Oude Waalsdorperweg 63
2597 AK Den Haag
Postbus 96864
2509 JG Den Haag

www.tno.nl

T +31 88 866 10 00
F +31 70 328 09 61

TNO-rapport

TNO 2021 R10090

Privacy bescherming bij niet-coöperatieve gezichtsherkenning

Datum	21 januari 2021
Auteur(s)	Drs. J.H.C. van Rest T. Attema MSc Dr. T. Timan Dr. ir. R.J.M. den Hollander Ing. G.P. van Voorthuijsen
Aantal pagina's	126 (incl. bijlagen)
Aantal bijlagen	12
Opdrachtgever	Nationale Politie, namens de samenwerkende partijen in het FieldLab Technologie en Data (Johan Cruijff ArenA, Gemeente Amsterdam) “Digitale Perimeter”
Projectnaam	Digitale Perimeter
Projectnummer	060.40162

Samenvatting

Dit rapport is de verslaglegging van het project gezichtsherkenning in de samenwerking tussen gemeente Amsterdam, politie, JC ArenA en TNO. Het beoogt inzicht te geven in de mate waarin het mogelijk is om met inachtneming van wettelijke, juridische, sociale en ethische kaders gezichtsherkenningstechnologie in te zetten ten behoeve van het waarborgen van de veiligheid in publieke ruimtes.

Gezichtsherkenning zonder medewerking van het subject noemen we niet-coöperatieve gezichtsherkenning. Dit is een potentieel krachtig hulpmiddel in het veiligheidsdomein. Het wordt door de politie toegepast. Naast bestaande toepassingen, worden in dit rapport verschillende andere potentiele toepassingen beschreven. Niet coöperatieve gezichtsherkenning vormt tevens inherente inbreuk van de privacy. Dat is nog meer het geval als bij een match een nadelige consequentie volgt. De beslissing om dergelijke technologie in te zetten voor een dergelijke niet-coöperatieve toepassing moet dus zorgvuldig worden afgewogen. Daarbij moet ook optimaal gebruik worden gemaakt van de laatste inzichten om de privacy te beschermen. Dit rapport geeft een overzicht van bestaande inzichten, en presenteert een aantal mogelijkheden om die privacy beter te beschermen.

Het is verstandig om privacy beschermende strategieën waar mogelijk af te dwingen in de (vorm van) technologie. Misbruik kan dan worden uitgesloten. Indien dat technisch onvoldoende mogelijk is dan kan dat (ook) op organisatorische wijze, bijvoorbeeld door regulering, opleiding, screening van personeel, instructies, audits, toezicht, etc. Aan die organisatorische maatregelen zijn soms ook weer technische voorwaarden verbonden. Het is bijvoorbeeld mogelijk om toezicht op het gezichtsherkenningsproces te faciliteren door in de architectuur af te dwingen dat bij iedere zoekopdracht een melding (zonder persoonsgegevens) naar de toezichthouder gaat.

In dit rapport wordt specifiek ingegaan op de mogelijkheid om de vergelijking van biometrische gezichtstemplates op een veilige wijze te verrichten. De technologie waarmee dit mogelijk is heet *multi-party computation* (MPC). Er is in deze studie een literatuuronderzoek uitgevoerd en een technisch experiment gedaan.

Meerdere onderzoeksgroepen doen al meer dan tien jaar actief onderzoek naar MPC. Uit de literatuur blijkt dat er verschillende relevante publicaties en patenten zijn. Daarnaast hebben verschillende bedrijven relevante volwassen technologie. Voorafgaand aan deze studie was het niet duidelijk of bestaande technologie voldoende snel kan werken.

Het technische experiment liet zien dat op reeds bekende varianten van MPC voor gezichtsherkenning technische verbeteringen mogelijk zijn die de snelheid ten goede komen. Het experiment liet tevens zien dat er afwegingen te maken zijn met betrekking tot de snelheid in relatie tot functionaliteit. Bijvoorbeeld bij de afweging of de maat voor de gelijkenis ook kan worden doorgegeven. Ook illustreert het experiment hoe de (afname in) snelheid relateert aan de grootte van de *matchlist* en het aantal zoekopdrachten. Of deze verbeteringen voldoende zijn, hangt af van de specifieke toepassing.

Kortom, het vergelijken van biometrische gezichtstemplates in het versleutelde domein, lijkt binnen bepaalde grenzen (afweging functionaliteit versus snelheid versus kwaliteit) mogelijk.

De technologie MPC lijkt niet alleen mogelijk, maar ook nuttig in specifieke toepassingsscenario's (denk aan proportionaliteit, subsidiariteit, traceerbaarheid en accountability). Dit is onderbouwd met twee concepten, één over evenementenveiligheid conform het idee van een digitale perimeter, en één over persoons- en bijbehorende objectbeveiliging in een rustige wijk.

Bij toekomstige implementaties is het aan te bevelen om MPC te combineren met andere maatregelen die het risico van de privacydreiging "vervorming" (i.e. foutieve herkenning en foutieve niet-herkenning) verkleinen zoals *managed analytics* (zie bijvoorbeeld sectie 5.4.4). Om de toets op subsidiariteit te kunnen uitvoeren, moet bij inzetbeslissingen ook alternatieve technologie zoals zachte biometrie (zie sectie 5.4.3) worden overwogen.

Een belangrijk aspect is de schaalbaarheid van de oplossing. Er bestaat geen standaard voor biometrische gezichtstemplates. Een MPC variant van gezichtsherkenning is in deze studie toegepast op een bepaald type gezichtstemplate. Daarbij is géén gebruik gemaakt van specifieke eigenschappen van deze template, dus in principe lijkt het mogelijk om de technologie ook voor andere templates te laten werken. Een ander aspect is interoperabiliteit. De vorm waarin de technologie nu is gemaakt, zou vereisen dat alle direct verbonden partijen met live- en met *enrollment* templates hetzelfde technische format zouden moeten gebruiken. Zowel voor wat betreft het gezichtstemplate, als voor wat betreft de "MPC-schil" daar omheen. Het kan wenselijk zijn om voor beiden een standaard te gaan nastreven, zodat er zo weinig mogelijk barrières zijn om deze vorm van veilige verwerking -waar nodig – toe te passen.

Deze studie heeft laten zien dat nieuwe technologie kan helpen om de privacy (beter) te beschermen. Met de informatie verzameld in dit rapport kan bij bestaande gezichtsherkenningstoepassingen getoetst worden of ze ook optimaal gebruik maken van nieuwe technologie om de privacy te beschermen. Aangezien er voortdurend technologische vernieuwingen beschikbaar komen, lijkt het nuttig om privacy- en gegevensbeschermingseffectbeoordeling niet een éénmalige toets te laten zijn, maar om die na verloop van tijd opnieuw te doen. Organisaties die nu niet-coöperatieve gezichtsherkenning toepassen (in het veiligheidsdomein) doen er dus goed aan om te controleren of hun toepassing in het licht van deze studie nog steeds "privacy by design" is.

Inhoudsopgave

1 Inleiding

De Johan Cruijff ArenA ambieert het beste, meest gastvrije, veiligste en duurzaamste stadion te zijn. De Gemeente Amsterdam wil gastvrij zijn, ze wil dat evenementen veilig verlopen en geen overlast generen voor buurtbewoners. De Nationale Politie wil de veiligheid dienen met actuele en efficiënte technieken op zo'n wijze dat ze passen bij en binnen de rechtsstaat. De Johan Cruijff ArenA, de Gemeente Amsterdam en de Nationale Politie werken samen in een FieldLab Technologie en Data, rondom het ArenA-terrein. Daarin worden innovatieve projecten uitgevoerd. Eén van die projecten gaat over het beschermen van de privacy bij het automatisch herkennen van gezochte personen.

1.1 Context

Gemeentes, de politie, private beveiligers en evenementenorganisaties zijn op verschillende manieren verantwoordelijk voor de veiligheid in de (semi-)openbare ruimte. Het gaat daarbij om handhaving van de openbare orde, object- en persoonsbeveiliging, evenementenveiligheid en bijvoorbeeld om *crowd management* bij evenementen. In ieder van deze contexten kan het nuttig zijn om specifieke personen tijdig te herkennen.

1.1.1 Menselijke gezichtsherkenning

Het herkennen van personen kan door mensen worden gedaan. Afbeeldingen van (gezichten van) gezochte personen kunnen worden verspreid onder politie en beveiligers. Ondernemers verspreiden foto's van bekende veelplegers via besloten berichtengroepen. Sommige mensen zijn erg goed in het herkennen van gezichten (Voskuil, 2019), maar dat zijn uitzonderingen. De meeste mensen moeten er tijd en moeite in steken om gezichten goed te onthouden. Een recent initiatief dat bij het herkennen van gezichten moet helpen is de ontwikkeling van KOPS, een app op de smartphone van agenten waarmee ze kunnen trainen (RTLNieuws, 2019). Het herkennen van gezichten door mensen is desondanks niet goed schaalbaar. Daarnaast kan het een negatief effect hebben op de ervaren werkdruk. Onder rustige omstandigheden is het erg saai werk, en onder hoog-risico omstandigheden kan het te veel mentale druk geven. Daarbij kan het omslachtig of zelfs onmogelijk zijn om de gegevens weer "uit de kantines en uit de hoofden" te verwijderen als de persoon niet meer gezocht wordt.

1.1.2 Automatische gezichtsherkenning

Sinds 2001 worden proeven gedaan met automatische gezichtsherkenning met niet-coöperatieve[1] scenario's[2]. Dat gebeurde bijvoorbeeld in 2001 bij de Super Bowl in de Verenigde Staten (Rogers, 2016). Snel werd duidelijk dat de ongecontroleerde omstandigheden van de toegangscontrole een desastreus effect

[1] Een niet-coöperatief scenario is een scenario waarin het subject van gezichtsherkenning geen belang heeft om mee te werken met het gezichtsherkenningssysteem. In een dergelijk scenario wendt die persoon bijvoorbeeld zijn gezicht niet actief naar de camera.

[2] In 2001 ontwikkelden twee computerwetenschappers een robuuste manier om gezichten te detecteren in beelden (MIT Review, 2015). Dat algoritme (Viola / Jones) vormde de aanzet voor veel nieuwe experimenten met gezichtsherkenning.

hadden op de prestaties. Afhankelijk van de configuratie van het systeem werd vrijwel niemand van de *watchlist* gevonden (dit duidt mogelijk op foutieve missers), of werd juist iedereen “herkend” op de watchlist (foutieve identificatie). In de jaren die volgden lieten proeven zien dat de kwaliteit van gezichtsherkenning gestaag vooruit ging – ook onder minder gecontroleerde omstandigheden. Dit gebeurde onder andere door gebruik te maken van neurale netwerken (Grother, Quinn, & Ngan, 2017). Daarnaast is meer bekend over (het controleren van) de operationele factoren die de kwaliteit van dit soort technologie beïnvloeden, zoals belichting, opnamehoek en cameraresolutie.

1.1.3 Toepassingen van niet-coöperatieve vormen van gezichtsherkenning

Tegenwoordig zijn allerlei toepassingen denkbaar. Dat kan gaan om het handhaven van gebieds- stadions- of openbaar vervoersverboden[3], of om het tijdig signaleren van gezochte criminelen of gemonitorde geradicaliseerde personen rond te-beschermen-personen, evenementen (zoals in het voorbeeld hierboven) of vitale objecten. De functionele eisen hangen sterk af van het toepassingsgebied. In sommige toepassingen gaat het om het herkennen van één of enkele personen, zoals bij het handhaven van een contactverbod. Maar het kan ook gaan om tientallen of zelfs honderden of duizenden personen. In het jaar 2015 waren bijvoorbeeld in Nederland 11.000 voortvluchtigen die ter opsporing gesignaleerd stonden (Schoenmakers, De Groot, Van Rooyen, Van Zanten, & Baars, 2017). De Nederlandse overheid zou bijvoorbeeld kunnen overwegen (na een analyse van proportionaliteit en subsidiariteit) om bij hun overheidsgebouwen herkenningstechnologie toe te passen om deze voortvluchtigen te detecteren als ze bijvoorbeeld een paspoort of rijbewijs komen vernieuwen.

1.1.4 Voorzichtigheid in de maatschappelijke discussie over gezichtsherkenning

In zijn brief aan de Tweede Kamer (Grapperhaus, 2019) stelt de minister van Justitie en Veiligheid dat

> *“een structurele inzet van een breed vertakt, realtime gezichtsherkenningstechnologie, waarbij mensen continu en overal in kaart worden gebracht, niet past bij de samenleving die ik voor sta.”*

De autoriteit persoonsgegevens (AP) heeft geïnventariseerd of en in welke sectoren camera’s met gezichtsherkenning worden gebruikt. In de resultaten zag de AP aanleiding om eind oktober 2020 een waarschuwing te publiceren tegen het gebruik van camera’s met gezichtsherkenning (Autoriteit Persoonsgegevens, 2020).

In diezelfde brief van de minister wordt ook bereidwilligheid aangegeven voor de verdere ontwikkeling van de technologie.

> *Ik sta open voor een verdere ontwikkeling van deze technologie. […] Bij de toekomstige inzet van gezichtsherkenningstechnologie zal veel afhangen van de wijze van inzet. Het gaat dan niet om het enkele feit dat er gezichtsherkenningstechnologie wordt ingezet, maar om hoe dat wordt gedaan en welke waarborgen worden ingebouwd om op een zorgvuldige wijze om te gaan met de inzet en de analyse.*

---

[3] In 2014 opperde de VVD in Amsterdam om door middel van pilots te onderzoeken of gezichtsherkenning leidt tot een hogere pakkans van overtreders van gebiedsverboden (VVD Amsterdam, 2014).

Ook in de wetenschappelijke gemeenschap (ACM, 2020) en bij toeleverende industrie leven zorgen. Bedrijven als Microsoft en IBM stellen de ethiek expliciet aan de orde en pauzeren in bepaalde landen de levering van gezichtsherkenning.

De rol van deze technologie in de samenleving verandert (Keymolen, Noorman, Van der Sloot, Cuijpers, & Koops, 2020). Zo was de toegangscontrole bij ADO Den Haag in een periode vanaf 2008 tot ongeveer 2012 gebaseerd geweest op een combinatie van bezit (kaartje) en biometrie (gezicht) om daarmee de beveiliging te verbeteren. Schiphol gebruikt irisherkenning in het *registered traveler program* Privium (Schiphol, 2020), en experimenteert nu ook met gezichtsherkenning voor alle vertrekkende passagiers in *Seamless Flow* (Schiphol, 2019). Burgers gebruiken smartphones met biometrie en onder de vlag van *smart cities* worden ook voor het veiligheidsdomein nieuwe concepten ontwikkeld. Sensoren worden via de *smart infrastructure* mogelijk ontsloten voor toepassingen van anderen, zoals van de politie of van een evenementenorganisatie. Dit roept vragen op gebied van privacy op.

Figuur 1 Fans van de Schotse club Celtic organiseerden in 2016 een protest tegen het vermeende voornemen van de politie om gezichtsherkenning te gebruiken rond voetbalwedstrijden (Daily Record, 2016).

Burgers hebben er begrip voor dat in bepaalde gebieden en onder bepaalde omstandigheden, gebruik wordt gemaakt van sensoren, mogelijk zelfs van biometrie (Snijders, Biesiot, Munnichs, & Van Est, 2019). Maar daar hoort wel bij dat dit op verantwoorde wijze gebeurt. In (Keymolen, Noorman, Van der Sloot, Cuijpers, & Koops, 2020) wordt daarom gesteld dat

> *… de samenleving nu de fundamentele vraag dient te stellen: "wat vinden wij wenselijk als het gaat om gezichtsherkenningstechnologie in onze democratische rechtsstaat?"*

Eén van hun adviezen is om daartoe de mogelijkheden van *privacy-by-design* te verkennen.

1.1.5 Gezichtsherkenning in een breder perspectief

Niet-coöperatieve gezichtsherkenning (NCG) kan worden gezien als een vorm van *sensing*. Sensing is het waarnemen met technische middelen ten behoeve van een (veiligheids)opvolging. Beleid over sensing is dus ook relevant voor gezichtsherkenning, zoals de brief van de minister van Justitie & Veiligheid (Van der Steur, 2015). TNO heeft eerder zeven richtinggevende sensing-principes beschreven voor de toepassing van sensing in het veiligheidsdomein, waaronder ook op gebied van privacy (Van Rest & Weima, 2018). Deze principes zijn daarom ook gebruikt om binnen deze studie de richting en aanpak te verfijnen (zie bijlage B “Sensing principes toegepast op niet-coöperatieve gezichtsherkenning”).

1.2 Probleemstelling

Het overwegen van het gebruik van gezichtsherkenning voor het herkennen van personen die bekend zijn bij een veiligheidsorganisatie brengt verschillende soorten uitdagingen met zich mee. Allereerst zijn over mogelijke toepassingen van gezichtsherkenning wel allerlei ideeën, maar die worden in het publieke debat typisch onder één noemer geschaard, zoals “live gezichtsherkenning” in de brief van de minister (Grapperhaus, 2019). Hierdoor is het lastig om de maatschappelijke waarde van gezichtsherkenning concreet te benoemen.

De tweede uitdaging met gezichtsherkenning is dat er concrete voorbeelden van specifieke privacydreigingen zijn waar bestaande gezichtsherkenningsystemen onvoldoende bestand tegen lijken te zijn. Bijvoorbeeld kan de verzameling templates van gezochte personen (de *watchlist*) -ondanks beveiligingsmaatregelen zoals toegangsbeheer en encryptie - via kwetsbaarheden in de ICT beveiliging toch uitlekken. De impact hiervan kan heel groot zijn, o.a. omdat mensen geen nieuw gezicht kunnen laten groeien. Een ander voorbeeld is wanneer disproportioneel veel mensen (herhaaldelijk) onterecht herkend worden (foutieve identificatie) waardoor zij onterecht geconfronteerd worden met een vorm van opvolging vanuit een veiligheidsorganisatie. Dit zijn slechts twee voorbeelden van privacydreigingen. In hoofdstuk 5 wordt hier verder op in gegaan.

Tenslotte is het wel duidelijk dat er verschillende soorten privacydreigingen zijn, zoals hierboven beschreven, het ontbreekt echter aan een compleet overzicht Wanneer verwezen wordt naar “de privacydreigingen” van (niet-coöperatieve) gezichtsherkenning, is het dus niet direct duidelijk welke specifieke dreigingen dan bedoeld worden. Daardoor is het onmogelijk om een compleet beeld te krijgen van de risico’s van gezichtsherkenning en om de kans en impact daarvan in te schatten. Ook is het dan niet mogelijk om te bepalen welk deel van het probleem eventuele tegenmaatregelen (zoals *privacy-enhanced technologies*) dan oplossen.

Deze uitdagingen werken als barrières om een inhoudelijk en zinnig gesprek te voeren over de (on)wenselijkheid van gezichtsherkenning, waardoor mogelijk kansen worden gemist om Nederland op een verantwoorde manier veiliger te maken.

1.3 Oplossingsrichting

De oplossingsrichting die in deze studie is verkend, bestaat uit verschillende onderdelen. Ten eerste geeft het onderzoek richting daar waar het gaat om het

bieden van overzicht en structuur op zowel toepassingsbieden (uitdaging 1), als privacydreigingen (uitdaging 3).

Voor een oplossing voor een aantal specifieke privacy dreigingen (deel van uitdaging 2) wordt inspiratie gezocht in een bestaande technische oplossingsrichting gebaseerd op veilige verwerking van gegevens: *Multi-party computation* (MPC).

Het lijkt verstandig om technologie die de privacy beschermt breed toegankelijk te maken, zoals in internationale standaarden. Het ligt dus voor de hand dat projectpartners de projectresultaten, bij gebleken succes, inbrengen in standaardisatieprocessen van surveillance- en gezichtsherkenningssystemen (ISO, EN). Het is daarom voor de oplossingsrichting van ondergeschikt belang of *unieke* technologie wordt ontwikkeld.

1.4 Doel en afbakening

Het doel van politie, gemeente en Johan-Cruijff Arena (JCA) met dit project is dat niet-coöperatieve gezichtsherkenningstechnologie in de vorm van een experiment en een demonstratie beschikbaar komt, waarbij die technologie de privacy van betrokkenen zo goed mogelijk beschermt. De verwachting is dat aan de hand van een dergelijke demonstratie het gesprek over het nut en de risico's van niet-coöperatieve gezichtsherkenning beter kan worden gevoerd, en dat niet-coöperatieve gezichtsherkenning in de openbare ruimte daardoor in specifieke situaties en na zorgvuldige afweging mogelijk een waardevolle, veilige en acceptabele maatregel kan worden.

In deze afweging is subsidiariteit een belangrijk principe. Dat betekent dat het *niet* gebruiken van gezichtsherkenning, maar in de plaats daarvan een lichter middel met een equivalent resultaat, ook altijd een optie moet zijn. Deze studie gaat daarom zijdelings ook in op dergelijke mogelijke lichtere alternatieven, waaronder manuele herkenning, en de technologie *herherkenning* – een vorm van zachte biometrie. Dit is een vorm van persoonsherkenning die het hele uiterlijk gebruikt, en juist niet het gezicht.

Andere vormen van niet-coöperatieve biometrische herkenning zijn buiten scope van deze studie, met name omdat zij erg weinig onderscheidend of moeilijk te meten lijken te zijn. Een voorbeeld daarvan is biometrie op basis van somatotype, d.w.z. de vorm van het lichaam. Ook het aflezen van andere informatie van het gezicht, zoals emotie, is buiten scope.

1.5 Onderzoeksvragen

Gezichtsherkenningstechnologie bestaat reeds. Ook is bekend hoe deze in het algemeen kan worden ingezet voor allerlei soorten toepassingen in de (semi-)openbare ruimte. Tevens is bekend dat er onderzoek wordt gedaan naar manieren om deze herkenningstechnologie toe te passen waarbij verschillende facetten van privacy extra worden beschermd. Dit gebeurt onder verschillende noemers zoals voor gezichtsherkenning bijvoorbeeld *secure face recognition*, *privacy enhanced face recognition*, en *privacy preserving face recognition*.

Niet bekend is welke verschillende soorten privacybedreigingen allemaal gemitigeerd kunnen worden door de inzet van dit soort technologieën bij de mogelijke toepassing van niet-coöperatieve gezichtsherkenning. Ook is niet bekend welke *privacy-preserving* (gezichts)herkenningstechnologie voldoende volwassen is (*technology readiness level*: TRL6+) om binnen korte termijn demonstrabel te zijn. En tenslotte is niet bekend welke functionele architectuur nodig is om dit in relevante contexten inzetbaar te maken. Daardoor is ook niet duidelijk of één overkoepelende architectuur mogelijk en wenselijk is voor alle relevante contexten, of dat dit meer een familie van privacy preserving gezichtsherkenningsarchitecturen of -standaarden moet zijn.

De kennisvragen voor (gezichts)herkenning zijn:

1 Welke operationele contexten zijn potentieel van belang als het gaat om (het maatschappelijke debat over) (gezichts)herkenning in de Nederlandse (semi-)openbare ruimte? Hoe zijn deze contexten gerelateerd aan de JCA?
2 A. Welke privacydreigingen moeten adequaat worden gemitigeerd als het gaat om (gezichts)herkenning in de Nederlandse (semi-)openbare ruimte? Gaat dat o.a. om het voorkomen van misbruik?
   B. Is er een gangbare conceptualisatie van privacy die geschikt is om dat mee te duiden?
3 A. Welke *privacy-preserving* (gezichts)herkenningstechnologie is voldoende volwassen (TRL6+) om tijdig operationeel inzetbaar te zijn? Is multi-party computation (MPC) hiervoor geschikt?
   B. Wat is de Intellectual Property (IP) situatie van relevante technologieën?
4 A. Welke kenmerken van de live situatie[4] zijn relevant?[5]
   B. Welke kenmerken kunnen automatisch worden bepaald uit het beeld of uit andere databronnen (zoals een weerstation)?
   C. Hoe kan die informatie worden gebruikt om te voorkomen dat onnodig veel mensen onterecht worden herkend, zonder dat het eenvoudig wordt om het (gezichts)herkenningssysteem te omzeilen?

De onderzoeksvragen worden beantwoord in hoofdstuk 5.

1.6 Leeswijzer

Dit rapport is de verslaglegging van het project gezichtsherkenning in de samenwerking tussen gemeente Amsterdam, politie, JCA en TNO. Het is daarmee ten eerste bedoeld als kennisborging en als gestructureerd kennisoverzicht voor eventueel vervolgstudies. Het is tevens de inhoudelijke verantwoording van de besteding van middelen.

Het is geschreven voor een inhoudelijk deskundig publiek, en gaat dus snel de diepte in. Voorkennis wordt verondersteld over politiewerk, beveiliging, gezichtsherkenning en privacy. Dit zijn specialismes op zich, en weinig mensen zijn specialist in ieder van deze onderwerpen. Daarom is benodigde voorkennis toegevoegd in bijlagen.

[4] De live situatie is de situatie waarin opnames van gezichten worden gemaakt van passanten die met gezichten uit eerder opgenomen enrolment beelden worden vergeleken.

[5] Denk hierbij o.a. aan kenmerken van het data-subject, omgevingsfactoren (zoals belichting en weersinvloeden) en kenmerken van de optische verwerkingsketen (zoals camerapositie).

Het rapport bevat geen gerubriceerde gegevens, en kan dus publiek worden gemaakt.

Hoofdstuk 1 geeft de introductie met daarin context, probleemstelling, oplossingsrichting, doel en afbakening. Hoofdstuk 2 introduceert de projectpartners als vertegenwoordigers van bepaalde belanghebbenden bij de uitdaging. Hoofdstuk 3 beschrijft de aanpak. Hoofdstuk 4 beschrijft de basis van de technologie van niet-coöperatieve gezichtsherkenning, mede in relatie tot mogelijke toepassingsgebieden. In hoofdstuk 5 wordt *privacy by design* toegepast op niet-coöperatieve gezichtsherkenning. Dat gebeurt achtereenvolgens door relevante privacy dreigingen te identificeren en evalueren, door daar privacy beschermende strategieën aan te koppelen, en door *privacy enhancing / preserving* technologie te identificeren en beschrijven. Hoofdstuk 6 gaat dieper in op de oplossingsrichting *multi-party computation* op basis van een literatuurstudie en marktscan, en een experiment met die technologie. Hoofdstuk 7 beschrijft twee toepassingsconcepten in meer detail, waaronder één voor een digitale perimeter, en één over bewaken en beveiligen. Hoofdstuk 8 beschrijft de evaluatie van de geschiktheid van een geopperde *use case* als context voor een praktijkexperiment. Hoofdstuk 9 bevat de antwoorden op de onderzoeksvragen en discussie, en hoofdstuk 10 bevat de conclusies en aanbevelingen.

1.7 Terminologie

In deze sectie wordt een selectie van relevante termen toegelicht. **Vetgedrukte** teksten worden elders in de terminologie gedefinieerd. Waar nodig wordt aangegeven dat de definitie specifiek is voor dit rapport.

De terminologie rond gezichtsherkenning en surveillance is vaak in het Engels. Waar mogelijk is in onderstaande tabel ook een Nederlandse term gegeven. Voor de leesbaarheid wordt in het rapport de in Nederland meest gangbare term gebruikt – soms dus de Engelse.

De termen *whitelist* en *blacklist* waren gangbare termen om positieve (toestemming) en negatieve (geen toestemming) herkenning mee te benoemen. Onder invloed van de maatschappelijke discussie over discriminatie is er in de zomer van 2020 een beweging op gang gekomen om dit soort beladen termen ook in de technische sector te vervangen door neutrale termen. Deze termen worden in dit rapport buiten deze sectie dus verder niet gebruikt.

Term	Definitie
Alarmresolutie	**Herkenningstechnologie** werkt niet perfect. De "**herkenning**" kan dus onterecht zijn. Alarmresolutie heeft als doel om de resterende onzekerheid zo veel mogelijk te verkleinen en te bepalen op welke manier het proces verder moet verlopen.
Authenticatie	(In dit project) Controle of een **identiteitsclaim** overeenkomt met de daadwerkelijke persoon. Authenticatie van identiteitsdocumenten is ook mogelijk, maar is wat anders: daarbij wordt de echtheid van het document gecontroleerd.

Term	Definitie
Autorisatie	Het bepalen of iemand recht heeft op toegang tot een locatie of asset.
Biometrie	Het meten aan het lichaam van levende wezens, typisch voor **herkennen** of **identificeren** van mensen.
Biometrisch template	Een technische beschrijving van een onderdeel van het menselijk lichaam, zoals het **gezicht**, bedoeld om (binnen een bepaalde context) uniek te zijn voor een specifiek persoon.
Chilling effect	Het subjectieve effect dat mensen door de aanwezigheid van veiligheidsmaatregelen bepaalde legitieme activiteiten niet meer durven te doen.
Coöperatief	Met medewerking (b.v. mensen moeten in de camera kijken).
Enrolment	Het vastleggen van een **gezicht** om daar op een ander moment personen mee te kunnen herkennen (NL. registratie).
Foutieve identificatie	Er wordt iemand foutief **herkend** als zijnde een specifiek iemand van de ***watchlist***. *(Engels: False positive)*.
Foutieve niet-identificatie	Iemand die wel op de ***watchlist*** staat, wordt gemist in de **herkenning**. *(Engels: False Negative)*.
Gezicht	De voorkant van het menselijk hoofd. Vanaf het voorhoofd tot en met de kin, inclusief de oren.
Gezichtsherkenning	Het **herkennen** van mensen middels hun **gezicht**. In dit rapport wordt daarmee geautomatiseerde gezichtsherkenning bedoeld, tenzij dat expliciet anders staat verwoord.
Gezichtstemplate	Een **biometrisch template** voor het **gezicht**.
Green lane	Een **toegangscontrole** concept dat uitgaat van een risico inschatting per bezoeker. Bezoekers met een lage risico inschatting worden minder streng gecontroleerd, wat voor hen het effect van een rij groene stoplichten geeft.
Herkennen	Iemand herkennen uit een eerdere verzameling van mensen. Bijvoorbeeld mensen die eerder zijn gezien. Herkennen omvat zowel **verifiëren** (1:1) als **identificeren** (1:N).
Herkenningstechnologie	Alle vormen van **biometrie**, zowel harde (onveranderlijke) als zachte (veranderlijke) **biometrie**.
Identificator	Een gegeven op basis waarvan iemand uniek **geïdentificeerd** kan worden.
Identificeren	(1) Iemand herkennen uit een eerdere groep (1:N). (2) Een unieke **identiteit** (2) aan iemand toekennen.

Term	Definitie
Identiteit	(1) Dat wat iets of iemand uniek maakt. (2) De formele en administratieve registratie van iemands **identiteit** (1). (3) De persoonlijke en intieme karaktereigenschappen van een persoon.
Identiteitsclaim	Een handeling of uitspraak waarmee een persoon expliciet (“ik ben Jeroen van Rest”) of impliciet (door een identiteitsbewijs te tonen) een bepaalde **identiteit (2)** claimt.
Indirecte identificator	Een combinatie van gegevens op basis waarvan iemand **geïdentificeerd** kan worden. Zie ook **identificator**.
Live situatie	Het moment dat een **voorbijganger** langs een **gezichtsherkenning**ssysteem loopt en het systeem een opname maakt waarmee de **voorbijganger** kan worden herkend tegen een eerdere ***enrolment*** opname.
Multi-factor authenticatie	Het **authentiseren** op basis van meerdere factoren. Factoren zijn typisch bezit (een pasje), kennis (een wachtwoord) of een uniek lichaam (**biometrie**). Andere (zwakkere) factoren kunnen ook informatiewaarde hebben zoals locatie (alleen op bepaalde plekken) of tijd (alleen in bepaalde periodes).
Negatieve herkenning	Het bijhouden van een lijst van mensen die geen toegang zouden mogen krijgen. Bijvoorbeeld voor criminelen of mensen met een omgevingsverbod. Voorheen: toepassen van een *blacklist.*
Niet-coöperatief	Zonder medewerking (bijv. mensen kijken niet in de camera).
Positieve herkenning	Het bijhouden van een lijst van mensen die wel toegang zouden mogen krijgen. Bijvoorbeeld voor medewerkers. Voorheen: toepassen van een *whitelist*.
Privacy	Het recht om met rust te worden gelaten (dus zonder bespieding of beïnvloeding).
relatie (verticale - , horizontale -)	*In dit project*: Een classificatie van **surveillance**toepassingen. In een horizontale relatie observeren burgers en bedrijven elkaar. In een verticale relatie observeert de overheid burgers of bedrijven.
Surveillance	Het nauwkeurig observeren van een proces of locatie met als doel bepaalde risico’s te beheersen. In de context van veiligheid gebeurt dit typisch ten behoeve van o.a. preventie, repressie en opsporing, of de voorbereiding daarop (zoals intelligence).

Term	Definitie
Toegangscontrole	Het **identificeren** en **authentiseren** van bezoekers, het verlenen van toegang van **geautoriseerde** bezoekers, en het stoppen van **ongeautoriseerde** bezoekers.
Verificatie	Het vaststellen dat een **identiteitsclaim** klopt.
Voorbijganger	*In dit project*: iemand die langs een gezichtsherkenningssysteem loopt. Dit kan een bonafide, malafide of neutraal persoon zijn. Dit kan een bezoeker van een evenement zijn, of een reiziger bij een grenspost.
Watchlist	Een verzameling **biometrische templates** van mensen. Afhankelijk van de consequentie van een **herkenning**, kan het gaan om **positieve** of om **negatieve herkenning**.

2 Belanghebbenden

In dit project werken gemeente Amsterdam, politie en de Johan-Cruijff Arena (JCA) samen, ieder vanuit eigen belangen.

2.1 Politie

De politie gebruikt gezichtsherkenning op dit moment voor opsporing van personen die verdacht worden van een strafbaar feit of op een andere manier betrokken waren bij een politieonderzoek (NU.nl, 2020). Daarnaast verwacht de samenleving dat de politie de (on)mogelijkheden van nieuwe technologie ook verkent (Snijders, Biesiot, Munnichs, & Van Est, 2019). De minister laat daar binnen kaders ook ruimte toe (Grapperhaus, 2019).

Dit levert een aantal spanningsgebieden op. De politie moet snelheid maken en houden in het oppakken en uitproberen van (nieuwe) herkenningstechnologie, maar daarbij wel vanaf het begin menselijke waarden zoals privacy goed meenemen. Naast die menselijke waarden, moet niet de technologie maar de eindgebruiker centraal staan. Beiden, zowel de eindgebruiker, als die technologie, kunnen in de toekomst in bepaalde toepassingsconcepten in eigendom zijn van partners van de politie[6], maar toch moet de politie daar bepaalde kwaliteitscriteria bij kunnen opleggen en toetsen. Dat leidt er toe dat de politie bij de inrichting van dit soort innovatieve concepten en het toepassen van nieuwe technologie zich een beeld vormt van zowel de benodigde als de verwachte kwaliteitsaspecten, en die ook proactief deelt met partners. Voor gezichtsherkenning gaat dat onder andere over de kans op fouten (*failure to capture*, foutieve identificatie, foutieve niet-identificatie), en de verdeling daarvan over bepaalde demografieën. Dergelijke spanningsgebieden stellen eisen aan de technologie en het kennisniveau van de betrokken organisaties.

2.2 Gemeente

De gemeente heeft verschillende rollen. In beginsel komt ze op voor de belangen van alle burgers. Leefbaarheid en veiligheid moeten daarin verenigd worden. Bijvoorbeeld is het typisch een uitgangspunt dat burgers zich onbespied èn veilig moeten voelen (in de stad).

De gemeente is het gezag waaronder de politie de openbare orde handhaaft. Ze geeft ook vergunningen uit voor bijvoorbeeld evenementen en demonstraties van bepaalde omvang. Binnen het stelsel bewaken en beveiligen is de lokale driehoek onder leiding van de burgemeester verantwoordelijk voor het bewaken en beveiligen in het decentrale domein.

[6] In een bepaald toepassingsconcept kan de eerste alertering van een failure-to-capture of van een match bijvoorbeeld naar een lokale (private) beveiliger gaan. Ook in dergelijke concepten kan de politie een verantwoordelijkheid hebben over de (kwaliteit van de) gezichtsherkenning.

2.3 Infrastructuurbeheerder en evenementorganisator (JCA)

De beheerder van een evenementenlocatie en de evenementenorganisator hebben vanuit een bepaalde missie (bijvoorbeeld aantrekkelijke evenementen organiseren) een commercieel belang. Dat uit zich in de verkoop van tickets, en van bezoekers die zich voldoende veilig en comfortabel voelen om ter plekke extra aankopen te doen. Daarbij moeten ze van te voren aannemelijk kunnen maken dat ze binnen de afspraken met de gemeente en politie een veilig evenement kunnen organiseren. Het kan voor die veiligheid nuttig zijn om onderscheid te maken in soorten bezoekers.

Wachttijden voor de toegang zijn over het algemeen een nadeel. Maar een langere wachttijd kan ook gebruikt worden om aan bepaalde doelgroepen te communiceren dat ze extra in de gaten worden gehouden. Dat kan nuttig zijn bij hoog-risico doelgroepen. Dit was bijvoorbeeld een element van het gezichtsherkennings-concept bij voetbalclub ADO Den Haag.

3 Aanpak

De aanpak van deze studie bestond uit een vijftal activiteiten:
A. Inventarisatie van relevante operationele contexten,
B. *Privacy by design* workshop,
C. Literatuur en marktscan MPC,
D. Experiment MPC,
E. Synthese.

Activiteit (D) was een experiment met *proof-of-principle* technologie. De concrete invulling van dat experiment hing af van de uitkomsten van de eerdere activiteiten (A-C).

3.1 Activiteit A: Inventariseren operationele context, scenario's en mogelijke use cases

Het inventariseren en benoemen van de operationele context met bijbehorende scenario's is van belang om te bepalen welke juridische kaders van toepassing zijn, welke privacy risico's (kans en impact) in die scenario's voorstelbaar zijn, en tevens om te bepalen welke privacy beheersende strategieën en use cases voor technologieën daar (dus) nodig en mogelijk zijn. Met die informatie is het vervolgens mogelijk om te bepalen of het mogelijk is om één generiek toepasbaar privacybeschermende (gezichts)herkenningscamera te ontwikkelen, of dat het meer een familie van systemen en standaarden moet zijn. Namelijk, als er twee of meer relevante toepassingsgebieden zijn met wezenlijk verschillende privacydreigingen, dan vereisen die dus verschillende oplossingen.

Deze operationele context is bepaald op basis van een interview en gesprekken met een innovatiemanager van de politie, op basis van ongerubriceerde informatie uit andere projecten, en op basis van documentstudie.

Een selectie van negen operationele contexten wordt beschreven in bijlage E. Dit beantwoordt onderzoeksvraag 1.

3.2 Activiteit B: Privacy by design workshop

Een privacy-by-design workshop werd gehouden met politie, gemeente Amsterdam en TNO om de belangrijkste privacy risico's en privacy preserving ontwerppatronen voor (gezichts)herkenning in de (semi-)openbare ruimte te inventariseren. Dit beantwoordt kennisvraag 2 en deels ook vraag 3[7].

[7] Deze activiteit heeft tevens kennis opgeleverd waarmee de gegevensbeschermingseffectbeoordeling (GEB) van de politie ten behoeve van eventuele experimenten ondersteund kan worden.

Van een aantal privacy risico's is reeds bekend dat ze relevant zijn. Daarom kon bij deze workshop ook al een aantal experimenten worden geopperd die invulling kunnen geven aan "*privacy by design*":

1 Privacy dreiging: lekken van biometrische data (naar een niet-vertrouwde partij)
   a. Experimenteerrichting: bewaar en werk met de (soft) biometrische data op een veilige manier waardoor partners niet elkaars brongegevens krijgen.
2 Privacy dreiging: onterechte aanwijzing van onschuldig persoon als gezocht persoon
   a. Experimenteerrichting: voer geen herkenning uit in slechte condities met kans op fouten.
   b. Experimenteerrichting: gebruik van soft biometrics als aanvulling op biometrische data.
3 Onterecht gebruik van gezichtsherkenning wanneer een lichter middel ook volstaat.
   a. Experimenteerrichting: gebruik van soft biometrics in plaats van biometrische data.

In bijlage J "Suggesties voor experimenten" zijn deze suggesties uitgebreider beschreven.

De concrete keuze voor een experiment (1a, 2a, 2b en / of 3a) is bij die workshop gedaan in overleg met projectpartners. TNO heeft hierbij gelet op de uitvoerbaarheid en de wetenschappelijke validiteit. Experiment 1a "*multi-party computation*" (MPC) is gekozen als onderwerp. De concrete variant daarvan hing nog wel af van de uitkomsten van de literatuurstudie in de volgende activiteit. Door deze keuze is onderzoeksvraag 4 (over factoren die de kwaliteit van gezichtsherkenning beïnvloeden) alleen op basis van literatuurstudie en parate kennis van experts beantwoord. Van deze workshop zijn slides en een verslag beschikbaar.

3.3 Activiteit C: Literatuur en marktscan

Deze activiteit omvat een beknopte literatuur- en marktscan naar de beschikbaarheid van relevante *privacy-preserving* (gezichts)herkennings-technologie op gebied van MPC. Dit beantwoordt dus een onderdeel van onderzoeksvraag 3 (over geschikte *privacy enhancing* technologie).

Een selectie van privacy-preserving gezichtsherkenningstechnologie en een inschatting van de maturiteit is te vinden in hoofdstuk 3. Relevante publicaties zijn als referenties opgenomen in deze studie.

3.4 Activiteit D: Experiment

Middels een experiment (TRL-4) is MPC gevalideerd op technische haalbaarheid, en is de praktische bruikbaarheid verkend. Hiermee kon onderzoeksvraag 3A in meer detail worden beantwoord voor wat betreft MPC. Het experiment maakte gebruik van *proof-of-principle* varianten van *privacy-preserving* technologie binnen een bijbehorend operationeel concept. Meer informatie over het experiment is te vinden in hoofdstuk 6.

Het experiment zou offline worden uitgevoerd op vooraf opgenomen videobeelden van de JCA. Echter, door de Corona pandemie is het niet mogelijk gebleken om een grote hoeveelheid mensen bij elkaar te brengen. Noch het bijeen brengen van proefpersonen, noch het bijeen brengen van publiek om de demonstratie van technologie te aanschouwen is binnen tijdsbestek van het project mogelijk gebleken. Het experiment heeft zich daarom beperkt tot een technisch experiment op computers van TNO met behulp van een publiek beschikbare database met gezichten van beroemdheden die specifiek voor dit soort onderzoeken is opgesteld. Voor technisch onderzoek op gebied van MPC is dat overigens voldoende nuttig: het maakt voor het beproeven van de snelheid en functionaliteit van dit soort technologie niet uit waar de beelden vandaan komen.

