabel-delen-van-data-voor-ai-toepassingen/.
13 https://www.digishape.nl/.

2.10 Slotvraag

Heeft u inhoudelijke tips en/of opmerkingen voor de Buyer group AI beeldherkenning met drones die u zou willen meegeven? Of nog overige zaken die u zou willen meegeven aan de organisatie van de Buyer Group?

Bovenstaande vraag resulteerde in verschillende slotopmerkingen bij de schriftelijke reacties. Vanwege de grote diversiteit en de aard ervan bevat deze rapportage geen bevindingen op hoofdlijnen. Wat er in ieder geval uit blijkt is dat het veld van beeldherkenning met drone data aan het begin van ontwikkeling staat en dat de deelnemende partijen in deze marktconsultatie met enthousiasme informatie erover delen. Vandaar:

Wij willen u heel hartelijk bedanken voor uw bijdrage aan deze marktconsultatie.

3 PROCEDURE VAN DE CONSULTATIE

3.1 Procedure

De consultatie is gepubliceerd op (onder andere) TenderNed als open consultatie. Dit houdt in dat alle ondernemers, kennisinstellingen, en inhoudsdeskundigen de gelegenheid hebben gehad om deel te nemen. De consultatie is uitgevoerd in twee stappen:

1) Schriftelijk: het beantwoorden van de vragen (zie hoofdstuk 2 uit het marktconsultatiedocument);
2) Plenaire bijeenkomst met (een deel van) de geïnteresseerde partijen.

Deelname aan deze consultatie was geheel vrijwillig en zal niet leiden tot enige voorrechten in het verdere verloop en betrokkenheid bij de Buyer Group, marktvisie, en/of eventuele toekomstige aanbesteding(en) bij deelnemers van de Buyer Group. Evenzo geldt het tegenovergestelde; niet-deelname in de consultatie zal niet leiden tot enige benadeling in het verdere verloop en betrokkenheid bij de Buyer Group, marktvisie, en/of eventuele toekomstige aanbesteding(en) bij deelnemers van de Buyer Group.

De Buyer Group hecht veel waarde aan een transparant proces. Het is niet de bedoeling om deelnemende partijen aan de consultatie te bevoordelen of niet-deelnemende partijen op achterstand te plaatsen. Vanuit dit perspectief zal het verslag na afloop met alle deelnemers worden gedeeld en het staat de Buyer Group vrij dit op elk ander gewenst moment ook met anderen te delen.

3.2 Planning

De marktconsultatie bestond uit de volgende onderdelen:

Datum	Activiteit
Donderdag 17 maart 2022	Publicatie van het consultatiedocument / uitnodiging tot deelname aan de consultatie van de Buyer Group
Donderdag 7 april 2022	Aanmelden voor deelname bijeenkomst (via buyergroups@pianoo.nl)
Dinsdag 12 april 2022	Uiterste datum voor het indienen van de antwoorden op de vragen (via buyergroups@pianoo.nl)
Dinsdag 19 april, 13.00-15.00 uur *en* donderdag 21 april, 10.00-12.00 uur	Bijeenkomst
Mei/juni 2022	Afronding van marktconsultatie en delen van afsluitend verslag

3.3 Contact

Bent u enthousiast over dit project en heeft u interesse om hierover verder met de Buyer Group van gedachten te wisselen of uw mening te uiten, of wilt u op de hoogte worden gehouden? In alle gevallen kunt u contact opnemen via buyergroups@pianoo.nl.

COLOFON

Een Buyer Group is een samenwerking van opdrachtgevers die een gedeelde marktvisie voor een specifieke product categorie ontwikkelen en deze binnen twee jaar implementeren in hun inkooppraktijk. Deze rapportage is op basis van een openbare marktconsultatie dat een breed draagvlak kent en stimuleert om verantwoorde oplossingen te ontwikkelen en in te kopen.

De Buyer Group Circulaire AI beeldherkenning met drones is ondersteund door PIANOo en NL AI Coalitie.

De Buyer Group wordt mogelijk gemaakt door ministerie van Binnenlandse Zaken en Koninkrijksrelaties (Directoraat-generaal Overheidsorganisatie).

Auteurs
Joris Krüse
Rolf Zeldenrust
Dit rapport is geschreven in samenwerking met de Buyer Group deelnemers.

Redactie
Iris Gouweloos
Anique van Leeuwen