3.5 Activiteit E: Synthese: verwerken resultaten in coherente potentiele toepassingen

Nadat alle deelresultaten bekend zijn, zijn de resultaten verwerkt in een tweetal coherente toepassingen. Deze zijn uitgewerkt in hoofdstuk 7.

4 Introductie van automatische gezichtsherkenning

In dit hoofdstuk wordt in de eerste twee secties beschreven wat gezichtsherkenning is (d.w.z. “wat is het”), en in de derde sectie hoe het kan bijdragen aan een veiliger samenleving (d.w.z. “hoe doet het deugd”). Met die kennis is het daarna in het volgende hoofdstuk mogelijk om in te gaan op de privacydreigingen en privacybeschermende maatregelen (d.w.z. “hoe deugt het?”).

4.1 De basis van automatische gezichtsherkenning

Automatische gezichtsherkenning is het automatisch herkennen van personen op basis van beelden.

De gezichtsherkenning kent tenminste twee fases: de *enrolment*-fase en de herkenningsfase. In de *enrolment* fase wordt een database gevuld met biometrische beschrijvingen van gezichten: gezichtstemplates. In de herkenningsfase worden biometrische templates van personen vergeleken met de templates in de database.

Technisch gezien, is een biometrisch template een rij (vector) van getallen. Het vergelijken van twee (of meer) gezichtstemplates gebeurt typisch door een afstand te berekenen tussen de twee betreffende vectoren. De betekenis van deze rij getallen is specifiek voor een algoritme. Er is géén internationale standaard voor een gezichtstemplate. Verschillende soorten gezichtsherkenningssoftware werken dus in principe technisch niet samen. Er zijn wel standaarden voor de beelden op basis waarvan de gezichtstemplates worden gemaakt.

In de context van gezichtsherkenning worden de volgende gegevens aangemerkt als persoonsgegevens:
- Opnames van gezichten en biometrische templates.
- Combinaties van persoonlijke attributen en overige metadata die -toepassing afhankelijk- specifiek zijn voor een individu of kleine groep personen: kleding, een sieraad, een bril, een tatoeage, gezichtsbeharing, huidskleur.
- Alle daaraan gekoppelde gegevens, zoals logs van hits en gekoppelde informatie, waaronder uiteraard ook eventuele andere identiteitsgegevens zoals identiteitsdocumenten.

Het maakt hierbij niet uit in welke fase deze gegevens worden verwerkt. Het gaat dus zowel om de *enrolment* gegevens, om de live opnames en ook om eventuele trainings- en evaluatiedata.

Meer achtergrondinformatie over gezichtsherkenning is te vinden in bijlage C.

4.2 Kenmerken van automatische gezichtsherkenning

Er zijn een aantal kenmerken waarin (de toepassing van) deze automatische gezichtsherkenningstechnologie kan variëren. Deze lijst is overgenomen uit tabel 1 van (NIST, 2017):

A. Positieve versus negatieve herkenning (dit rapport gaat over negatieve herkenning)

B. Grootte van database
C. Zoekkarakteristieken (bijv. het aantal voorbijgangers dat ook op de *watchlist* staat)
D. Geaccepteerde foutieve identificatie ratio (En. *false positive*)
E. Benodigd mensenwerk
F. Tijdsschaal voor opvolging
G. Controle over omgeving
H. Juridische gronden en relaties.

In deze sectie worden deze kenmerken verder toegelicht.

4.2.1 Positieve en negatieve herkenning

De eerste variatie is die tussen positieve of negatieve herkenning[8]. Deze studie gaat over negatieve herkenning.

Er is een verband tussen de consequentie van een match en de mate van medewerking die het subject typisch verleent aan een gezichtsherkenningssysteem. Met het kenmerk “positieve of negatieve herkenning”, kan worden beschreven dat deze relatie betrekking heeft op het (impliciete) doel van de passant, die hij uit in een (impliciete) claim. Dat werkt als volgt.

Bij positieve herkenning wordt er een positieve claim gedaan. Positieve herkenning veronderstelt een claim van het subject dat hij / zij *wel* op de lijst staat - ook als die lijst maar één item bevat. Dat is typisch een identiteitsclaim (“ik ben Jan Janssen”), maar kan ook een lidmaatschapsclaim zijn (“ik ben lid van de groep mensen die recht heeft op toegang”). Bij positieve herkenning volgt na een identificatie typisch toestemming om iets te doen, bijvoorbeeld medewerkers die toegang krijgen tot een gebouw. De claim is expliciet als hij wordt ondersteund door een ander token, zoals het noemen van een naam, of het tonen van een identiteitsdocument of bewijs van lidmaatschap. De claim is impliciet als de claim alleen blijft bij het feit dat de persoon op dat moment op die plek is. Omdat de persoon die de claim doet er baat bij heeft om herkend te worden, kan het redelijk zijn om een bepaalde mate van medewerking met de gezichtsregistratie en een leereffect te verwachten. Positieve herkenning leidt dus typisch tot coöperatieve herkenning. Aan de andere kant, er zullen ook geautoriseerde gebruikers zijn die er neutraal (vanuit verveling, afgeleid zijn) of vijandig tegenover staan (bijv. vanuit balorigheid). Positieve herkenning wordt typisch geïmplementeerd middels een verificatie, oftewel 1:1 vergelijking.

Bij negatieve herkenning is de claim negatief: ik ben *niet* een bepaalde persoon of ik behoor niet tot een bepaalde groep, en ik sta dus *niet* op de *watchlist*. Bij negatieve herkenning volgt na een identificatie typisch géén toestemming of zelfs een alarmopvolging, bijvoorbeeld van hooligans die geen toegang mogen krijgen tot voetbalstadion. Negatieve claims zijn typisch impliciet. Als de betreffende persoon terecht op de lijst staat, dan heeft hij / zij géén baat bij herkenning, en is het dus redelijk om een gebrek aan medewerking, mogelijk zelfs frustratie van het gezichtsherkenningsproces te verwachten. Negatieve herkenning leidt dus typisch

[8] NIST hanteert het onderscheid tussen negatieve en positieve claims (NIST, 2017).

tot niet-coöperatieve herkenning. Negatieve herkenning wordt typisch geïmplementeerd middels een “identificatie”[9], oftewel 1:N vergelijking.

4.2.2 Grootte van database

De tweede relevante variatie is die van de grootte van databases. Een *watchlist* waarop maar één persoon staat (zoals bij een contactverbod), zal in beginsel tot veel minder foutieve identificaties leiden omdat er nu eenmaal minder items op de lijst staan waarmee een persoon verward kan worden. Naarmate de *watchlist* langer wordt zal het aantal foutieve identificaties stijgen (zoals bij het zoeken naar 10.000 voortvluchtige veroordeelde criminelen (Schoenmakers, De Groot, Van Rooyen, Van Zanten, & Baars, 2017)).

4.2.3 Zoekkarakteristiek

De derde relevante variatie is die van de zoekkarakteristieken. Dit gaat over de voorbijgangers. Iedere voorbijganger zal immers een zoekopdracht vereisen in het systeem. Het gaat ten eerste om het aantal voorbijgangers die op de *watchlist* staan (niet te verwarren met het aantal personen die op de *watchlist* staan), en ten tweede om het aantal voorbijgangers die daar niet op staan. Met die twee getallen samen is ook de a-priori-kans te bepalen dat een voorbijganger op de *watchlist* staat. Dit is bepalend voor de frequentie van de verschillende soorten opvolging die nodig zijn. Naarmate er meer mensen langs het systeem lopen die niet in de database zitten, zullen er meer foutieve identificaties zijn, waarvoor (afhankelijk van de toepassing) ook meer opvolging georganiseerd moet worden.

4.2.4 Geaccepteerde foutieve identificatie ratio

De vierde relevante variatie is de geaccepteerde foutieve identificatie ratio (En. *false positive*). Die kan voor sommige toepassingen hoog zijn als het acceptabel is om de hoeveelheid opvolging daar ook voor te organiseren. Dit kan het geval zijn indien er een acute hoge dreiging is en er dus veel op het spel staat. Dit kan zelfs afhangen van het betreffende subject, dus voor sommige personen op de *watchlist* kan een andere foutieve identificatie worden geaccepteerd dan voor anderen.

Hier zijn verschillende -zelfs tegengestelde- redeneringen mogelijk[10]. Bijvoorbeeld, aan de ene kant: voor grote criminelen kan een slechtere foutieve identificatie ratio acceptabel zijn: het is belangrijk om ze te pakken, en het is acceptabel als we daarom ook de verkeerde personen stoppen. Maar, aan de andere kant: voor grote criminelen is alleen een betere foutieve identificatie ratio acceptabel, omdat de opvolging vanwege veiligheidsoverwegingen zwaarder moet zijn. We willen dus wel zeker zijn dat het om de juiste persoon gaat.

4.2.5 Benodigd mensenwerk

De vijfde relevante variatie is de hoeveelheid benodigd mensenwerk. Dit omvat zowel de mensen die de opvolging verzorgen na identificaties (zowel foutief als correct) en die daartoe ook paraat moeten staan als er geen identificaties zijn, alsook de mensen die het herkenningsproces eventueel (continue) moeten monitoren op correcte toepassing en gebruik. Het is immers technisch gezien

---

[9] Met identificatie wordt hier niet persé bedoeld de koppeling van een fysiek persoon aan zijn unieke identiteit. Het kan ook gaan om een veel beperktere groep mensen waarbinnen iemand herkend wordt. Bijvoorbeeld wanneer het gaat om het herkennen van iemand die mogelijk op een bepaalde plek ook eerder is gezien.

[10] Het zijn slechts dat: redeneringen. In dit project is dit niet juridisch getoetst.

triviaal om gezichtsherkenning te omzeilen door een vermomming of door gezichtsbedekkende kleding of een masker te dragen, door de andere kant op te kijken of door de optica te verstoren met lichten.

4.2.6 Tijdsschaal voor opvolging

De zesde relevante variatie is de tijdsschaal waarop een opvolging nodig is. Dat kan typisch variëren van seconden tot uren. Dit heeft ook te maken met de hoeveelheid mensenwerk. Immers, indien menselijke opvolging snel moet gebeuren dan moet die mens dus al paraat staan (te wachten). Deze tijdsschaal is ook relevant om te bepalen omdat sommige soorten privacy-beschermende maatregelen veel rekenwerk vereisen (zoals encryptie) wat zich kan vertalen in significante doorlooptijd. Snelheid van opvolging (een vorm van veiligheid) kan dan dus op gespannen voet komen te staan met privacy.

4.2.7 Controle over omgeving

De zevende relevante variatie is de mate waarin de omgeving onder controle is van de beheerder van het gezichtsherkenningssysteem. Omgevingsfactoren bepalen in grote mate de technische en operationele werking. Hieronder vallen factoren zoals belichting, blikvangers en bewegwijzering. Voor sommige toepassingsscenario's zijn de optische keten (de camera en de datatransmissie en -opslagapparatuur) niet bij dezelfde partij in beheer als het gezichtsherkenningssysteem. Dit is typisch het geval indien de optische keten voor meerdere doeleinden wordt gebruikt. De kwaliteitseisen aan de optische keten hangen af van het doel waarvoor ze gebruikt worden, en die moeten dan dus ook voor gezichtsherkenning worden vastgesteld en gemonitord.

4.2.8 Juridische gronden en relaties

Er ook verschillende juridische gronden mogelijk voor gezichtsherkenning. In hoofdlijnen kan het ten eerste gebeuren op basis van strafrecht. Dat zijn de regels waar een Nederlander zich aan moet houden, en of en hoe hij gestraft moet worden als hij dat niet doet. Dan gaat het bijvoorbeeld over de overheid die een misdaad wil voorkomen of oplossen.

Ten tweede kan het gebeuren op basis van bestuursrecht. Op basis van artikel 151c van de gemeentewet kan de burgemeester ter bescherming van de openbare orde besluiten tot toezicht middels camera's. Op basis van de vreemdelingenwet worden van vreemdelingen en van mensen die asiel hebben aangevraagd onder andere gezichten verwerkt.

Bij deze eerste twee soorten recht is sprake van een *verticale relatie* tussen een overheid "boven" die een burger "beneden" in de gaten houdt.

Ten slotte kan het ook op basis van burgerlijk recht, typisch wanneer een burger of private organisatie gezichtsherkenning toepast op een (andere) burger. Denk bijvoorbeeld aan toegangscontrole met *watchlists* van klanten die zich eerder misdragen hebben (Autoriteit Persoonsgegevens, 2020). Als het gaat om de inzet van surveillance-achtige technologie, zoals biometrie, dan heet dit een *horizontale relatie*. Deze relatie is het onderwerp van (Keymolen, Noorman, Van der Sloot, Cuijpers, & Koops, 2020).

In sommige potentiele toepassingen van niet-coöperatieve gezichtsherkenning kan er een beroep worden gedaan op meerdere juridische gronden. Bijvoorbeeld wanneer een private partij wordt overvallen door een crimineel, en deze opnames - via een private beveiliger (privaatrecht) - aan de politie worden overgedragen die er nog in de heterdaadfase mee op zoek gaat naar de verdachte (strafrecht). Of wanneer de beveiligingscamera's van een evenement (privaatrecht) worden gebruikt om bepaalde gezochte personen vroegtijdig te signaleren (strafrecht). Of wanneer de beveiligingscamera's van een OV-bedrijf (privaatrecht) worden gebruikt om een vermist persoon te vinden zonder vermoeden van criminaliteit (bestuursrecht).

4.3 Potentiele toepassingen voor niet-coöperatieve gezichtsherkenning

Het "beschermen van de privacy bij het niet-coöperatief herkennen van gezochte personen" is nogal abstract. Het is dus nodig om potentiele toepassingen duidelijk te omschrijven. Dat omvat het doel van de toepassing (bijvoorbeeld persoonsbeveiliging), eventuele alternatieven (bijvoorbeeld het herkennen met minder invasieve technieken) en overige context (bijvoorbeeld het juridisch kader). Met die omschrijvingen is het mogelijk om een goede demonstratiecontext te kiezen (zie hoofdstuk 7). Bijvoorbeeld, als de gekozen *privacy enhancing* technologie meer rekenintensief is naarmate er meer gezichten moeten worden vergeleken, dan is het dus verstandig om te beginnen met een toepassing waarin de grootte van de database laag is, het aantal passanten ook laag, en waar er (relatief) veel tijd is om een interventie te plegen. Ook is het met een dergelijk overzicht mogelijk om duidelijkheid te geven in de maatschappelijke discussie over gezichtsherkenning (zie onderzoeksvraag 1).

Daarom worden in bijlage E de volgende potentiële toepassingen[11] beschreven:

A. Object- en persoonsbeveiliging
B. Opsporing heterdaadfase
C. Opsporing *most wanted*
D. Beschermen *soft targets*
E. Handhaving contact-, winkel, OV- stadion- of gebiedsverbod
F. Toegangscontrole evenement
G. Private vervoerder richting beveiligd gebied
H. Green lane
I. Monitoren openbare (private) online platformen.

Tabel 1 beschrijft de indicatieve kenmerken van deze toepassingen. Hierbij wordt gebruik gemaakt van de structuur van (NIST, 2017) die in de vorige sectie 4.2 is geïntroduceerd.

Er zijn allerlei toepassingen denkbaar voor gezichtsherkenning in de (semi-) openbare ruimte. Ieder van de deze varianten heeft tenminste een niet-coöperatief element. Sommige hebben daarnaast ook een coöperatief element (zie kolom "Pos./Neg." in tabel 1). Iedere rij in deze tabel is uniek. Dus dit zijn wezenlijk verschillende soorten toepassingen.

[11] Dit zijn slechts *potentiële* toepassingen. Een daadwerkelijke toepassing zou bijvoorbeeld ook een ethische afweging en een uitgewerkte de juridische basis en kader vereisen, en die zijn niet in het kader van deze studie gedaan.

Tabel 1 Toepassingen van niet-coöperatieve gezichtsherkenning. De namen van de kolommen verwijzen naar de kenmerken die geïntroduceerd zijn in de vorige sectie. De waardes in de cellen zijn ruwe indicaties.

	Pos. / Neg.	Grootte (# personen)	Zoek	Accep. Fout	Mens	Tijd	Contr.	Jur.
A. Object- en persoonsbeveiliging	Beide	10	Laag-hoog	Laag-Hoog	10	Sec.	Niet / Wel	Strf
B. Opsporing heterdaadfase	Neg	1	Hoog	Hoog	10	Min.	Niet	Strf
C. Opsporing *most wanted*	Neg	1000	Hoog	Laag	1	Min.-Dgn	Niet	Strf / Best
D. Beschermen *soft targets*	Neg	100	Hoog	Laag	10	Sec.	Niet	Strf
E. Handhaving contact-, winkel, OV- stadion, of gebiedsverbod	Neg	10	Laag-Hoog	Laag	1	Min.	Niet	Burg
F. Toegangscontrole evenement	Beide	100	Hoog	Laag	1	Sec,	Wel	Burg
G. Private vervoerder richting beveiligd gebied	Beide	100	Hoog	Laag	1	Min.-Uren	Wel	Strf / Burg
H. Green lane	Beide	100	Hoog	Laag	1	Sec.	Wel	Strf / Burg
I. Monitoren openbare (private) online platformen	Neg	1000+	Zeer Hoog	Hoog	100	Dgn./ Wkn	Niet	Strf / Burg

Deze manier om toepassingen van gezichtsherkenning te beschrijven wordt ook gebruikt in de uitwerking van twee concepten in hoofdstuk 7.

5 Privacy by design voor niet-coöperatieve gezichtsherkenning

In dit hoofdstuk wordt beschreven wat privacy-by-design is, en hoe het concreet kan helpen om de privacy[12] te beschermen van mensen die te maken krijgen met niet-coöperatieve gezichtsherkenning. In dit hoofdstuk gaat het niet over specifieke potentiele toepassingen. Dat komt in hoofdstuk 7 aan de orde.

5.1 De achtergrond van privacy by design bij gezichtsherkenning

Privacy-by-design is een ontwerpstrategie die als doel heeft gedurende de hele levenscyclus van een systeem rekening te houden met privacy. Het is een vorm van *value-sensitive* design, een meer algemene ontwerpfilosofie die ook rekening houdt met andere menselijke waarden (Hoepman, 2014) (Van Rest, Boonstra, Everts, van Rijn, & Van Paassen, 2012).

Automatische gezichtsherkenning kan worden toegepast in verschillende toepassingsscenario's. De proportionaliteit en subsidiariteit van het inzetten van gezichtsherkenning verschillen per toepassingsscenario. Het is dus niet mogelijk om een eenduidige lijst van privacy vereisten op te stellen die voor alle toepassingsscenario's geldt. Beter is het om een flexibele aanpak te kiezen die voor verschillende toepassingsscenario's inzetbaar is.

In 2000 is de ontwikkeling van *privacy-by-design* begonnen (Van Rest, Boonstra, Everts, van Rijn, & Van Paassen, 2012). Dit gebeurde als reactie op het toenemende gebruik van ICT om persoonsgegevens te verwerken, en het inzicht dat menselijke waarden daarbij stelselmatig pas (te) laat in het ontwikkelproces werden meegenomen. De Canadese Ann Cavoukian introduceerde zeven *privacy-by-design* principes die als inspiratie hebben gediend voor de huidige AVG (Cavoukian, 2009). Hoewel niet zo gedetailleerd als andere modellen, zijn ze wel bruikbaar om op hoofdlijnen te beschrijven hoe in een bepaald concept met privacy wordt omgegaan (zie bijvoorbeeld verderop in hoofdstuk 7).

Nederlandse organisaties, zoals de registratiekamer, KPN en TNO zijn voorlopers geweest in die ontwikkeling. KPN ontwikkelde bijvoorbeeld de *privacam*-patenten (Van Delft, 2001), dit is een serie *privacy-enhancing technologies* (PET) die speciaal bedoeld zijn om bij het gebruik van biometrie meer rekening te houden met privacy en data bescherming[13]. Het gaat bijvoorbeeld om het detecteren van persoonsgegevens in beelden (zoals gezichten of kentekens), om die automatisch uit het beeld te knippen en apart beveiligd op te slaan. Een ander voorbeeld van die *privacam* patenten is het versleutelen van beeldgegevens op zo'n manier dat ze alleen bruikbaar kunnen worden gemaakt als meerdere partijen hun eigen sleutel inleggen.

[12] Sectie C.4 "Persoonsgegevens bij gezichtsherkenning" van bijlage C beschrijft welke soorten persoonsgegevens betrokken zijn bij niet-coöperatieve gezichtsherkenning.

[13] De PrivaCam patenten zijn bij de overname van het onderzoeksdeel van KPN in handen gekomen van TNO.

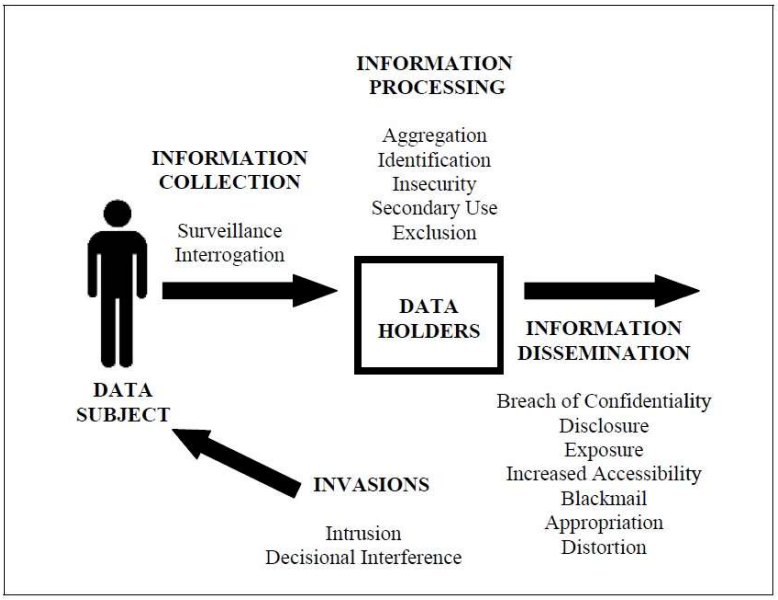
5.2 Privacydreigingen bij niet-coöperatieve gezichtsherkenning

Er zijn verschillende soorten privacy. Sommigen daarvan worden inherent door niet-coöperatieve gezichtsherkenning bedreigd. Dit is één van de redenen waarom het fundamenteel onmogelijk is om gezichtsherkenning in de (semi-) openbare ruimte te verenigen met de waarde “zonder privacy te schenden”. Anderen soorten privacy worden echter niet door de inzet van dit soort technologie bedreigd, of alleen afhankelijk van de specifieke manier van inzetten. Om hier in meer detail over te kunnen redeneren, wordt in deze sectie niet-coöperatieve gezichtsherkenning beschouwd vanuit verschillende soorten privacy.

5.2.1 Identificatie van privacydreigingen

Er zijn meerdere soorten privacydreigingen. In eerder onderzoek over privacy-by-design heeft TNO o.a. gebruik gemaakt van de taxonomie van privacydreigingen van Solove. Ook in ander recent onderzoek over horizontale relaties in gezichtsherkenning (burgers en bedrijven die elkaar observeren) is deze taxonomie gebruikt (Keymolen, Noorman, Van der Sloot, Cuijpers, & Koops, 2020). In deze taxonomie zijn zestien soorten privacydreigingen geïdentificeerd en gegroepeerd in vier categorieën (Solove, 2005) (zie figuur 2). In het Nederlands zijn de vier categorieën:

- Informatieverzameling (surveillance of ondervraging),
- Informatieverwerking (bijv., combineren, identificeren),
- Informatie disseminatie (bijv., publicatie of lekken),
- Inbreuk (beslissingen of indringing).



Figuur 2 Solove's vier soorten privacy bedreigingen.

Het is dus niet voldoende om slechts één van de privacy risico’s te beheersen. Zo is er bijvoorbeeld veel verschil tussen de volgende twee stellingen:

- We kunnen gezichtsherkenning nu ook veilig doen met partners / op platformen die we niet volledig kunnen vertrouwen.
- We beheersen de relevante soorten privacy dreigingen zo goed mogelijk.

Bij de eerste stelling wordt slechts één soort privacy dreiging geadresseerd, en bij de tweede stelling allemaal.

Ieder van deze privacydreigingen moet afdoende worden beheerst, in relatie tot het doel (proportionaliteit) en alternatieven (subsidiariteit). Daarom is het van belang om direct breder te kijken naar alle relevante privacy-risico's. In de volgende subsectie wordt de taxonomie van Solove gebruikt om de privacydreigingen verder uit te werken voor niet-coöperatieve gezichtsherkenning.

5.2.2 Evaluatie van privacydreigingen bij niet-coöperatieve gezichtsherkenning

In deze sectie worden de privacydreigingen met behulp van Solove's model geëvalueerd voor niet-coöperatieve gezichtsherkenning (NCG).

Hierbij wordt voor elke dreiging beschreven of de dreiging inherent is aan gezichtsherkenning en of de dreiging nog afhangt van de uitvoering en wijze van gebruik. (Zie tabel 2). Als de privacy dreiging (deels) inherent is, dan is het voor dat deel alleen mogelijk om de betreffende privacy dreiging te mitigeren door gezichtsherkenning niet toe te passen. Als de privacy dreiging (ook) afhankelijk is van uitvoering en gebruik, dan is er ruimte om de privacy dreiging in enige mate te mitigeren terwijl nog wel sprake is van niet-coöperatieve (gezichts)herkenning.

Tevens wordt beschreven of het een voorstelbaar risico is dat de betreffende privacy dreiging zich bij automatische negatieve gezichtsherkenning in Nederland zal manifesteren indien er géén aanvullende maatregelen worden getroffen (technisch of organisatorisch).

Tenslotte is beschreven of het betreffende risico ook wordt opgepikt in de maatschappelijke discussie. Dit wordt gedaan aan de hand van drie documenten:

- [M]: De brief van de minister[14] (Grapperhaus, 2019),
- [H]: Het rapport over horizontaal gebruik van gezichtsherkenning[15] (Keymolen, Noorman, Van der Sloot, Cuijpers, & Koops, 2020), en
- [A] Een verklaring van de vereniging ACM[16] (ACM, 2020).

Bovenstaande is in een overzicht samengebracht en te zien in tabel 2.

[14] De minister heeft in zijn brief over gezichtsherkenning in het algemeen (dus niet specifiek voor niet-coöperatieve toepassingen) enkele privacydreigingen benoemd (Grapperhaus, 2019).

[15] De privacydreigingen bij horizontale toepassing (burgers en bedrijven onderling) zijn geïnventariseerd in sectie 5.3 van (Keymolen, Noorman, Van der Sloot, Cuijpers, & Koops, 2020).

[16] De *Association for Computing Machinery* (ACM) is een toonaangevend wereldwijde wetenschappelijk community van wetenschappers. Zij verklaren dat "de technologie [gezichtsherkenning] te vaak duidelijk bias vertoont … en dat het gebruik van dit soort technologie in situaties waar het bekend of aannemelijk is dat het indruist tegen de rechten van de mens, onmiddellijk gepauzeerd moet worden." (ACM, 2020).

Tabel 2 Privacy dreigingen van niet-coöperatieve gezichtsherkenning (NCG). Ook is beschreven of ze inherent zijn, of afhankelijk van uitvoering en gebruik, en of ze leiden tot een voorstelbaar risico. Ook is aangegeven of ze zijn genoemd in de brief van de minister (M), in het rapport over horizontaal gebruik van gezichtsherkenning (H), of door de wetenschappelijke vereniging ACM (A).

	Inherent aan NCG	Manifestatie afhankelijk van uitvoering en gebruik	Voorstelbaar risico	Genoemd door …
Surveillance	X	X	Ja	M
Ondervraging			Nee	
Aggregatie	X	X	Ja	
Identificatie		X	Ja	H
Onveiligheid	X	X	Ja	M; H
Secondair gebruik		X	Ja	H
Uitsluiting		X	Ja	H
Inbreuk op vertrouwelijkheid		X	Ja	
Openbaring		X	Nee	
Blootstelling		X	Nee	
Vergrote toegankelijkheid			Nee	
Chantage			Nee	
Toe-eigening			Nee	
Vervorming	X	X	Ja	M; A; H
Indringing		X	Ja	
Beslissings-interferentie	X	X	Ja	

Zoals te lezen in tabel 2 zijn sommige privacy dreigingen inherent aan de toepassing van gezichtsherkenning. Er zijn ook privacy dreigingen waarvan de manifestatie afhankelijk is van de wijze van uitvoering, en waarvan een hoge verwachting is dat die zich in Nederland manifesteren indien geen maatregelen worden getroffen. Voor alle privacy dreigingen is het belangrijk en zinnig om privacy beschermende strategieën te ontwerpen.

Het is verstandig om die strategieën waar mogelijk af te dwingen in (de vorm van) de technologie. De kans op misbruik kan dan stelselmatig(er) worden verkleind. Indien het technisch onvoldoende mogelijk is, dan kan dat (ook) op organisatorische wijze, bijvoorbeeld door regulering, opleiding, screening van personeel, instructies, audits, toezicht, etc. Aan die organisatorische maatregelen zijn soms ook weer technische voorwaarden verbonden. Het is bijvoorbeeld mogelijk om toezicht op het gezichtsherkenningsproces te faciliteren door in de architectuur af te dwingen dat bij iedere zoekopdracht een melding (zonder persoonsgegevens) naar de toezichthouder gaat.

Een uitgebreide onderbouwing over de privacy dreigingen van automatische niet-coöperatieve (gezichts)herkenning is beschreven in bijlage G “Privacydreigingen van niet-coöperatieve gezichtsherkenning”. De relatie tussen privacydreigingen en de *privacy enhancing* technologie *multi-party computation* volgt in sectie 6.4.

5.3 Privacy-beschermende strategieën

Een privacy beschermende strategie is een aanpak die helpt om de privacy te beschermen. Het kan gebaseerd zijn op technologie, en ook op organisatorische of procesmatige ingrepen. Hoepman heeft een overzicht gepresenteerd van mogelijke

privacy beschermende strategieën (Hoepman, 2014). Dat overzicht vormt de inspiratie voor de manieren waarop de privacy dreigingen die in de vorige sectie geïdentificeerd zijn, gemitigeerd kunnen worden. Hier volgt een verkort overzicht:

- **Minimalisatie** gaat over het verwerken van persoonsgegevens van minder mensen, minder persoonsgegevens per persoon, of minder gevoelige persoonsgegevens.
- Door persoonsgegevens uit verschillende contexten **gescheiden** van elkaar te verwerken, wordt het lastiger om ze onbedoeld met elkaar te combineren.
- Door persoonsgegevens in minder detail (dus **geabstraheerd**) te verwerken, wordt als het ware “uitgezoomd” van de individuele persoon.
- Door persoonsgegevens te **verbergen**, kan worden voorkomen dat ze in verkeerde handen vallen.
- Door mensen te **informeren** over het gebruik van hun gegevens kunnen privacydreigingen worden beheerst, met name *uitsluiting*.
- Door mensen **controle te geven** over het gebruik van hun gegevens, kan met name het risico van *uitsluiting* worden beheerst.
- Door **af te dwingen** dat afspraken over privacy – ook met derden- worden nageleefd, kunnen de andere strategieën versterkt worden.
- Door gebruiksgegevens te verzamelen en extern te rapporteren kan worden **aangetoond** of privacy-beschermende strategieën werken.

Dit is in bijlage I “Privacy-beschermende strategieën” uitgebreid beschreven. Voor dreigingen die in de vorige sectie tenminste als *voorstelbaar* zijn ingeschat, bevat de volgende tabel een overzicht van mogelijke beheersmaatregelen. De relatie tussen privacybeschermende strategieën en de *privacy enhancing* technologie *multi-party computation* volgt in sectie 5.4.5.

Tabel 3 Overzicht van privacy dreiging en hoe ze beheerst kunnen worden. Een + betekent dat de strategie kan helpen om de betreffende dreiging te mitigeren. Een 0 betekent dat dat niet het geval is. Een +/- betekent dat er zowel mitigerende als verergerende effecten zijn. Een i betekent dat het effect alleen bereikt wordt in combinatie met een andere strategie.

	Minimaliseer	Scheid	Abstraheer	Verberg	Informeer	Geef controle	Dwing af	Toon aan
Surveillance	+	0	0	0	0	0	i	i
Aggregatie	+	+	+	+	0	0	i	i
Identificatie	+	+	+	+	0	0	i	i
Onveiligheid	+	+	+	+	0	0	i	i
Secondair gebruik	+	+	+	+	0	0	i	i
Uitsluiting	0	0	0	0	+	+	i	i
Inbreuk op vertrouwelijkheid	+	+	+	+	0	0	i	i
Vervorming	+/-	0	-	0	0	0	i	i
Beslissingsinterferentie	+	+	+	+	+/-	+/-	i	i

Uit dit overzicht zijn de volgende deelconclusies te trekken. De strategieën *informeer* en *geef controle* zijn de enigen die direct helpen de privacy dreiging van *uitsluiting* te beheersen. Afdwingen en aantonen helpen wel, maar alleen indien deze samen met andere strategieën toegepast worden.

De strategieën minimaliseer, scheid, abstraheer en verberg hebben breed effect op de privacy dreigingen. De strategie *minimaliseer* helpt als enige ook op de privacy dreiging van surveillance. Daarom is *select before you collect* ook voor niet-coöperatieve gezichtsherkenning een nuttig principe.

Vervorming, de kans dat het gezichtsherkenningssysteem fouten maakt, is een belangrijke privacy dreiging die prominent in de maatschappelijke discussie terug komt als het gaat om foutieve herkenning. Twee privacy-beschermende strategieën kunnen een negatief effect hebben op vervorming als ze er toe leiden dat onderscheidende informatie niet kan worden gebruikt: *minimaliseer* en *abstraheer*. Indien goed toegepast of indirect hebben *minimaliseren*, *afdwingen* en *aantonen* een positief effect op 'vervorming'. Het gaat dan om het weigeren van lage kwaliteit live beelden, en het afdwingen en aantonen van minimale kwaliteitscriteria aan de optische keten en relevante omgevingsfactoren.

Beslissingsinterferentie bij malafide personen is het afschrikkingseffect. Dat wordt typisch niet als een privacy-dreiging beschouwd. Bij bonafide personen heet het *chilling effect*. Van alle relevante privacy dreigingen lijkt beslissingsinterferentie / *chilling effect* door de meeste strategieën te beschermen te zijn. Indien het chilling effect als een belangrijke privacy dreiging wordt gezien, dan lijkt daar dus veel aan te doen te zijn.

5.4 Innovatieve privacy beschermende technologieën

Er bestaan reeds verschillende relevante privacy beschermende technologieën (En. *privacy preserving technologies - PETs*) die invulling kunnen geven aan de privacy-beschermende strategieën. In deze sectie worden een aantal innovatieve technologieën beschreven die tot extra bescherming kunnen leiden. De effectiviteit van onderstaande PETs hangt af van de manier waarop ze worden geïntegreerd met gezichtsherkenning en de wijze van gebruik. Dit is geen uitputtende verzameling.

5.4.1 Gezichtsherkenning via een online platform

Platformtechnologie is technologie die producenten van informatie koppelt aan consumenten van informatie[17]. Een database met biometrische templates, zoals die uit een *enrolment* proces komen, kan ook 'zulke informatie zijn. Dit soort platformtechnologie kan[18] helpen om de volgende privacy beschermende strategieën te implementeren: *informeer*, *geef controle*, *dwing af* en *toon aan*. Gezamenlijk helpen ze vooral de privacydreiging *uitsluiting* te beheersen indien het platform de beslissing om gezichten te delen in handen geeft van de eigenaar van het gezicht.

Centrale identiteitsplatformen zoals DigiD hebben informatie over wachtwoorden en telefoonnummers waarmee 2-factor authenticatie (kennis en bezit) kan worden verzorgd. DigiD is van de Nederlandse overheid. Er zijn ook private bedrijven en

[17] Bekende voorbeelden van platformtechnologie zijn Uber en AirBnb. Ook *password* managers zoals LastPass of authenticatie systemen zoals DigiD zijn voorbeelden van platformtechnologie.

[18] De potentie is er, maar bestaande technologie, producten en dienstverlening lijken nog niet alle mogelijkheden volledig te benutten in relatie tot gezichtsherkenning.

zelfs stichtingen (Mozilla) die online *password managers* aanbieden. Het lijkt logisch en technisch gezien goed mogelijk om dit soort platformen ook met een derde factor, een biometrisch template, uit te breiden.

Commerciële bedrijven zijn in dat gat gesprongen. Keymolen heeft bijvoorbeeld het Nederlandse bedrijf 20FACE als casus onderzocht (Keymolen, Noorman, Van der Sloot, Cuijpers, & Koops, 2020). 20FACE biedt de dienst aan om een biometrisch gezichtstemplate op veilige wijze te bewaren, en te ontsluiten[19]. De eindgebruiker bepaalt zelf welk beeld voor *enrolment* wordt aangeboden aan 20FACE. 20FACE haalt daar het biometrisch template uit (vermoedelijk in een standaard format) en verwijdert het beeld. Het template wordt versleuteld opgeslagen.

Het ter beschikking stellen van dit template aan dienstverleners (20FACE noemt dit *endpoints*) gebeurt alleen op verzoek van een *endpoint* en alleen met toestemming van de betreffende persoon. Deze weet daardoor zeker dat het alleen met zijn toestemming kan worden gebruikt, en wordt daar dus ook over geïnformeerd. De dienstverlener krijgt het biometrisch template en kan zelf een vergelijking uitvoeren.

Een alternatief zou zijn als het *endpoint* het live beeld naar het platform (zoals 20FACE) verstuurt, en dat dit platform de vergelijking doet. Nog een alternatief is dat deze vergelijking zelf in het versleutelde domein gebeurt, eventueel met hulp van nog een partij. Dan zouden het *endpoint* en het platform helemaal geen persoonsgegevens hoeven uit te wisselen (zie ook sectie 5.4.2).

De bestaande dienst van 20FACE ondersteunt een 1-op-1 vergelijking, omdat het alleen over verificatie gaat nadat de persoon op andere wijze al een identiteit heeft geclaimd. Anders weet het platform immers niet welk gezichtstemplate met dat van de dienstverlener moet worden vergeleken. In theorie kan de eigenaar van een *watchlist* aan het platform vragen om biometrische templates op hun platform te zetten. Er is dan echter geen relatie vanuit het platform naar de betreffende persoon (data subject), en dus geen mogelijkheid om te vragen om diens toestemming.

Bedrijven die zelf al biometrische templates van hun klanten hebben, zouden kunnen overwegen om (ook) van de diensten van bedrijven (zoals 20FACE) gebruik te maken. Ze worden ontzorgd voor wat betreft het maken en beheren van een biometrisch gezichtstemplate, maar ze zouden zich afhankelijk maken van een extra schakel in de keten. Voor klanten die al gebruik maakten van biometrie om zich te authentiseren bij dienstverleners betekent het dat ze dan ook een dienst afnemen van zo'n platform. Hun gezichtstemplate wordt nog steeds alleen ingezet op basis van hun consent, net zoals dat was toen het biometrisch template in bezit was van de betreffende dienstenaanbieder. Wel is het voor klanten eenvoudiger om het overzicht te bewaren en om hun toestemming in te trekken.

Dit soort platformen kunnen aantrekkelijke doelwitten worden voor criminele organisaties en statelijke actoren. De beveiliging van hun systemen zal aan hoge eisen moeten voldoen, inclusief de koppelvlakken waarmee ze met afnemers

[19] Informatie over 20FACE is beschikbaar op hun website, op YouTube en in diverse andere openbare bronnen. Ook is er in 2020 mailcorrespondentie met een vertegenwoordiger van 20FACE geweest. Een duidelijke en compacte beschrijving van hun dienst is te vinden in een *white paper* die via hun website is te verkrijgen (20FACE, 2020).

communiceren. Het is in deze studie niet uitgezocht welke eisen dat zouden moeten zijn.

5.4.2 Intelligence-on-the-edge

Een normaal surveillance systeem bestaat typisch uit meerdere camera's, centrale opslag en een centrale verwerkingseenheid. In de centrale verwerkingseenheid kan bijvoorbeeld gezichtsherkenning worden gedaan op de beelden.

Edge-computing of *intelligence-on-the-edge* plaatst een deel van die verwerking fysiek in (of vlakbij) de camera. Voor bepaalde toepassingen hoeven er dan minder beelden over het netwerk te worden verstuurd. De verwerking op de camera kan beelden filteren, zoals die waar geen gezichten op staan. Ze kan ook metadata maken dat in plaats van beelden kan worden verstuurd, zoals een biometrisch gezichtstemplate.

Intelligence-on-the-edge kan dus helpen om de privacy beschermende strategie *verberg* te implementeren. Dit helpt vooral de privacydreigingen *onveiligheid*, *secondair gebruik* en *inbreuk op vertrouwelijkheid* te beheersen. Het vereist overigens ook typisch minder bandbreedte, wat een voordeel kan zijn met betrekking tot snelheid en kosten.

Er zijn meerdere fabrikanten van camera's die *intelligence-on-the-edge* bieden. Bijvoorbeeld, een bekende fabrikant van camera's met *intelligence-on-the-edge* is AXIS. Bij hen heet dit het Axis Camera Application Platform (ACAP) (AXIS, 2020). AXIS heeft een *plug-in* model op hun camera's. Daarmee kunnen andere bedrijven software maken die draait op AXIS camera's. Eén van die andere bedrijven is Hampentech. Zij maken gezichtsherkenningssoftware die draait op de AXIS camera. Volgens de website van Hampentech:

> *"The frontend camera software performs edge processing and transfers proprietary*[20] *data, instead of video stream, back to the face recognition server for face matching. This edge processing configuration significantly reduces the network bandwidth, especially when a large number of frontend cameras are deployed."* (Hampentech, 2020)

In deze uitvoering staat de *watchlist* dus niet op de camera op straat. Deze staat op een aparte server en kan daarmee fysiek beveiligd worden. De "proprietary" data wordt dus vanuit de camera naar die server gestuurd. Het is uit deze bron niet duidelijk wat er in die data zit: het biometrisch template, een uitsnede van het beeld (in pixels), of beide. Wat deze bron ook niet vermeldt, is of de *intelligence-on-the-edge* achter een adequate beveiligingsschil zit ten opzichte van de centrale server. Het is wel aannemelijk dat er een basisvorm van informatiebeveiliging op zal zitten, en dat extra eventueel benodigde maatregelen wel te organiseren zijn.

Het is technisch gezien ook denkbaar dat de *watchlist* ook op of nabij de camera wordt geplaatst. Dat vereist wat meer opslagruimte, maar niet buitensporig veel. Het voordeel is dat gezichtstemplates van voorbijgangers niet meer de camera hoeven te verlaten. Er zijn echter twee nadelen. Ten eerste is de privacy van de mensen op de *watchlist* minder goed beschermd. Hun gezichtstemplate komt

---

[20] *Proprietary* betekent formeel "gepatenteerd", maar het kan ook betekenen dat het niet-standaard is. *Proprietary* wordt wel eens gebruikt om te suggereren dat data niet kan worden gelezen, maar dat hoeft niet te kloppen. Het is geen verstandige vorm van informatiebeveiliging – hooguit een vorm van *security through obscurity*.

immers letterlijk “op straat” te hangen. Ten tweede de conformiteit met beveiligingseisen. *Watchlists* bevatten typisch gerubriceerde informatie (want van gezochte criminelen) waarop dus zware fysieke en logische (cyber)informatiebeveiligingseisen van toepassing zijn. Eén van de belangrijkste eisen is typisch een vorm van fysieke afscherming, en dat is met een camera in de (semi-)openbare ruimte per definitie niet mogelijk. Dit leidt vermoedelijk tot de tegenstelling dat een *watchlist on-the-edge* alleen veilig lijkt voor laag-risico toepassingen, maar dat niet-coöperatieve gezichtsherkenning niet in proportie lijkt te zijn met een laag-risico toepassing.

De mate waarin het gebruiken van een dergelijk product de privacy beschermt, is beperkt. Ten eerste, het is hoogst aannemelijk bij dit soort producten dat de complete video stream alsnog ook uitgelezen kan worden. Dat zal immers een software-instelling zijn die makkelijk veranderd kan worden. Ten tweede, op basis van het biometrisch template is het nog steeds mogelijk om het gezicht (in pixels) te reconstrueren (zie sectie G.3.1.1 in bijlage G). Ten derde, als de *watchlist* niet in de camera zit, dan zal nog steeds een template van elk gezicht worden opgestuurd naar de centrale server, wat verschillende privacydreigingen met zich meebrengt. En tenslotte, met behulp van de uitvoer van zo’n *edge-camera* kan ook een kwaadwillende zoeken in verzamelingen gezichten. *Proprietary* dataformaten of algoritmes hoeven daar geen enkele bescherming tegen te bieden.

De enige privacy winst is dus dat informatie over persoonlijke attributen (kleding, haardracht, bril) rond het gezicht niet meer beschikbaar is, en de achtergrond waartegen de live opname van het gezicht is gemaakt. Het biometrisch gezichtstemplate is nog steeds te achterhalen, inclusief mogelijk ook het gezicht in pixels.

Maar in combinatie met andere technieken waarbij het biometrisch template en (daarmee) het gezicht (in pixels) verborgen worden, zoals MPC, kan het in specifieke *use cases* nuttig zijn. Het verdient dus wel een plaatsje in de grotere toolbox van *privacy-preserving* gezichtsherkenning.

5.4.3 Soft biometrics: herherkenning

Er zijn verschillende soorten *soft biometrics* die gebruikt kunnen worden om mensen te herkennen uit kleinere groepen[21]. Dit kan helpen om invulling te geven aan privacy beschermende strategieën als *minimaliseer* en *abstraheer*. Het komt er op neer dat gezichten, als directe indicatoren, niet altijd proportioneel en subsidiair zijn, en dat voor die gevallen mogelijk betere alternatieven zijn.

In deze sectie wordt daarom ingegaan op *herherkenning*: het herkennen van mensen in beelden aan indirecte identificatoren zoals kleding, haardracht, en andere persoonlijke attributen, maar *niet* het gezicht[22]. Herherkenning staat ook bekend onder de naam *forensic search* omdat het historisch gezien vooral voor forensische toepassingen toepasbaar bleek. Met de verbetering van de rekenkracht en opslagruimte wordt herherkenning ook toepasbaar voor live situaties, bijvoorbeeld als onderdeel van een video-gebaseerd tracking-systeem. De soorten factoren die de accuratesse bepalen van herherkenning komen grotendeels overeen met gezichtsherkenning, zoals belichting, opnamehoek en resolutie.

[21] Zie bijlage C, sectie C.7 voor een toelichting op *soft biometrics*.
[22] Ook bij herherkenning wordt dus gebruik gemaakt van (mogelijk gevoelige) persoonsgegevens.

Herherkenning is echter minder volwassen en gestandaardiseerd dan gezichtsherkenning. In de context van politie is herherkenning vooral bekend bij opsporingsorganisaties.

De politie heeft samen met de KMar in de Experimenteeromgeving Bewaken en Beveiligen geëxperimenteerd met tracking-technologie afkomstig van een Engels bedrijf. Daar ging het om het signaleren van afwijkend gedrag van bezoekers op beveiligd terrein. TNO ontwikkelt herherkenning ook voor specifieke toepassingen (Bouma, Borsboom, Den Hollander, Landsmeer, & Worring, 2012) en heeft er een patent op.

Ook voor herherkenning zijn trainingsdatasets nodig, waar ook persoonlijke data in zit. Het algemene uiterlijk van mensen is nodig om te trainen, maar het gezicht van mensen is voor herherkenning mogelijk niet nodig in deze trainingsdatasets. Specifiek voor herherkenning is daarom ook verkend wat de impact zou zijn op de prestaties van het (op verschillende manieren) weglaten van gezichten. Die impact blijkt laag te zijn (Van Rooijen, et al., 2020). Het trainen van dit soort algoritmes kan dus vaak ook zonder gezichtsdata.

Naast het algemene werk over herherkenning heeft TNO ook onderzoek gedaan naar de detectie van specifieke persoonlijke attributen zoals brillen en gezichtsbeharing (Bouma, et al., 2018) dat ook voor herherkenning gebruikt kan worden.

5.4.4 Managed analytics

De correcte werking van gezichtsherkenning hangt af van meerdere technische en operationele factoren zoals afstand tot het gezicht, occlusie, oriëntatie van het gezicht t.o.v. de camera, resolutie en belichting. Als die factoren bekend zijn, dan kan er worden ingegrepen als ze buiten operationele marges dreigen te raken (bijv. als het donker wordt automatisch bijlichten), en kan de verwachting over de prestaties worden bijgesteld op basis van modellen die voorspellen hoe goed gezichtsherkenning werkt onder uitdagende omstandigheden. Zo is er bijvoorbeeld kennis over gezichtsherkenning bij lage resolutie beelden (Li, Prieto, Mery, & Flynn, 2018), en over gezichtsherkenning bij occlusie (Su, Yang, Guo, & Yang, 2015). Uiteraard moet daarbij worden opgelet dat er geen technieken worden gebruikt die niet-bestaande data “invullen”.

Voor het vaststellen van relevante omgevingsvariabelen voor gezichtsherkenning heeft NIST een face recognition- evaluatie geïnitieerd van software die die beeldkwaliteitsparameters bepaalt, zoals *“… focus, illumination, distortion, and noise, and also subject-related properties like head-pose, facial expression, and eyeglasses effects*” (NIST, 2019).

Technologie kan helpen om technische en operationele factoren die de kwaliteit van gezichtsherkenning (en van andere vormen van beeldverwerking) beïnvloeden, automatisch te bepalen. Technologie kan vervolgens ook automatisch acties uitvoeren om die kwaliteit binnen gewenste marges te houden. Dit kan een invulling geven aan de privacy strategieën *minimaliseer*, *abstraheer* en *dwing af*.

TNO heeft een raamwerk voorgesteld onder de noemer *Managed Analytics* die bovenstaande functies verenigt. Dit raamwerk zou als het ware “om” reguliere

gezichtsherkenning heen kunnen worden toegepast als een soort management-laag. Mogelijk kan dit helpen om een *video-analytics* systeem semi-automatisch te beheren (Den Hollander, et al., 2017), waardoor het onder allerlei uitdagingende omstandigheden met een van te voren bepaalde minimale accuratesse zou blijven werken.

5.4.5 Vergelijken van biometrische templates op een veilige wijze

Het gebruiken van versleuteling bij het vergelijken van biometrische templates is al jaren een actief onderzoeksgebied. Het idee hierbij is om twee biometrische templates met elkaar te vergelijken terwijl die templates op een specifieke manier met semi-willekeurige getallen zijn bewerkt. De afstandsmaat verandert door de specifieke manier niet, maar het is onmogelijk geworden voor de deelnemende partijen om de biometrische templates van de andere partij te achterhalen. Dus de partij met de live beelden kan de biometrische templates van de partij met de enrolment templates niet achterhalen, en vice versa. Daarom heet dit soort versleuteling (*secure*) *multi-party computation* (MPC). MPC kan een invulling geven aan de privacy strategie '*verberg'*. Daarmee is het mogelijk om meerdere privacy dreigingen te helpen beheersen, waaronder bijvoorbeeld *secondair gebruik* en *inbreuk op vertrouwelijkheid*. Ook kan MPC helpen om *Dwing af* en *Toon aan* te implementeren.

Een variant van MPC voor kentekenherkenning was de initiële aanleiding voor deze studie, en lijkt dus ook bij nadere beschouwing nuttig. Ook is de toepasbaarheid door recente wetenschappelijk en technologische ontwikkelingen fors vergroot. In de workshop is daarom besloten om MPC verder te verkennen. In het volgende hoofdstuk wordt daar verder op in gegaan.

Er zijn ook andere varianten van veilige verwerking die mogelijk relevant zijn voor de bescherming van privacy bij gezichtsherkenning. Bijvoorbeeld *federated learning* kan worden gebruikt om een model te trainen op basis van gedistribueerde data. Mogelijk kan dit gebruikt worden om gezichtsherkenningsalgoritmes op veiliger wijze te trainen.

6 Multi-party computation voor niet-coöperatieve gezichtsherkenning

In dit hoofdstuk wordt *multi-party computation* (MPC) toegelicht, en worden de resultaten van een laboratoriumexperiment met MPC en gezichtsherkenning beschreven.

6.1 Wat is multi party computation?

Multi-party computation (MPC) is een verzameling versleutelingstechnieken die meerdere partijen in staat stellen om op veilige wijze input (informatie) te verwerken die gedistribueerd is over de partijen. Iedere partij bezit dus slechts een deel van de input. In het geval van gezichtsherkenning bevat één partij de gezichtstemplates vanuit live beelden, en een andere partij de *enrolment* templates. Veilige verwerking betekent hier dat de partijen alleen de uitkomst van de verwerking leren (identificatie of geen identificatie, eventueel met of zonder afstandsmaat), en dat alle andere informatie over de gevoelige brongegevens (de gezichtstemplates) verborgen blijven voor de andere partijen.

6.1.1 Een introductie op MPC

Het volgende voorbeeld illustreert hoe een eenvoudige berekening met behulp van MPC op een veilige wijze kan gebeuren. Stel er zijn drie personen die willen weten hoeveel ze samen gemiddeld verdienen, maar geen van hen wil informatie prijsgegeven over de hoogte van het eigen salaris. De normale berekening is:

$$\text{Gemiddeld Salaris} = (\text{Salaris A} + \text{Salaris B} + \text{Salaris C}) / 3$$

De MPC-variant hiervan kan als volgt gaan.

Stap 1:
Iedere partij A, B en C verdeelt zijn eigen salaris in drie willekeurige delen. Dan is ieder van deze delen op zich betekenisloos. Maar ze tellen wel op tot het eigen salaris. Dus het getal dat het salaris van partij A voorstelt (stel: 60.000) wordt verdeeld in:

- –23.000 (negatief);
- 17.000;
- 66.000.

Deze drie getallen opgeteld geven weer 60.000. Dat betekent dat het gemiddelde van dit ene getal niet is veranderd.

Stap 2:
Twee van deze aldus verkregen getallen worden gedeeld met de andere twee partijen. Partij A deelt bijvoorbeeld het getal 17.000 met partij B, en het getal 66.000 met partij C. En de andere partijen doen hetzelfde met hun verkregen getallen. Iedere partij heeft dus op zich betekenisloze getallen gekregen, maar nog steeds geldt dat voor het geheel van alle -nu negen- getallen, het gemiddelde niet veranderd is.

Stap 3:
De partijen tellen drie getallen bij elkaar op. Bijvoorbeeld partij A telt bij elkaar op:

- het ene getal dat partij A zelf heeft achtergehouden en dat "afkomstig" is van zijn eigen salaris: - 23.000 (negatief),
- het ene getal van partij B, en
- het ene getal van partij C.

Stap 4:
De partijen delen de resultaten van Stap 3. De partij die alle drie de totalen ontvangt kan deze bij elkaar op tellen en delen door 3. Het resultaat is – nog steeds – het gemiddelde van de drie originele salarissen.

Het is al lange tijd bekend dat iedere soort wiskundige berekening op een dergelijke veilige manier verwerkt kan worden. Echter, er zijn kosten aan verbonden in termen van benodigde rekenkracht en communicatie-overhead. Daarnaast, MPC is maatwerk. Bovenstaand voorbeeld werkt bijvoorbeeld alleen omdat het hier gaat om een gemiddelde, en omdat deze MPC-verwerking zo is ontworpen dat die uitkomst er niet door wordt veranderd. Het bepalen van het minimum is met deze werkwijze bijvoorbeeld niet mogelijk. Bij gezichtsherkenning gaat het echter niet om het bepalen van een gemiddelde van twee (of meer) gezichtstemplates.

6.1.2 MPC en niet-coöperatieve gezichtsherkenning

Bij gezichtsherkenning gaat het om het bepalen van een gelijkenismaat, inclusief eventueel het vergelijken van die maat tegen een bepaalde grenswaarde. Dat vereist een andere, meer complexe vorm van MPC.

In 2009 heeft de TU Delft samen met o.a. Philips onderzoek gedaan naar een protocol om gezichtsherkenning te doen op een server die niet kan worden vertrouwd (Erkin & Toft, 2009). Dit lukte wel, maar er was significante computerkracht voor nodig. Het Franse IDEMIA (toen Morpho) beschreef in 2013 de toenmalige stand van de techniek en beschreef hoe versimpelde afstandsmaten tussen biometrische templates minder rekenkracht vroegen, ten koste van accuratesse (Bringer, Chabanne, & Patey, 2013). In 2016 en 2019 hebben verschillende groepen Chinese wetenschappers gepubliceerd over de mogelijkheid om de benodigde rekentaken te delegeren naar een derde partij zoals een cloud-provider (Xiang & Xu, 2016) (Guo, Xiang, & Li, 2019). Een terugkerend thema in deze onderzoeken is dat verschillende partijen betrokken zijn bij de uitvoering van gezichtsherkenning, die elkaar niet volledig kunnen vertrouwen. In het voorbeeld van de Chinese onderzoeken ging het bijvoorbeeld om een *cloud-provider*, eigenaar van de *watchlist* en eigenaar van te verifiëren beelden.

6.2 Literatuurstudie MPC en gezichtsherkenning

De functie van de literatuurstudie was om te bepalen hoe "volwassen" de combinatie van MPC en gezichtsherkenning is. Daarmee kon worden ingeschat of dit een "levensvatbare" route is in deze studie, en kan later eventueel worden besloten met welke partijen moet worden samengewerkt om een succesvolle implementatie te realiseren. Om de literatuurstudie structuur te geven, zijn een vijftal hypotheses opgesteld. Deze staan hier in aflopende volgorde van "volwassenheid", met de bijbehorende deelconclusies uit de literatuurstudie.

1 Hypothese: Het is een volwassen familie van oplossingen die reeds in standaarden (ISO, EN) is uitgewerkt.
   a. Antwoord: **Vermoedelijk niet.** Er is een tiental standaarden bij ISO geïdentificeerd[23] waar mogelijk MPC in beschreven staat[24]. Er is nog geen bestaande standaard geïdentificeerd die MPC, zoals in deze studie gezocht, expliciet beschrijft. In ISO/IEC JTC 1/SC 27 zijn twee Working Groups (WG) die relevant zijn, en waar relevante standaardisatie inspanning verwacht mag worden:
      i. WG2 over Cryptography and security mechanisms
      ii. WG5 over Identity management and privacy technologies

2 Hypothese: Er is een bestaand volwassen product (incl. open source) voor de combinatie van MPC en NCG (die door een bedrijf in NL kan worden geleverd).
   a. Antwoord: **Vermoedelijk niet.** Er zijn vier bedrijven geïdentificeerd[25] die mogelijk een volwassen NCG product kunnen leveren: NEC, Idemia, D-ID en 20FACE. Géén van hen voert dit (herkenbaar) als een product. Van deze bedrijven hebben alleen NEC (NEC, 2018) en Idemia (Bringer, Chabanne, & Patey, 2013) expliciet gepubliceerd over MPC. Het is echter zeer gespecialiseerde technologie die vermoedelijk alleen als onderdeel van een groter geheel wordt gepositioneerd.Daardoor kan het lastig zijn om er specifieke informatie over te vinden. Het is dus aannemelijk dat dit rijtje niet compleet is. In dit project hebben we bestaande projecten en producten van bedrijven niet ontleed om te bekijken of ze MPC-achtige functionaliteit bieden. Ook is nog geen goede zoektocht gedaan naar patenten.

3 Hypothese: Er zijn organisaties die een (deel)oplossing voor MPC en gezichtsherkenning in realistische omstandigheden hebben beproefd.
   a. Antwoord: **Ja, de hypothese klopt.** Er zijn verschillende organisaties die projecten doen, waaronder in de Europese onderzoeksprogramma's (bijv. PROTECT, ADDPRIV, D-ID, PARIS en P5). Er zijn publicaties die overzichten presenteren van de *state of the art* (Bringer, Chabanne, & Patey, 2013) (Toli, 2018). Er is in deze studie niet uitgezocht of er al proeven zijn gedaan in realistische omstandigheden.

4 Hypothese: MPC voor gezichtsherkenning is nog in onderzoeksfase, het is nog te vroeg voor experimenten, demo's of producten.
   a. Antwoord: **Nee, dit klopt niet.** Experts wijzen op de recente voortgang in dit onderzoeksgebied die de toepasbaarheid zou moeten vergroten.

5 Hypothese: Het is een nieuw onderwerp. Dit project is de eerste keer dat naar MPC voor gezichtsherkenning wordt gekeken.
   a. Antwoord: **Deze hypothese klopt niet.** Er is al vijftien jaar onderzoek naar MPC voor gezichtsherkenning (Bringer, Chabanne, & Patey, 2013) (Toli, 2018), inclusief patenten.

Samenvattend, er zijn verschillende publicaties en meerdere patenten van meerdere onderzoeksgroepen die mogelijk relevant zijn. Verschillende bedrijven hebben mogelijk relevante volwassen technologie. Zeker is dat er tenminste twee

[23] Een voorbeeld van een mogelijk relevante standaard is ISO/IEC 24745:2011 Information technology — Security techniques — Biometric information protection.
[24] De resultaten van dit project kunnen mogelijk in een ISO standaard ingebracht worden. Dat vereist consensus van een brede groep experts van meerdere organisaties.
[25] Dit zijn gezichtsherkenningsbedrijven die in dit onderwerp actief zijn geweest. Mogelijk zijn er ook MPC-bedrijven die een toepassing rond gezichtsherkenning kunnen maken.

onderzoeksgroepen zijn (Idemia en NEC) die al meer dan tien jaar actief onderzoek naar MPC doen. Het is denkbaar dat MPC al in relevante internationale standaarden is opgenomen, maar daar is geen bewijs voor gevonden.

Tegen deze achtergrond is het nuttig om zelf ook praktijkervaring op te doen. Immers, zelfs als zou blijken dat MPC voor gezichtsherkenning al in standaarden zit, dan nog is het nodig om te weten welke vormen van MPC mogelijk zijn, welke kennis het vereist om het te implementeren en te gebruiken, en wat dit betekent (denk aan snelheid en verwerkingskracht) voor de beoogde use cases.

6.3 Experiment MPC en niet-coöperatieve gezichtsherkenning

Het doel van het experiment was om technisch inzicht te krijgen in MPC voor gezichtsherkenning. Het gaat dan met name om de benodigde verwerkingstijd en dataoverdracht van enkele moderne MPC algoritmes voor gezichtsherkenning. Hierbij is ook gekeken naar het effect van het verbergen van de afstandsmaat in de uitkomst van de verwerking[26], en naar de kracht van de gebruikte versleuteling (in lengte van de sleutel in bits)[27].

Deze experimenten zijn in laboratorium-condities gedaan. Hiervoor is een bestaand *open source* algoritme gebruikt (King, Dlib-ml: A machine learning toolkit, 2009). Dat algoritme is getraind met een publiek gezichtsherkenningsmodel (King, dlib face recognition resnet model v1, 2017). Dit aldus getrainde gezichtsherkenningsalgoritme is vervolgens toegepast op een publieke dataset met gezichten (University of Massachusetts, 2007)[28]. Hiermee is *technology readiness level* (TRL) 4 bereikt.

In het experiment zijn twee bestaande algoritmes als inspiratie genomen (Erkin & Toft, 2009) (Toli, 2018), en deze zijn door het MPC-team van TNO verbeterd. Per vergelijking leverde dat snelheidswinst en een reductie in de hoeveelheid informatieoverdracht.

Tabel 4 Resultaten voor een enkele vergelijking per oplossing.

	Erkin (face+fingerprint+iris)	Toli	TNO variant
Seconden	0,125	0,16	0,0088
Uitgewisselde MB	0,022	1,68	0,016

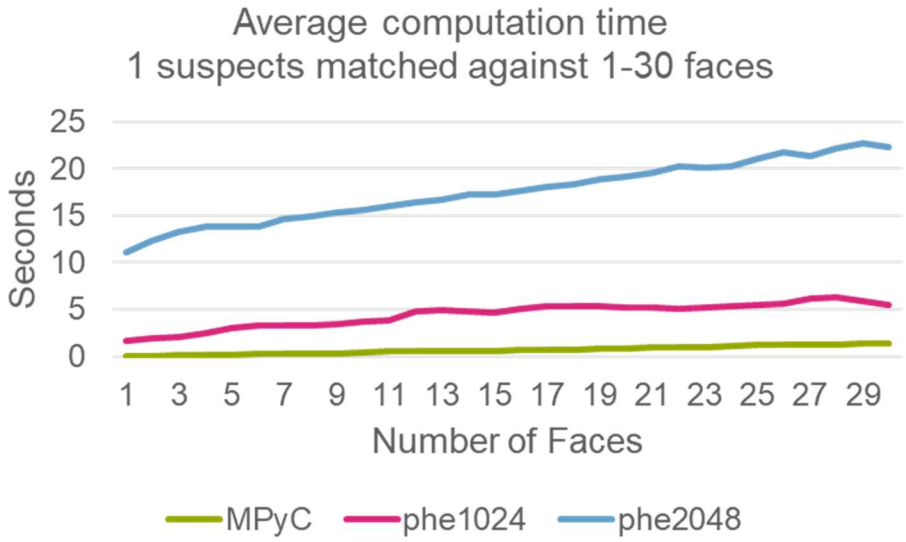
In figuur 3 is te zien dat, naarmate de *watchlist* groter wordt, de verwerkingen langer duren. Ook is te zien dat naarmate de versleuteling zwaarder is (lengte van de sleutel in bits), de verwerking langer duurt. In het meest veilige geval, bij een

[26] Met de afstandsmaat is het in theorie mogelijk om informatie over de inhoud van de *watchlist* te achterhalen. Zie bijlage G, sectie G.3.1.4.

[27] Sommige rubriceringsniveaus vereisen een bepaalde mate van versleuteling, soms uitgedrukt in de lengte van de versleutelingssleutel.

[28] Voor het trainen van modellen zijn grote, en hele goede kwaliteit datasets nodig. Er zijn geen dergelijke datasets bekend van gezichten die conform de GDPR zijn opgesteld. Het is niet mogelijk om die hoeveelheid en kwaliteit in een studie als deze te bereiken. Ook de modellen en gezichten die in deze studie zijn gebruikt zijn vermoedelijk *niet* conform de GDPR samengesteld. Dat betekent dat mensen geen toestemming hebben gegeven voor het gebruik van hun gezicht in de database, en dat er geen bewaartermijn is vastgesteld. Het is wèl mogelijk om een beperktere dataset te maken voor demonstratie doeleinden. Dat is alsnog de bedoeling bij de JCA. Zie daarvoor hoofdstuk 8.

sleutellengte van 2048 bits, is een reactietijd van tientallen seconden haalbaar bij een *watchlist* van enkele tientallen gezichten. Dat kan al voor veel soorten toepassingen nuttig zijn. In sectie 4.3 worden categorieën van toepassingen beschreven, daar zitten er meerdere bij waarvoor een reactietijd van minuten of langer nog steeds nuttig is, zoals bij opsporing, of handhaving van een gebiedsverbod.



Figuur 3 Prestaties van drie varianten van één van de algoritmes bij een groter wordend aantal gezichtstemplates op de *watchlist*.

In bijlage K “Experiments on multi-party computation for non-cooperative facial recognition” wordt meer informatie gegeven over de (uitkomsten van) de experimenten met MPC voor gezichtsherkenning.

Met een keuze voor een bepaald toepassingsgebied wordt het mogelijk om keuzes te maken in de specifieke MPC algoritmes, en wordt de bruikbaarheid ook meer duidelijk. Dit omvat beperkingen in tijd, de wijze van toezicht en de acceptatie van organisatorische en technische complexiteit.

Bijvoorbeeld, een extra partij kan helpen om de computationele complexiteit te verkleinen, maar dat kan als een ongewenste organisatorische complexiteit worden gezien. Aan de andere kant kan die rol ook goed aansluiten bij de rol van een toezichthouder: een partij die het aantal verwerkingen wil kunnen monitoren, maar niet de inhoud of resultaten.

Er kan een eerlijkheidsprincipe geldig zijn met betrekking tot wie bepaalde dure verwerkingen uitvoert. Als het technisch gezien eenvoudiger is indien er aan de kant van de *watchlist* meer verwerking plaats vindt, maar de partij met de live beelden er het meeste baat bij heeft, dan kan dit principe geschonden zijn.

Het gebruik van MPC vereist meer complexe verwerking en communicatie. Dit verlaagt de leerbaarheid, “onderhoudbaarheid”, en uitlegbaarheid en daardoor ook de *accountability* van betrokken partijen met betrekking tot het gebruik van deze technologie. Hoe makkelijk is het voor een niet technisch onderlegd persoon om te verifiëren dat de juiste vorm van MPC is gebruikt, in een actuele software versie,

goed geconfigureerd en dat het (dus) daadwerkelijk de gewenste mate van bescherming biedt? Deze vraag is niet uniek voor MPC, en er zijn bestaande manieren om dit te adresseren.

Een belangrijk aspect is de schaalbaarheid van de oplossing. Er is geen standaard voor gezichtstemplates. Een MPC variant van gezichtsherkenning is nu toegepast op een bepaald type gezichtstemplate. Daarbij is géén gebruik gemaakt van specifieke eigenschappen van deze template, dus in principe lijkt het mogelijk om de technologie ook voor andere templates te laten werken. Een ander aspect is interoperabiliteit. De vorm waarin de technologie nu is gemaakt, zou vereisen dat alle direct verbonden partijen met live- en met enrolment templates hetzelfde technische format zou moeten gebruiken. Zowel voor wat betreft het gezichtstemplate, als voor wat betreft de “MPC-schil” daar omheen. Het kan wenselijk zijn om voor beiden een standaard te gaan nastreven, zodat er zo weinig mogelijk barrières zijn om deze vorm van veilige verwerking – waar nodig – toe te passen.

6.4 Hoe helpt MPC bij niet coöperatieve gezichtsherkenning tegen bepaalde soorten privacy dreigingen?

MPC kan bij niet-coöperatieve gezichtsherkenning helpen om meerdere privacy beschermende strategieën te implementeren: “Verberg”, “Dwing af” en “Toon aan”. Zoals eerder beschreven in tabel 3 in sectie 5.3 over “Privacy-beschermende strategieën” kunnen deze strategieën *in het algemeen* helpen om meerdere privacy dreigingen te mitigeren, waaronder .

MPC is echter een bepaalde variant van deze strategie, en in dit rapport wordt ook een specifieke vorm van MPC verkend. In deze sectie wordt daarom beschreven hoe die bescherming specifiek voor deze vorm van MPC uitpakt.

Deze specifieke vorm van MPC kan de volgende privacy dreigingen als volgt helpen mitigeren:

- Aggregatie:
  - *Verberg*: Deze vorm van MPC helpt voorkomen dat andere verwerkingen kunnen worden gedaan met de gezichtstemplates. Bijvoorbeeld is het niet meer mogelijk om een template van een *enrolment* situatie te veranderen (bijv. met als mogelijk doel het te verbeteren) op basis van een gezichtstemplate van de *live* situatie.
- Onveiligheid:
  - *Verberg:* Deze vorm van MPC helpt deels tegen onveiligheid. Biometrische templates staan weliswaar nog wel leesbaar bij iedere partij, maar in de communicatie zitten ze niet, en de andere partij krijgt de eigen templates niet. Wat een partij niet bezit, kan vanuit hem ook niet lekken.
- Secondair gebruik:
  - *Verberg*: Deze vorm van MPC helpt niet tegen secondair gebruik. Er is geen mechanisme dat een tweede vergelijking onmogelijk maakt, of dat toetst voor welk doel een specifieke verwerking wordt gedaan. Het natuurlijk mogelijk om gezichtstemplates en beelden na een vergelijking te verwijderen, maar dat is niet inherent aan MPC.

  - *Toon aan*: Het is mogelijk om een afschrift van iedere verwerking (zonder inhoudelijke informatie) naar een toezichthoudende partij gaat, die dan een toets zou kunnen doen.
- Inbreuk op vertrouwelijkheid:
  - *Verberg*: Deze vorm van MPC helpt tegen inbreuk op vertrouwelijkheid. De partij waarmee wordt samengewerkt krijgt geen directe inzage in de eigen gezichtstemplates. Een variant van de onderzochte vorm geeft ook geen afstandsmaten van *matches* terug, en maakt het daarmee ook moeilijk om daar de gezichtstemplates uit af te leiden (zie ook sectie C.4.1).
- Beslissingsinterferentie
  - *Verberg:* Het is onduidelijk of deze vorm van MPC helpt tegen beslissingsinterferentie. Aan de ene kant worden mensen daadwerkelijk gecontroleerd tegen een *watchlist*. Aan de andere kant kunnen bepaalde privacy dreigingen (zoals inbreuk op vertrouwelijkheid) hierdoor beter worden beheerst.

De strategieën *Toon aan* en *Dwing af* kunnen ook worden ingevuld buiten de functionaliteit van de technologie zelf. Bijvoorbeeld kan de broncode openbaar worden gemaakt (Toon aan), of kan dit zelfs verplicht worden gesteld voor bepaalde toepassingen (Dwing af).

MPC kan vooral helpen tegen bepaalde “modussen operandi” van bepaalde soorten “dreigende actoren”: van allerlei soorten onbetrouwbare medewerkers van de eigen organisatie of van partners, zoals onvoorzichtige, fanatieke, discriminerende of corrupte medewerkers. Ook kan het helpen tegen privacy dreigingen afkomstig van zware criminelen of terroristen of van statelijke actoren. Zie bijlage D “*Design basis threat* voor gezichtsherkenning” voor een beschrijving van dergelijke “dreigende actoren”.

6.5 Functioneel ontwerpen ter inspiratie voor vervolgexperiment

Op basis van de inzichten uit dit rapport, en de resultaten van het MPC experiment is het mogelijk om enkele functionele ontwerpen voor vervolgexperimenten te formuleren. Deze ontwerpen geven inzicht in de vraag “hoe verder?”

In alle varianten kan het nuttig zijn om ook te verkennen of MPC ook kan werken met versleutelde gezichtstemplates. Dan zijn de gezichtstemplates dus niet alleen in de communicatie veilig, maar ook in de lokale opslag. Dat helpt tegen de dreiging *Onveiligheid*.

6.5.1 MPC voor gezichtsherkenning

Voor een vervolgexperiment kan de functionele architectuur van het laboratoriumexperiment worden hergebruikt. MPC voor gezichtsherkenning zou dan in de back-office draaien van twee (of drie) deelnemende partijen.

6.5.2 Federatieve MPC voor gezichtsherkenning en voor andere modaliteiten

Het is ook denkbaar dat er meerdere partijen met *watchlists* zijn, en meerdere partijen met live camera’s. Dan zou een *many-to-many* variant van MPC nodig zijn. Dit kan nuttig zijn in omgevingen waar meerdere hoog-risico objecten (van verschillende beheerders) bij elkaar staan. Wellicht is dit (ook) toepasselijk in een *smart city* context.

Wellicht is het nuttig om te verkennen of dit in bepaalde omstandigheden direct voor een variatie aan datatypen nuttig is, zoals ook kentekens, afgelegde routes en telefoongegevens.

6.5.3 MPC en Managed Analytics on the edge

Een volgend ontwerp zou kunnen zijn om gezichtsherkenning, MPC en een paar belangrijke beeldkwaliteitsdetectoren (bijv. kijkhoek op gezicht, resolutie op gezicht en contrast op gezicht) op een camera te implementeren. Daarmee wordt het mogelijk om data direct bij de bron te verrijken, eventueel te filteren en te metadateren. *Failure-to-capture* kan dan gedetecteerd worden, en slechte kwaliteit opnames kunnen gefilterd worden waardoor te veel loze alarmen voorkomen kunnen worden. In dit ontwerp is het niet nodig om de *watchlist* op de camera te plaatsen. Hierdoor worden mogelijk ook toepassingen mogelijk in het hogere risicospectrum.

Deze camera zou dan werken via MPC (bij voorkeur middels een open ISO standaard) samen met een back-end. Op dat back-end kan dan een *watchlist* staan, of dat kan een koppelvlak zijn naar een online platform met gezichtstemplates (zie sectie 5.4.1) voor een 1-op-1 verificatie. MPC zou bijvoorbeeld in een vorm kunnen zijn (eventueel optioneel) waarbij een derde partij toezicht kan houden op het proces van zoekopdrachten (niet de inhoud, alleen het proces), en die een soort dodemansknop heeft om het zoekproces stil te leggen.

7 Concepten voor niet-coöperatieve gezichtsherkenning

In dit hoofdstuk zijn twee concepten uitgewerkt. Eén voor evenementenveiligheid conform het idee van een Digitale Perimeter, en één voor persoons- en bijbehorende objectbeveiliging. In deze concepten wordt van alle eerder geïntroduceerde vormen van *privacy enhancing* technologie (zie sectie 5.4) gebruik gemaakt.

Aan het einde van ieder concept wordt met behulp van de zeven principes van Cavoukian (Cavoukian, 2009) op hoofdlijnen beschreven hoe in het concept met privacy wordt omgegaan.

7.1 Concept voor digitale perimeter: adaptieve herkenning bij evenementen

Hier volgt een beschrijving van een mogelijke concept voor het onderdeel gezichtsherkenning onder de noemer *adaptieve herkenning*. Dit operationele concept is bedacht voor evenementenveiligheid en geïnspireerd door het concept van de digitale perimeter (Digitale Perimeter, 2020):

> *De UEFA stelt als een van de veiligheidseisen dat er tijdens de EK 2021 wedstrijden een hek rondom alle stadions komt, zo ook rondom de Johan Cruijff ArenA. Het project de Digitale Perimeter is* […] *gestart om deze veiligheidseisen van de UEFA na te streven en tegelijk op een innovatieve manier vorm te geven aan service en veiligheid tijdens grote evenementen in de Johan Cruijff ArenA. Het idee van de Digitale Perimeter is om te onderzoeken of een zogenaamd 'virtueel hek' het gebied in te gaten kan houden met behulp van o.a. camera's en sensoren. Zodat we wel aan de veiligheidseisen kunnen voldoen, maar geen fysiek hek hoeven te plaatsen.*

Het concept is een combinatie van toepassingsscenario's D (beschermen *soft targets*) en E (toegang tot evenementen), en kan indien nodig ook scenario H (Green Lane) ondersteunen.

7.1.1 Motivatie

Adequate toegangscontrole is een essentieel onderdeel van evenementenveiligheid. Toegangscontrole heeft een duale functie: het faciliteren van toegang voor geautoriseerde personen, en tegelijkertijd het weren van ongeautoriseerde personen. De aanwezigheid van grote groepen mensen leidt daarnaast tot het risico van *soft targets*. Toegangscontrole die onvoldoende capaciteit heeft, eventueel in combinatie met piekbelasting rond de aanvang of afronding van evenementen, vergroot dat risico omdat er dan wachtrijen ontstaan. Een digitale perimeter kan gezichtsherkenning gebruiken om deze risico's te helpen beheersen, en om geautoriseerde bezoekers zo goed mogelijk te faciliteren.

Het fysieke hek dat de UEFA vereist heeft ook als functie om passanten die binnen het stadion niets te zoeken hebben, buiten te houden. Als dat fysieke hek er niet is, dan kunnen passanten dus direct tot aan de buitenmuur van de JCA komen. De digitale perimeter zou wel eens een duidelijke visuele geografische markering (bijv.

een lijn op de grond) kunnen vereisen om de werking van gezichtsherkenning duidelijk te maken aan bezoekers en passanten.

Het gebruiken van gezichtsherkenning in de openbare ruime, en in combinatie met publiek-private samenwerkingen, leidt tot allerlei privacy risico's. Daarom kan het nuttig zijn om extra technische waarborgen in te bouwen, zoals in dit concept is gedaan.

7.1.2 Uitgangspunten

Uitgangspunten van dit mogelijke operationeel concept zijn:

- Er zijn drie soorten kwaliteit opnames gedefinieerd voor herkenning. Te weten:
  - *Redelijk*, geschikt voor herherkenning (zachte biometrie)
  - *Goed*, geschikt voor gezichtsherkenning *in flow* (hard biometrie) van meewerkende subjecten.
  - *Extra goed*, geschikt voor gezichtsherkenning *in flow* (harde biometrie) van neutrale subjecten onder moeilijke operationele en technische omstandigheden.
- Alle mensen moeten live worden gedetecteerd.
  - Van mensen die hun gezicht actief onttrekken aan het systeem worden dan geen opnames gemaakt die geschikt zijn voor herkenning. Maar er moet dus wel een live detectie zijn dat er een persoon is.
  - Van iedereen die de digitale perimeter binnen komt, wordt binnen bepaalde tijd een *redelijke* live opname vastgelegd, dus tenminste geschikt voor herherkenning.
    - Optioneel wordt van mensen die zich daar aan proberen te onttrekken de identiteit gecontroleerd.
  - Live beelden die zelfs *goed* zijn of *extra goed* worden apart opgeslagen, zodat ze niet per ongeluk kunnen worden gebruikt om iemand te identificeren waarvan dat niet nodig is. Deze beelden blijven echter wel beschikbaar, voor het geval er plotseling een verhoogde dreiging ontstaat.
- Van iedereen die een beveiligd object binnen wil, zijn goede enrolment opnames beschikbaar, dus tenminste geschikt voor gezichtsherkenning.
- Van iedereen met een verhoogd risicoprofiel, zijn extra goede enrolment opnames beschikbaar, dus tenminste geschikt voor gezichtsherkenning onder lastige omstandigheden.
- Het aantal foutieve identificaties wordt zo veel mogelijk beperkt door alleen te zoeken in beelden die kwalitatief goed genoeg zijn.
- De normale bewaartermijnen worden toegepast. Er worden dus geen beelden direct verwijderd.
- Het systeem moet adaptief zijn in relatie tot een verhoogde dreiging.
- Het systeem moet kunnen dienen als privacy-preserving koppelvlak ook voor detectiemiddelen tegen andere dreigingen. Dan verzorgt het de signaalattributie voor die andere detectiemiddelen.

7.1.3 Functionele onderdelen

Dit operationeel concept omvat zowel automatische gezichtsherkenning als automatische herherkenning (zie sectie 5.4.3). We noemen dit "adaptieve herkenning" omdat de kwaliteit van de herkenning proportioneel is gemaakt ten opzichte van het concrete gedrag van de bezoeker en de situatie (bijv. de dreiging).

Ook wordt er gebruik gemaakt van *managed analytics* om te weten of de kwaliteit van de beelden en dus van de biometrische gezichtstemplates voldoende goed is (zie sectie 5.4.4).

Er wordt gebruik gemaakt van online platformen voor bezoekers die zelf meer controle over hun *enrolment* gezichtstemplate willen hebben (zie sectie 5.4.1).

Er wordt gebruik gemaakt van vergelijking in het versleutelde domein om te voorkomen dat partijen onnodig persoonlijke informatie moeten delen (zie sectie 5.4.2).

7.1.4 Algemene werking

In deze sectie wordt de basiswerking beschreven, en de werking in een aantal varianten. Het concept werkt in de openbare ruimte, dus ook wordt beschreven hoe het concept werkt voor passanten die verder niets met de situatie te maken hebben.

7.1.4.1 Basiswerking

Voorafgaand aan een bezoek worden mensen die naar binnen willen uitgenodigd om een screening en een *goede* enrolment opname te komen maken. De screening wordt gemaakt bijvoorbeeld op basis van eerdere overtredingen van de huisregels, of van signalen verkregen via *social media*. Bezoekers die een verhoogd risicoprofiel hebben worden uitgenodigd om *extra goede* enrolment opnames te komen laten maken.

De screening kan opnieuw worden gedaan als er nieuwe informatie beschikbaar is over de bezoeker. Dat kan aanleiding zijn om de bezoeker uit te nodigen om alsnog *extra goede* enrolment opnames te laten maken.

Bezoekers van de digitale perimeter rond de JCA worden live in beeld gebracht. Van iedere unieke bezoeker wordt een digitaal “trackrecord” aangemaakt. Daar worden de beelden aan gekoppeld die er van de bezoeker zijn. Daar wordt middels *managed analytics* bepaald hoe goed die beelden per unieke bezoeker zijn. Die kwaliteit wordt uitgedrukt in termen van geschiktheid voor herkenning: voor menselijke herkenning, voor automatische gezichtsherkenning en voor herherkenning. Er wordt op dit moment nog geen verband gelegd met identiteit. Op het moment dat het gedurende enige tijd niet lukt om redelijke opnames van iemand te maken, dan meldt het systeem dat automatisch, en dan kan er worden besloten om die persoon aan te spreken. Middels een bodycam kan er dan alsnog een opname van de persoon worden gemaakt, en indien passend kan er om een identiteitsbewijs worden gevraagd.

Op het moment dat een persoon toegang wil tot een beveiligd object, zoals de JCA, dan doet hij dat door een toegangsbewijs te presenteren (bijv. een ticket of RFID kaart) en een identiteitsclaim. Het toegangsbewijs wordt geauthentiseerd. Die identiteitsclaim kan ook in het toegangsbewijs verwerkt zitten (bijv. een ticket op naam) of dat kan via een online platform worden gedaan (zie sectie 5.4.1). Op basis van het toegangsbewijs, eventueel in combinatie met de identiteitsclaim, kan een verband met het bijbehorende risicoprofiel worden bepaald. Dat risicoprofiel kan zelfs dan nog worden aangepast op basis van actuele informatie, zoals het gedrag tijdens dit bezoek of recente communicatie op *social media*. Vervolgens wordt geverifieerd of er van die bezoeker al voldoende goede (dus beter dan redelijke)

live opnames zijn, passend bij zijn risico profiel: *goed* voor mensen met een normaal risico profiel, *extra goed* voor bezoekers met een hoog risicoprofiel.

Als er reeds adequate opnames zijn, dan kunnen die worden vergeleken in het versleutelde domein met het biometrisch template van een platform. Als de biometrische verificatie klopt, en de bezoeker is geautoriseerd om binnen te komen, dan kan de bezoeker direct doorlopen. Hierdoor worden voor die bezoekers wachtrijen voorkomen.

Als er nog geen goede beelden zijn, dan worden er ter plekke alsnog goede opnames gemaakt. Dit zal wat extra tijd kosten (seconden), dat bij grote aantallen mensen mogelijk tot langere wachttijden kan leiden. Deze wachttijd vormt de stimulans voor bezoekers om mee te werken aan het afstaan van adequate beelden.

7.1.4.2 Verhoogde dreiging vooraf bekend

Soms wordt een evenement georganiseerd tijdens een verhoogde dreiging waarbij de dreiging afkomstig is van bekende personen, althans er zijn biometrische gezichtstemplates beschikbaar in de vorm van een *watchlist*. Dan moet het systeem dus anders werken in het gebied rond het object. Dan kan er een extra beveiligingsring worden gemaakt waar alle bezoekers tegen die *watchlist* worden gecontroleerd. Dat gebeurt door gebruik te maken van camera's van de ArenA of de gemeente, die middels *multi-party computation* communiceren met een *watchlist* van de politie[29]. Daardoor wordt voorkomen dat beelden en biometrisch gezichtstemplates tussen hen worden uitgewisseld. Bezoekers waar geen goede opname van kan worden gemaakt (*failure to capture*), worden gesignaleerd voor alternatieve (manuele) inspectie. Uiteraard vereist dit de aanwezigheid van voldoende goede camera's in dat gebied.

Uiteraard moeten ook gezichten van mensen die zich van te voren hebben aangemeld met een gezicht, langs de *watchlist* worden gehaald. Anders wordt dat een achterdeur waarlangs een crimineel alsnog binnen kan komen.

7.1.4.3 Ad hoc verhoogde dreiging

Als er plotseling een verhoogde dreiging is binnen de digitale perimeter, bijvoorbeeld vanwege een aanslag, dan is het mogelijk om direct een lijst van de beste beelden van aanwezigen te genereren, met daaraan gekoppeld hun huidige locatie. De digitale perimeter kan dan ook als een vertrekfilter worden ingericht, waarbij het systeem een alertering geeft bij iedereen die het gebied verlaat waar géén goede beelden van zijn gemaakt. Vluchtende terroristen staan dan ofwel goed op beeld, ofwel worden aangemerkt waardoor beveiligingspersoneel het handelingsperspectief krijgt om hen extra te controleren.

7.1.4.4 Incident achteraf reconstrueren

Als er een incident is geweest, dan worden de opnames gebruikt om de betrokkenen te herkennen en / of te lokaliseren. Afhankelijk van de kwaliteit van de opnames, worden deze direct aangeboden aan de best-passende algoritmes. Dus als de opnames van hoge kwaliteit zijn (frontaal, voldoende resolutie en goede belichting) dan wordt er (ook) gezocht in de gezichtsopnames. Als de opnames van

[29] Het is een onderzoeksvraag of dit schaalbaar te maken is naar de grote aantallen mensen die op evenementen af kunnen komen.

lagere kwaliteit zijn, dan wordt er alleen gezocht in de opnames die geschikt zijn voor herherkenning met herherkennings-software. *Managed analytics* helpt om die schifting snel te maken waardoor het opsporingsproces zo gericht mogelijk gebeurt. Indien nodig kunnen alsnog alle beelden worden bekeken.

7.1.4.5 Gezichtsherkenning als koppelsysteem

Er kunnen allerlei soorten dreigingen zijn. Ook dreigingen waarbij van te voren geen gezicht of identiteit bekend is. Of dreigingen waarbij een niet-coöperatieve toepassing van gezichtsherkenning niet proportioneel is. Dreigingen kunnen en moeten dus op allerlei manieren worden gedetecteerd. Bijvoorbeeld middels sniffers om sporen van explosieven te detecteren. Of middels menselijke observatie van afwijkend gedrag. Als er een indicator van een dreiging wordt vastgesteld middels zo'n andere manier, dan is het verstandig dat die in de context van eventueel andere indicatoren kan worden bekeken. Het is ook zeer wenselijk om de tijd en ruimte te kunnen kiezen waar het beste een interventie wordt gepleegd. Maar dat vereist dat het subject kan worden gevolgd, of later kan worden herkend. Dat kan ook een functie van de adaptieve herkenningstechnologie.

Dan wordt de herkenning dus niet gebruikt om iemand tegen een *watchlist* te vergelijken, maar om te bepalen of een observatie aan een persoon kan worden gekoppeld aan een andere observatie van diezelfde persoon. De herkenningstechnologie functioneert dan als onderdeel van een koppelvlak tussen andere systemen, en verzorgt wat in jargon *signaalattributie* heet: het toekennen van een observatie (signaal) aan een persoon.

7.1.4.6 Verwerking van persoonsgegevens minimaliseren van passanten

Dit concept speelt zich af in de openbare ruimte binnen de digitale perimeter. Het alternatief voor deze digitale perimeter was om dit af te sluiten met een fysiek hek. Zonder een dergelijk hek zijn er dus ook passanten in de digitale perimeter die verder niets met de situatie te maken hebben. Dat kunnen omwonenden zijn, mensen die in die omgeving komen om te recreëren, winkelen of voor ongerelateerd werk. Die mensen komen in beeld van het (gezichts)herkenningssysteem.

Passanten die niet bij de JCA naar binnen willen, hebben geen enrolment gedaan. Als zij zich toch in de digitale perimeter begeven dan wordt er een redelijke kwaliteit opname van hen gemaakt, net zoals die nu ook wordt gemaakt op basis van het bestaande cameratoezicht.

Als ze willen, dan zouden ze op een positieve *watchlist* kunnen worden geplaatst. Als de gezichtsherkenning hen dan herkend, dan zullen ze niet worden aangesproken. Mogelijk zijn er ook andere juridische gronden op basis waarvan positieve *watchlists* kunnen worden gemaakt van passanten (zie ook bijlage I, sectie I.1.3.2).

7.1.5 Juridische inbedding

Beelden worden opgeslagen conform de normale bewaartermijnen. Daardoor is het mogelijk om iedere automatische stap ook door mensen te controleren, alvorens de operationele en juridische consequenties daarvan getrokken worden. Beelden worden beveiligd opgeslagen, en toegang tot de beelden is conform de normale toegangsregimes. Dus politie heeft toegang tot beelden in de openbare ruimte, plus

beelden van JCA (en andere objecten) op basis van een vrijwillige overdracht van JCA / evenementorganisator (zoals AJAX), of op basis van vordering. JCA / AJAX hebben toegang tot beelden van camera's op eigen domein, en tot beelden van camera's van domeinen waar ze een gerechtvaardigd zaaksbelang hebben. ARENA heeft toegang tot ticketting-gegevens. Politie heeft toegang tot camera's in het openbare gebied voor openbaar toezicht, *watchlists* van gezochte personen en tot opsporingsgegevens.

7.1.6 Privacy by design

Bovenstaand concept geeft een concrete invulling aan de zeven principes van privacy by design:

- Proactive not reactive; preventative not remedial: Het concept moet helpen privacy incidenten te voorkomen en niet (alleen maar) daarop te reageren. Dit gebeurt doordat er alleen data wordt verzameld waar een legitieme behoefte voor is.
- Privacy as the default: De gezichtssensor stuurt *by default* geen persoonsgegevens. Er is zelfs geen optie om dat te veranderen.
- Privacy embedded into design: Privacy zit geïntegreerd in het ontwerp van het concept. Privacy wordt niet afgedwongen middels simplistische deeloplossingen ("alles crypteren", of "alleen de politie heeft toegang").
- Full functionality – positive-sum, not zero-sum: Legitieme doelen kunnen normaal bereikt worden, namelijk het beveiligen van een te beveiligen persoon (TBP) ook thuis.
- End-to-end security – full lifecycle protection: Er worden geen live beelden opgeslagen. Die kunnen dus ook niet op straat komen te liggen.
- Visibility and transparency – keep it open: Het concept hangt niet af van verborgen kennis (*security through obscurity*) en kan daardoor gepubliceerd worden. Daarmee kan het ook gereviewed worden en zwakke kanten kunnen verbeterd worden. Het is ook transparant, en kan dus aan passanten en bezoekers worden uitgelegd.
- Respect for user privacy – keep it user-centric: bezoekers hoeven zich geen zorgen te maken dat er data van hen bewaard blijft. Er is geen handeling van hen nodig om data te verwijderen.

7.1.7 Benodigde soort gezichtsherkenning

De volgende tabel beschrijft volgens de structuur van sectie 4.2, welke soort gezichtsherkenning nodig is voor dit concept.

Tabel 5 Indicaties van de soort gezichtsherkenning die nodig is voor dit concept.

Aspecten van gezichtsherkenning	Indicatie
Positieve of negatieve herkenning	Beide
Grootte van database	Enkele honderden, waaronder bijvoorbeeld geradicaliseerde personen en / of gezochte criminelen.
Zoekkarakteristiek	Het aantal passanten bij een evenement is groot in verhouding tot het aantal mensen op de *watchlist*.
Acceptatie van foutieve identificatie	De acceptatie van foutieve identificatie (En. *false positive*) is laag. Mensen willen op tijd naar binnen en zullen weinig geduld hebben. Mogelijk zijn er ook complicerende factoren zoals alcohol inname en / of mobiliteitsissues waardoor mensen een lage tolerantie voor vertraging hebben. De druk op de operatie zal door het evenement al hoog zijn, ook daar zal een lage tolerantie voor veel handwerk zijn.

Aspecten van gezichtsherkenning	Indicatie
Benodigd mensenwerk	Het aantal mensen dat nodig is om de opvolging te verrichten kan in de tientallen lopen. Dat zit wellicht voor een groot deel in reeds benodigde capaciteit bij *crowd management*, handhaving openbare orde en bij toegangscontrole.
Tijdsschaal voor opvolging	Na een identificatie kan er relatief weinig tijd zijn voor opvolging. Dit hangt af van de dreiging die van de herkende persoon af komt.
Controle over omgeving	In beginsel is de omgeving buiten het evenement in beheer bij publieke partijen. De evenementorganisator zal er vanuit zijn gerechtvaardigd belang ook beperkte controle over hebben, zoals bijvoorbeeld camera's gericht op de toegang tot het evenement.
Juridische gronden en relaties	Er is zowel sprake van horizontale relaties tussen de evenementorganisator en zijn bezoekers, als verticale relaties tussen de politie en passanten.

7.2 Concept voor persoons- en bijbehorende objectbeveiliging in een rustige wijk

Hier volgt een beschrijving van een mogelijke operationeel concept voor het onderdeel gezichtsherkenning onder de noemer *gezichtsensor*. Dit operationele concept is bedacht voor persoonsbeveiliging van een te beveiligen persoon (TBP) bij zijn woning. Het zou dan worden toegepast in het publiek toegankelijke deel van de observatiering rond zijn woning. Het gaat dan dus typisch om de openbare weg in een rustige wijk, in die delen waar gezichten goed in beeld kunnen worden gebracht. Het is gebaseerd op toepassingsgebied A "Persoons- en objectbeveiliging".

7.2.1 Motivatie

Sommige dreigingen op TBP's bestaan uit specifieke personen die een aanslag op hem of haar willen plegen. Dat kan gaan om zware criminelen, terroristen of zelfs statelijke actoren. Van een deel van dergelijke dreigende personen heeft de politie een biometrisch gezichtstemplate van voldoende kwaliteit. Het beschermen van de TBP kan gebeuren door menselijke beveiligers in de buurt te stationeren. Dat vergt veel capaciteit, en maakt die beveiligers kwetsbaar. Automatische gezichtsherkenning kan een oplossingsrichting zijn voor dit probleem.

Echter, het plaatsen van gezichtsherkenning in de openbare ruimte kan op politieke en maatschappelijke weerstand stuiten. Daarom kan het nuttig zijn om extra technische waarborgen in te bouwen, zoals in dit concept is gedaan.

7.2.2 Uitgangspunten

Uitgangspunten van dit operationele concept zijn:

- Het moet onmogelijk zijn voor fanatieke of corrupte medewerkers om meer te doen met het herkenningssysteem dan hier is beschreven. *Mission creep* en *function creep* moeten dus worden voorkomen, en het systeem moet dus minimale flexibiliteit hebben.
- De sensor moet zo onderhoudsvrij mogelijk zijn. Als de sensor niet goed meer werkt, dan moet hij dat melden, inclusief informatie over waar dit vermoedelijk door komt. In deze informatie mogen geen persoonsgegevens zitten.
- De sensor is alleen bedoeld als alertering voor bewaken en beveiligen, niet bedoeld voor opsporing.

- De sensor hangt op straat en is dus in beginsel gevoelig voor diefstal, vandalisme en voor gerichte sabotage. Er mogen geen persoonsgegevens op de sensor opgeslagen staan – ook niet in versleutelde vorm.
- Mogelijk is de sensor niet in eigendom van de beheerder van de *watchlist* (d.w.z. niet van de politie).
- De *watchlist* bevat gerubriceerde informatie. Deze mag niet fysiek op de sensor staan – ook niet in versleutelde vorm.
- Het systeem mag niet onnodig gevoelig zijn voor aanvallen tegen de gezichtsherkenning door *presentation attacks*. Een *failure to capture* moet altijd worden gemeld.

7.2.3 Functionele onderdelen

Dit operationeel concept omvat automatische gezichtsherkenning, en dan in een vorm dat er geen live beelden of live gezichten uit komen. Er komen alleen uit de sensor:

- indicaties van failure-to-capture;
- hits met een verwijzing naar een gezicht op een watchlist en een gelijkenisscore.

In beide gevallen (zowel *failure to capture* als hits) komen er ook technische maten voor de kwaliteit van het beeld mee, zoals maten voor belichting, contrast, beweging, etc (zie sectie C.6 in bijlage C). Hiermee kunnen de kwaliteit van het systeem, en de wijze van gebruik, inclusief de kans op aanvallen tegen het systeem, worden bepaald (zie sectie 5.4.4).

Deze informatie wordt via *multi party computation* gecombineerd met informatie op de *watchlist*. Door het op deze manier te doen, is het niet nodig om persoonlijke informatie naar een andere partij te sturen.

Er is dus geen sprake van opslag van de live beelden, noch in de sensor, noch in de server. Daarom is de term *camera* niet in de naam gebruikt, en is die vervangen door *sensor*. Dit kan wel misleidend overkomen, omdat er natuurlijk wel degelijk een camera in moet zitten om de live beelden te maken. Ter vergelijking, het is algemeen bekend dat kenteken*camera's* ook echt beelden mee kunnen sturen ter verificatie.

7.2.4 Algemene werking

In deze sectie wordt de basiswerking beschreven, en de werking in een aantal varianten. Het concept werkt in de openbare ruimte, dus ook wordt beschreven hoe het concept werkt voor passanten die verder niets met de situatie te maken hebben.

7.2.4.1 Basiswerking

Passanten en bezoekers van de TBP worden in beeld gebracht door de gezichtssensor. Deze maakt een biometrisch template van ieder gezicht, en verstuurt dit versleuteld naar een server. Deze versleuteling is zodanig dat de server, of een afluisterende partij, noch het biometrisch template, noch het gezicht kan reconstrueren. Hier wordt een voor van *multi-party computation* voor gebruikt.

De server bepaalt een afstandsmaat tot ieder gezicht op de *watchlist*. Bij voldoende gelijkenis wordt het betreffende gezicht op de *watchlist*, en de afstandsmaat, aan een eindgebruiker, typisch een persoonsbeveiliger van de politie, gepresenteerd.

Afhankelijk van de tijd die het kost om van de gezichtssensor bij de TBP te komen, en van de tijd die het kost om een interventie te plegen, moet dit binnen enkele seconden werken, anders is het mogelijk te laat om een aanval af te wenden.

Als er een foutieve identificatie is, dan wordt de interventie dus onnodig gepleegd. Dat kan een onnodige evacuatie betekenen, of een onnodige interactie met een onschuldige passant of bezoeker. Het nadeel van deze werking is dat het niet mogelijk is om de automatische herkenning handmatig te controleren aan de hand van de live beelden – ook niet achteraf. Als dat toch gewenst is, dan kan er natuurlijk wel een reguliere toezichtscamera bij worden geplaatst.

Als er een *failure-to-capture* is dan kan dat per ongeluk komen, maar ook doordat een malafide persoon zijn gezicht opzettelijk verbergt. Een *failure-to-capture* moet dus zeer serieus worden genomen. Dit kan reden zijn om alsnog live beelden door te zetten, eventueel van een andere camera, of om voor de zekerheid toch een interventie te plegen. Een alternatieve oplossingsrichting is om ook een *white-list* op basis van andere identificatoren te gebruiken. Bijvoorbeeld het MAC-adres van smartphones van bekende passanten. Dit kan uiteraard ook (weer) omzeild worden door tegenstanders, en brengt uiteraard weer andere privacy risico's met zich mee.

7.2.4.2 Verwerking van persoonsgegevens minimaliseren van passanten

Dit concept speelt zich af in de openbare ruimte in een woonwijk. Er zijn dus ook passanten die verder niets met de situatie te maken hebben. Dat kunnen omwonenden zijn, mensen die in die omgeving komen om te recreëren, winkelen of voor ongerelateerd werk. Dat kunnen ook medebewoners zijn, bonafide gasten – genodigd of ongenodigd, bezorgers, deurverkopers, collectanten, etc. Die mensen komen in beeld van het (gezichts)herkenningssysteem.

Als deze mensen willen (consent), dan zouden ze eventueel op een positieve *watchlist* kunnen worden geplaatst. Dat kan gaan om medebewoners en vertrouwde en / of aangekondigde gasten. Als de gezichtsherkenning hen dan herkend, dan zullen ze niet worden aangesproken. Mogelijk zijn er ook andere juridische gronden op basis waarvan positieve *watchlists* kunnen worden gemaakt van passanten (zie ook bijlage I, sectie I.1.3.2).

7.2.5 Juridische inbedding

Het beveiligen van sensoren om een persoon thuis te beveiligen, is geregeld in de circulaire die het stelsel bewaken en beveiligen beschrijft. Daarin is sprake van een rijksdomein en een decentraal domein. Onder beide domeinen kunnen personen in hun woning in een willekeurige woonplaats beveiligd worden. Daar kan ook gezichtsherkenning bij worden gebruikt.

Als het wenselijk is om zo weinig mogelijk camera's op straat te hebben, dan kan er worden verkend of de functionaliteit ook kan worden gecombineerd met reguliere overzichtscamera's.

7.2.6 Privacy-by-design

Bovenstaand concept geeft een concrete invulling aan de zeven principes van *privacy by design*:

- *Proactive not reactive*; preventative not remedial: Het concept moet helpen privacy incidenten te voorkomen niet (alleen maar) daarop te reageren. Dit

gebeurt doordat er alleen data wordt verzameld waar een legitieme behoefte voor is.
- *Privacy as the default*: De gezichtssensor stuurt *by default* geen persoonsgegevens. Er is zelfs geen optie om dat te veranderen.
- *Privacy embedded into design*: Privacy zit geïntegreerd in het ontwerp van het concept. Privacy wordt niet afgedwongen middels simplistische deeloplossingen ("alles crypteren", of "alleen de politie heeft toegang").
- *Full functionality – positive-sum, not zero-sum*: Legitieme doelen kunnen normaal bereikt worden, namelijk het beveiligen van een TBP ook thuis.
- *End-to-end security – full lifecycle protection*: Er worden geen live beelden opgeslagen. Die kunnen dus ook niet op straat komen te liggen.
- *Visibility and transparency – keep it open*: Het concept hangt niet af van verborgen kennis (*security through obscurity*) en kan daardoor gepubliceerd worden. Daarmee kan het ook gereviewed worden en zwakke kanten kunnen verbeterd worden. Het is ook transparant, en kan dus aan passanten en bezoekers worden uitgelegd.
- *Respect for user privacy – keep it user-centric*: bezoekers hoeven zich geen zorgen te maken dat er data van hen bewaard blijft. Er is geen handeling van hen nodig om data te verwijderen.

Bovenstaande houdt nog geen rekening met een positieve *watchlist* zoals beschreven in sectie 7.2.4.2. Er is dus wellicht meer privacy bescherming mogelijk, maar dat vereist een zorgvuldige afweging.

7.2.7 Benodigde soort gezichtsherkenning

De volgende tabel beschrijft volgens de structuur van sectie 4.2, welke soort gezichtsherkenning nodig is voor dit concept.

Tabel 6 Indicaties van de soort gezichtsherkenning die nodig is voor dit concept.

Aspecten van gezichtsherkenning	Indicatie
Positieve of negatieve herkenning	Alleen negatieve herkenning
Grootte van database	Enkelen tot enkele tientallen, hooguit de personen die specifiek tegen deze te beveiligen persoon een dreiging vormen.
Zoekkarakteristiek	Het aantal passanten in een rustige wijk is laag, ook in verhouding tot een klein aantal mensen op de *watchlist*.
Acceptatie van foutieve identificatie	De acceptatie van foutieve identificatie (En. *false positive*) zou relatief hoog kunnen zijn, mits de alarmresolutie geen grote inbreuk vormt op de privacy van bonafide voorbijgangers (i.e. bij foutieve identificaties).
Benodigd mensenwerk	Er is een klein aantal mensen nodig voor opvolging, mits de opvolging efficiënt kan worden georganiseerd.
Tijdsschaal voor opvolging	Na een identificatie kan er relatief weinig tijd zijn voor opvolging. Dit hangt af van de onmiddellijke dreiging die van de herkende persoon af komt. Afhankelijk van de lokale situatie kan gezichtsherkenning bijvoorbeeld vooral voor het verzamelen van intelligence over voorbereidende handelingen van tegenstanders worden ingezet. Dan is er typisch meer tijd dan in een acute situatie.
Controle over omgeving	In beginsel is de omgeving in een wijk waar mensen wonen in beheer bij publieke partijen. Er kan een lage tolerantie zijn voor beveiligingsmiddelen (camera's en middelen voor alarmresolutie) bij omwonenden.
Juridische gronden en relaties	Er is vooral sprake een verticale relatie tussen de politie en passanten.

8 Geschiktheid JCA voor demonstratie van niet-coöperatieve gezichtsherkenning

Uitgangspunt bij deze studie was dat JCA een geschikte omgeving biedt voor demonstratie van nieuwe technologie. Voor dit project ging die geschiktheid met name om de volgende aspecten. Ten eerste dat het een omgeving is die fysiek kan worden afgesloten van de openbare ruimte. “Bijvangst”, i.e. verwerking van persoonsgegevens van mensen die daar geen toestemming voor hebben gegeven, kan daarmee worden uitgesloten. Ook kan worden uitgesloten dat het experiment verstoord wordt door tegenstanders, of dat gevoelige apparatuur onvoldoende beveiligd in de openbare ruimte komt te staan.

Ten tweede of de technische infrastructuur voldoende geschikt is voor niet-coöperatieve gezichtsherkenning. En ten derde dat de context en de processen die er zich afspelen zich lenen voor een niet-coöperatief gezichtsherkenningsscenario. In dit hoofdstuk worden deze laatste twee aannames uitgewerkt, getoetst aan de hand van opnames die bij de JCA zijn gemaakt en wordt op de validiteit van de aannames gereflecteerd.

8.1 Contextuele en operationele geschiktheid van JCA voor niet-coöperatieve gezichtsherkenning

In deze sectie wordt beschreven op welke manier, en in welke mate, de context en operationele processen bij de JCA geschikt zijn om de effecten van niet-coöperatieve gezichtsherkenning mee te demonstreren.

8.1.1 De context

De JCA is een aansprekende omgeving die primair als evenementengebied wordt gezien. Met wat welwillendheid kan ze ook worden gezien als een omgeving die representatief is voor een “soft target” in de context van contra-terrorisme, en als een “te beveiliging object” in de context van bewaken en beveiligen.

De JCA en het omliggende gebied is regelmatig het toneel van grote evenementen, waaronder uiteraard betaald voetbal. De politie-inzet die bij evenementen gepleegd wordt, is onderwerp van maatschappelijk debat (Adang, et al., 2014) – ook bij betaald voetbal.

De inzet van politie bij betaald voetbal is dalende bij een gelijkblijvend aantal incidenten. Deze efficiëntieverbetering is te danken aan gerichte maatregelen, betere samenwerking en snellere opvolging (Adang, et al., 2014). Ook het terugdringen van incidenten is een prioriteit. Tien jaar geleden betrof dat spreekkoren, tegenwoordig gaat de aandacht uit naar het illegaal afsteken van vuurwerk (Politie, 2018). Dit soort incidenten gebeuren zowel binnen als buiten het stadion. Daartoe is een veiligheids-infrastructuur ingericht o.a. bestaande uit camera’s in de omgeving, bij de toegangen en in de JCA zelf. De camera’s bij de toegang en binnen de JCA zijn van de JCA zelf, en in beheer bij één en de zelfde afdeling bij JCA.

Voetbalhooligans zijn op de hoogte van de maatregelen, en nemen tegenmaatregelen. Het bedekken van het gezicht, en het wisselen van kleding

worden al lange tijd gebruikt om de pakkans te verkleinen, en dus sancties of vervolging te voorkomen.

8.1.2 Het operationele proces

In deze sectie worden het huidige en voorgestelde operationele proces vergeleken.

8.1.2.1.1 Het huidige operationele proces

Als er zich nu een incident binnen de JCA voordoet, dan kan JCA zelf de beelden van het incident, de beelden van de toegangsregistratie en de toegangsadministratie met elkaar combineren om de daders te achterhalen. Het is daarvoor niet nodig om eventuele gezichtstemplates met andere organisaties te delen.

Als het incident crimineel van aard is, dan kan de eigenaar van de beelden (AJAX, JCA of een evenementenorganisatie) de beelden uit eigen beweging verstrekken aan de politie. Desgewenst, kan de politie deze ook vorderen bij JCA. Daarna kan de politie in die beelden -al of niet met behulp van automatische gezichtsherkenning - gaan zoeken naar betrokkenen.

Dit lijken legale, proportionele, bestaande, en geaccepteerde oplossingen voor het geschetste probleem.

8.1.2.1.2 Het voorgestelde operationele proces

In het programmaplan van de digitale perimeter (v20190329) is het volgende concept opgenomen dat een ander operationeel werkproces beschrijft.

> *Stel: men fotografeert iedereen bij binnenkomst, met een referentie naar het kaartje, op grond waarvan men entree verkrijgt (aannemende dat daaruit de relevante verwijzing naar de identiteit te herleiden is). Uit de foto's worden de meet- en verhoudingsgetallen berekend die de gelaatsvergelijkingssoftware nodig heeft om te kunnen vergelijken. Die vergelijking vindt niet plaats, maar de waardes worden uitgedrukt in een uniek, maar op zichzelf niet direct inhoudsvol getal. De foto's worden onmiddellijk daarna automatisch vernietigd; zij zijn niet langer nodig om bedoelde getallen te berekenen.*
>
> *Als nu later zich een incident voordoet, en daarvan worden beelden gemaakt die door de aard en inhoud wel kunnen worden bewaard en gebruikt, dan kan daaruit ook een waarde worden berekend, die dan vergeleken kan worden met de waarden bij binnenkomst. Zo kan een relatie worden gelegd naar een ticket. Natuurlijk is dat niet sluitend, maar wel een waarschijnlijke relatie, die kan worden onderzocht alvorens dingen te doen die de privacy meer schenden.*

8.1.3 Opmerkingen bij context en voorgestelde operationele proces

Zowel bij de context als het voorgestelde operationele proces zijn opmerkingen te maken. Achtereenvolgens over de noodzaak, de proportionaliteit, over de aan- of afwezigheid van een stimulans voor bezoekers om mee te werken, en over de belangen van burgers.

Aangezien de JCA en de politie nu al data (mogen) uitwisselen, en dit doen op een geaccepteerde manier, is de noodzaak om daar automatische gezichtsherkenning bij te gebruiken moeilijk hard te maken. Het voegt immers een complexiteit toe die

(voor JCA en politie) operationeel gezien geen meerwaarde lijkt op te leveren *voor die toepassing*.

Doordat JCA nu geen niet-coöperatieve gezichtsherkenning gebruikt, en dat in dit fictieve concept wel zou doen, wordt de proportionaliteitsvraag relevant. Wegen de privacy inbreuken van het gebruik van niet-coöperatieve gezichtsherkenning op tegen het illegaal afsteken van vuurwerk? Een hier aan gerelateerde vraag gaat over subsidiariteit. De JCA en de politie hebben op dit moment andere manieren om daders van incidenten te herkennen. Daders komen vaak uit een relatief kleine en reeds bekende groep mensen. Gebruikmakend van beelden kunnen daders ook “handmatig” herkend worden, wat misschien een lichter middel is tegen dat probleem. Kortom, hoe proportioneel en hoe subsidiair is het dat JCA niet-coöperatieve gezichtsherkenning zou gebruiken om bijvoorbeeld het illegaal afsteken van vuurwerk tegen te gaan? Het antwoord op die vragen is niet evident, wat het illustratief gehalte van de demo kan overschaduwen. Het publiek van de demonstratie moet immers over dit potentiele ethische bezwaar tegen deze opzet heenstappen om de toegevoegde waarde van MPC te zien.

Malafide bezoekers ervaren een stimulans om niet mee te werken: dan kunnen ze later immers ook niet aan een incident worden gekoppeld. Het scenario omvat geen identiteitsverificatie bij de toegangscontrole, zoals bij ADO Den Haag is gedaan (Holst & Vellekoop, 2020). En dit (toegangscontrole op basis van biometrie) is bij JCA niet geïmplementeerd. Daar komt dus geen stimulans vandaan om mee te werken. Het scenario beschrijft niet of er een andere stimulans wordt toegepast. Vermoedelijk moet het scenario dus worden uitgebreid met een controle op een goede kwaliteit opname van een gezicht. Een identiteitscheck is daarbij niet nodig.

In het voorgestelde operationele concept is op een bijzondere manier sprake van *enrolment*. Er is immers geen sprake van toegangscontrole op basis van gezichtsherkenning, hooguit (zie vorige paragraaf) op basis van gezichts*detectie*. In het voorgestelde concept worden de beelden van de toegang direct omgezet in een gezichtstemplate, en worden direct daarna deze (*enrolment*) beelden verwijderd. Het doel daarvan is om “secondair gebruik” door (in dit geval) de JCA onmogelijk te maken. Secondair gebruik moet in deze context opgevat worden als gebruik niet zijnde voor beveiligingsdoeleinden of voor handhaving van de huisregels. Een voorbeeld van secondair gebruik zou bijvoorbeeld gebruik voor marketingdoeleinden kunnen zijn. In de maatschappelijke discussie (zie sectie 5.2.2) wordt dat weliswaar gezien als een voorstelbaar risico, maar niet als een groot risico. Daardoor lijkt de demonstratiewaarde van dit onderdeel van het concept (het verwijderen van de beelden) in de operationele context van JCA relatief beperkt.

Een automatische biometrische vergelijking levert een gesorteerde lijst van gelijkende gezichten op, waarbij de echte match er niet echt bij hoeft te zitten. Weliswaar is dit een aanwijzing om de zoekruimte (van alle bezoekers) te verkleinen, maar vervolgens is een handmatige verificatie nog steeds nodig. En juridisch is deze handmatige verificatie zelfs vereist omdat er geen automatische beslissing mag worden genomen zonder handmatige verificatie. De achtergrond hiervan is dat burgers zich moeten kunnen verweren tegen fouten van het systeem en tegen misbruik. Dat wordt veel moeilijker (onmogelijk?) als de beelden bij de toegang zijn verwijderd. Deze lijst met gezichten moet dus door mensen worden

gecontroleerd, en de meest gelijkende(n) (volgens mensen) moet(en) vervolgens worden geselecteerd en opgevolgd. Maar dit is niet meer mogelijk als de beelden van de toegang weg zijn gegooid. De ene privacy beschermende strategie (“Minimaliseer”) komt dan op gespannen voet te staan met andere (“Geef controle” en “Toon aan”). Het kan ook een functie van dit concept zijn om die spanning ook te illustreren, maar dat is een extra nuance die ook voor verwarring kan zorgen.

8.2 Technische geschiktheid van JCA voor demonstratie van gezichtsherkenning

De technische geschiktheid van de JCA voor demonstratie van gezichtsherkenning gaat over de bestaande camera-infrastructuur en de verlichting. Die zijn immers a priori niet bedoeld voor gezichtsherkenning.

Om de technische geschiktheid te bepalen, en alvorens eventueel onnodige investeringen te doen, zijn op 7 februari 2020 een schouw uitgevoerd en zijn testopnames gemaakt. De voorbereiding en resultaten zijn apart beschreven in een memo (Van Voorthuijsen & Den Hollander, Opnames voor gelaatsvergelijking in Johan Cruijff Arena, 2020).

Uit deze schouw en testopnames is gebleken dat bestaande camera’s bij de ingang niet geschikt zijn. In de meest optimale condities en instellingen zijn de bestaande camera’s op de tribune slechts matig geschikt voor niet-coöperatieve gezichtsherkenning. Dat betekent alleen bij daglicht, scherp ingezoomd, en (ze staan op een paal) bij weinig wind. Dat betekent dat de zekerheid waarmee bruikbare opnames kunnen worden gemaakt, laag is.

Om die zekerheid te vergroten kan er gebruik worden gemaakt van bestaande internationale standaarden (ISO/IEC 30137-1:2019), NIST rapporten (NIST, 2019) en ADO-experimenten. Deze kennis is ten behoeve van nieuwe opnames verzameld in een aparte memo (Van Voorthuijsen, Den Hollander, & Van Rest, Memo Opnames JCA, 2020).

Er kan ook nog worden gekozen om deze uitdaging te koppelen aan de privacy-dreiging van “vervorming” (bijv. foutieve herkenning). Immers, onder deze matige condities bestaat er een risico op allerlei soorten (systematische) fouten in de herkenning (zie bijlage G.3.7). Dat biedt wellicht een interessante omgeving om ook de privacy-beschermende technologie “*managed analytics*” in te ontwikkelen en beproeven. Die heeft immers als functie om de technische en operationele condities van een gezichtsherkenningssysteem te monitoren om zodoende te helpen binnen bepaalde prestatiemarges te blijven (zie sectie 5.4.4).

Er zijn (tijdelijke) aanpassingen nodig aan de camera-infrastructuur en belichting bij de JC ArenA voor demonstraties op dit gebied.

9 Beantwoording onderzoeksvragen en discussie

In dit hoofdstuk worden de onderzoeksvragen beantwoord.

9.1 Beantwoording onderzoeksvragen

1 Welke potentiele operationele contexten zijn van belang als het gaat om (het maatschappelijke debat over) (gezichts)herkenning in de Nederlandse (semi-)openbare ruimte? Hoe zijn deze contexten gerelateerd aan de JCA?

In bijlage E zijn voorbeelden van potentiele toepassingen gegeven van gezichtsherkenning. Hierbij is gefocust op niet-coöperatieve gezichtsherkenning:

A. Object- en persoonsbeveiliging
B. Opsporing heterdaadfase
C. Opsporing *most wanted*
D. Beschermen *soft targets*
E. Handhaving contact-, winkel, OV- of gebiedsverbod
F. Toegangscontrole evenement
G. Private vervoerder richting beveiligd gebied
H. Green lane
I. Monitoren openbare (private) online platformen

Als er een logische relatie is met andere soorten gezichtsherkenning, dan is die meegenomen. Deze voorbeelden zijn niet bedoeld als doelsituatie van dit project. Het zijn slechts beschrijvingen van potentiele toepassingen waarmee het gesprek over niet-coöperatieve gezichtsherkenning beter kan worden gevoerd.

In hoofdstuk 7 zijn twee operationele concepten voor “Object- en persoonsbeveiliging” en voor “evenement veiligheid” in een digitale perimeter uitgewerkt met privacy beschermende strategieën en technologieën. Deze uitwerkingen zijn nuttig om het gesprek over gezichtsherkenning en de bescherming van privacy daarbij beter te kunnen voeren.

De JCA is beperkt geschikt als demonstratieomgeving voor deze studie. Het is een veilige omgeving waar kan worden uitgesloten dat persoonsgegevens worden verwerkt van mensen die daar geen toestemming voor hebben gegeven. Ook kan worden uitgesloten dat het experiment verstoord wordt door tegenstanders, of dat gevoelige apparatuur onvoldoende beveiligd in de openbare ruimte komt te staan. Er zijn echter opmerkingen te plaatsen bij het voorgestelde operationele concept dat als context voor een demonstratie bij JCA kan dienen. Die gaan bijvoorbeeld over de proportionaliteit van gezichtsherkenning bij betaald voetbal, en over de bruikbaarheid en juridische haalbaarheid indien beelden worden verwijderd (waardoor handmatige verificatie en verweer niet mogelijk zijn). Verder wordt bij de JCA nu geen toegang verleent op basis van biometrische kenmerken. Hierdoor is er, juist voor malafide bezoekers, geen stimulans om mee te werken aan goede gezichtsopnames. Het kan voor het publiek van een demonstratie lastig zijn om de toegevoegde waarde van MPC in deze context te zien. Daarom kan het nodig zijn

om het concept ten behoeve van de demonstratie uit te breiden met biometrische toegangscontrole bij de ingang van de JCA[30].

2 *Welke privacydreigingen moeten adequaat worden gemitigeerd als het gaat om (gezichts)herkenning in de Nederlandse (semi-)openbare ruimte? Gaat dat o.a. om het voorkomen van misbruik? Is er een gangbare conceptualisatie van privacy die geschikt is om dat mee te duiden?*

In sectie 5.2 is een analyse en overzicht gegeven van alle privacydreigingen zoals Solove die heeft geïdentificeerd. Van de zestien privacydreigingen van Solove, zijn er vijf inherent aan automatische gezichtsherkenning en daardoor onvermijdelijk: *surveillance*, *aggregatie*, *onveiligheid*, *vervorming* en *beslissingsinterferentie*. Wel is het mogelijk om deze dreigingen deels te verkleinen. Van nog eens vier van die zestien soorten privacydreigingen is het denkbaar dat ze zich in de Nederlandse context manifesteren. Dan gaat het ook om dreigingen zoals *identificatie*, *inbreuk op vertrouwelijkheid, uitsluiting*, *indringing* en *secondair gebruik*. Misbruik valt in deze context onder wat Solove *secondair gebruik* noemt.

3 *Welke privacy-preserving (gezichts)herkenningstechnologie is voldoende matuur (TRL6+) om tijdig operationeel inzetbaar te zijn? Is multi-party computation (MPC) hiervoor geschikt? Wat is de IP situatie van relevante technologieën?*

Er zijn in deze studie vijf soorten *privacy enhancing* technologieën verkend: gezichtsherkenning via een online platform, *intelligence-on-the-edge*, herherkenning, *managed analytics* en *multi-party computation*. De maturiteit van deze technologieën varieert van de definiërende fase (TRL3) voor *managed analytics* tot en met operationeel (TRL9) voor gezichtsherkenning via een online platform.

In deze studie is gefocust op *multi-party computation* (MPC).Er is voldoende literatuur aangetroffen om aan te nemen dat MPC nog niet operationeel inzetbaar is, maar wel op korte termijn mogelijk een voldoende volwassen oplossing kan bieden om bepaalde privacy dreigingen te helpen mitigeren in specifieke *use cases*. Het gaat dan om *onveiligheid*, *secondair gebruik*, en *inbreuk op vertrouwelijkheid* (zie sectie 6.4). De IP situatie van de onderzochte vorm van MPC (bijv. of er patenten zijn) is verkend (zie sectie 6.2), maar nog niet uitputtend onderzocht. MPC bevindt zich daarmee op TRL5.

Het lijkt verstandig om MPC te combineren met andere privacy-preserving technologie om ook de andere relevante privacydreigingen zo goed mogelijk te mitigeren, zoals technologie die de kans op foutieve herkenning, en de kans op *biases* verkleint.

Het ligt niet voor de hand dat platformtechnologie (zie sectie 5.4.1) ook bij niet-coöperatieve gezichtsherkenning de dreigingen van “inbreuk op vertrouwelijkheid” en van “secondair gebruik” kan helpen mitigeren. MPC kan misschien wel de veiligheid van dit soort diensten op gebied van *coöperatieve* gezichtsherkenning

[30] Dit is alleen nodig ten behoeve van de demonstratie, en heeft niets te maken met de normale toegangscontrole bij de JCA.

helpen verbeteren door de communicatie tussen het platform en de partij die de identificatie doet veiliger te maken (zie sectie 5.4.1).

Het lijkt verstandig om technologie die de privacy beschermt breed toegankelijk te maken, zoals in internationale standaarden. Daar lijkt nu nog geen sprake van te zijn. De benodigde stappen daarvoor zijn …
(1) … een technisch en daarna …
(2) … operationeel werkende referentie-architectuur die …
(3) … aantoonbaar de gewenste maatschappelijke effecten levert, en vervolgens …
(4) … tot voldoende draagvlak leidt bij betrokkenen voor het opnemen in een standaard.

4 Welke kenmerken van de live situatie zijn relevant? Welke kenmerken kunnen automatisch worden bepaald uit het beeld of uit andere databronnen (zoals een weerstation)? Hoe kan die informatie worden gebruikt om te voorkomen dat onnodig veel mensen onterecht worden herkend, zonder dat het eenvoudig wordt om het (gezichts)herkenningssysteem te omzeilen?

NIST heeft een duidelijke opsomming van dit soort kenmerken gegeven in (NIST, 2019). Voorbeelden zijn occlusie, contrast op het gezicht en resolutie op het gezicht. In diezelfde studie verkent NIST in welke mate software in staat is om dergelijke factoren automatisch te meten. In eerder werk is beschreven dat het concept *managed analytics* een framework biedt om deze informatie te gebruiken om te voorkomen dat onnodig veel mensen onterecht worden herkend (Den Hollander, et al., 2017). Het is nog niet uitgezocht hoe dat in een concreet live en operationeel systeem ingezet kan worden. Wel zijn twee operationele toepassingsconcepten beschreven die als context voor een experiment kunnen dienen.

9.2 Discussie

Het beschermen van privacy bij de toepassing van een inherent invasieve technologie zoals gezichtsherkenning vereist de combinatie van meerdere disciplines, zorgvuldigheid en integraliteit. Als aan die voorwaarden is voldaan, dan kunnen sommige oplossingen vrij eenvoudig zijn.

Op het moment dat de business case van beveiligingsmaatregelen verdwijnt omdat de dreiging niet meer speelt (zoals bij ADO Den Haag), dan bestaat het risico dat praktijkkennis over bijgaande privacy-preserving technologieën (zoals de manier hoe bezoekersprofielen werden gebruikt bij ADO om de werking gezichtsherkenning te reguleren) verloren gaat. Daarom is in dit rapport wat meer tekst gebruikt om dat soort kennis expliciet te maken, te borgen en te ontsluiten.

In hoofdstuk 7 zijn twee operationele concepten uitgewerkt. Ze verschillen met name in de mate van flexibiliteit die ze ondersteunen. Dat thema, flexibiliteit, staat op gespannen voet met beheersbaarheid. Mission creep, het onbeheerst uitbreiden van de functionaliteit van een systeem, moet worden voorkomen. Maar een genuanceerd en goed geïnformeerd besluitvormingsproces zou het ook mogelijk moeten maken om de functie van gezichtsherkenningssystemen op beheersbare

wijze aan te passen aan nieuwe omstandigheden. Dat komt zowel veiligheid, leefbaarheid als betaalbaarheid ten goede.

Deze studie heeft laten zien dat nieuwe technologie kan helpen om de privacy (beter) te beschermen. Met de informatie verzameld in dit rapport kan bij bestaande gezichtsherkenningstoepassingen getoetst worden of ze ook optimaal gebruik maken van nieuwe technologie om de privacy te beschermen. Aangezien er voortdurend technologische vernieuwingen beschikbaar komen, lijkt het nuttig om privacy- en gegevensbeschermingseffectbeoordeling niet een éénmalige toets te laten zijn, maar om die na verloop van tijd opnieuw te doen.

Multi-party computation is een verzameling van technologieën waarmee veilige verwerking van gegevens kan worden gedaan. In het veiligheidsdomein zijn ook andere toepassingen denkbaar. Wellicht is het nuttig om voor afgebakende onderdelen van dit domein clusters van MPC-oplossingen te ontwikkelen. Zoals bijvoorbeeld voor bewaken en beveiligen, of voor opsporing.

De rol van de inlichtingendiensten is in dit rapport niet belicht. Het risico dat privacy preserving technologieën hun legitieme behoeften onmogelijk maken, speelt vaker, bijvoorbeeld bij de encryptie van informatiedragers en communicatiekanalen. Het vijfde principe van *privacy-by-design* luidt*:* “full functionality – positive-sum, not zero-sum. Oftewel: legitieme doelen -dus ook die van geheime diensten- kunnen normaal bereikt worden. In dit rapport zijn geen toepassingen opgenomen waarin zij een expliciete rol hebben. Mogelijk gaat dat nog wel nodig blijken.

10 Conclusies en hoe verder

10.1 Hoofdconclusie

De taxonomie van Solove en de strategieën van Hoepman helpen om structuur te brengen in de probleem- en oplossingsruimte. Het is daardoor beter mogelijk om de privacydreigingen van niet-coöperatieve gezichtsherkenning te benoemen, te analyseren, en om oplossingen te bespreken. Het was hierdoor bijvoorbeeld mogelijk om enkele bronnen in het maatschappelijk debat te analyseren op verwijzingen naar concrete privacy dreigingen (zie sectie 5.2.2). Een goed begrip van de werking, de functie, van de operationele en van de juridische context van gezichtsherkenning is daarbij ook nodig. Bijvoorbeeld het begrip dat er wel een technisch verschil is tussen een opname van een gezicht, en een biometrisch template op basis van die opname, maar dat beiden juridisch gezien als gevoelige persoonsgegevens worden beschouwd.

In specifieke toepassingsscenario's (denk aan proportionaliteit, subsidiariteit, traceerbaarheid en accountability) kan de onderzochte vorm van MPC nuttig zijn om de privacydreigingen "onveiligheid", "inbreuk op vertrouwelijkheid" en "secondair gebruik" te beheersen. Dat kan mogelijk nog verbeterd worden door gezichtstemplates direct vanuit versleuteling te vergelijken.

Het vergelijken van biometrische gezichtstemplates op een veilige wijze met behulp van *multi-party computation* lijkt binnen bepaalde grenzen (afweging functionaliteit versus snelheid versus kwaliteit) technisch gezien mogelijk in de vorm van een demonstratiemiddel.

Mocht er te zijner tijd aan implementatie worden gedacht, dan lijkt het verstandig om dat te combineren met aanvullende maatregelen die het risico van "vervorming" (zoals *biases* en foutieve identificatie (En. *false positives*) verkleinen (zie bijvoorbeeld sectie 5.4.4). Om de toets op subsidiariteit te kunnen uitvoeren, moet bij inzetbeslissingen ook alternatieve technologie zoals zachte biometrie (zie sectie 5.4.3) worden overwogen.

10.2 Overige conclusies

Er is vermoedelijk niet één eenduidige configuratie en architectuur voor een *privacy preserving* gezichtsherkenningssensor mogelijk die voor alle mogelijke soorten toepassingen optimaal is. Ten eerste, de toepassingen lijken daarvoor te verschillend (zie sectie 4.3), mede in relatie tot de eisen aan verwerkingskracht en communicatieoverhead die voor MPC-varianten nodig lijken (zie sectie 6.3). Ten tweede dekt MPC slechts een deel van de privacydreigingen af.

Het is wel mogelijk gebleken om twee verschillende potentiële concepten te bedenken (zie hoofdstuk 7). Eén is een uitwerking van het begrip "Digitale perimeter", bedoeld voor evenementenveiligheid in een stedelijke omgeving. De tweede gaat over object- en persoonsbeveiliging in een rustige wijk. Daarvoor is het nodig gebleken om eerst een verzameling strategieën (Hoepman, 2014) en een "gereedschapskist" aan privacy-beschermende technologieën (*soft-biometrics*, platformtechnologie, *intelligence-on-the-edge multi-party computation* en *managed*

analytics) te definiëren. Daarmee konden voor deze twee soorten toepassingen de functionele eisen van twee soorten *privacy-preserving* gezichtsherkenningscamera's worden opgesteld (zie sectie 6.5), waarbij het tevens aannemelijk is gemaakt dat ze (ten behoeve van een demonstratie) ook gerealiseerd zouden kunnen worden.

Er is in overleg met de deelnemers aan dit project gekozen om voor MPC enkele verdiepende experimenten uit te voeren op TRL4. De hypothese dat MPC een bijdrage kan leveren aan het voorkomen van onnodige verspreiding van persoonsgegevens is in een laboratorium gevalideerd. De experimenten wijzen uit dat MPC voor bepaalde toepassingen van niet-coöperatieve gezichtsherkenning toegevoegde kan hebben.

Er is bestaand onderzoek naar MPC bij TNO, waarop verder kan worden gebouwd. Ook is er internationaal onderzoek, maar voor zover we in dit project hebben kunnen uitzoeken, is er nog geen commercieel verkrijgbaar product of dienst die het beschikbaar stelt. Wel zijn er vermoedelijk bedrijven die, met wat ondersteuning, een dergelijke oplossing kunnen implementeren. MPC lijkt dus zowel mogelijk, nuttig als vernieuwend, en daardoor dus innovatief.

De eerste resultaten geven inzicht in de snelheid, complexiteit en ontwerpafwegingen van MPC voor gezichtsherkenning. Een demonstratie met gezichten van vrijwilligers (expliciete consent) lijkt goed mogelijk.

Er zijn aanpassingen nodig aan de bestaande camera-infrastructuur en belichting bij de JC ArenA voor experimenten of demonstraties op dit gebied. Uit proefopnames is gebleken dat de camera's bij de ingang niet geschikt zijn. Zelfs in de meest optimale condities en instellingen zijn de camera's op de tribune slechts matig geschikt voor niet-coöperatieve gezichtsherkenning. Dat betekent bij daglicht, scherp ingezoomd, en (ze staan op een paal) bij weinig wind. Hier is vanuit bestaande internationale standaarden (ISO/IEC 30137-1:2019), NIST rapporten (NIST, 2019) en ADO-experimenten kennis over die nog toegepast kan gaan worden bij / door JCA.

10.3 Hoe verder

Op basis van de inzichten van dit rapport en de resultaten van het MPC experiment is een aantal functionele ontwerpen voorgesteld voor een "privacy preserving gezichtsherkenningssensor" (zie sectie 6.5). Die zouden in een vervolg de basis kunnen vormen voor een technisch, operationeel èn mogelijk ook een sociaal experiment.

Het lijkt verstandig om voor de demonstratie een zorgvuldige beschrijving van de (fictieve) operationele context te maken. Bij voorkeur is de demonstratie niet alleen technisch inspirerend, maar ook operationeel, ethisch en juridisch. Het lijkt verstandig om daarin onder andere een stimulans voor medewerking aan een goede enrolment opname te verwerken.

Onderdeel van een demonstratie zou ook kunnen zijn aan te tonen hoe goed het concept en de concrete technologie bestand zijn tegen aanvallen op de privacy van betrokkenen. Voor wat betreft de functionaliteit van MPC vallen daar twee aspecten

onder. Ten eerste, welke persoonsgegevens (of mogelijk andere gevoelige gegevens) vallen er uit de uitkomst van de verwerking af te leiden? Wordt de privacy dan wel echt goed beschermd?

En ten tweede, wijkt de daadwerkelijk implementatie af van het theoretische protocol en / of van de hier beschreven functionele specificatie, en valt die afwijking te exploiteren door een aanvaller? Met name dat tweede kan in de vorm van *red teaming* oefeningen, eerst “droog”, in een workshop of op papier, en als er technologie beschikbaar is, ook met technische middelen daarop. Als het later ook tot een pilot komt of zelfs implementatie, dan kan het nodig zijn om opnieuw dergelijke testen uit te voeren, waarbij ook *social engineering* kan worden ingezet tegen medewerkers.

Bij een vervolg richting implementatie (dus voorbij een demonstratie) zou er aandacht moeten zijn voor de volgende aspecten:

- Tegen welke privacydreigingen wil je gaan beschermen en hoe kun je als opdrachtgever afdwingen dat er geen zwaktes ontstaan (bijv. door afwijkingen in de implementatie)?
- Wordt de camera-infrastructuur aangepast / gemanaged om goede opnames te krijgen (gezichtsherkenning met kleine foutratio) of kan men operationeel omgegaan met een hoge foutratio?
- Juist malafide bezoekers zullen niet-coöperatief zijn en proberen gezichtsopnames te vermijden. Wordt er een stimulans ingebracht om mee te werken aan goede opnames?
- Hoe belangrijk is het in de toepassing dat mensen zich kunnen verweren tegen de consequenties van een foutieve match? Kunnen ze daarbij gebruik maken van alle opnames die van hen gemaakt zijn?

Een belangrijk aspect is de schaalbaarheid van de oplossing. Er bestaat nu geen standaard voor gezichtstemplates. Een MPC variant van gezichtsherkenning is in deze studie toegepast op een bepaald type gezichtstemplate. Daarbij is géén gebruik gemaakt van specifieke eigenschappen van deze template, dus in principe lijkt het mogelijk om de technologie ook voor andere templates te laten werken. Een ander aspect is interoperabiliteit. De vorm waarin de technologie nu is gemaakt, zou vereisen dat alle direct verbonden partijen met live- en met enrolment templates hetzelfde technische format zou moeten gebruiken. Zowel voor wat betreft het gezichtstemplate, als voor wat betreft de “MPC-schil” daar omheen. Het kan wenselijk zijn om voor beiden een standaard te gaan nastreven, zodat er zo weinig mogelijk barrières zijn om deze vorm van veilige verwerking -waar nodig – toe te passen.

De mogelijke toepassingen en operationele concepten zijn hier opgenomen om het gesprek over de wenselijkheid van gezichtsherkenning, en daarbij het nut, de werkbaarheid en de realiseerbaarheid van privacy beschermende strategieën en *privacy preserving* technologieën beter bespreekbaar te kunnen maken. Mogelijk ontstaat er interesse om te verkennen of dit soort technologie, samen met het bedrijfsleven, daadwerkelijk kan worden ontwikkeld. Hoe het ecosysteem er uit zou kunnen zien om dat te doen, en waar dat in zou kunnen landen, is nog niet duidelijk. In hoofdstuk 7 van (Keymolen, Noorman, Van der Sloot, Cuijpers, & Koops, 2020) worden bijvoorbeeld een aantal reguleringsopties voor de Nederlandse overheid gepresenteerd, waaronder het compleet verbieden van

gezichtsherkenning. Het lijkt verstandig om de bevindingen van deze studie, en dat van andere relevante studies, bij privacy-auditeurs en bij de regelgevende overheid onder de aandacht te brengen. Het lijkt verstandig om daarbij zowel gezichtsherkenning in de fysieke omgeving mee te nemen, alsook die in de virtuele online omgeving (zie toepassing I “Monitoren openbare (private) online platformen”).

11 Referenties

20FACE. (2020). The future of access management: privacy-proof face recognition - Whitepaper. 20FACE.

ACLU. (2020, mei 28). *ACLU Sues Clearview AI*. Opgehaald van ACLU.org: https://www.aclu.org/press-releases/aclu-sues-clearview-ai

ACM. (2020, June 30). *USTPC Facial Recognition Tech Statement.* Opgehaald van ACM.org: https://www.acm.org/binaries/content/assets/public-policy/ustpc-facial-recognition-tech-statement.pdf

Adang, O., Van Arkel, D., Esman, C., Van Oorschot, W., Stronks, S., & Vos, H. (2014). *Politie en evenementen.* Den Haag: Boom Lemma Uitgevers.

Adler, A. (2003). Sample images can be independently restored from face recognition templates. . *CCECE 2003-Canadian Conference on Electrical and Computer Engineering. Toward a Caring and Humane Technology (Cat. No. 03CH37436)* , 1163-1166. Opgehaald van http://sce.carleton.ca/faculty/adler/publications/2003/adler-2003-ccece-restore-face-recognition-templates.pdf

Autoriteit Persoonsgegevens. (2020, oktober 29). *AP: Pas op met camera's met gezichtsherkenning*. Opgehaald van Autoriteit Persoonsgegevens: https://autoriteitpersoonsgegevens.nl/nl/nieuws/ap-pas-op-met-camera%E2%80%99s-met-gezichtsherkenning

Autoriteit Persoonsgegevens. (2020, Mei). *Zwarte Lijst*. Opgehaald van Autoriteit Persoonsgegevens: https://autoriteitpersoonsgegevens.nl/nl/zelf-doen/zwarte-lijst

AXIS. (2020). *AXIS Camera Application Platform*. Opgehaald van AXIS: https://www.axis.com/support/developer-support/axis-camera-application-platform

Bacchi, U. (2020). *Face for sale: Leaks and lawsuits blight Russia facial recognition*. Opgehaald van trust.org: https://news.trust.org/item/20201109090922-3k4a5/

Bouma, H., Borsboom, S., Den Hollander, R. J., Landsmeer, S. H., & Worring, M. (2012). Re-identification of persons in multi-camera surveillance under varying viewpoints and illumination. *Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense XI (Vol. 8359, p. 83590Q). International Society for Optics and Photonics*.

Bouma, H., Joosten, B., Kruithof, M., de Boer, M., Ginsca, A., Labbe, B., & Vuong, Q. T. (2018). Flexible image analysis for law enforcement agencies with deep neural networks to determine: where, who and what. *Counterterrorism, Crime Fighting, Forensics, and Surveillance Technologies II (Vol. 10802, p. 108020R). International Society for Optics and Photonics*.

Bringer, J., Chabanne, H., & Patey, A. (2013). Privacy-preserving biometric identification using secure multiparty computation: An overview and recent trends. *IEEE Signal Processing Magazine, 30(2)*, 42-52.

Cavoukian, A. (2009). *Privacy by design: The 7 foundational principles. .* Information and privacy commissioner of Ontario, Canada, 5.

Chibba, M., & Stoianov, A. (2014). *On Uniqueness of Facial Recognition Templates.* NTIA US Department of Commerce. Opgehaald van

https://www.ntia.doc.gov/files/ntia/publications/uniqueness_of_face_recognition_templates_-_ipc_march-2014.pdf

ClearView. (2020). *ClearView*. Opgehaald van ClearView: https://clearview.ai/

Daily Record. (2016). *Celtic fans protest against facial recognition by unveiling banners at Parkhead*. Opgehaald van Daily Record: https://www.dailyrecord.co.uk/sport/football/football-news/celtic-fans-protest-against-facial-7232589

Davies, B., Innes, M., & Dawson, A. (2018). *An evaluation of South Wales Police's use of Automated Facial Recognition.* Cardiff, UK: Cardiff University.

Den Hollander, R. J., Bouma, H., Van Rest, J. H., ten Hove, J. M., Ter Haar, F. B., & Burghouts, G. J. (2017). Automatically assessing properties of dynamic cameras for camera selection and rapid deployment of video content analysis tasks in large-scale ad-hoc networks. *Counterterrorism, Crime Fighting, Forensics, and Surveillance Technologies (Vol. 10441, p. 1044108). International Society for Optics and Photonics*.

Digitale Perimeter. (2020). Opgehaald van Gemeente Amsterdam: https://www.amsterdam.nl/wonen-leefomgeving/innovatie/de-digitale-stad/digitale-perimeter/

ECP. (2020). *Begeleidingsethiek bij de politie - Verslag van de eerste ethiektafels.* Leidschendam: ECP.

Erkin, Z. F., & Toft, T. (2009). Privacy-preserving face recognition. *International symposium on privacy enhancing technologies symposium*, 235-253.

Flight, S. (2013). *Cameratoezicht en design.* DSP Groep.

Grapperhaus, F. (2019, November 20). Waarborgen en kaders bij gebruik gezichtsherkenningstechnologie. Den Haag. Opgehaald van https://www.rijksoverheid.nl/documenten/kamerstukken/2019/11/20/tk-waarborgen-en-kaders-bij-gebruik-gezichtsherkenningstechnologie

Grother, F., Quinn, G., & Ngan, M. (2017). *NIST Interagency Report 8173: Face In Video Evaluation (FIVE) Face Recognition of Non-Cooperative Subjects.* NIST. Opgehaald van https://nvlpubs.nist.gov/nistpubs/ir/2017/NIST.IR.8173.pdf

Guo, S., Xiang, T., & Li, X. (2019). Towards Efficient Privacy-Preserving Face Recognition in the Cloud. *Signal Processing, 164*, 320-328.

Hampentech. (2020). *Face Recognition ACAP*. Opgehaald van http://hampentech.com/face-recognition-acap/

Hoepman, J. (2014). Privacy design strategies. *IFIP International Information Security Conference*, 446-459.

Holst, M. v., & Vellekoop, L. (2020, mei 15). Interview Beveiliging Ado Den Haag. (J. v. Rest, Interviewer)

Homburg, G., Schreijenberg, A., & Stouten, J. (2014). *Zoekmiddelen bij urgente persoonsvermissingen.* Amsterdam: Regioplan.

Jain, A. K., Flynn, P., & Ross, A. A. (2007). *Handbook of biometrics. .* Springer Science & Business Media.

Janssen, A., Kool, L., & Timmer, J. (2015). *Dicht op de Huid.* Den Haag: Rathenau Instituut.

JRC, ERNCIP. (2019, August). *Early Warning Zones.* Opgehaald van European Reference Network for Critical Infrastructure Protection (ERNCIP): https://erncip-project.jrc.ec.europa.eu/networks/tgs/fencing

Keymolen, E., Noorman, M., Van der Sloot, B., Cuijpers, C., & Koops, B.-J. Z. (2020). *Op het eerste gezicht.* Tilburg: TILT, Universiteit van Tilburg.

King, D. E. (2009). Dlib-ml: A machine learning toolkit. *The Journal of Machine Learning Research 10*, 1755-1758.

King, D. E. (2017). *dlib face recognition resnet model v1*. Opgehaald van GitHub: https://github.com/davisking/dlib-models/blob/master/dlib_face_recognition_resnet_model_v1.dat.bz2

Li, P., Prieto, L., Mery, D., & Flynn, P. (2018). Face recognition in low quality images: a survey. . *arXiv preprint arXiv:1805.11519.*

Lucas, T., & Henneberg, M. (2015). Are human faces unique? A metric approach to finding single individuals without duplicates in large samples. *Forensic Science International, 257*, 514-e1.

Mai, G., Cao, K., Yuen, P. C., & Jain, A. K. (2018). On the reconstruction of face images from deep face templates. . *IEEE transactions on pattern analysis and machine intelligence, 41(5)*, 1188-1202.

Meester, R., Preneel, B., & Wenmackers, S. (2019). Reply to Lucas & Henneberg: Are human faces unique?. *Forensic science international, 297*, 217-220.

MIT Review. (2015, februari 16). *The Face Detection Algorithm Set to Revolutionize Image Search*. Opgehaald van MIT Technology Review: https://www.technologyreview.com/2015/02/16/169357/the-face-detection-algorithm-set-to-revolutionize-image-search/

Mohanty, P. (2008). *USA Patentnr. 60594187.* Opgehaald van http://patft.uspto.gov/netacgi/nph-Parser?Sect1=PTO1&Sect2=HITOFF&d=PALL&p=1&u=%2Fnetahtml%2FPTO%2Fsrchnum.htm&r=1&f=G&l=50&s1=8,165,352.PN.&OS=PN/8,165,352&RS=PN/8,165,352

NEC. (2018, November). *Introduction of NEC's Secure Computing Technology.* Opgehaald van NEC.com: https://www.nec.com/en/global/rd/technologies/201805/pdf/mpc_introduction.pdf

NIST. (2017). *Face in Video Evaluation (FIVE) - Face Recognition of Non-Cooperative Subjects.* NUS Department of Commerce.

NIST. (2019). *Face Recognition Vendor Test, Part 3: Demographic Effects.* NIST.

NIST. (2019). *FRVT Quality Assessment*. Opgehaald van NIST: https://www.nist.gov/programs-projects/frvt-quality-assessment

NIST. (2020, May 10). *Face Recognition Vendor Test (FRVT) Ongoing*. Opgehaald van NIST: https://www.nist.gov/programs-projects/face-recognition-vendor-test-frvt-ongoing

NU.nl. (2020, januari 31). *Onduidelijk hoe vaak gezichtsherkenning bij politie leidt tot aanhoudingen*. Opgehaald van NU.nl: https://www.nu.nl/tech/6025903/onduidelijk-hoe-vaak-gezichtsherkenning-bij-politie-leidt-tot-aanhoudingen.html

Politie. (2018). *Jaarverslag Voetbal en Veiligheid Seizoen 2017/2018.* Politie.

Poole, R. (2008). *Toward Risk-Based Aviation Security Policy.* Los Angeles: OECD. Opgehaald van https://www.itf-oecd.org/sites/default/files/docs/dp200823.pdf

Ramachandra, R., & Busch, C. (2017). Presentation attack detection methods for face recognition systems: A comprehensive survey. *ACM Computing Surveys (CSUR), 50(1)*, 1-37.

Rogers, K. (2016, February 7). *That Time the Super Bowl Secretly Used Facial Recognition Software on Fans*. Opgehaald van Vice: https://www.vice.com/en_us/article/kb78de/that-time-the-super-bowl-secretly-used-facial-recognition-software-on-fans

RTLNieuws. (2019, June 26). *Nieuwe app helpt agenten gezichten te herkennen*. Opgehaald van RTLNieuws: https://www.rtlnieuws.nl/editienl/artikel/4754646/nieuwe-app-helpt-agenten-gezichten-te-herkennen

Schiphol. (2019, February 18). Schiphol start proef voor boarden met gezichtsherkenning. Opgehaald van https://nieuws.schiphol.nl/schiphol-start-proef-voor-boarden-met-gezichtsherkenning/

Schiphol. (2020). Privium. *Live life in the fast lane*. Opgeroepen op 2020, van https://www.schiphol.nl/nl/privium/

Schoenmakers, Y., De Groot, I., Van Rooyen, A., Van Zanten, J., & Baars, J. (2017). *De onvindbaren.* Deventer: Vakmedianet.

Snijders, D., Biesiot, M., Munnichs, G., & Van Est, R. (2019). *Burgers en sensoren.* Den Haag: Rathenau.

Solove, D. J. (2005). A taxonomy of privacy. *U. Pa. L. Rev. 154*, 477.

Su, Y., Yang, Y., Guo, Z., & Yang, W. (2015). Face recognition with occlusion. *2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR)*.

Toli, C. (2018). *Secure and Privacy-Preserving Biometric Systems.* Leuven: KU Leuven.

TRESSPASS. (2019). *TRESSPASS*. Opgehaald van TRESSPASS: https://cordis.europa.eu/project/id/787120

University of Massachusetts. (2007). *Labeled Faces in the Wild (LFW) dataset.* Opgehaald van University of Massachusetts: http://vis-www.cs.umass.edu/lfw/

Van Delft, D. (2001, April 18). *Camera beschermt privacy* . Opgehaald van NRC: https://www.nrc.nl/nieuws/2001/04/28/camera-beschermt-privacy-7539929-a605166

Van der Steur, G. (2015, November 24). BRIEF VAN DE MINISTER VAN VEILIGHEID EN JUSTITIE (Nr. 594). Den Haag.

Van Rest, J. (2019). Werkdocument sensing binnen politiewerk 2022 en verder.

Van Rest, J. v., Peeters, M., Smits-Clijsen, E., Sternheim, A., & Wessels, M. (2020). *Professioneel Profileren: een methode voor het ontwerp en gebruik van profielen door de politie.* Den Haag: TNO.

Van Rest, J., & Weima, I. (2018). *De toekomst van sensing voor veiligheid.* Den Haag: TNO.

Van Rest, J., Boonstra, D., Everts, M., van Rijn, M., & Van Paassen, R. (2012). Designing privacy-by-design. *Annual Privacy Forum*, 55-72.

Van Rooijen, A., Bouma, H., Pruim, R., Baan, J., Uijens, W., & Van Mil, J. (2020). Anonymized person re-identification in surveillance cameras. *Counterterrorism, Crime Fighting, Forensics, and Surveillance Technologies IV (Vol. 11542, p. 115420A). International Society for Optics and Photonics*.

Van Voorthuijsen, G., & Den Hollander, R. (2020). *Opnames voor gelaatsvergelijking in Johan Cruijff Arena.* Den Haag: TNO.

Van Voorthuijsen, G., Den Hollander, R., & Van Rest, J. (2020). *Memo Opnames JCA.* Den Haag: TNO.

Voskuil, K. (2019, June 16). Superherkenner Robert heeft al meer dan 250 criminelen herkend. *Algemeen Dagblad*. Opgehaald van https://www.ad.nl/binnenland/superherkenner-robert-heeft-al-meer-dan-250-criminelen-herkend~a9ee10d6/

VVD Amsterdam. (2014, February 27). *Gebiedsverboden handhaven door slimme camera's met gezichtsherkenning*. Opgehaald van VVD Amsterdam:

https://www.vvdamsterdam.nl/nieuws/21039/gebiedsverboden-handhaven-door-slimme-camera-s-met-gezichtsherkenning

Wiebes, M. (2018, August 22). *De politie vraagt uw hulp*. Opgehaald van KIVI: https://www.kivi.nl/nieuws/artikel/de-politie-vraagt-uw-hulp

Xiang, C. T., & Xu, Q. (2016). Privacy-preserving face recognition with outsourced computation. *Soft Computing, 20(9)*, 3735-3744.

XP-DITE. (2012). *XP-DITE*. Opgehaald van XP-DITE: https://cordis.europa.eu/project/rcn/104801/en

A Review

Datum	Versie	Review door
Najaar 2019	“Discussiestuk”	TNO kernteam, Mark Wiebes
Voorjaar 2020	“Theorie en focus op MPC”	TNO kernteam
Zomer 2020	“Resultaten MPC verwerkt”	TNO kernteam
Najaar 2020	“Externe versie”	Mark Wiebes, Henri Bouma, Eefje Smits-Clijsen

Bijlage K “Experiments on multi-party computation for non-cooperative facial recognition” is geschreven door het MPC team o.l.v. Thomas Attema, en gereviewed door Jeroen van Rest en Henri Bouma.

B Sensing principes toegepast op niet-coöperatieve gezichtsherkenning

In (Van Rest & Weima, De toekomst van sensing voor veiligheid, 2018) zijn zeven sensing-principes gedefinieerd voor toepassing van sensing in het veiligheidsdomein. Deze principes hebben geholpen om, binnen de afbakening van de vraag, te bepalen hoe het doel bereikt zou moeten worden. In deze bijlage is die redenatie uitgeschreven per sensing-principe.

1 **Privacy-by-design en security-by-design voor sensing.** De vorm waarin niet-coöperatieve gezichtsherkenning in 2020 als technologie op de markt beschikbaar is, is over het algemeen niet direct te rijmen met dergelijke ontwerpprincipes (En. *value sensitive design*). De uitdaging van deze studie was om te verkennen in hoeverre het mogelijk is om een dergelijke vorm van niet-coöperatieve gezichtsherkenningstechnologie alsnog te maken. Daarnaast is er een grote rol weggelegd voor het specificeren van het specifieke inzetconcept. Immers, alleen in die context kunnen afwegingen zoals proportionaliteit en subsidiariteit worden gemaakt.
2 **Metadateren en filteren bij de bron.** Dit is een sensing-specifieke variant van het eerste principe. Door gezichten al in de sensor te detecteren, is direct duidelijk of en in welke mate die sensordata nuttig is, bewaard moet blijven (of juist niet) en beschermd moet worden. Er is daar een reflectie over geschreven als het gaat om *intelligence-on-the-edge* in sectie 5.4.2, en een aanzet tot een functioneel ontwerp beschreven in sectie 6.5.3.
3 **Sensordata van voldoende kwaliteit**. Het is verstandig om mogelijke factoren die leiden tot fouten van de gezichtsherkenning (*failure to capture*, foutieve identificatie en foutieve niet-identificatie) proactief te monitoren. Zoals belichtingsomstandigheden, pose, en beweging van de camera. Daarmee wordt het mogelijk om de benodigde kwaliteit van gezichtsherkenning proactief te beheersen. Daardoor wordt het ook duidelijk welke factoren dus nog op andere manieren beheerst moet worden. Daar is in sectie 5.4.4 technologie voor geopperd die ook kan worden ingezet om bepaalde kwaliteit af te dwingen.
4 **Sensing slechts als het nodig is.** Ook dit is een sensing-specifieke variant van het eerste principe over privacy-by-design. In deze studie zijn verschillende toepassingen gedefinieerd die een zeer tijdelijk karakter hebben, zoals de opsporing in een heterdaadfase (zie sectie E.2) of evenementenbeveiliging (zie sectie E.6 en sectie 7.1). In die gevallen is het inherent aan de toepassing dat gezichtsherkenning standaard uit staat, en pas aan gaat als er een verhoogd risico is waarbij niet-coöperatieve gezichtsherkenning kan helpen het risico te beheersen. In het systeemconcept dat is bedacht is het mogelijk om een derde partij de technische mogelijkheid te geven om de gezichtsherkenning te monitoren, aan te zetten of zelfs te stoppen (zie secties I.7 en I.8).
5 **Gebruikmaken van andermans sensoren.** Dit principe zit in de kern van de uitdaging. Hoe kan een partij die op zoek is naar bekende personen, dat doen in beelden van sensoren van een andere partij? De technologie *multi-party computation* (zie hoofdstuk 6) doet precies dat.
6 **Nieuw sensing concept ontwikkelen.** De uitdaging waar in deze studie aan gewerkt is, raakt meerdere organisaties en de samenleving als geheel. Het

innovatieproces moet dus met meer zorg omgeven worden dan wanneer het alleen slechts interne (efficiëntie) winst op zou leveren. De samenwerking met gemeente, politie, JCA en TNO was nodig om dat proces goed in te kunnen richten.

7 **Sensing waarderen op effecten**. (Gezichts)herkenning is een middel waarmee politie en veiligheidspartners bepaalde maatschappelijke effecten kunnen pogen te bereiken. Het beheersen van negatieve bijeffecten hoort hier ook bij. De manier waarop (gezichts)herkenning wordt ingezet (denk aan bejegening en aan vermijdbaarheid) maakt daarin veel uit, dus (opnieuw) zijn de inzetconcepten essentieel. Tijdig een compleet beeld krijgen van maatschappelijke effecten maakt het ook mogelijk om de motivatie voor eventuele implementatie op tijd te formuleren.

C De basis van automatische gezichtsherkenning

In deze bijlage wordt de basis van gezichtsherkenning beschreven.

C.1 Het gezicht

Het menselijk gezicht wordt gedefinieerd als de voorkant van het hoofd vanaf voorhoofd tot en met de kin en inclusief de oren. Ogen, neus en mond horen hier dus bij, maar niet de binnenkant van de mond of neus.

Gezichten van twee verschillende personen kunnen in principe van nature identiek zijn[31]. Er is geen biologische reden waarom gezichten uniek zouden zijn. Het is weliswaar mogelijk om mathematisch te beredeneren dat in een bepaalde beperkte groep mensen, enkelvoudigheid aannemelijk te maken is naarmate een gezicht in meer detail is beschreven (Lucas & Henneberg, 2015). Echter, deze redenering gaat niet op in de complexe realiteit (Meester, Preneel, & Wenmackers, 2019). Vooralsnog moet er van worden uitgegaan dat dubbelgangers (voor wat betreft hun gezichten) op natuurlijke wijze bestaan.

Met behulp van plastische chirurgie kunnen gezichten worden veranderd. Daarbij kunnen zowel expres als per ongeluk gelijkenissen ontstaan met andere mensen. Er is geen mechanisme dat stuurt op uniciteit.

C.2 De afbeelding van het gezicht en het template

Het gezicht kan op verschillende manieren worden bemonsterd om een afbeelding te krijgen. Variatie is er in:

- de modaliteit: zichtbaar licht, infrarood, mmwave
- de ruimtelijke dimensies: 2-dimensionaal en 3-dimensionaal
- de spatiele resolutie van de bemonstering
- de tijdsduur en frequentie van de bemonstering, variërend van praktisch instantaan voor een foto, tot een willekeurige lengte bij video.

In de praktijk zijn allerlei ethische, operationele en technische beperkingen waar rekening mee moet worden gehouden. De inhoud van dit rapport is valide voor alle mogelijke variaties. De experimenten zijn gedaan met 2-d kleuren foto's van gangbare resolutie.

Van het gezichtsopname kan een biometrisch gezichtstemplate worden berekend. Technisch gezien, is een biometrisch template een rij (vector) van getallen. Het vergelijken van twee (of meer) gezichtstemplates gebeurt typisch door een afstand te berekenen tussen de twee betreffende vectoren. De betekenis van deze rij getallen is specifiek voor een algoritme. Er is géén internationale standaard voor een gezichtstemplate. Verschillende soorten gezichtsherkenningssoftware werken

[31] Dat is een belangrijk verschil met bijvoorbeeld kentekens. Die zijn immers kunstmatig en expres ontworpen om uniek te zijn. Dat is relevant in deze studie omdat de inspiratie voor dit onderzoek kwam uit een concept dat gebaseerd was op kentekens.

dus in principe technisch niet samen. Er zijn wel standaarden voor de beelden op basis waarvan de gezichtstemplates worden gemaakt.

C.3 Het proces en de functionele onderdelen

Er zijn meerdere stappen nodig om dit te laten werken:

- Parametrisatie: Het bepalen van de parameters van het algoritme is nodig om de specifieke werking te bepalen. Gezichtsherkenningspakketten komen typisch ook met een standaard instelling, waardoor deze stap technisch gezien optioneel is[32]. Tegenwoordig worden dit soort parameters geleerd op basis van een aparte verzameling gezichten waarvan bekend is welke van dezelfde persoon zijn. Dit heet een trainingsdataset.
- Enrolment: het aanmelden van gezichten en andere relevante gegevens in de database van het systeem.
- Herkennen: Het herkennen van live gezichten van “voorbijgangers” tegen gezichten in de database. Dit gebeurt door een afstandsmaat te berekenen tussen iedere combinatie van biometrische templates, en die om te zetten naar een maat voor gelijkenis – typisch uitgedrukt in percentage. Hierbij wordt in de terminologie onderscheid gemaakt tussen verificatie en identificatie. In dit rapport gaat het alleen over identificatie ten behoeve van negatieve herkenning (zie ook 4.2.1 ”Positieve en negatieve herkenning”).
  - Een terechte herkenning heet een correcte identificatie (ook wel “hit”).
  - Een foutieve herkenning heet een foutieve-identificatie (En. *false identification*).
  - Een terechte niet-herkenning heet een correcte niet-identificatie .
  - Een foutieve niet-herkenning heet een foutieve niet-identificatie.
- Opruimen: gezichten en bijbehorende gegevens worden weer verwijderd in de database van het systeem.

Zowel de enrollment als de herkenning gebeuren typisch in drie stappen:

1 Er wordt een opname gemaakt van het gezicht van een persoon. Dit kan volcontinue (als in video), of door een externe *trigger*, zoals het aanbieden van een pas.
2 Gezichtsdetectie: in het beeld wordt een gezicht gedetecteerd. Fouten in deze stap heten failure-to-capture (als een gezicht wordt gemist) en false detection (als er onterecht een gezicht wordt gedetecteerd waar er geen is).
3 Gezichtscodering: de relevante karakteristieken van een gezicht worden vastgesteld en deze worden in een gezichtstemplate geplaatst

Herkenningstechnologie werkt niet perfect. Mensen kunnen dus onterecht niet worden herkend, en een “herkenning” dus onterecht kan zijn. Na een herkenning volgt er dus een essentiële stap die in jargon “alarmresolutie” heet. Alarmresolutie heeft dus als doel om de resterende onzekerheid zo veel mogelijk te verkleinen en te bepalen op welke manier het proces verder moet verlopen. Dat kan in een aantal stappen verlopen. Een eerste stap kan bijvoorbeeld neerkomen op het maken van

[32] Het is tegenwoordig bekend dat de standaard configuratie van een gezichtsherkenningssysteem vertekende resultaten kan geven. Zie bijlage H “Vertekeningen (biases) bij niet-coöperatieve gezichtsherkenning”. Industrie claimt dat ze dit aan het verbeteren zijn.

een betere gezichtsopname. Bijvoorbeeld vanuit een betere opnamehoek en / of een kleinere afstand, zoals met behulp van een *bodycam*. Het ligt ook erg voor de hand om als onderdeel van alarmresolutie om identiteitsdocumenten te vragen en te controleren.

C.4 Persoonsgegevens bij gezichtsherkenning

De AVG en WPG[33] definiëren persoonsgegevens als volgt:

> *alle informatie over een geïdentificeerde of identificeerbare natuurlijke persoon ("de betrokkene"); als identificeerbaar wordt beschouwd een natuurlijke persoon die direct of indirect kan worden geïdentificeerd, met name aan de hand van een identificator zoals een naam, een identificatienummer, locatiegegevens, een online identificator of van een of meer elementen die kenmerkend zijn voor de fysieke, fysiologische, genetische, psychische, economische, culturele of sociale identiteit van die natuurlijke persoon;*

Daar bovenop zijn er in de AVG ook *gevoelige* persoonsgegevens gedefinieerd met daarbij ook biometrische gegevens (artikel 9 AVG)[34]. De verwerking daarvan is in principe verboden, en alleen toegestaan voor bepaalde uitzonderingen.

Als we deze definities toepassen op automatische gezichtsherkenning, dan gaat het om de *enrollment* gegevens, om de live opnames en ook om trainings- en evaluatiedata.

Wat een identificator is, hangt af van de mogelijkheden van de verwerker van de betreffende gegevens, en ook van de partijen waarmee hij samenwerkt. Als hij zelf toegang heeft tot informatie of verwerkingskracht waarmee een gegeven alsnog als een identificator kan worden gebruikt, dan is het dus een identificator. En andere partijen die in een samenwerking zitten met die partij, moeten die informatie dan zelf ook als persoonsgegeven beschouwen – ook als ze het zelf niet als identificator kunnen gebruiken. Dus als een private beveiliger een unieke sleutel heeft van een biometrisch template (bijv. “template nr 511”) en samenwerkt met de politie die de bijbehorende *watchlist* heeft, dan moet die private beveiliger die unieke sleutel ook als persoonsgegeven beschouwen – ook als ze zelf géén toegang heeft tot die *watchlist*.

Achtereenvolgens worden hieronder identificatoren en overige persoonsgegevens in meer detail behandeld.

C.4.1 Directe identificatoren bij gezichtsherkenning

Het is evident dat de afbeelding (pixels) van het gezicht van een persoon een directe identificator is. Ook als er meerdere personen met hetzelfde gezicht bestaan

---

[33] De WPG verwijst voor de definitie van persoonsgegevens naar de AVG (WPG, artikel 1, lid m).

[34] AVG Artikel 9 “Verwerking van bijzondere categorieën van persoonsgegevens”. *Verwerking van persoonsgegevens waaruit ras of etnische afkomst, politieke opvattingen, religieuze of levensbeschouwelijke overtuigingen, of het lidmaatschap van een vakbond blijken, en verwerking van genetische gegevens,* biometrische gegevens met het oog op de unieke identificatie van een persoon*, of gegevens over gezondheid, of gegevens met betrekking tot iemands seksueel gedrag of seksuele gerichtheid zijn verboden.* – In het volgende lid van dit artikel worden tien uitzonderingen op dit verbod gegeven.

(tweelingen, *look-alikes*). Ook is het daarvan afgeleide biometrische template een directe identificator (AVG, artikel 9). Beiden zijn immers kenmerkend voor de fysieke identiteit van de persoon. Ze zijn overigens ook uitwisselbaar voor elkaar. Van een afbeelding van een gezicht is een biometrisch template te maken, en van een biometrisch template is een herkenbare afbeelding van een gezicht te maken (zie sectie G.3.1.1) (Chibba & Stoianov, 2014).

In sectie 4.2.2. van (Keymolen, Noorman, Van der Sloot, Cuijpers, & Koops, 2020) beweert de Privacy en Public Policy manager van Facebook dat biometrische templates niet door politie te gebruiken zijn, omdat de biometrische templates niet interoperabel[35] zijn. Dat statement is niet volledig. Ten eerste is het met toegang tot de software van Facebook mogelijk om met behulp van die templates op internet gezichten te vinden die daar -volgens de software van Facebook- op lijken. Vervolgens kunnen daar templates van worden gemaakt met een ander gezichtsherkenningssoftware. Ten tweede is een gebrek aan interoperabiliteit geen informatiebeveiliging. Het is immers simpelweg een kwestie van het maken van een koppelvlak, om de biometrische templates wel degelijk direct te kunnen gebruiken.

Voor de duidelijkheid, en nogmaals, ook zonder deze uitwisselbaarheid is een biometrisch template een identificator, en dus een persoonsgegeven.

C.4.2 Indirecte identificatoren bij gezichtsherkenning

Daarnaast zijn er ook indirecte identificatoren (zie ook sectie C.7 “Harde en zachte biometrie”). Dat zijn gegevens die tezamen een identificator vormen. Dus bijvoorbeeld een stuk kleding, een sieraad, een bril, een tatoeage, gezichtsbeharing, huidskleur, een tijdstip en locatie en een ruwe opname van iemand die daar in de buurt een treinticket koopt waarop die attributen ook te zien zijn, vormen tezamen ook een identificator. Welke combinaties van gegevens er kunnen bestaan, hangt van de toepassing af. Om dit te onderscheiden van gezichten, noemen we dit persoonlijke attributen.

Gezichtsherkenning is vaak ook gekoppeld aan andere gegevens, zoals aan identiteiten of identiteitsdocumenten. Uiteraard zijn dat ook identificatoren.

C.4.3 Overige persoonsgegevens bij gezichtsherkenning

Overige persoonsgegevens zijn alle gegevens die aan een identificator gekoppeld kunnen worden. Dat kan dus gaan om gebruiksgegevens en metadata, logs met hits, en ook andere gekoppelde gegevens. Bij een niet-coöperatieve toepassing kan het bijvoorbeeld nuttig zijn om te registreren of de persoon vuurwapengevaarlijk is, of de reden waarom hij gezocht wordt. Dit zijn dan dus ook persoonsgegevens. Ook als deze niet apart als gegeven opgeslagen staan, maar evident zijn uit het doel van het betreffende systeem.

---

[35] Hier wordt *technische* interoperabiliteit bedoeld. Dat is de capaciteit van technische systemen om zonder beperkingen technisch samen te werken. Voor gezichtsherkenning zijn internationale standaarden gemaakt om biometrische templates te beschrijven en tussen systemen uit te kunnen wisselen. Bijvoorbeeld “ISO/IEC 19794-5:2005 biometric data interchange format standard for face image data”.

C.5 Rollen bij gezichtsherkenning

Deze processen kunnen tezamen door één partij worden uitgevoerd die zichzelf volledig vertrouwt. Ook is het mogelijk dat er twee of meer partijen samenwerken die tegelijkertijd een belang hebben om informatie voor partners te verbergen. Daar zijn veel denkbare varianten van, bijvoorbeeld:

- De partij die de *enrollment* beelden maakt, wil voorkomen dat die beelden kunnen worden gereconstrueerd uit de biometrische templates.
- De partij die de database beheert wil voorkomen dat de biometrische templates gedeeld worden met de partij die de live beelden maakt.
- De partij die de live beelden maakt wil voorkomen dat die beelden herkenbaar bij de eigenaar van de database komt.

C.6 Factoren die de accuratesse beïnvloeden

Gezichtsherkenning werkt niet 100% accuraat. Zeker niet op basis van opnames die in minder-gecontroleerde omstandigheden zijn gemaakt, zoals rond een incident, zoals in een menigte en zelfs niet bij een toegangscontrole waar het subject een belang heeft om mee te werken.

Er zijn meerdere factoren die de kwaliteit van gezichtsherkenning in significante mate bepalen. Ten eerste hoeft een gezicht niet uniek te zijn (zie sectie C.1).

Daarnaast zijn er een aantal technische factoren (NIST, 2019):

- Kwaliteit van de beelden uit de *enrolment* en uit de live opnames
  - Occlusie (denk aan gezicht bedekkende kleding en maskers[36], maar ook aan drukke mensenmassa's)
  - Opnamehoek
  - Belichting op het gezicht: sterkte en verdeling
  - Resolutie op het gezicht
  - Contrast
  - Bewegingsonscherpte door beweging van subject ten opzichte van camera en door beweging van achtergrond ten opzichte van camera.
- Kwaliteit software
  - Kwaliteit gezichtsdetectie
    - Kwaliteit van testdata.
  - Kwaliteit gezichtsherkenning
    - Kwaliteit van testdata.
- Grootte van de database.

Er zijn ook operationele factoren die deze technische factoren sterk beïnvloeden:

- Begeleiding / toezicht op de *enrolment* en op de live opnames.
- Attitude van subjecten ten opzichte van *enrolment* en van live opnames.

[36] De Corona pandemie wordt mede bestreden door het gebruik van gezichtsmaskers. Die hebben invloed op de prestatie van gezichtsherkenning. NIST heeft op 1 mei 2020 aangekondigd om daar onderzoek naar te doen.

C.7 Harde en zachte biometrie

Gezichtsherkenning wordt gezien als een vorm van *harde biometrie* (Jain, Flynn, & Ross, 2007). Harde biometrie[37] is de verzameling biometrische kenmerken die niet kan worden veranderd aan een persoon omdat ze genetisch bepaald zijn: ze zijn permanent[38]. Er is geen enkele garantie dat gezichten uniek zijn (zie ook sectie C.1). De vraag of er ooit technologie komt die *alle* mensen ter wereld puur op basis van hun gezicht uit elkaar kan houden, is dus niet sluitend te beantwoorden. Naast gezichten horen hier ook vingerafdrukken en DNA bij. Juridisch gezien wordt dit soort biometrie gezien als een directe identificator (zie ook sectie C.4).

Ter vergelijking, *zachte biometrie* omvat herkenningskenmerken die wel kunnen worden veranderd – desnoods middels training of conditionering, zoals de manier hoe iemand zich beweegt of kleedt, of als de kenmerken die minder uniek beschrijvend voor een persoon (Jain, Flynn, & Ross, 2007). Juridisch gezien wordt dit soort biometrie gezien als een indirecte identificator als ze samen met andere gegevens alsnog kenmerkend zijn voor een individu of kleine groep mensen (zie ook sectie C.4).

C.8 Attitude ten opzichte van gezichtsherkenning

In dit rapport maken we onderscheid tussen drie soorten attitude die voorbijgangers kunnen hebben in het herkenningsproces ten opzichte van de toepassing van gezichtsherkenning:

- Vriendelijk: ze zullen in beginsel mee willen werken. Eventuele fouten komen door onkunde, fysieke beperkingen of onbegrip, maar niet door onwil. Dit zijn typisch mensen die een belang hebben dat er een goede opname van hen wordt gemaakt.
- Neutraal: ze zullen niet mee of tegen willen werken. Eventuele fouten kunnen ook komen door onwil om mee te werken, maar worden niet opzettelijk veroorzaakt. Dit zijn typisch mensen die denken dat ze geen belang hebben bij de kwaliteit van de gezichtsdetectie, of die niet weten of zich niet beseffen dat er gezichtsherkenning plaats vindt.
- Vijandelijk: ze zullen niet meer willen werken. Fouten kunnen ook komen uit onwil om mee te werken waarbij het subject probeert ze opzettelijk te veroorzaken. Dit zijn mensen die een belang hebben bij een slechte kwaliteit gezichtsdetectie. Zie ook bijlage D “*Design basis threat* voor gezichtsherkenning” voor voorbeelden van dit soorten mensen, en de volgende sectie voor methoden om dat te bereiken.

---

[37] Er zijn verschillende definities van harde en zachte biometrie in omloop. Zachte biometrie wordt ook wel gedefinieerd als die kenmerken die een mens op natuurlijke wijze kan gebruiken. Of als die kenmerken die niet uniek zijn voor een persoon. In dit rapport gebruiken we *permanentie* als definitie omdat die het meest relevant is voor een toepassingsgebied met een lerende tegenstander.

[38] Gezichten kunnen ook veranderen door veroudering, plastische chirurgie, make-up en tatoeages.

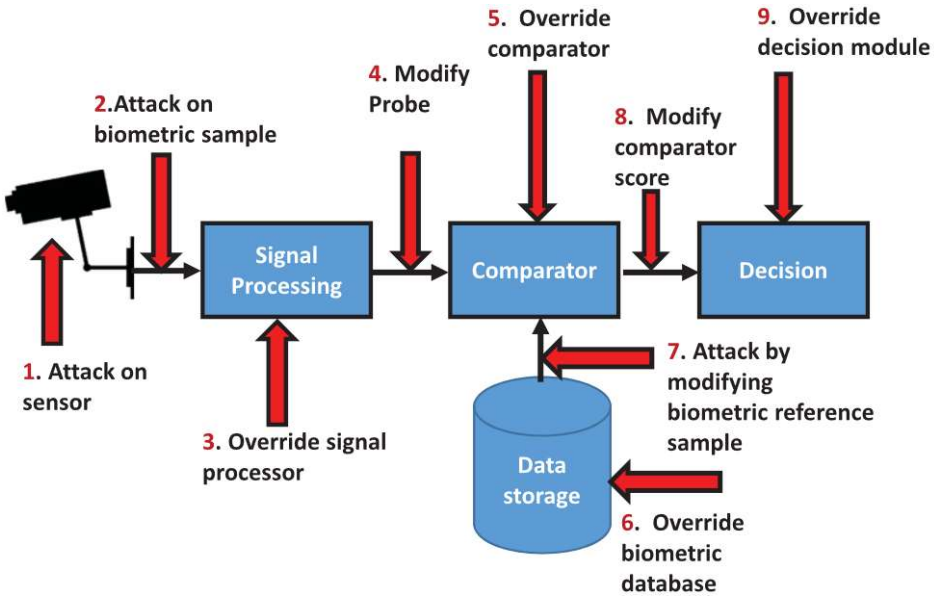
C.9 Aanvallen tegen gezichtsherkenning

Er zijn verschillende manieren voor een vijandelijke partij om gezichtsherkenning te frustreren. Voor het doel van dit rapport is het voldoende om ter illustratie een aantal voorbeelden te geven. Zie ook bijlage D "*Design basis threat* voor gezichtsherkenning" voor een beschrijving van actoren die dit soort aanvallen zouden kunnen uitvoeren.

Aanvallen tegen de correcte werking kunnen bedoeld zijn om (tijdelijk) de identiteit van een ander aan te nemen. Als dit gebeurt door het presenteren van valse kenmerken, dan heet dit *spoofing*. Een voorbeeld is het presenteren van een masker dat op iemand anders lijkt. Dit komt dus neer op het genereren van een foutieve niet-identificatie. In de context van niet-coöperatieve gezichtsherkenning, zal de tegenstander daar op uit zijn. Hij wil immers *niet* herkend worden als iemand die op een *watchlist* staat.

Aanvallen die werken door andere kenmerken aan te bieden, zoals spoofing, heten *presentation attacks*. Een *presentation attack* kan er ook simpelweg op gericht zijn om een *failure to capture* te genereren. Dus er voor te zorgen dat de camera überhaupt geen gezicht detecteert[39]. Voorbeelden hiervan werken bijvoorbeeld door camouflerende kleding of make-up. Het gebruik van gezicht bedekkende kleding is hier ook een voorbeeld van, mits het met dit doel wordt gedragen.

Er zijn ook allerlei aanvallen denkbaar tegen het systeem zelf (zie figuur 4). Daaronder behoren ook aanvallen tegen de database waarin de *watchlist* zelf opgeslagen zit.



Figuur 4 Verschillende soorten aanvallen tegen het biometrisch herkenningssysteem (Ramachandra & Busch, 2017).

[39] Een *presentation attack* moet niet worden verward met een systeem dat niet in staat is om mensen van bepaalde demografieen (denk aan huidskleur) goed te bemonsteren. Zie ook sectie H "Vertekeningen (biases) bij niet-coöperatieve gezichtsherkenning".

D *Design basis threat* voor gezichtsherkenning

In deze bijlage wordt een *design basis threat* (DBT) op hoofdlijnen beschreven. Een DBT is een expliciete beschrijving van de dreiging. In deze DBT wordt gefocust op de “dreigende actor”. Een dreigende actor is een persoon of een groep personen die een dreiging vormen. Deze actoren worden ook gebruikt om de toepassingen van gezichtsherkenning elders in dit rapport mee te verduidelijken. Niet coöperatieve gezichtsherkenning (NCG) heeft op drie manieren met dreigende actoren te maken. Figuur 5 illustreert hoe deze dreigende actoren te maken hebben met risico’s.

- **Maatregel**: NCG kan een maatregel (M1) tegen bepaalde dreigende actoren (uit risico I1) zijn;
- **Doelwit**: Dreigende actoren (uit risico I2) kunnen de correcte werking van NCG bedreigen:
  - een dreigende actor kan gezichtsherkenning misbruiken (restrisico R1);
  - een dreigende actor kan de correcte werking van gezichtsherkenning tegen gaan (restrisico R2).

Figuur 5 Risico reductie overzicht van niet-coöperatieve gezichtsherkenning. Initiële risico’s I1 en I2 worden (ten dele) gemitigeerd door maatregel M1: de niet-coöperatieve gezichtsherkenning. Als het goed werkt, dan levert dat Finaal restrisico F1 op. Daarnaast kan het systeem worden misbruikt. Dat levert restrisico R1 op. Bij de aanwezigheid van malafide actoren die zich tegen de correcte werking wenden, levert dat ook restrisico R2 op.

Met betrekking tot NCG als doelwit: een aantal privacydreigingen veronderstellen de aanwezigheid van een concrete actor die door zijn (niet) handelen de privacy bedreigt. Het is nuttig om die actoren expliciet, tot op zeker detailniveau, te beschrijven. Deze beschrijvingen kunnen bijvoorbeeld worden gebruikt om tijdens test- en gebruiksfases de spelers van *red-teaming* operaties mee te ontwerpen.

Met betrekking tot NCG als maatregel: in de potentiele toepassingen (zie bijlage E) wordt gebruik gemaakt van onderstaande dreigende actoren.

Ieder van de dreigende actoren wordt hieronder kort geïntroduceerd. Vervolgens wordt NCG als maatregel tegen de betreffende actor beschreven, en wordt beschreven hoe de actor NCG als doelwit kan zien.

D.1 Vandaal

Deze actor is gedreven door opportunisme en sensatie. Persoonlijke eer ten opzicht van een sociale groep is een grote motivatie. Ze treden daarom vaak in groepen op. Hooligans horen hier ook bij. Modus operandi bestaat typisch uit geweld tegen spullen, maar kan ook richting willekeurige (kwetsbare) voorbijgangers gericht zijn. In extreme gevallen kan het gaan richting zwaar geweld tegen concurrerende sociale groepen waar ook enige voorbereiding voor nodig kan zijn.

Maatregel: Over het algemeen zal het niet proportioneel zijn om niet-coöperatieve gezichtsherkenning te gebruiken tegen vandalisme. Maar als het met grote impact (bijv. grof geweld) gebeurt, dan wellicht wel.

Doelwit: Camera's die op straat staan kunnen doelwit worden van vandalisme. Dat zal over het algemeen gaan om (gerichte) vernieling, niet om diefstal.

D.2 Kleine crimineel

Deze actor is gedreven door geldelijk gewin. Ze treden typisch alleen of in kleine aantallen op. Modus operandi bestaat uit diefstal van spullen en beroving. In extreme gevallen kan geweld worden gebruikt, tot en met doodslag aan toe en bijbehorende (voorbereidende) handelingen.

Maatregel: Dit is een gevarieerde groep, waardoor de proportionaliteit van het gebruik van niet-coöperatieve gezichtsherkenning niet eenduidig te beschrijven is.

Doelwit: Camera's die op straat staan kunnen doelwit worden van diefstal, maar daar is relatief eenvoudig tegen te beveiligen. Het zijn immers camera's. Een groter risico is dat ze onbruikbaar worden gemaakt om zodoende ongestoord hun misdaad te kunnen plegen. En een nog groter risico is *spoofing*.

D.3 Zware crimineel / terrorist

Deze actor is gedreven door fors geldelijk gewin of het bereiken van een maatschappelijk of politiek doel. De kern treedt typisch alleen of in kleine aantallen op, maar er kan een uitgebreid crimineel netwerk omheen zitten. Modus operandi kan bestaan uit grof geweld inclusief meervoudige moord en ontwrichting van de maatschappij en bijbehorende (voorbereidende) handelingen.

Maatregel: De inzet van bijzondere opsporingsmiddelen kan proportioneel zijn tegen deze soort actoren. Wellicht geldt dat ook voor niet-coöperatieve gezichtsherkenning.

Doelwit: Met name zware criminelen kunnen moeite doen om de werking van niet-coöperatieve gezichtsherkenning tegen te gaan. Bijvoorbeeld door camera's te

saboteren, of door te proberen de *watchlist* te achterhalen om te weten of zij gezocht worden. Zij kunnen ook de middelen en kennis verzamelen om dat te doen. Voor (zelfmoord) terroristen is het minder logisch om daar in te investeren. Ten eerste kunnen ze proberen zo snel te zijn dat ze de interventie vóór blijven, en ten tweede is er soms sprake van een mate van zelfopofferingsgezindheid, van waaruit ze een hoge pakkans accepteren, als ze maar hun aanval hebben kunnen plegen.

D.4 Statelijke actor

Deze actor is gedreven door nationale belangen van bijvoorbeeld geopolitieke of economische aard. Dit soort doelen kunnen een bedreigingen vormen voor de nationale veiligheid. Er zijn statelijke actoren met zeer veel kennis en informatie, en met een agenda op zeer lange termijn. Dit zijn qua potentiele schade de gevaarlijkste tegenstanders. De modus operandi kan zeer gevarieerd zijn. Spionage, infiltratie en ook propaganda en *hybride oorlogsvoering* zijn denkbaar.

Maatregel: Het lijkt proportioneel niet-coöperatieve gezichtsherkenning in te zetten tegen agenten van vijandelijke staten in Nederland.

Doelwit: Statelijke actoren kunnen het nuttig vinden om gezichtsherkennings-systemen in Nederland te saboteren, of om ze voor eigen doeleinden (verkeerd) te laten werken. Ook kan het nuttig zijn voor hen te proberen de *watchlist* te achterhalen om te weten of zij “in beeld” zijn. Het werven van corrupte medewerkers behoort ook tot de mogelijkheden. Ook het beïnvloeden van de *supply chain* zou kunnen. Bijvoorbeeld het hacken van de gezichtsherkennings-systemen van een leverancier uit het eigen land, die deze gehackte hard- en software aan Nederland levert. Via propaganda kunnen activisten worden beïnvloed om zich tegen gezichtsherkenning te keren.

D.5 Activist

Deze actor is gedreven door het bereiken van een kleiner of groter maatschappelijk of politiek doel. Ze kunnen zowel alleen, als in grote groepen optreden. Over het algemeen werken ze op legitieme wijze middels onthullingen in de media en demonstraties. Daarbij kunnen ze gebruik maken van infiltratie, hacking en stelen van (gerubriceerde) informatie, waaronder ook persoonsgegevens. Demonstraties kunnen uit de hand lopen, en / of zich richten tegen hulpdiensten en hun middelen. In extreme gevallen kan een activist zich qua modus operandi gaan bewegen richting een terrorist.

Maatregel: De inzet van niet-coöperatieve gezichtsherkenning is vermoedelijk over het algemeen niet proportioneel tegen deze groep, maar er kunnen uitzonderingen zijn voor extreme gevallen.

Doelwit: Of activisten zich tegen niet-coöperatieve gezichtsherkenning keren zal sterk afhangen van hun attitude ten opzichte van een handhavende overheid. De beelden van demonstranten in Hong Kong die lantaarnpalen om ver trekken waar (vermeend) gezichtsherkenning in zit, zijn bekend. Maar dit zijn uitzonderingen. In

steden als Den Haag en Amsterdam zijn iedere dag protesten en demonstraties. Verreweg de meeste verlopen zonder schade aan de toezichtinfrastructuur.

D.6 Onvoorzichtige medewerker

Deze actor is gedreven door een gebrek aan motivatie of alertheid, mogelijk gecombineerd met een verkeerde werkdruk. Er is typisch geen sprake van opzet, maar er kan wel sprake zijn van nalatigheid. De mogelijke modus operandi is enorm gevarieerd. Die kan bestaan uit onzorgvuldig omgaan met authenticatie middelen (bijv. wachtwoorden). Het kan ook indirecter, als het gaat om slecht (ICT) onderhoud op de systemen. Een camera systeem dat niet goed wordt onderhouden, zal slechte beelden gaan geven, wat tot lagere nauwkeurigheid en dus fouten gaat leiden. Deze actor kan bij de “eigen” organisatie werken, maar ook bij een partnerorganisatie, zoals bijvoorbeeld bij een cloud-dienstverlener. Deze actor kan een aantrekkelijk doelwit zijn voor de een andere dreigende actor: de corrupte medewerker.

Maatregel: Over het algemeen zal het niet proportioneel zijn om niet-coöperatieve gezichtsherkenning te gebruiken tegen onvoorzichtige medewerkers. Het zal ook niet subsidiair zijn omdat er vermoedelijk allerlei andere, minder ingrijpende, maatregelen mogelijk zijn omdat het een medewerker betreft.

Doelwit: Automatische gezichtsherkenning kan bij grote mensenstromen tot veel handmatig werk leiden in de opvolging. Zeker als het systeem niet goed is geïnstalleerd of geconfigureerd. Als die werkdruk niet goed wordt beheerst, dan kan een onvoorzichtige medewerker tot forse problemen leiden. Dit kan in theorie een negatieve spiraal worden als hoge werkdruk tot (nog) slechter onderhoud leidt.

D.7 Fanatieke medewerker

Deze actor is gedreven door professionele erkenning, mogelijk ook door geld, sociale druk vanuit een bepaalde cultuur of persoonlijke motieven om zijn werk zo “goed” mogelijk te doen. Er is typisch geen kwade opzet in het spel, maar wel eenzijdigheid. Mogelijk zijn werkprocessen niet goed ontworpen of beschreven. De mogelijke modus operandi is enorm gevarieerd. Die kan bestaan uit het toevoegen van mensen aan de *watchlist* die er niet op horen, of het instellen van verkeerde grenswaardes waardoor er op te veel mensen een hit komt. Het kan gaan om een ongepast strenge opvolging die niet past bij de (on)zekerheid die er nog is, of om het koppelen van onnodig veel informatie aan de *watchlist*. Deze actor kan bij de “eigen” organisatie werken, maar ook bij een partnerorganisatie.

Maatregel: Over het algemeen zal het niet proportioneel zijn om niet-coöperatieve gezichtsherkenning te gebruiken tegen fanatieke medewerkers. Het zal ook niet subsidiair zijn omdat er vermoedelijk allerlei andere, minder ingrijpende, maatregelen mogelijk zijn omdat het een medewerker betreft.

Doelwit: Automatische gezichtsherkenning kan in potentie enorm krachtig zijn. Zowel het afschrikwekkend vermogen op malafide personen, als het effect van de daadwerkelijke herkenning van malafide personen die terecht op de *watchlist* staan

is in potentie groot. Medewerkers kunnen door die kracht verleid worden om het vaker in te zetten of om er meer betrouwbaarheid aan toe te kennen dan gerechtvaardigd is.

D.8 Discriminerende medewerker

Deze actor is gedreven door een discriminerend motief gericht tegen bepaalde, of tegen andere bevolkingsgroepen van een andere leeftijd, geslacht, etniciteit, etc. Naast de handelingen die een onvoorzichtige en fanatieke medewerker kunnen doen, kunnen discriminerende medewerkers ook expres handelingen doen die het systeem saboteren. Bijvoorbeeld om hele andere mensen op de *watchlist* te zetten, of om mensen van de *watchlist* af te halen.

Maatregel: Over het algemeen zal het niet proportioneel zijn om niet-coöperatieve gezichtsherkenning te gebruiken tegen discriminerende medewerkers. Het zal ook niet subsidiair zijn omdat er vermoedelijk allerlei andere, minder ingrijpende, maatregelen mogelijk zijn omdat het een medewerker betreft.

Doelwit: Discriminerende medewerkers kunnen een eigen motief hebben om de werking van een gezichtsherkenningssysteem te beïnvloeden. Daar boven op, kunnen malafide organisaties interesse hebben in discriminerende mensen die toegang hebben tot gezichtsherkenning en hen proberen voor zich te winnen. Het kan gaan om zware criminelen die op geldelijk gewin uit zijn, terroristen met een ideologisch motief, en ook om statelijke actoren die de werking van een democratische rechtstaat willen ondermijnen.

D.9 Corrupte medewerker

Deze actor is gedreven door vanuit financieel belang. Hij is mogelijk omgekocht door een criminele organisatie, een terroristische cel of een vijandelijke staat. Er is zeker sprake van kwade opzet. Naast de handelingen die een onvoorzichtige en fanatieke medewerker kunnen doen, kunnen corrupte medewerkers ook expres handelingen doen die het systeem saboteren. Bijvoorbeeld om hele andere mensen op de *watchlist* te zetten[40], of om mensen van de *watchlist* af te halen.

Maatregel: Over het algemeen zal het niet proportioneel zijn om niet-coöperatieve gezichtsherkenning te gebruiken tegen corrupte medewerkers. Het zal ook niet subsidiair zijn omdat er vermoedelijk allerlei andere, minder ingrijpende, maatregelen mogelijk zijn omdat het een medewerker betreft.

Doelwit: Corrupte medewerkers zullen een externe (financiële) prikkel nodig hebben om de werking van het gezichtsherkenningssysteem te gaan beïnvloeden. Malafide organisaties kunnen interesse hebben in medewerkers die tegen financiële vergoeding toegang kunnen verschaffen tot gezichtsherkenning. Het kan gaan om zware criminelen die op geldelijk gewin uit zijn, terroristen met een ideologisch

[40] Volgens een recent bericht in de media is er een incident geweest in Moskou waarbij corrupte politieagenten tegen betaling willekeurige gezichten langs de gezichten database haalden, en de locaties en tijdstippen van hits teruggaven (Bacchi, 2020).

motief, en ook om statelijke actoren die de werking van een democratische rechtstaat willen ondermijnen.

E Potentiele toepassingen voor niet-coöperatieve automatische gezichtsherkenning

In deze bijlage zijn verschillende potentiele toepassingen verzameld. De bronnen daarvoor zijn heel divers. Uiteraard de deelnemers aan het project, de gemeente Amsterdam en de politie. Ook andere studies die recent zijn gepubliceerd zoals (Keymolen, Noorman, Van der Sloot, Cuijpers, & Koops, 2020) en (JRC, ERNCIP, 2019) hebben inspiratie gegeven. Een aantal van deze toepassingen passen wellicht binnen een “smart city” context. De meeste toepassingen gaan over de fysieke omgeving. Eén toepassing gaat over het monitoren van (private) online platformen.

E.1 Toepassing A: Object- en persoonsbeveiliging

De politie en / of KMar verzorgen object- en persoonsbeveiliging van diverse hoog-risico objecten en personen tegen vandalen, activisten, kleine en grote criminelen, terroristen en statelijke actoren. Gezichtsherkenning kan nuttig zijn. Onder het stelsel bewaken en beveiligen mag de overheid beveiligingsmaatregelen inzetten, onder andere bewakingscamera’s.

Gezichtsherkenning vereist de inzet van camera’s. Deze camera’s kunnen van de KMar of politie zelf zijn. In een stedelijke omgeving kunnen ze ook in beheer zijn van andere partijen, waaronder (private) gebouw- of gebiedsbeheerders of de gemeente. Ieder van deze beheerders willen in controle blijven over waar gezichtsherkenning op wordt toegepast, ook (of met name) als ze dat niet zelf doen.

De politie mag en wil hun *watchlists* (per te beveiligen object / persoon) typisch niet overdragen aan de gemeente en gebouwbeheerders. Dit scenario is verder uitgewerkt in een concept op operationeel niveau in sectie 7.2 “Concept voor persoons- en bijbehorende objectbeveiliging in een rustige wijk”.

Specifiek voor objecten die onder vitale infrastructuur vallen, heeft DG HOME van de Europese Commissie een studie gevraagd en gekregen over de wenselijkheid en mogelijkheden van gezichtsherkenning. Het gaat dan specifiek om het gebruik van gezichtsherkenning in de observatiering rond dergelijke objecten. Zie bijlage F "Gezichtsherkenning in de observatiering van vitale infrastructuur”.

In deze toepassing kan het gaan om verticale relaties, horizontale relaties en combinaties daarvan.

E.2 Toepassing B: Opsporing heterdaadfase

Kort na een melding van een incident, bijvoorbeeld een overval of beroving, of na een zedenmisdrijf, wanneer de betrokkenen nog op enige manier “in beeld” kunnen zijn, is er sprake van de heterdaadfase. Dit heet ook wel “manhunt”. In die fase is de pakkans het grootst. De politie kan in die heterdaadfase een belang hebben om de mogelijkheid te hebben om rond specifieke plaatsen-delict gezichten vast te

leggen en te vergelijken met de opnames van het incident (een ad hoc *watchlist*). De gemeente en beheerders van infrastructuur willen in controle blijven over de camera's waar herkenningstechnologie op wordt toegepast. Maar de politie mag en wil dergelijke ad hoc *watchlists* mogelijk niet overdragen aan de gemeente en infrastructuurbeheerders.

Een variant van deze toepassing (B-evenement) is als het incident plaats vindt binnen een privaat beheerd beveiligd gebied. Dan heeft de evenementbeheerder zowel de beelden van het incident, als de beelden van de toegang zelf in handen.

In deze toepassing kan het gaan om verticale relaties, horizontale relaties en combinaties daarvan.

E.3 Toepassing C: Opsporing "most wanted" en vermiste personen

Politie en andere veiligheidsorganisaties voeren opsporingsonderzoeken uit naar gevaarlijke criminelen, terroristen en statelijke actoren. Het is bekend dat zij zich veilig wanen en zich tussen het normale publiek in de (semi-)openbare ruimte bewegen: op straat, in het openbaar vervoer, etc. Het kan nuttig zijn voor dergelijke onderzoeken om op bepaalde plaatsen waar de kans significant is dat deze gezochte criminelen passeren, gezichten vast te leggen en te vergelijken met een *watchlist*. De gemeente en beheerders van infrastructuur willen in controle blijven over de camera's waar herkenningstechnologie wordt toegepast. Maar de politie mag en wil de *watchlists* niet overdragen aan de gemeente en infrastructuurbeheerders.

Een specifieke variant hiervan gaat om gezichtsherkenning aan de grens. De Schengen grenscode staat toe om gezichten van bekende gezochte personen in het *Schengen Information System* (SIS) op te slaan, en om daarmee gezochte personen te herkennen bij grensposten.

Het kan ook gaan om het terugvinden van vermiste personen zonder vermoeden van criminaliteit. De burgemeester (bestuursrecht) kan camerabeelden daartoe willen inzetten (Homburg, Schreijenberg, & Stouten, 2014).

In deze toepassing gaat het typisch om een verticale relatie tussen overheid en burger, maar het kan voorkomen dat de overheid gebruik wil maken van bestaande camera's van bijvoorbeeld een privaat OV bedrijf, zoals van een luchthaven of spoorbeheerder[41].

E.4 Toepassing D: Beschermen soft targets

Er is de afgelopen jaren sprake van een terroristische dreiging op soft targets, locaties waar grote hoeveelheden mensen bijeen zijn. Terroristen maken zich minder zorgen over herkenning tijdens een aanslag. Recente aanslagen laten zien dat zij meer rekenen op verrassing, snelheid en slagkracht. Ook is er geleerd dat in

[41] Na een terroristische aanslag op een kerstmarkt in Berlijn vluchtte de dader onder andere via de trein via Nederland naar Italie. De *manhunt* die zich op zo'n moment afspeelt kan worden ondersteunt middels niet-coöperatieve gezichtsherkenning.

het geval ze zelf niet omkomen bij de aanslag, ze een volgende aanslag willen plegen. Omdat de opsporingsdiensten hen dan op de hielen zitten, doen ze dat noodgedwongen op korte termijn, en dus relatief nabij. De modi operandi waar gemeenten zich het meest zorgen over maken, maken gebruik van makkelijk verkrijgbare middelen zoals auto's, messen en vuurwapens. Tegen inrijdende voertuigen worden bij drukke plaatsen extra fysieke barrières geplaatst. Maar dat helpt niet tegen uitgestegen terroristen. De politie en de KMar hebben de mogelijkheid om bij hoge dreigingen gebruik te maken van *watch-lists* van bekende geradicaliseerde personen, en om deze personen te signaleren middels cameratoezicht.

De politie heeft biometrische gezichtstemplates van veelplegers en / of van gevaarlijke criminelen. Dit zijn geclassificeerde gegevens. Voor dit doel, en bij groet hoeveelheden mensen, is het wellicht denkbaar dat de politie toegang krijgt tot live beelden met voldoende kwaliteit voor gezichtsherkenning van camera's die er staan voor handhaving openbare orde.

De locatiebeheerder of evenementorganisator (zoals JCA) heeft typisch, net als andere hoog-risico objecten, ook cameratoezicht. Het is ook in hun belang dat gevaarlijke criminelen worden herkend als ze proberen het evenement binnen te komen. Echter, enerzijds kan de evenementorganisators kan niet om gaan met geclassificeerde gegevens zoals een *watchlist* van geradicaliseerde personen. Anderzijds wil de evenementorganisator misschien liever niet dat alle gezichtsopnames van bezoekers naar de politie gaan.

In deze toepassing gaat het typisch om een verticale relatie tussen overheid en burger, maar het kan voorkomen dat de overheid daartoe gebruik wil maken van bestaande camera's van bijvoorbeeld een privaat OV bedrijf.

E.5 Toepassing E: Handhaving contact-, winkel-, OV-, stadion- of gebiedsverbod

Er zijn sinds 2007 in Nederland verschillende experimenten geweest voor de handhaving van winkel, OV- en gebiedsverboden met behulp van gezichtsherkenning (VVD Amsterdam, 2014) (Janssen, Kool, & Timmer, 2015). Het gaat hier typisch om kleine criminelen die herhaaldelijk winkeldiefstal of zwartrijden hebben gepleegd. Het kan ook om zwaardere vergrijpen gaan inclusief *stalking* en zedenmisdrijven.

De camera's die de live beelden maken, kunnen in eigendom zijn van een private partij, zoals de winkelier, de uitbater van een winkelcentrum of een OV-bedrijf. Zij kunnen in bepaalde omstandigheden een *watchlist* bijhouden van gezichten van ongewenste klanten (Autoriteit Persoonsgegevens, 2020).

Ook individuele burgers zouden een *watchlist* kunnen hanteren. Dat zou bijvoorbeeld manier kunnen zijn om een contactverbod met een stalker te ondersteunen middels bijvoorbeeld een ring-deurbel[42]. In (Keymolen, Noorman,

[42] Als dit contactverbod door de overheid wordt gehandhaafd, dan is het typisch een invulling van Toepassing A: Object- en persoonsbeveiliging.

Van der Sloot, Cuijpers, & Koops, 2020) wordt een bekende bellentrekker uit de buurt als voorbeeld gegeven.

Een alternatief voor een *watchlist* in handen van private partijen, is dat de gezichten bij een overheidspartij in beheer blijven (zoals de politie). De winkelier – of althans zijn beveiliger – krijgen slechts een melding als er iemand van de *watchlist* wordt gesignaleerd.

Verboden bij tijdgebonden evenementen, zoals voetbalwedstrijden, blijken ook zonder biometrie gehandhaafd te kunnen worden. Bijvoorbeeld door mensen zich op dat moment elders te laten melden. Daardoor lijkt gezichtsherkenning voor stadionverboden *an sich* niet subsidiair.

In deze toepassing kan het gaan om verticale relaties, horizontale relaties en combinaties daarvan.

E.6 Toepassing F: Toegangscontrole evenement

In 2007 is bij ADO Den Haag de ontwikkeling gestart van het nieuwe stadion met het terugdringen van voetbalvandalisme als topprioriteit. In nauw overleg met gemeente, politie, industrie en kennisinstellingen is daar een breed scala aan innovaties verkend, waaronder ook ten behoeve van het detecteren en herkennen van ongewenst gedrag (denk aan spreekkoren) en van personen die bij incidenten betrokken raakten.

De basis voor dit beveiligingsconcept was 2-factor authenticatie (bezit ticket-op-naam en gezichtsherkenning[43]) bij de toegangscontrole voor bepaalde doelgroepen. Dit gebeurde op basis van consent (*opt-in*) van de bezoeker. Op die manier weet de betaald voetbal organisatie (BVO) van die groepen zeker wie er die dag binnen is, en hoe ze er die dag, althans bij het moment van toegang, uit zien.

Supporters met een hoger dreigings-profiel moesten in dat concept extra goed op de *enrolment* opname lijken. Daardoor werd bij hen extra streng gecontroleerd op identiteitsfraude. Dit is een eenvoudige variant van een *green lane* concept (zie E.8). Stadionverboden kunnen hiermee ook worden gehandhaafd middels een *watchlist* op basis van biometrie.

[43] Voor toegangsbeveiliging zijn internationale standaarden en *good practices* beschreven. Multi-factor authenticatie is zo'n *good practice*. Dit betekent dat meerdere (bij ADO twee) onafhankelijke bewijzen van iemands identiteit gecombineerd worden om zodoende een grotere zekerheid te krijgen. Of dat subsidiair is, hangt er van af of minder veilige alternatieven equivalente veiligheid opleveren. Of 1- of multi-factor authenticatie proportioneel zijn, hangt af van de risico-exposure van het evenement. In (Keymolen, Noorman, Van der Sloot, Cuijpers, & Koops, 2020) wordt gesteld dat het gebruik van gezichtsherkenning bij bijvoorbeeld een sportcentrum zou kunnen worden vervangen door een clubkaart. Dat is uitlegbaar omdat het risico van een sportcentrum kleiner lijkt (gederfde inkomsten van één bezoeker) dan die van een evenementorganisator en van de gemeente. De binnenkomst van een ongewenste bezoeker bij een betaald voetbalwedstrijd kan immers leiden tot grotere inkomstenderving (bijv. een boete n.a.v. een ongewenst spreekkoor), incidenten in mensenmassa's en openbare orde problemen.

Kennis over sensoriek, *privacy-by-design* en risico-sturing zijn daar vanaf het eerste moment meegenomen in het ontwerpproces. Goede camera's in de omgeving van het stadion, bij de toegangen en in het stadion bleken essentieel. In de beleving van de directeur Stadion, Veiligheid & Wedstrijdorganisatie heeft gezichtsherkenning geholpen om een gedragsverandering bij supporters te bereiken, waardoor het systeem nu niet meer nodig is. ADO heeft geleerd over de gebruikerservaring, zowel bij de *enrolment* als bij de toegang bij wedstrijden, en ook over het belang van goede camera's en goede belichting (Holst & Vellekoop, 2020).

Later zijn ook proeven gedaan bij andere Nederlandse voetbalverenigingen met gezichtsherkenning die alleen gebruik maakten van minder gecontroleerde *enrolment* opnames. Dit werd alleen gedaan voor de handhaving van stadionverboden waarbij de betreffende bezoekers alleen een stimulans ervoeren om *niet* mee te werken met het systeem. Immers, als ze niet herkend werden, dan konden ze naar binnen. Voor zo ver bekend hebben die proeven tot nu toe nog niet geleid tot implementatie.

In sectie 4.1 van (Keymolen, Noorman, Van der Sloot, Cuijpers, & Koops, 2020) zijn voor de casus ""organisatie van evenementen" het Amerikaanse bedrijf Zenus, en het Nederlandse bedrijf 20FACE geïnterviewd. In die interviews ligt de focus op het ondersteunen van de beveiliging door check-in en authenticatie middels gezichtsherkenning, en op allerlei andere functies voor commerciële doeleinden. Dit soort gespecialiseerde diensten maken het mogelijk voor evenementorganisatoren om de biometrische authenticatie uit te besteden aan een andere organisatie (zie ook sectie 5.4.1 "Gezichtsherkenning via een online platform"). Op het gebruik van niet-coöperatieve gezichtsherkenning bij evenementen, of hoe die integreren in het toegangscontrole proces wordt in dat rapport niet ingegaan.

Ook de JCA en het omliggende gebied is regelmatig het toneel van grote evenementen, waaronder uiteraard betaald voetbal. De politie-inzet die bij evenementen gepleegd wordt, is onderwerp van maatschappelijk debat (Adang, et al., 2014) – ook bij betaald voetbal.

De inzet van politie bij betaald voetbal is dalende bij een gelijkblijvend aantal incidenten. Deze efficiëntieverbetering is te danken aan gerichte maatregelen, betere samenwerking en snellere opvolging (Politie, 2018). Ook het terugdringen van incidenten is een prioriteit. Tien jaar geleden betrof dat spreekkoren, tegenwoordig gaat de aandacht uit naar het illegaal afsteken van vuurwerk (Politie, 2018). Dit soort incidenten gebeuren zowel binnen als buiten het stadion. Daartoe is een veiligheids-infrastructuur ingericht o.a. bestaande uit camera's in de omgeving, bij de toegangen en in de ARENA zelf.

Voetbalhooligans zijn zich bewust van veiligheidsmaatregelen, en nemen zelf tegenmaatregelen. Het bedekken van het gezicht, en het wisselen van kleding worden al lange tijd gebruikt om de pakkans te verkleinen, en dus sancties of vervolging te voorkomen. Dit zijn voorbeelden van *afwijkend gedrag*, gedrag dat onderdeel is van een modus operandi, en dat tevens onderscheidend is van het gedrag van bonafide supporters.

In deze toepassing kan het gaan om verticale relaties, horizontale relaties en combinaties daarvan. Veel evenementenorganisatoren zijn private ondernemingen, zoals betaald voetbalorganisaties en *dance*feestorganisatoren. Ook zijn er evenementen die door de overheid wordt georganiseerd, bijvoorbeeld nationale evenementen zoals de herdenking op de Dam. In bepaalde sectoren beheert de politie wel degelijk ook beelden van bekende dreigende personen, bijvoorbeeld als die personen ook buiten het private evenement voor problemen zorgen. Zoals de politie doet in het Centraal Informatiepunt Voetbalvandalisme.

E.7 Toepassing G: Grenswacht richting internationaal evenement

Op grote internationale evenementen, zoals een EK of WK voetbal, olympische spelen of bij een Eurovision songfestival, komen grote hoeveelheden buitenlanders af. Daar zullen vooral bonafide bezoekers zijn, maar ook deel niet bonafide. Dat kan bijvoorbeeld gaan om activisten of vandalen. Politie, KMar en beheerders van grensposten kunnen samenwerken om malafide bezoekers zo spoedig mogelijk te signaleren. Een publieke grenswachtorganisatie, zoals de KMar, kan voor het grenscontrole proces gebruikmaken van een biometrisch systeem van de private operator van een grenspost, zoals van het bedrijf Schiphol. In geval van Schiphol is dat straks bijvoorbeeld het *seamless flow* systeem. De politie en de KMar hebben via SIS, Interpol en Europol gezichtstemplates van malafide internationale bezoekers. De beheerder van het beveiligd gebied, zoals bijvoorbeeld JCA, heeft mogelijk ook nog gezichtstemplates van (VIP-)bezoekers. Het is in ieders belang dat bepaalde soorten supporters tijdig worden onderkend zodat passende service en / of opvolging kan worden verzorgd. Echter, het delen van gezichtstemplates tussen vervoerder (bijv. Schiphol) en beheerder van bestemming (bijv. JCA) is ongewenst.

In deze toepassing gaat het om een mix van horizontale en verticale relaties.

E.8 Toepassing H: Green lane

In verschillende contexten zijn voorstellen gedaan voor zogenaamde *green lanes*. Een *green lane* is een toegangscontrole concept waarbij bezoekers anders worden gecontroleerd afhankelijk van een van-te-voren ingeschat risicoprofiel. Het idee is dat het hierdoor mogelijk wordt om meer proportionele controles toe te passen. Als er weinig hoog-risico bezoekers zijn, dan levert dit dus een veel grotere doorstroming op en voor laag-risico bezoekers een betere beleving en doorstroming. Die betere beleving en doorstroming is waar de naam *green lane* vandaan komt. Hoog-risico bezoekers kunnen proportioneel strenger worden gecontroleerd. Voor bezoekers waarvoor geen risico inschatting kan worden gemaakt kunnen altijd nog de standaard controles worden toegepast. Deze twee categorieën ervaren dan geen *green lane*.

Bij verkeerd ontwerp en gebruik van *green lane* concepten kan het restrisico onnodig hoog zijn. Als het hoog-risico bezoekers lukt om zich voor te doen als laag-risico bezoeker, dan ondergaan ze immers onterecht een relaxtere check. Een *good practice* bij *green lanes* is dus om zowel profielen te maken voor bonafide bezoekers, als voor malafide bezoekers (Van Rest, Peeters, Smits-Clijsen,

Sternheim, & Wessels, 2020), en om bezoekers alleen uit te sluiten van een zwaardere check als ze zowel een hit hebben op ten minste één bonafide profiel, en tegelijkertijd op geen enkel malafide profiel een hit hebben.

Door slimme combinaties van screenings en checks te ontwerpen, kunnen sector-specifieke *green lanes* worden ontworpen. Dit soort concepten zijn voorgesteld voor de beveiliging van burgerluchtvaart (Poole, 2008) (XP-DITE, 2012), voor grenscontroles (TRESSPASS, 2019), voor de toegangscontrole van vitale infrastructuur (JRC, ERNCIP, 2019), en ook in de *digitale perimeter* in Amsterdam.

Green lanes gaan typisch uit van een expliciet model van het aanbod van bezoekers (bijv. van reizigers), gebaseerd op profielen. Een profiel is een combinatie van indicatoren op basis waarvan een inschatting kan worden gemaakt van welke doelgroepen de betreffende bezoekers allemaal lid is. Indicatoren kunnen bijvoorbeeld iets zeggen over *bezit* (zoals cash geld, of wapens), over iets *kunnen* (zoals vechten), *fysiologische toestand* (zoals ziekte), iets *willen* (zoals intentie), en uiteraard ook *identiteit* (zoals lijken op iemand die op een *watchlist* staat). De betrouwbaarheid en tijdigheid waarmee indicatoren kunnen worden ingeschat kan veel variëren tussen verschillende toepassingsgebieden. Uiteraard moet het ontwerp van de *green lane* daar rekening mee houden.

De *green lane* is als zodanig geen concrete toepassing, maar meer een *type* toepassing. Van de andere toepassingen lijkt met name de toepassing “toegangscontrole evenement” interessant om in een *green lane* variant uit te werken. De effectiviteit, doorstroming en ethische effecten van *green lane* configuraties kunnen met behulp van simulaties berekend en ingeschat worden, waardoor (hoog risico) situaties kunnen worden geanalyseerd die in de praktijk moeilijk te reproduceren zijn, en waardoor praktijkexperimenten meer efficiënt kunnen worden uitgevoerd. In twee van de hierboven genoemde projecten zijn dat soort simulaties ook daadwerkelijk gebouwd en gebruikt.

Er zijn verschillende mogelijke functies van gezichtsherkenning in een *green lane* concept. Ten eerste het herkennen van bezoekers tussen de verschillende contactmomenten. Het kan verstandig zijn om dat te initiëren met een andersoortige identiteitsclaim, zoals het aanbieden van een digitaal token middels een smartphone o.i.d. Ten tweede, en zoals hierboven al benoemd, de niet-coöperatieve herkenning van (mogelijk) ongewenste personen.

E.9 Toepassing I: Monitoren openbare (private) online platformen

Op diverse online platformen kunnen mensen en organisaties afbeeldingen plaatsen met daarop ook gezichten. Een aantal daarvan hebben ook de mogelijkheid om die beelden “openbaar” ter beschikking te stellen. Facebook, LinkedIn, Twitter, Pinterest en Instagram zijn wellicht de meest bekende voorbeelden.

Ten eerste is het voor de betreffende platformen zelf mogelijk om gezichten van door henzelf gezochte personen te detecteren. Bijvoorbeeld voor de handhaving van huisregels, zoals het weren van mensen wiens account is geblokkeerd.

Het is ook voor buitenstaanders mogelijk om de openbare delen van deze (private) platformen te monitoren en te doorzoeken. De functie zou kunnen zijn om die online “openbare ruimte” (op private platformen) ook in de gaten te houden op het gebruik er van door gezochte personen. Er zijn diverse bedrijven die dat al jaren doen, zoals bijvoorbeeld de bedrijven Coosto en PublicSonar, en ook bij de politie het internet Recherche Netwerk (iRN) en opvolgers daarvan.

Deze vorm van monitoring is ook mogelijk met behulp van een *watchlist*. In 2017 is het Amerikaanse ClearView opgericht. ClearView schraapt dit soort platformen af voor afbeeldingen van, in hun geval, specifiek gezichten, en vermarkt dit alleen naar politieorganisaties (ClearView, 2020). In 2019 kwam ClearView in de media, en kwam er forse kritiek op het business model (waaronder het verwerken van persoonsgegevens zonder expliciete toestemming van betreffende mensen), de transparantie, mogelijke bias in hun algoritmes, en de manier (of het gebrek daar aan) waarop mensen (waarvan een foto is geschraapt) hun beelden konden inzien of laten verwijderen (ACLU, 2020).

F Gezichtsherkenning in de observatiering van vitale infrastructuur

Het Directoraat Generaal for Migration and Home Affairs (DG HOME) van de Europese Commissie heeft aan een werkgroep van het *European Reference Network for Critical Infrastructure Protection* (ERNCIP) gevraagd om een expertopinie over het gebruik van biometrie, met name gezichtsherkenning ten behoeve van *early warning zones*, een andere naam voor een digitale perimeter. In 2019 heeft die werkgroep daar een onafhankelijke bureaustudie over gepubliceerd (JRC, ERNCIP, 2019).

Dit onderzoek wijst onder andere op het belang van fallback-scenario's indien de (gezichts)herkenning niet goed werkt. Bijvoorbeeld indien de tegenstander effectieve tegenmaatregelen inzet. Ook wordt gewezen op het belang van goede, en goed onderhouden sensor-infrastructuur, en worden er suggesties gegeven hoe dat met behulp van bepaalde intelligente algoritmes geborgd kan worden. Een actueel en uitgebreid overzicht van relevante standaarden maakt ook onderdeel uit van het rapport, net als een uitgebreide analyse van de Europese juridische kaders.

G Privacydreigingen van niet-coöperatieve gezichtsherkenning

In deze bijlage worden de privacydreigingen van niet-coöperatieve gezichtsherkenning beschreven. Er wordt gebruik gemaakt van de taxonomie van Solove (Solove, 2005) en ook van de beschrijvingen van actoren die een dreiging kunnen vormen tegen de correcte werking van een gezichtsherkenningssysteem (zie bijlage D "*Design basis threat* voor gezichtsherkenning").

G.1 Informatie verzameling

In deze categorie privacy dreigingen gaat het om verschillende manieren om informatie te verzamelen. Onder informatie verzameling heeft Solove twee privacy dreigingen gedefinieerd: "ondervraging" en "surveillance".

G.1.1 Surveillance

De privacy dreiging surveillance is inherent aan gezichtsherkenning. Het is voor gezichtsherkenning immers nodig om gezichten van mensen te observeren.

Deze privacy dreiging neemt toe wanneer er op meer locaties of vaker surveillance wordt toegepast, en wanneer er meer mensen onder surveillance komen. De onwenselijkheid die de minister uitspreekt van een "*structurele inzet van een breed vertakt, realtime gezichtsherkenningstechnologie, waarbij mensen continu en overal in kaart worden gebracht*" valt onder deze privacy dreiging.

Maatregelen die de surveillance dynamisch uitzetten, bijvoorbeeld als er geen sprake is van een dreiging of als er niets is om te beschermen, dragen dus bij aan het mitigeren van deze privacy dreiging.

G.1.2 Ondervraging

Ondervraging is minder relevant voor het onderhavige onderwerp omdat dit voor gezichtsherkenning op zich niet nodig is. Het kan wel zijn dat een identificatie (zowel een correcte als een foutieve) leidt tot een ondervraging.

G.2 Informatie verwerking

In deze categorie privacy dreigingen gaat het over activiteiten die reeds verzamelde informatie verwerken.

G.2.1 Aggregatie

De privacy-dreiging aggregatie is het combineren van meerdere stukjes informatie die gezamenlijk een scherper beeld geven van een persoon. Dit is inherent aan gezichtsherkenning: bij een correcte identificatie wordt een biometrisch template gecombineerd met een live opname. Het is vervolgens mogelijk om informatie van het *enrolment* moment te combineren met het live moment.

Aan de ene kant is dit de kernfunctie van gezichtsherkenning en lijkt het daarom niet mogelijk om dit privacy risico kleiner te maken. Aan de andere kant, het is denkbaar dat er te veel informatie van beide momenten met elkaar wordt gecombineerd. Dat kan bijvoorbeeld gebeuren als de originele enrolmentfoto ook in de database staat. Daarop kunnen bijvoorbeeld tekenen staan van sociale identiteit (zoals kleding of haardracht) die dan ook op het moment van herkenning beschikbaar zijn. Het verwijderen of althans loskoppelen van de enrolment foto van de biometrische template in de database bevordert dus dit soort privacy.

G.2.2 Identificatie

Identificatie is het koppelen van informatie aan individuen. Het lijkt een tautologie dat gezichtsherkenning leidt tot identificatie. Maar het verschil is dat het voor typische niet-coöperatieve toepassingen niet nodig is om iemand te identificeren, i.e. te koppelen aan één identiteit uit 7 miljard mensen, maar alleen maar om de persoon te herkennen uit een eerdere verzameling van mensen. Het probleem kan zijn dat gezichtsherkenning te uniek is, waardoor het voor sommige toepassingen krachtiger is dan een bruikbaar alternatief. Daardoor is het in dergelijke toepassingen dus niet subsidiair.

Er zijn alternatieven voor de harde biometrie gezichtsherkenning. Zachte biometrie omvat het herkennen van personen op hun algemene uiterlijk zoals kleding, haardracht en andere grove uiterlijkheden die in enige vorm veranderlijk zijn. Daarom kijken we in dit rapport breder dan alleen gezichtsherkenning, en wordt er ook gekeken naar herherkenning (zie sectie 5.4.1). Daar waar herkenning of herkenningstechnologie staat, bedoelen we in dit document zowel harde als zachte biometrie.

G.2.3 Onveiligheid

Onveiligheid is het ongeautoriseerde gebruik van persoonsgegevens. Dit kan zowel per ongeluk als door opzet gebeuren (zoals door *hacking*). Dit kan dus worden gedaan door allerlei soorten actoren, waaronder onvoorzichtige, fanatieke, en corrupte medewerkers, activisten en statelijke actoren. Deze privacy dreiging is relevant zodra persoonsgegevens bestaan. Het is dus zeker relevant voor gezichtsherkenning.

Deze dreiging is extra relevant voor bepaalde soorten biometrie omdat de impact van het manifesteren van de dreiging groter is:

- voor *harde biometrie*: mensen kunnen immers geen nieuw gezicht groeien (anders dan bijvoorbeeld de kleding die iemand draagt).
- voor biometrie met hoge uniciteit: mensen kunnen er uit grotere groepen mee herkend worden (anders dan bijvoorbeeld alleen de afstand tussen de ogen).
- Voor biometrie die op natuurlijke wijze kan worden gebruikt: mensen gebruiken zelf ook gezichten om andere mensen mee te herkennen (anders dan bijvoorbeeld iris).

Gezichtsherkenning is zowel hard als uniek en kan op natuurlijke wijze gebruikt worden. De privacy dreiging *onveiligheid* is dus extra relevant voor gezichtsherkenning.

G.2.4 Secondair gebruik

Secondair gebruik is wanneer persoonsgegevens worden gebruikt voor andere doeleinden dan waarvoor ze oorspronkelijk zijn verzameld. Het kan typisch door een fanatieke medewerker worden gedaan. Solove koppelt dit zowel aan situaties waar de data is verzameld op basis van expliciete toestemming van het data subject, als aan situaties waar een expliciete handeling van het data subject logischerwijs tot het verzamelen van persoonsgegevens zou leiden op een andere juridische grond, zoals bij het aanvragen van een rijbewijs.

Secondair gebruik is niet inherent vereist voor automatische gezichtsherkenning, maar het is wel een relevante privacy dreiging. Ten eerste wanneer *enrolment* opnames worden gebruikt die oorspronkelijk voor een ander doel zijn verzameld. Bijvoorbeeld het gebruiken van rijbewijsfoto's van criminelen om ze te herkennen op straat. Ten tweede wanneer live opnames worden gebruikt die oorspronkelijk voor een ander doel zijn verzameld. Zoals opnames ten behoeve van de biometrische verificatie bij toegangscontrole van een evenement hergebruiken voor marketing doeleinden, of om ook gezochte criminelen in te herkennen.

Ook is het denkbaar dat de uitkomst van de gezichtsherkenning wordt hergebruikt voor een ander doel. Bijvoorbeeld als een zware crimineel binnen korte tijd op twee verschillende plekken wordt herkend, waartoe hij per definitie ook een snelheidsovertreding moet hebben begaan, en er ook daarvoor opvolging gestart wordt. Dit is een voorbeeld van surveillance-bias: omdat iemand nauwkeurig in de gaten wordt gehouden, wordt er ook meer (illegaals) van hem of haar gezien.

G.2.5 Uitsluiting

Uitsluiting is de privacy dreiging dat het data subject niet wordt betrokken bij de verwerking van zijn gegevens. Hij wordt er niet van op de hoogte gesteld en heeft er geen controle over. Dit is niet inherent vereist voor gezichtsherkenning, en hangt dus van de uitvoering en toepassing af. Het is een reële privacy dreiging.

Voor veel mogelijke toepassingen is het realistisch te vereisen dat de persoon die op de *watchlist* komt te staan, daarover geïnformeerd wordt. Het is bijvoorbeeld erg lastig om een goede enrolment foto te maken zonder medewerking van een adequaat geïnformeerd subject.

Voorbijgangers kunnen kiezen of ze zich in het betreffende gebied begeven, en ze kunnen kiezen of ze hun gezicht daadwerkelijk tonen (zie echter ook sectie G.4.1 over de privacydreiging Indringing). Voor de meeste toepassingen van gezichtsherkenning hebben voorbijgangers op basis van hun gedrag dus controle over hun interactie met het systeem.

Desondanks is het met moderne technologie technisch relatief eenvoudig om een gezichtsherkenningssysteem te implementeren waarbij de mensen die op de *watchlist* staan daar geen idee van hebben, laat staan dat ze er aan hebben meegewerkt. Ook is het technisch haalbaar dat in bepaalde semi-gecontroleerde omstandigheden van een groot deel van neutrale voorbijgangers, zonder dat zij dat weten, gezichten worden verzameld.

G.3 Informatie verspreiding

In deze categorie privacy dreigingen gaat het om verschillende manieren om persoonlijke informatie te ontsluiten.

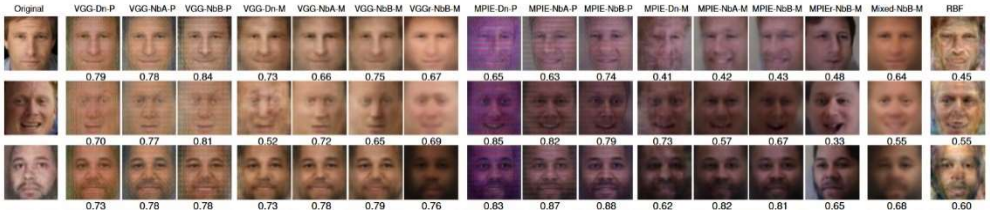
G.3.1 Inbreuk op vertrouwelijkheid

Het verstrekken van biometrische templates aan partijen die het strikt gezien niet nodig hebben, heet in Solove's taxonomie een inbreuk op vertrouwelijkheid. Dit is een type privacy-dreiging dat ontstaat omdat er simpelweg onvoldoende is nagedacht over betere alternatieven[44]. Daar kunnen verschillende varianten van zijn.

G.3.1.1 Reconstrueren van gezicht uit biometrisch template

Deze privacy dreiging kan optreden indien de partij die *enrolment* foto's heeft gemaakt en daar biometrische templates van maakt, geen vertrouwen kan hebben in de partij die de biometrische templates vervolgens gaat beheren of verwerken. De privacy dreiging is dan dat de tweede partij gezichten in pixels kan reconstrueren uit biometrische templates – ook als deze op onbekende manier zijn gemaakt (Chibba & Stoianov, 2014).

Er zijn verschillende manieren om van een biometrisch template een gezicht te reconstrueren. De eerste manier vereist een biometrisch template met bekende en betekenisvolle kenmerken, zoals de afstand tussen de ogen, etc. Als het format van het template bekend is, dan kan het feitelijk als een bouwtekening gelezen worden. De tweede manier heeft dat niet nodig. Die vereist alleen maar toegang tot de software in uitvoerbare vorm (dus niet de broncode) om gezichten te vinden in een willekeurige database. Door herhaaldelijk kunstmatig gemaakte gezichten of - eenvoudiger nog- biometrische gezichtstemplates aan te bieden aan die software en de scores te gebruiken om die kunstmatige gezichten gericht te veranderen, is het mogelijk om op zoek te gaan naar het kunstmatige gezicht dat met de hoogste score lijkt op het echte gezicht (USA Patentnr. 60594187, 2008). Het maakt daarvoor niet uit of de templates binnen die software nog versleuteld worden (Adler, 2003), of werken op basis van neurale netwerken (Mai, Cao, Yuen, & Jain, 2018).



Figuur 6 Plaatjes van gezichten kunnen worden gereconstrueerd uit gezichtstemplates. Links de originele foto. Daarnaast gereconstrueerde foto's uit het bijbehorende biometrisch template.

[44] Deze dreiging lijkt op die van "onveiligheid", waarbij persoonsgegevens per ongeluk of door kwade opzet worden gelekt (zie sectie 3.1.2.3 over Onveiligheid). In deze sectie gaat het echter om het verstrekken van persoonsgegevens omdat het inherent is aan het systeemconcept.

G.3.1.2 Onnodig verstrekken van templates

Het kan gaan om het verstrekken van templates aan partijen die het strikt gezien niet nodig hebben.

In het voorbeeld van de kentekenherkenning (zie bijlage L) speelde deze dreiging indien Spanje niet alle kentekens met Nederland wilde delen, en Nederland niet alle kentekens met Spanje. Ook speelt dit risico indien de herkenning wordt uitgevoerd op een verwerkingsplatform waarbij de dienstverlener niet volledig vertrouwd wordt.

Ook in de context van de digitale perimeter zijn er verschillende gezichtsherkenningstoepassingen denkbaar waarbij die situatie speelt, zelfs als er geen rekening hoeft te worden gehouden met een aparte dienstverlener die mogelijk niet te vertrouwen is. In bijlage E zijn verschillende voorbeelden gegeven van situaties waarbij de verstrekker van de *watchlist* en de eigenaar van de beelden elkaar mogelijk niet volledig vertrouwen, c.q. er zelfs wettelijke barrières zijn om gegevens uit te wisselen. Het is echter ook mogelijk dat de eigenaar van de *watchlist* dezelfde is als de eigenaar van de beelden.

G.3.1.3 Onnodig verstrekken van (biometrische templates uit) live beelden

Het verstrekken van live beelden aan de beheerder van de *watchlist* kan onderdeel zijn van een systeemconcept. Echter, die beheerder heeft de live beelden helemaal niet nodig. Hij heeft zelfs de biometrische templates van voorbijgangers niet nodig. In veel toepassingen hoeft hij strikt gezien alleen maar te weten of er een identificatie was. Het is afhankelijk van de toepassing of hij ook de identificatie-score moet weten, of de live opname (bijvoorbeeld ter menselijke verificatie) of de *enrolment* opname (opnieuw bijvoorbeeld ter menselijke verificatie).

Deze privacy dreiging speelt onder andere in zogeheten *smart city* context. Een gedachtelijn daarin is om de producent van data (i.e. van live beelden) los te koppelen van de gebruiker van die dat (i.e. de beheerder van de *watchlist*) die eventueel opvolging wil plegen. Deze privacy dreiging ontstaat indien die koppelingen niet strak worden beheerd, en er onterechte koppelingen ontstaan.

G.3.1.4 Biometrische templates op watchlist afleiden uit identificatie scores

Indien het systeemconcept de afstand maat / gelijkenis scores teruggeeft, dan zit daar informatie in die een malafide partij kan gebruiken om de inhoud van de database te onthullen (zie ook sectie C.4.1). Deze dreiging wordt vermoedelijk kleiner naarmate de database groter wordt.

G.3.2 Openbaring

Openbaring is de privacy dreiging waarbij persoonlijke informatie openbaar wordt gemaakt. Dit is niet inherent aan automatische gezichtsherkenning.

Het is een relevante privacy dreiging indien de opvolging publiekelijk gebeurt en het bekend is dat die opvolging komt door de gezichtsherkenning. Dan is daarmee dus openbaard dat de betreffende persoon op de *watchlist* staat en -afhankelijk van de toepassing- dus gezocht werd.

G.3.3 Blootstelling

Blootstelling is als diep persoonlijke gegevens zoals die betreffende naaktheid of fysiologische lichaamsprocessen onthuld worden. Dit is niet inherent aan automatische gezichtsherkenning.

Er kunnen wel gezichtsherkenningssystemen bestaan die ook fysiologische kenmerken gebruiken zoals de verdeling van bloed(aders) in het gezicht. Ook zijn er alternatieve herkenningssystemen die werken op basis van de vorm van het volledige naakte lichaam, zoals die op basis van somatotype.

G.3.4 Vergrote toegankelijkheid

Deze privacy dreiging gaat over het vergroten van de toegankelijkheid van reeds publieke persoonsgegevens. Dit is niet inherent aan gezichtsherkenning.

G.3.5 Chantage

Chantage gaat over het creëren van een machtsverhouding tussen twee personen op basis van de dreiging van het vrijgeven van informatie. Dit is niet inherent aan gezichtsherkenning.

G.3.6 Toe-eigening

Toe-eigening is de privacy dreiging dat identiteit of beeltenis van de één worden gebruikt voor de doelen van een ander. Dat is niet inherent aan automatische gezichtsherkenning.

G.3.7 Vervorming

Een belangrijk privacy-risico bij gezichtsherkenning in de (semi-)openbare ruimte is wat Solove vervorming (En. *distortion*) noemt. Vervorming komt er op neer dat er een verkeerd beeld van een persoon wordt gegeven. Dat is wat een herkenningssysteem doet als iemand onterecht wordt herkend als iemand anders: een foutieve identificatie.

In de context van gezichtsherkenning is vervorming een groot probleem. De kans dat het gebeurt is relatief groot. Ten eerste zijn gezichten niet expres verschillend. Anders dan kentekens zijn gezichten niet kunstmatig gemaakt, maar het organische resultaat van evolutie en allerlei omgevingsfactoren (zie sectie C.1).

Ten tweede zijn er allerlei complicerende omgevingsfactoren bij live opnames voor niet-coöperatieve herkenning (zie sectie C.6). Het is bekend welke omgevingsfactoren leiden tot slechte opnames, en daarmee tot lage detecties en lage herkenning. In 2017 heeft de South Wales police een serie proeven gedaan met de state of the art gezichtsherkenning bij voetbalwedstrijden (beste uit de face-recognition vendor test van NIST (NIST, 2020)). Laag kwaliteit opnames en niet-meewerkende supporters hebben er echter toe geleid dat onnodig veel mensen onterecht werden “herkend”. De prestaties waren zelfs dermate slecht dat midden in de reeks proeven de leverancier verzocht werd om het algoritme te verbeteren. In plaats daarvan heeft de leverancier alleen maar de kwaliteitseisen aan het input-beeld verhoogd, waardoor van minder supporters het gezicht werd vergeleken met

de *watchlist* (Davies, Innes, & Dawson, 2018). Die supporters glipten dus door de biometrische controle heen. “Vijandige” supporters konden dus zelf het gezichtsherkenningssysteem relatief eenvoudig omzeilen. Deze proeven hebben in het Verenigd Koninkrijk tot politieke en maatschappelijke discussie geleid. Dit probleem speelt met name in minder-gecontroleerde omstandigheden zoals wanneer er geen stimulans is voor mensen om mee te werken met het gezichtsherkenningssysteem.

De impact van vervorming varieert van de toepassing. Het kan bij een niet-coöperatief scenario groot zijn. Bijvoorbeeld, bij een eventuele fysieke opvolging van een identificatie (bijv. iemand aanspreken) moet worden uitgegaan van de mogelijkheid dat het systeem het juist heeft, met alle bejegeningsconsequenties van dien (hoe benader je bij een groot evenement iemand die erg veel lijkt op een gevaarlijk persoon?). Op basis van alternatieve methoden (zoals eerst het manueel controleren van de live opname tegen de *enrolment* opname, en het (ogenschijnlijk) routinematig vragen van een identiteitsbewijs) kan worden gepoogd de situatie te verduidelijken.

Een specifiek aandachtpunt bij vervorming betreft oneerlijke vervorming, oftewel een systematische fout die afhankelijk van niet-relevante aspecten (zoals demografische verschillen) herhaaldelijk tot verkeerde uitkomsten leidt. Dat wordt apart beschreven in bijlage H “Vertekeningen (biases) bij niet-coöperatieve gezichtsherkenning”.

G.4 Inbreuken

In deze categorie privacy dreigingen gaat om het recht met rust te worden gelaten.

G.4.1 Indringing

De privacy dreiging van indringing gaat over het betreden van iemand persoonlijke of zelfs intieme domein: zijn huis, zijn persoonlijke ruimte om zijn lichaam en ook zijn geest. Dit is bij automatische gezichtsherkenning niet inherent nodig. Het is niet nodig om iemands woning te betreden, om in zijn persoonlijke ruimte te komen (anders dan bij bijvoorbeeld vingerafdrukherkenning op basis van contactsensoren) of om iemands gedachten te profileren.

Er kan wel sprake zijn van bijvangst die met deze privacy dreiging te maken heeft. Indien een opname wordt gemaakt van een gezicht, dan wordt de gezichtsuitdrukking vanzelfsprekend ook meegenomen. De gezichtsuitdrukking kan indicatief zijn voor mentale constructen zoals emoties, cognitieve druk of attitude. Het is dus van belang dat biometrische templates de gezichtsuitdrukking niet vastleggen indien ze strikt voor herkenning bedoeld zijn.

De kwaliteit van de gezichtsopnames hangt af van de resolutie en het contrast op het gezicht. Die zijn typisch beter naarmate de optische sensor dichter op het gezicht staat, waarmee de optische sensor voor bepaalde toepassingen dus bij voorkeur in de persoonlijke ruimte van de persoon staat. Die persoonlijke ruimte is onder te verdelen in ringen, en beslaat (cultuur- en situatieafhankelijk) niet meer dan enkele meters. Met moderne optische technologie is die afstand eenvoudig te

bereiken. Maar er is tegenwoordig een toenemend aantal draagbare sensoren in gebruik (smartphones en bodycams) die behoorlijk dicht op het subject komen.

G.4.2 Beslissingsinterferentie

Dit is de mate waarin individuen vrij zijn om beslissingen te nemen inzake hun privé zaken. Er zijn verschillende soorten beslissingen die hier een rol spelen. Een gezichtsherkenningssysteem is beter voor privacy naarmate het deze soorten beslissingen *minder* beïnvloedt.

Ten eerste gaat dit om de beslissing om medewerking te verlenen aan het maken van de *enrolment* opname. Er zijn situaties denkbaar waarbij het subject een belang heeft om toch in enige mate aan gezichtsherkenning met negatieve herkenning mee te werken. Bijvoorbeeld, verslaafden aan gokspelen besluiten op sterke momenten om zich aan te melden voor de *watchlist* van casino's. Dit helpt hen om te stoppen met gokken. Over het algemeen is dit in toepassingsscenario's voor de politie anders. Malafide subjecten hebben immers géén belang om een *enrolment* opname te laten maken: daarmee is immers de kans immers kleiner dat ze hun malafide doelen kunnen realiseren.

Ten tweede de beslissing om naar een omgeving te gaan waarvan bekend is dat er gezichtsherkenning wordt toegepast. Het is denkbaar dat bonafide personen vanwege de aanwezigheid van gezichtsherkenning een drempel voelen om een locatie te bezoeken. Dit heet het *chilling effect*: het niet doen van legitieme dingen vanwege de perceptie van negatieve consequenties van risicobeheersingsmaatregelen[45]. De inhoud van die perceptie zal afhangen van wat er gecommuniceerd wordt over wat er verzameld wordt, hoe dat gebruikt wordt en hoe de privacy-risico's worden beheerst. Naar het *chilling* effect van cameratoezicht lijkt weinig onderzoek te zijn gedaan (Flight, 2013).

Voor malafide personen geldt uiteraard dezelfde afweging. Maar indien zij niet gaan dan heet het afschrikwekkende werking, en dat is soms een zeer groot (doch bij lokale maatregelen zoals gezichtsherkenning slechts lokaal) en zeer gewenst effect. Dit is immers een preventieve werking. Criminaliteit kan zich als gevolg van lokale maatregelen verplaatsen, maar er zijn aanwijzingen – en het is aannemelijk – dat daar geografische grenzen aan zitten, en dat de criminaliteit dus deels echt voorkomen wordt.

Ten derde is er de beslissing om bij de live toepassing bij de sensor actief mee te werken aan een goede opname. Er zijn allerlei nudging-strategieën bedacht om mensen te verleiden bewust even goed in de camera te kijken. Velen daarvan zijn zeer effectief. Ook is er veel mogelijk met blikvangers nabij de sensor die op het onbewuste inwerken.

---

[45] In (Keymolen, Noorman, Van der Sloot, Cuijpers, & Koops, 2020) wordt een ruimere definitie van *chilling* effect gebruikt waarin geen onderscheid wordt gemaakt tussen legitiem en illegitiem handelingen, c.q. tussen bonafide en malafide personen. Het *chilling* effect op malafide personen, waardoor ze geen illegitieme handelingen doen wordt in dit TNO rapport afschrikking genoemd.

Aan de hand van deze beschrijvingen is het duidelijk dat deze specifieke privacy dreiging het minst optreedt bij een verborgen gezichtsherkenningssysteem. Er zijn toepassingsscenario's denkbaar waarbij het van groot belang is dat malafide personen niet weten dat er gezichtsherkenning is, bijvoorbeeld in de context van heimelijke observatie in grote opsporingsonderzoeken. Uiteraard leidt dat tot de privacydreiging *uitsluiting*.

H Vertekeningen (biases) bij niet-coöperatieve gezichtsherkenning

Vertekeningen (*biases*) kunnen leiden tot een (beeld van een) oneerlijk werkend system en (daardoor) een gebrek aan draagvlak voor de toepassing van niet-coöperatieve gezichtsherkenning. Een vertekening is een verkeerd beeld van een situatie. Een voorbeeld is wanneer gezichtsherkenning een verschillende nauwkeurigheid heeft bij verschillende demografieën, zoals etniciteit, leeftijd en geslacht (NIST, 2019).

Biases kunnen zichzelf versterken. Bijvoorbeeld indien foto's van veroordeelde criminelen (met een onder- / oververtegenwoordiging van bepaalde demografieën) gebruikt worden om gezichtsherkenningsalgoritmes te verbeteren. Het is dus verstandig om alert te zijn op het identificeren, voorkomen en mitigeren van *biases*[46].

Er is een spanningsveld tussen eerlijkheid, en effectiviteit van een niet-coöperatief gezichtsherkenningssysteem. Als er grote, ongebalanceerde datasets vrij beschikbaar zijn, dan kunnen die gebruikt worden om voor bepaalde demografieën hele nauwkeurige algoritmes te ontwikkelen. Maar die zijn dan dus minder eerlijk. In concrete toepassingsgebieden kan dat dus leiden tot een lagere pakkans van criminelen (van bepaalde demografieën).

Er wordt aan verschillende oplossingsrichtingen gedacht. Eén is om de trainingsdatasets "aan de voorkant" te cureren. Dus om ze zodanig samen te stellen dat ze een representatief beeld geven van de (betreffende) demografie. Dat kan worden gezien als een voorbeeld van de strategie *dwing af*. Een andere is om ze achteraf te testen, zoals NIST heeft gedaan (NIST, 2019). Dat is een voorbeeld van de privacy-beschermende strategie *toon aan*. Nog een oplossingsrichting kan wellicht zijn om de privacy-beschermende strategie *minimaliseer* toe te passen, en om een gezichtstemplate te ontwikkelen dat inherent géén informatie bevat over etniciteit, leeftijd en geslacht (zie ook sectie I.1.1).

[46] Er is een breed debat in Westerse samenlevingen over *biases* in relatie tot demografische eigenschappen van mensen. Wetenschappelijke (ACM, 2020), maatschappelijke, politiek en industriële partijen hebben opgeroepen tot een pauze in het toepassen er van.

I Privacy-beschermende strategieën

In deze sectie wordt beschreven hoe deze strategieën voor gezichtsherkenning zouden kunnen uitwerken in relatie tot de privacy dreigingen van bijlage G “Privacydreigingen van niet-coöperatieve gezichtsherkenning”. De structuur is gebaseerd op (Hoepman, 2014).

I.1 Minimaliseer

Minimalisatie gaat over het verwerken van persoonsgegevens van minder mensen, minder persoonsgegevens per persoon, of minder gevoelige persoonsgegevens. Deze maatregelen kunnen helpen om de privacydreigingen surveillance, identificatie, aggregatie of secondair gebruik van persoonsgegevens te voorkomen of verkleinen.

I.1.1 Minder persoonsgegevens per persoon

De beelden (pixels) hoeven niet altijd (allemaal) te worden opgeslagen. Dit kan zowel gelden voor training van algoritmes, voor *enrolment* als voor live beelden.

Voor de training van herherkenningsalgoritmes is het bijvoorbeeld niet altijd nodig om daar ook gezichten in beeld te hebben (Van Rooijen, et al., 2020). Mogelijk is er ook een equivalente vorm te ontwikkelen die onnodige ruwe data (pixels) weglaat bij het trainen van gezichtsherkenningsalgoritmes.

Een mogelijk nadeel van het niet-opslaan van beelden van enrolment is dat daardoor de traceerbaarheid van de totstandkoming van een zoekresultaat verminderd of zelfs onmogelijk wordt. Bijvoorbeeld, indien de *enrolment* foto wordt verwijderd, is het niet meer mogelijk om te controleren of een identificatie fout of correct was aan de hand van die *enrolment* foto. Uiteraard kan dat nog steeds wel aan de hand van het echte subject wanneer een live opname van hem is gemaakt. Hier moeten twee belangrijke kanttekeningen worden gemaakt. Ten eerste dat uit een biometrisch template ook weer een beeld kan worden gemaakt. Ten tweede dat een biometrisch template zonder bijbehorend beeld, zelf ook een identificator en dus een persoonsgegeven is, nota bene een *gevoelig* persoonsgegeven.

Andere persoonlijke attributen dan de gezichtskenmerken in *enrolment* en/of live beelden kunnen worden weggelaten door slimme beeldverwerking. Bijvoorbeeld kleding, bril, en sieraden. Hierdoor wordt alleen het gezicht opgeslagen.

Als het dan nodig is om het beeld op te slaan, dan kan dat beeld zo getransformeerd worden dat de gezichtsuitdrukking niet meer te zien is. Een geautomatiseerd biometrisch systeem dat voor herkenning of identificatie wordt gebruikt heeft typisch last van de gezichtsuitdrukking, dus informatie over de expressie zit over het algemeen niet in het biometrisch template.

Wellicht is het ook mogelijk om een standaard voor een biometrisch gezichtstemplate te ontwikkelen die inherent géén informatie bevat over etniciteit,

geslacht en leeftijd van een persoon. De vraag wordt dan hoeveel onderscheidende informatie zo'n template dan nog kan bevatten, en voor welke toepassingen dat nog voldoende is.

I.1.2 Minder gevoelige persoonsgegevens

Soms is het niet nodig om een hele hoge nauwkeurigheid te hebben. Met name als het subject geen mogelijkheid (geen tijd, geen verstopplek) heeft om zijn uiterlijk te veranderen, dan kan het voldoende zijn om gebruik te maken van indirecte identificatoren (zie sectie C.4). Bijvoorbeeld kunnen persoonlijke attributen worden gebruikt, zoals kleding en haardracht (d.w.z. zachte biometrie). Dit kan nuttig zijn om mensen te herkennen tussen kort opéénvolgende momenten. Bijvoorbeeld om *signaalattributie* te verzorgen in een digitale perimeter rond een evenement (zie ook sectie 7.1.4.5 "Gezichtsherkenning als koppelsysteem" als onderdeel van het "Concept voor digitale perimeter: adaptieve herkenning bij evenementen").

I.1.3 Persoonsgegevens van minder personen

In deze sectie zijn drie varianten beschreven waarmee bepaalde soorten mensen worden uitgesloten van de live gezichtsherkenning.

I.1.3.1 Personen die een bepaald kenmerk hebben

Als de database alleen gezichten van personen met een bepaald kenmerk bevat, dan hoeven gezichten van mensen die dat kenmerk niet hebben, niet verzameld en verwerkt te worden. Bijvoorbeeld als de identiteit van personeel dat op de betreffende locatie werkt middels een pas is geauthentiseerd, dan hoeven er van hen wellicht geen gezichten te worden gecontroleerd tegen de *watchlist*. Hier moet bij worden opgemerkt, dat dat kenmerk dan wel met grote nauwkeurigheid moet kunnen worden vastgesteld. Anders kan dat een achterdeur worden voor tegenstanders. In dit voorbeeld, als een tegenstander aan een medewerkerspas kan komen waarmee hij kan voorkomen dat zijn gezicht tegen de *watchlist* wordt vergeleken, èn waarmee hij binnen mag komen, dan is er een gat in de beveiliging.

I.1.3.2 Personen die op een equivalente positieve watchlist staan

Bij toepassingen van niet-coöperatieve gezichtsherkenning in de openbare ruimte zijn er ook voorbijgangers die verder niets met de situatie te maken hebben. Dat kunnen omwonenden zijn, om mensen die ter plekke komen om te recreëren, winkelen of voor ongerelateerd werk in die omgeving. Die mensen komen in beeld van een eventueel (gezichts)herkenningssysteem.

Het is mogelijk om hen op een equivalente positieve *watchlist* te zetten. Dat is een *watchlist* die bedoeld is om bonafide voorbijgangers met rust te laten. Meer specifiek, om te voorkomen dat mensen geconfronteerd worden met de eventuele consequenties van een foutieve match met iemand op de negatieve *watchlist*.

Bijvoorbeeld kunnen mensen waarbij een foutieve match is gebleken, op hun eigen verzoek op een dergelijke positieve *watchlist* worden gezet.

Ook kan dit een uitkomst zijn in situaties waar de beeldregistratie ook voor andere doeleinden wordt gebruikt dan voor live herkenning. Bijvoorbeeld voor intelligence

of opsporingsdoeleinden. Met een dergelijke positieve *watchlist* kan mogelijk worden uitgesloten dat persoonsgegevens van dergelijke personen onnodig worden verwerkt. Uiteraard is de verwerking ten behoeve van de positieve *watchlist* zelf ook een verwerking van persoonsgegevens. Het hangt van het concept en de context af of daar een goede juridische basis (inclusief proportionaliteit en subsidiariteit) voor kan worden gevonden. Soms kan consent (zoals in het voorbeeld hierboven) een optie zijn. Dat is echter lastig indien de toepassing van gezichtsherkenning niet openbaar bekend is. Dit zou wellicht nader uitgezocht moeten worden.

I.1.3.3 Personen wiens live opname niet van voldoende kwaliteit is

Als de live opname van een persoon van onvoldoende kwaliteit is, dan heeft het weinig zin om die aan te bieden aan een gezichtsherkenningssysteem. De kans op fouten, en dus op het privacyrisico *vervorming* wordt dan te groot. Mensen van wie geen goede kwaliteit opname is gemaakt, volgen dan een ander proces. De factoren die de kwaliteit bepalen (zoals belichting), en de kwaliteitskenmerken zelf (bijv. contrast) kunnen deels uit beelden zelf ook weer bepaald worden. Zie daarvoor bijvoorbeeld sectie 5.4.4 over *managed analytics*. Ook dit mag natuurlijk niet leiden tot een achterdeur in het systeem. In plaats van een goede opname of biometrisch template van een live gezicht, moet er dan vermoedelijk in de plaats een signaal worden gegeven dat er een persoon is gedetecteerd waarvan dus geen goede opname is gemaakt (d.w.z. *failure to capture)*. Dat moet vervolgens leiden tot een minstens zo goede opvolging. Bijvoorbeeld een staande-houding waarbij alsnog een opname van het gezicht wordt gemaakt middels een mobiel device, en / of waarbij identiteitsdocumenten worden gevraagd.

I.2 Scheid

Door persoonsgegevens uit verschillende contexten gescheiden van elkaar te verwerken, wordt het lastiger om ze onbedoeld met elkaar te combineren. Dit kan helpen om de volgende privacydreigingen te helpen beheersen: aggregatie, onveiligheid.

De eerste soort scheiding, en tevens de meest fundamentele, is die waarin de biometrische templates van de *enrolment* nooit in dezelfde handen komen als die van de live beelden. Dat klinkt paradoxaal voor een herkenningssysteem. Met *multi-party encryption* is het toch mogelijk om een verwerking te doen waarbij de partijen die ieder invoer leveren, elkaars invoer niet leren kennen.

De volgende soort scheiding is tussen de opslag van *enrolment* beelden en de opslag en verwerking van biometrische templates. Dit helpt voorkomen dat voor-mensen-herkenbare beelden onbedoeld beschikbaar komen. Dit is geen krachtige soort scheiding, omdat het met behulp van biometrische templates wel mogelijk is om gezichten te reconstrueren. Alleen persoonlijke attributen die niet onderdeel zijn van het biometrisch template (kleding, bril, etc) blijven wel verborgen. In een variatie van deze soort scheiding worden de *enrolment* beelden wel ontsloten indien het systeem een identificatie heeft gemaakt en er behoefte is aan menselijke controle.

De derde soort scheiding is relevant indien twee *watchlists* voor verschillende toepassingen op dezelfde camera worden gebruikt. Het risico is dat ze onbedoeld met elkaar vermengd kunnen raken. De strategie *scheid* schrijft voor om ze dan in verschillende (logische) ICT omgevingen te verwerken. Bijvoorbeeld, in een *smart city* omgeving is het mogelijk om diverse producenten van live beelden middels een *smart-city*-koppelvlak te koppelen aan meerdere *watchlist* beheerders, iedere koppeling met eigen doelbinding en juridisch kader. Als dat niet zorgvuldig gebeurt, dan kan de eigenaar van de eerste *watchlist* onbedoeld de hits krijgen van de tweede *watchlist*.

De vierde soort scheiding is relevant als een herkenningssysteem zowel gebruik maakt van harde biometrie / directe identificatoren, zoals gezicht, als van zachte biometrie / indirecte identificatoren, zoals persoonlijke attributen zoals kleding. Het risico is dan dat het geheel niet subsidiair en / of niet meer proportioneel is. Door de harde biometrie te scheiden van de zachte biometrie wordt het mogelijk om afhankelijk van het dreigingsniveau één van beiden of allebei in te zetten. Bijvoorbeeld, als er een laag dreigingsniveau is, dan worden alleen minder scherpe beelden gebruikt waarmee mensen alleen kunnen worden herkend uit een kleine groep. Maar als de dreiging hoger wordt, dan worden ook hogere resolutie beelden ontsloten, waarmee mensen ook kunnen worden geïdentificeerd. Zie ook sectie 7.1.4.3 “Ad hoc verhoogde dreiging”.

I.3 Abstraheer

Door persoonsgegevens in minder detail (dus geabstraheerd) te verwerken, wordt als het ware “uitgezoomd” van de individuele persoon. Door gezichten of andere biometrische kenmerken alleen zodanig gedetailleerd (resolutie, variatie in kijkhoek) waar te nemen als nodig is voor de betreffende toepassing – en niet meer gedetailleerd dan dat. Als het gaat om gezichtsherkenning, dan lijkt dit op één van de minimaliseer-strategieën (zie ook sectie I.1.2).

Als er te veel wordt geabstraheerd, dan kunnen er te veel fouten worden gemaakt, en dan kan de privacy dreiging “vervorming” optreden. Tevens dreigt daarbij het risico dat er van unieke herkenning wordt overgegaan naar profilering: iedereen die voldoet aan een aantal minder-unieke kenmerken (“bildragend”, “bepaald type kleding”, etc.) kan dan “lijken” op iemand die op de *watchlist* staat. Dat hoeft op zich niet erg te zijn, maar professioneel profileren is een aparte specialisatie (Van Rest, Peeters, Smits-Clijsen, Sternheim, & Wessels, 2020).

I.4 Verberg

Door persoonsgegevens te verbergen, kan worden voorkomen dat ze in de verkeerde handen vallen. Het helpt onder andere tegen de privacydreiging *onveiligheid*, *secondair gebruik* en *inbreuk op vertrouwelijkheid*. Oftewel dat ze mogelijk in verkeerde handen komen.

Door het hele herkenningsproces alleen achter beveiligingsschillen te laten plaatsvinden, kunnen beelden, biometrische templates en informatie over persoonlijke attributen worden verborgen. Een specifieke variant hiervan is door het

herkenningsproces op de camera te doen, achter een beveiligingsschil. Dat heet *intelligence-on-the-edge* (zie sectie 5.4.2).

Maar om mensen te herkennen, moet informatie gecombineerd worden die zich mogelijk in verschillende domeinen bevindt: biometrische templates van live beelden, en die van *watchlists*. Er zijn moderne versleutelingstechnologieën waarmee het mogelijk is om dat te doen (zie sectie 5.4.2).

I.5 Informeer

Door mensen te informeren over het gebruik van hun gegevens kunnen privacydreigingen worden beheerst, met name *uitsluiting*. Voor gezichtsherkenning kunnen bijvoorbeeld aparte notificaties worden gemaakt, of camera's die er duidelijk anders uit zien, die duidelijk maken dat het niet (alleen) om algemeen cameratoezicht gaat.

Mensen kunnen worden geïnformeerd dat ze op een *watchlist* staan. Dat is bij kleine criminelen ook verplicht als het op basis van de AVG gebeurt (Autoriteit Persoonsgegevens, 2020). Voor grotere criminelen of voor statelijke actoren kan dit het onderzoek hinderen.

Mensen kunnen ook ter plekke worden geïnformeerd door displays te gebruiken waarop voorbijgangers zien dat (en hoe goed) hun gezichten gedetecteerd worden. Dit heeft als bijkomend voordeel dat vriendelijke voorbijgangers daarmee de informatie hebben waarmee ze eventueel hun gezicht beter kunnen aanbieden. Dit kan bij toegangscontrole acceptabel zijn, waar mensen een belang hebben om goed herkend te worden. Bij niet-coöperatieve gezichtsherkenning in de openbare ruimte is het vermoedelijk echter zeer ongewenst om er op die manier (regelmatig) mee geconfronteerd te worden.

Informeren gaat nog een stap verder als mensen worden geïnformeerd dat ze zijn herkend in een live beeld – en door welke organisatie. Dat zou ook kunnen helpen om persoonsverwisseling te voorkomen: als iemand onterecht herkend is, dan kan hij direct de eigenaar van de *watchlist* verwittigen met bewijs dat hij het niet was, en dat het systeem dus een fout heeft gemaakt.

I.6 Geef controle

Door mensen controle te geven over het gebruik van hun gegevens, kan met name het risico van *uitsluiting* worden beheerst. Als mensen zelf controle hebben over de inzet van hun biometrisch template, dan kan het niet worden gebruikt zonder hun goedkeuring.

Dit vereist ten eerste de mogelijkheid om voor een alternatief te kiezen. Voor betekenisvolle controle moeten mensen de mogelijkheid hebben om ook gebruik te maken van bepaalde diensten zonder daarbij in aanraking te komen met gezichtsherkenning. Bijvoorbeeld door ook aan te bieden dat mensen zich met behulp van identiteitsdocumenten identificeren. Uiteraard moet dat niet ten koste

gaan van de veiligheid. Een 2-factor beveiligingssysteem levert nu eenmaal meer veiligheid dan een 1-factor systeem.

Een alternatief is door mensen te vragen of ze hun *enrolment* opname willen inzetten om ergens binnen te komen (zie sectie 5.4.1). Dat kan werken in een *authenticatie* scenario, maar niet bij scenario waar van te voren geen identiteit wordt geclaimd. Immers, dan is niet te bepalen met wiens biometrisch template het live biometrisch template moet worden vergeleken. Niet-coöperatieve toepassingen hebben daar dus niets aan.

I.7 Dwing af

Door af te dwingen dat afspraken over privacy – ook met derden- worden nageleefd, kunnen de andere strategieën versterkt worden. Echter, als de privacy afspraken niet goed zijn, dan voegt het afdwingen er van maar beperkt iets toe.

Het afdwingen van afspraken heeft ook het nadeel dat systemen minder flexibel worden. Als omstandigheden veranderen, bijvoorbeeld door een nieuw soort dreiging, dan kan het nodig zijn om gezichtsherkenning anders in te zetten.

Door te controleren dat afspraken over het verwijderen van beelden en biometrische templates worden nageleefd, kan *onveiligheid* en een aantal andere privacydreigingen worden beheerst.

Door de kwaliteit van de optische keten, ook (of juist) als die in beheer is bij anderen, op geautomatiseerde wijze continu te monitoren, kan tijdig worden ingegrepen als die ondermaats dreigt te worden. Daarmee kan een minimale nauwkeurigheid worden gehandhaafd, wat de dreiging van *vervorming* kan beheersen. Een andere “afdwingende” maatregel om systematische vervorming (bias) in het gezichtsherkenningsalgoritme tegen te gaan, kan zijn om de trainingsdatasets zorgvuldig samen te stellen zodat ze een representatief beeld geven van de betreffende demografie.

Door iedere zoekopdracht automatisch te melden bij een onafhankelijke toezichthouder kan toezicht worden afgedwongen.

Door iedere zoekopdracht via een centraal punt te laten lopen, kan er een schakelaar worden gemaakt waarmee gezichtsherkenning uit kan worden gezet. Deze schakelaar kan aan een onafhankelijk toezichthouder worden gegeven, en eventueel ook aan het betreffende data subject.

I.8 Toon aan

Door gebruiksgegevens te verzamelen en extern te rapporteren kan worden aangetoond of privacy-beschermende strategieën werken. Deze strategie heeft alleen indirect effect. Als de privacy afspraken niet goed zijn, dan voegt het aantonen van hun beperkte werking maar beperkt iets toe.

Door gegevens te rapporteren over correcte identificaties, *failure-to-capture* en over foutieve niet-identificaties, kan de effectiviteit van de gezichtsherkenning beter onderbouwd worden.

Door gegevens te rapporteren over het aantal voorbijgangers waarvan hun gezicht is gecontroleerd tegen het aantal *watchlists*, kan de impact op de samenleving beter worden onderbouwd, zowel in positieve zin (bijdrage aan veiligheid), als in negatieve zin.

Door informatie te rapporteren over *data-breaches* kan de werking van privacy beschermende maatregelen worden aangetoond.

J Suggesties voor experimenten

Dit zijn de experimenten die bij de privacy-by-design workshop zijn geopperd. Er is toen gekozen voor experiment 1a. Die is vervolgens verder uitgewerkt.

1a. Voer een experiment uit met gezichtsherkenning in het versleutelde domein. Onderzoeksvragen: Kan gezichtsherkenning met een kwalitatief goed algoritme op deze opnames in het versleutelde domein worden uitgevoerd? Is de berekening in het versleutelde domein efficiënt genoeg voor deze scenario's?

2a. Controleer de omgevingscondities (weer, belichting, resolutie, opnamehoek) van een opname voordat er een herkenning wordt uitgevoerd. Onder slechte condities moeten of de condities worden aangepast (door beheerder, operationeel medewerker of bezoeker) of moet de herkenning niet worden gedaan. Onderzoeksvragen: Kunnen de omgevingscondities goed genoeg uit de opnames en andere sensoren (zoals een weerstation) worden bepaald? Hoe sterk wordt de kans op het missen van een gezocht persoon beïnvloed door het tegengaan van onterechte herkenning?

2b. Onderzoek of met soft biometrics (met name kledingkleur en -samenstelling) een extra verificatie in aanvulling op de gezichtsherkenning kan worden gedaan, bijvoorbeeld als de originele gezichtsopnames van de groep personen (achteraf) niet beschikbaar zijn (zie risico 1). Onderzoeksvragen: Kan met geanonimiseerde enrolment opnames (dus geen gezicht) en een actuele opname van de gehele gezochte persoon, een betere menselijke verificatie van de herkenningsresultaten afkomstig van gezichtsherkenning worden gedaan? Kan geautomatiseerde vergelijking o.b.v. soft biometrics hier werken en heeft het meerwaarde in dit scenario?

3a. Onderzoek of met soft biometrics (met name kledingkleur en -samenstelling, eventueel gezichtseigenschappen) de herkenning (zonder gezichtsherkenning) goed genoeg kan worden gedaan. Onderzoeksvragen: kunnen er betrouwbaar soft biometrische eigenschappen worden bepaald in de opnames? Hoe goed kan geautomatiseerde vergelijking o.b.v. soft biometrics hier werken en heeft het meerwaarde in dit scenario?

K Experiments on multi-party computation for non-cooperative facial recognition

This Annex presents the results of experiments done on multi-party computation for non-cooperative facial recognition. The Annex is in English because of the international composition of this part of the research team.

K.1 Introduction

Facial recognition systems process extremely sensitive personal information and, if used incorrectly, have the potential to severely violate the privacy of individuals. For this reason, it is of the utmost importance to process this information securely.

However, in some situations additional privacy or confidentiality constraints apply. These situations typically involve multiple parties that cannot simply share their data. For example, one party recording a live video stream of passersby at a high risk location, and another party have a watch-list composed of biometric templates of wanted dangerous individuals.

K.1.1 Multiparty Computation

Multiparty computation (MPC) is a collection of cryptographic techniques that allows multiple parties to *securely* evaluate functions on input data that is distributed amongst the parties. Secure evaluation means that parties are guaranteed to learn *only* the outcome of the computation, i.e. all additional information about the parties' private input values remains secret to the other parties.

It has long been known that any computable function can be computed securely, i.e. in an MPC manner. However, securely processing information comes at a cost. Applying MPC can increase the communication costs between parties, when compared to a straightforward evaluation of the function under consideration. Similarly, it can reduce the computational efficiency significantly.

Moreover, there is not a single one-size-fits-all MPC solution. MPC is a field of research that studies various cryptographic techniques, all with their own advantages and disadvantages. For this reason, applying MPC requires careful secure algorithmic engineering.

The following properties allow us to distinguish various MPC protocols. We note that many of these protocols are related and are therefore subject to various trade-offs. For instance, a strong security model will have a negative impact on the efficiency of the protocol.

- **Number of Parties**: Typically MPC protocols are tailored either for a two party setting or for an $n \geq 3$ party setting. In the context of facial recognition systems the two party setting seems natural. However, for efficiency or security reason we can decide to consider a 3-party protocol by including an *untrusted* third party.

- **Adversarial Model**: The adversarial model specifies the capabilities of a malicious party against whom the protocol must be secure. For example, it specifies the number of parties an adversary can corrupt.
- **Efficiency**: The main efficiency bottlenecks of MPC protocols are the computational complexity, communication complexity and the round complexity. Depending on the exact scenario one might, for instance, sacrifice the computational complexity in order to decrease the communication complexity of a protocol.
- **Computation Model**: Typically MPC protocols assume that the function that is to be computed securely is modelled in terms of some universal set of operations, such as additions and multiplications. However, different MPC protocols use different universal sets of operations or computational models. In turn, these computational models might significantly impact the efficiency of an MPC solution. Additionally, depending on the MPC protocol, some operations (e.g., linear operations) are more easily dealt with than other (e.g., non-linear operations).
- **Load Balance:** MPC protocols contain multiple parties that are required to participate. This participation implies that the parties perform computations. For some MPC protocols this computational load is evenly distributed amongst the parties (symmetrical), while for other MPC protocol the majority of the computational tasks is performed by one of the parties (client-server setting.

One can identify the following main categories of MPC solutions:

- **Garbled circuits:** Typically tailored for the secure evaluation of Boolean circuits in the two party setting. In this approach, one party obfuscates the computation or `garbles' the circuit to guarantee security, while the other party evaluates this garbled circuit.
- **Homomorphic encryption:** Relying on public-key cryptography and, thereby, computational assumptions. This approach is typically suited for the secure evaluations of linear (arithmetic) functions in a client-server model, i.e. one party performing the computations on encrypted input data from the other parties. Non-linearities require additional techniques and incur additional costs. In this approach, the private input values are encrypted and computations on this encrypted data are performed homomorphically.
- **Secret-Sharing:** This setting typically requires at least 3 parties for the secure evaluation of arithmetic circuits. Moreover, the number of communication rounds depends on the multiplicative depth of the arithmetic circuit. Additional techniques exist to apply this approach in a 2-party setting. In this approach, all parties `secret-share' their data and computations are subsequently performed on the shares.

K.1.2 Facial Recognition Algorithms

We describe, on a high level, the workings of (simplified) non-cooperative facial recognition algorithms. Facial recognition algorithms extract from each person in each image a template. Templates are abstractions of `faces' that are used to compare different face images. A template is a vector of real number, of a certain dimension. By scaling the coefficients and rounding them we can assume the template to be an integer vector. This transformation is more convenient for the application of MPC and does not limit its applicability. Extracting templates of

different images of the same person results in similar templates, more precisely, templates with a small Euclidean distance to each other. Note that we must use the same scaling factor for all templates.

From this observation facial recognition algorithms can be constructed. We consider the two algorithms described below. Both use a vector of 128 values to describe a face. They differ only in that Algorithm 1 outputs a score, smaller scores indicate better identifications, whereas Algorithm 2 outputs a Boolean, obtained by comparing this score to a threshold. This minor difference can have a significant impact on the performance of the MPC solution. Namely, comparison are costly operations when conducted in an MPC manner. On the other hand the output of Algorithm 1 reveals more information about the private input values than the output of Algorithm 2. Namely, from enough output scores, the secret input vector can be reconstructed (see section G.3.1.1). For this reason we take both algorithms into consideration.

A non-cooperative face recognition algorithm compares a sequence of live images to multiple images on the watchlist (1:N).

Algorithm 1	**Algorithm 2**
Input: - Template 1: $\boldsymbol{x} = (x_1, \dots, x_{128}) \in \mathbb{Z}^{128}$ - Template 2: $\boldsymbol{y} = (y_1, \dots, y_{128}) \in \mathbb{Z}^{128}$ Output: - Score: $s \in [0, \infty)$ Procedure: - $s = \left\|\|\boldsymbol{x} - \boldsymbol{y}\|\right\|^2 = \langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y} \rangle$	Input: - Template 1: $\boldsymbol{x} = (x_1, \dots, x_{128}) \in \mathbb{Z}^{128}$ - Template 2: $\boldsymbol{y} = (y_1, \dots, y_{128}) \in \mathbb{Z}^{128}$ - Threshold: T Output: - Boolean: $s \in \{0,1\}$ Procedure: - $s = \left(\left\|\|\boldsymbol{x} - \boldsymbol{y}\|\right\|^2 < T\right)$

K.2 Related Work

Privacy-preserving face recognition, (Erkin & Toft, 2009)
This publication presents a solution based on the Paillier cryptosystem. It is similar to our solution based on the same cryptosystem. The main difference is that [Erkin] also does a secret comparison to a threshold. Our secret sharing based solution outperforms the one presented in this work with regards to both computation time and communication costs (see Table).

A Privacy-Preserving Model for Biometric Fusion, (Toli, 2018)
This publication focuses on multimodal user authentication, thus each user template is composed of face (1024 bits), fingerprint (4096 bits) and iris (2048 bits). This differs from our approach where only the face biometric is considered. This solution leverages Shamir secret sharing to perform secure MPC. Our secret share based solution outperforms the one presented in the publications with regards to both computation time and communication costs (see Table).

Table 1 Results for a single authentication/identification per solution

	Erkin (face+fingerprint+iris)	**Toli**	**TNO solution**
Seconds	0,125	0,16	0,0088
Exchanged MB	0,022	1,68	0,016

K.3 Use Case

We consider the situation where two parties each have a template and wish to run one of the above algorithms securely, i.e. obtaining the output without revealing any additional information on the private input value. The threshold of Algorithm 2 is considered to be public input.

This can be considered as a representation of the situation where one party owns the camera images and another party owns a watchlist. They wish to process this information without revealing it to the other parties.

Scalability of the solutions is explored by increasing the number of templates of one of the parties.

K.4 Solution Architecture

The facial recognition algorithm involves integer arithmetic, i.e. arithmetic over $\mathbb{Z}$. Since the computations only involve additions and multiplications, they can be approximated well by arithmetic over a finite field with a large enough cardinality. For this reason, we consider two MPC solutions: 1) based on the homomorphic Paillier encryption scheme and 2) based on the secret-sharing based MPC implementation MPyC.

K.4.1 Homomorphic Encryption (Paillier)

The Paillier based implementation of Algorithm 1 is described in Figure 1. Its security relies on the computational RSA assumption. Moreover, the resulting MPC protocol is passively secure, i.e. secure under the assumption that all parties correctly follow the protocol. The protocol is a 2-party protocol. Since our alternative approach is more efficient we do not consider Algorithm 2 in this setting.

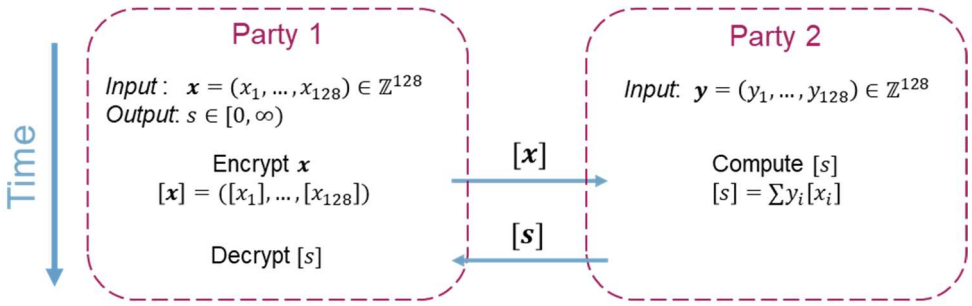


Figure 1 Paillier Based Algorithm 1.

K.4.2 Secret Sharing Based MPC

The secret sharing based implementation of Algorithm 1 is described in Figure 2 and the secret sharing based implementation of Algorithm 2 is described in Figure 3. Because the underlying MPC protocol requires at least three parties, the protocol requires a computation helper. This helper is a third party that does not supply private input to the computation and that does not learn any information about the private input values of parties A and B. We must, however, trust that this helper does not collude with party A or with party B. The protocol is unconditionally secure, i.e. it does not rely on computational hardness assumption. Moreover, the solution has passive security. The need for a computation helper can be eliminated at the cost of relying on computational assumptions and introducing a computationally expensive off-line phase.

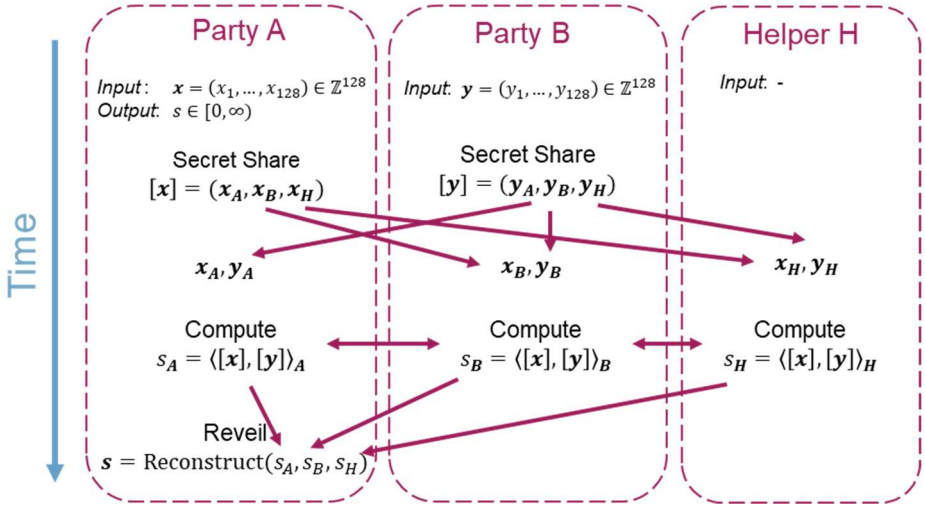


Figure 2 Secret Sharing Based Algorithm 1.

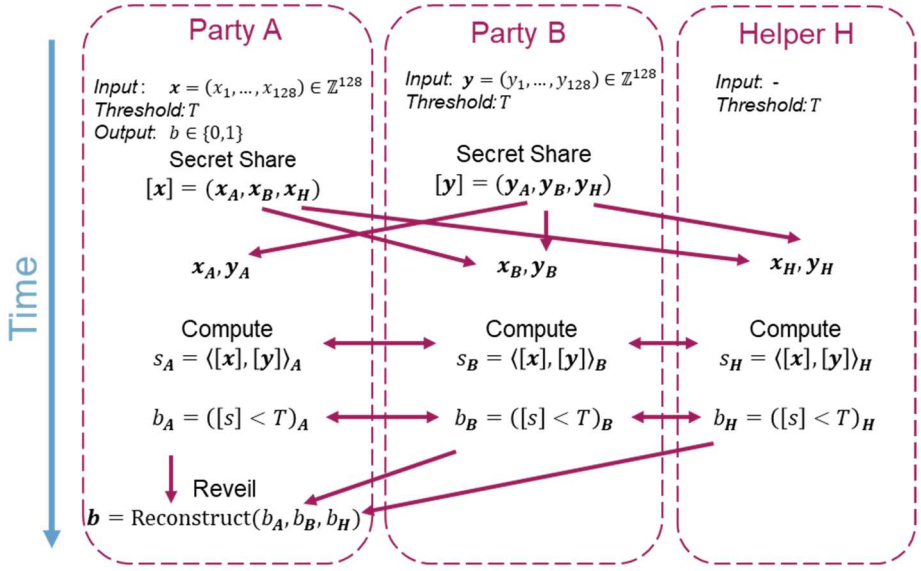


Figure 3 Secret Sharing based Algorithm 2.

K.5 Performance

K.5.1 Paillier vs MPyC

The described MPC solutions have been implemented in proof-of-concept Python implementations, using the Paillier and MPyC libraries. The performances of these implementations have been evaluated by running all three parties on one machine, thereby having negligible communication costs. The number of templates (feature vectors) of Party 2 (B) is increased to determine the scalability of these solutions. The performance of the Algorithm 1 solutions is depicted in Figure 4, where the Paillier cryptosystem has been considered with two different public key sizes (1024 and 2048 bits).

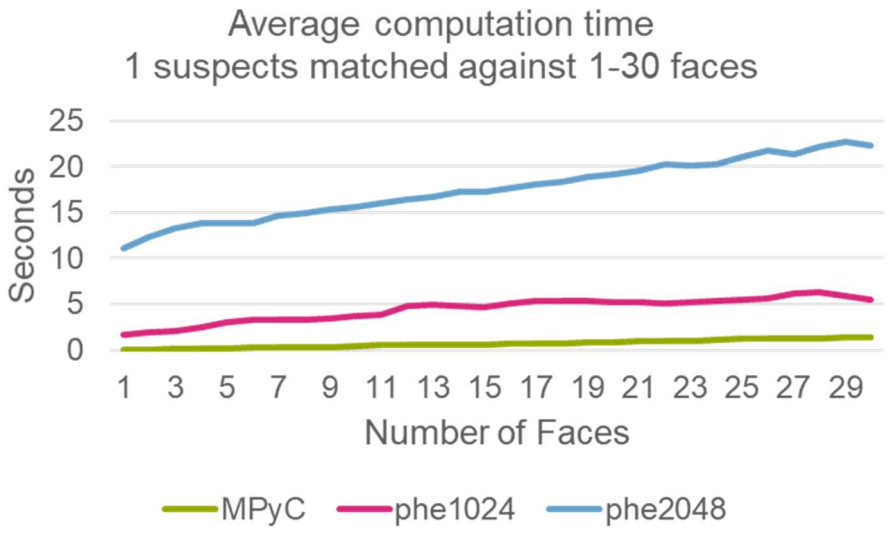


Figure 4 Algorithm 1 Performance Test.

The computational complexity of the secret sharing based (MPyC) solution significantly outperforms the Paillier based implementation. This can be explained by all the public key operations (encryption) that are required when applying the Paillier encryption scheme. For this reason, further performance evaluations were run for the MPyC implementation. These results are depicted in Figure 5 and Figure 6.

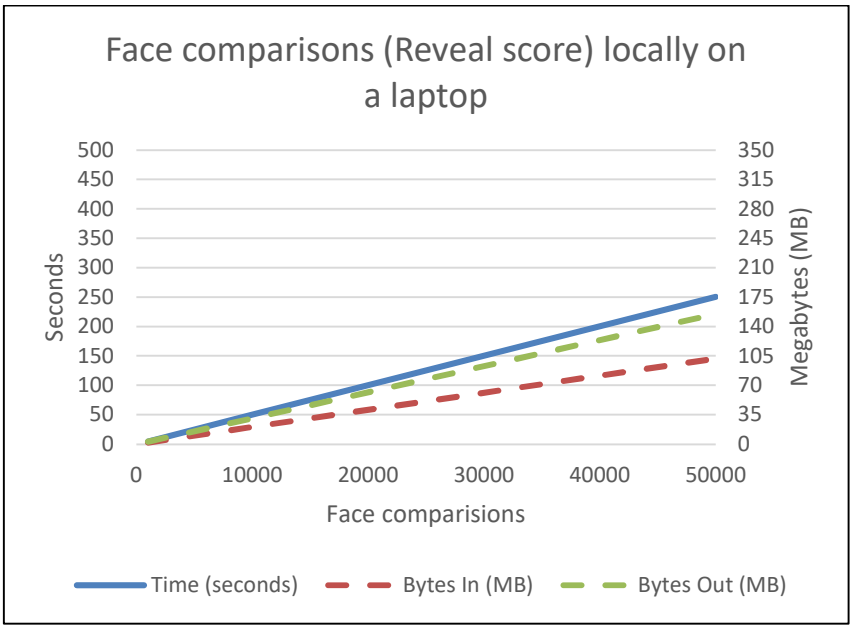


Figure 5 Performance Tests MPyC Based Algorithm 1.

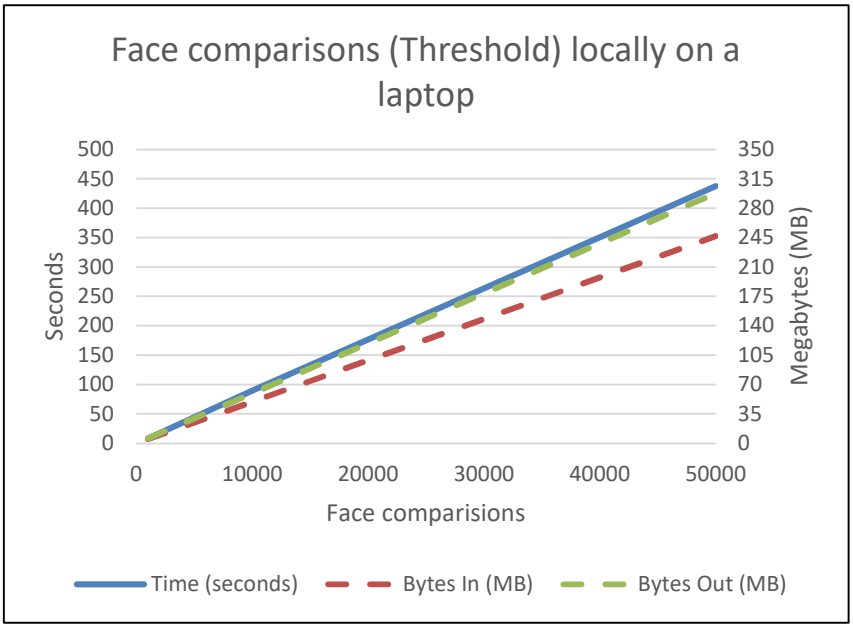


Figure 6 Performance Tests MPyC Based Algorithm 2.

K.5.2 MPC output policy

The performance of a chosen MPC protocol would vary based on the chosen output strategy, I.e. what the computation would reveal to the parties.
Within the facial recogniton scenario the considered output policies were:

- Score: Reveal all scores (Algorithm 1).
- Threshold: Reveal only scores below threshold (Combination of Algorithm 1 and 2).

- Threshold & hide: Binary reveal if score whether score is below threshold (Algorithm 2).
- Sort: Sort the scores but reveal only indices associated with output.
- Threshold & sort: Sort and reveal only scores below threshold ranked.
- Threshold & sort & hide: Sort and reveal the sorted order of items below threshold.

The sorting of score values is a standard operation when using a large face database.
How each strategy impacts the overall algorithm computational time is represented in Figure 7. The graph shows that sorting is a computational demanding operation in a MPC protocol. This can be mitigated by sorting a shorter list of elements with scores below a predefined threshold.

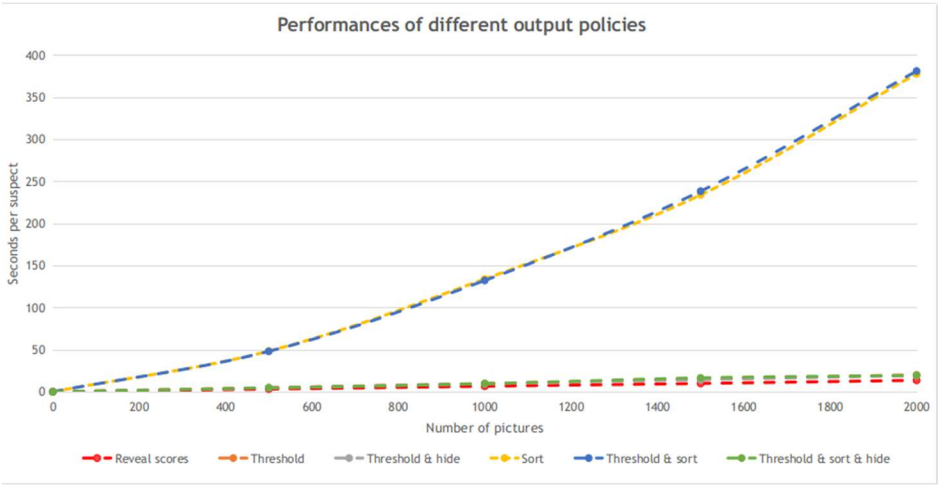


Figure 7 Output policy performance.

K.6 Discussion

With more information about the application domain, it is possible to determine the usefulness of the explored solution directions. This includes time constraints, the type of oversight and the accepted level of organizational and technical complexity. For example, the role of a helper can be seen as an undesirable organizational complexity. On the other hand, it may also align the responsibilities of an oversight party. One that must be able to report on the number of comparisons done, but not on the contents or results. There are many domains where such an oversight role is desired.

As another example, there may be a fairness principle involved with regards to who has to take upon himself the burden of computation, maintenance, etc. If the solution requires more computation on the side of the watchlist, but the provider of the live images benefits the most, this principle may be violated.

Using MPC constitutes more complex computation and communication overhead. This reduces the learnability, maintainability, explainability and thereby also the accountability regarding the technology itself. How easy is it for non-technical

stakeholder (e.g. a policymaker) to verify that the technology uses the proper kind of MPC, and an up-to-date version of it, that it is configured as it should, and that it indeed provides the required level of protection? This question is not unique for MPC, but for all sorts of ICT systems, and can be addressed in an equivalent manner.

K.7 Next Steps

We enumerate a number of improvements/modifications that can be applied to the above mentioned solutions.

- Removing the Computation Helper. The secret sharing based solution can also be implemented without a computation helper. This would degrade the security from unconditional to computational and introduce a preprocessing phase. The preprocessing phase is independent of the input values and can be conducted in an offline manner. The online computations will still be performed in an unconditionally secure and efficient manner.
- Performance improvement of Paillier Encryption, by introducing a preprocessing phase certain computationally intensive tasks can be performed prior to receiving input values. These performance improvements might make the Paillier approach compatible with the secret sharing based approach.
- There might be other facial recognition algorithms or metrics to consider. Alternative recognition metrics can possibly have a smaller impact on the computational complexity.
- The performance tests can be improved by distributing the parties over different machines (and locations) and thereby including communication costs.
- We followed a modular approach in which templates are extracted from images and afterwards processed by the MPC implementation. These different modules can be integrated into a single solution.
- Additional design considerations:
  - The datasets of the different parties might not have the same level of classification. For this reason different input data might have to be treated differently.
  - Computational burden should mostly be with the party that benefits most from the solution.

L Reflectie op de inspiratie voor de uitdaging

In (Wiebes, 2018) werd de uitdaging geformuleerd of de wijze waarop kentekenherkenning (ANPR) werd ingezet voor de handhaving van rijtijdenwet, ook nuttig en mogelijk is voor gezichtsherkenning. In een workshop met de ECP is vervolgens in meer detail beargumenteerd waarom dit idee ook mogelijk en nuttig zou zijn voor gezichtsherkenning (ECP, 2020). Met de informatie verzameld in dit rapport kan op die uitdaging, en op die inspiratie gereflecteerd worden. Deze reflectie illustreert de hoofdconclusie in sectie 10.1.

L.1 De casus handhaving rijtijdenwet middels ANPR

Er zijn vrachtwagenchauffeurs die beroepsmatig vanuit Zuid-Europa naar Nederland rijden. Daarbij kunnen ze de rijtijdenwet overtreden. Dat kan worden gesignaleerd door het tijdstip van passage bij de respectievelijke landgrenzen met elkaar te vergelijken.

Een (privacy-)probleem in die context is dat het juridisch en privacy-ethisch onwenselijk is om kentekengegevens *voor dit doel* tussen lidstaten uit te wisselen. Daarom is het volgende concept bedacht.

Kentekens van vrachtwagens worden bij de grensovergang van Spanje naar Frankrijk geregistreerd en versleuteld opgeslagen. De oorspronkelijke, onversleutelde kentekens, worden verwijderd. Bij de grensovergang Nederland in worden ook kentekens geregistreerd en *in het versleutelde domein* vergeleken met de kentekens uit Spanje. Het effect hiervan is dat Nederland dus alleen kan opvolgen op basis van kentekens die (binnen een korte tijdsduur) zowel op de grensovergang in Spanje, als op de grensovergang in Nederland waargenomen zijn. In dit concept krijgt data (in dat geval een kenteken) pas een betekenis als de hypothese van een verdenking (van overtreding van de rijtijdenwet) waarheid wordt. Andere versleutelde kentekens die Nederland van Spanje ontvangt zijn dus niet als kenteken te gebruiken (Wiebes, 2018).

L.2 Kentekens en gezichten

Kentekens zijn ontworpen om uniek te zijn. Een detectie van twee identieke kentekens kan daardoor als een indicator van fraude en van ernstiger zaken worden gezien. Gezichten zijn echter niet inherent uniek, denk bijvoorbeeld aan tweelingen. Dat kan adequate herkenning van personen lastiger of in bepaalde gevallen zelfs onmogelijk maken. De opvolging na een identificatie (zowel correct als foutief) moet daar dus rekening mee houden. Het is daarbij een barrière voor adequate opvolging, voor transparantie en voor *accountability* als de enrolment opname niet gebruikt kan worden om de live opname handmatig mee te vergelijken, omdat die *enrolment* opname is weggegooid.

Afbeeldingen van gezichten en biometrische templates zijn gevoelige persoonsgegevens. Kentekens zijn dat niet. Daardoor vereisen gezichtsafbeeldingen en biometrische gezichtstemplates meer terughoudendheid

bij toepassing, en betere gegevensbescherming. Bij het selecteren van een toepassing moeten proportionaliteit en subsidiariteit worden gewogen. Het identificeren van een mogelijke toepassing luistert in die zin dus “nauwer”.

Voor kentekens is er een standaard om de data mee te beschrijven, o.a. gebaseerd op ISO 3166 voor de landencode. Er is geen standaard om een biometrisch gezichtstemplate te beschrijven. Eventuele oplossingen zijn dus leverancier-specifiek. MPC kan alleen werken als alle partijen hetzelfde data-format gebruiken. Het ontbreken van een standaard kan een barrière opleveren voor grootschalige toepassing van MPC voor (niet-coöperatieve) gezichtsherkenning.

L.3 Privacy dreigingen

De kentekens die in Spanje en Nederland verzameld werden, zouden ook (door het andere land) voor andere doeleinden gebruikt kunnen worden. Door ze zo snel mogelijk te verwijderen, en door ze alleen versleuteld ter beschikking te stellen, kunnen deze privacy risico’s worden gemitigeerd.

Het analyseren in het versleutelde domein helpt met name tegen de privacy dreigingen “Secondair gebruik” (zie sectie G.2.4), en “Inbreuk op de vertrouwelijkheid” (zie sectie G.3.1). Het helpt ook tegen “Onveiligheid” (zie sectie G.2.3), maar daar zijn eenvoudiger maatregelen voor denkbaar.

Er zijn toepassingen denkbaar voor niet-coöperatieve gezichtsherkenning waar deze privacy dreigingen ook denkbaar zijn (zie sectie 5.2.1). Uit documentstudie is echter niet gebleken dat dit door overheid of de wetenschappelijke gemeenschap ook zo wordt gezien (zie sectie 5.2.2). Alleen (Keymolen, Noorman, Van der Sloot, Cuijpers, & Koops, 2020) maken melding van de dreiging van “Secondair gebruik”.

De kans dat vervorming voorkomt in het kentekenscenario is erg laag, omdat kentekens en de weginfrastructuur (denk aan belijning, belichting) en verkeersregels (denk aan maximum snelheid) zodanig zijn dat het eenvoudig is gemaakt om kentekens te herkennen. Ook is te beargumenteren dat het belang voor overtreders relatief laag is om niet herkend te worden. Het gaat om een (stevige) boete en vertraging. Er mag dus verwacht worden dat er geen grote inspanning (die ook tijd kost) zal worden verricht om herkenning te voorkomen.

Voor gezichtsherkenning wordt juist de dreiging “Vervorming” als relevant gezien. Het is denkbaar dat bij toepassingen waar gezichtsherkenning proportioneel is, het belang van de gezochte persoon ook groter is om niet gevonden te worden. De kans dat hij of zij probeert de werking van het systeem te frustreren kan daardoor ook groter zijn.

Het is in dat licht relevant te benoemen dat verwerking in het versleutelde domein in combinatie met het verwijderen van de oorspronkelijke *enrolment* opnames, het risico verbonden aan “Vervorming” vergroot. Immers, er kunnen minder geavanceerde afstandsmaten worden gebruikt, en foutieve identificaties kunnen niet worden gecontroleerd tegen de oorspronkelijke opnames.