De (R)evolutie van Digitaal Bewijs

dr. ir. J. Henseler
Lector Digital Forensics
& E-Discovery

De (R)evolutie van Digitaal Bewijs

Lectorale rede, zoals uitgesproken door dr. ir. J. Henseler
op 21 november 2017 op Hogeschool Leiden.

Colofon
Rede, in verkorte vorm uitgesproken in opdracht van Hogeschool Leiden op 21 november 2017.

Auteur: dr. ir. J. Henseler
Uitgever: Hogeschool Leiden
Vormgeving: UWMERK!WAARDIG, Rotterdam
Omslag illustratie: afgeleid van "smart city and wireless communication network, Internet of Things, abstract image visual" (auteur chombosan / Shutterstock.com)
Drukwerk: Tuijtel, Hardinxveld-Giessendam
ISBN: 978-90-821382-2-1

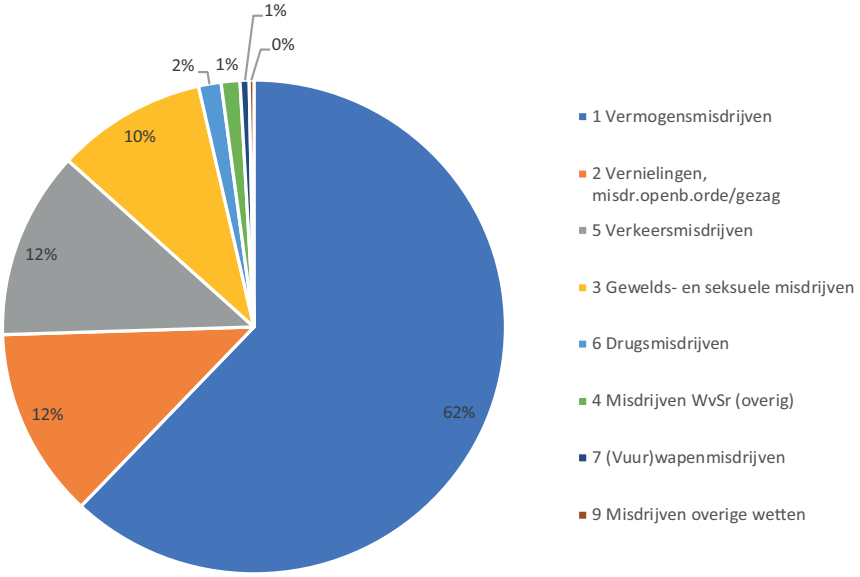
Inhoudsopgave

1. Inleiding

In 2010 heb ik mijn lectorale rede gehouden met de titel 'Op zoek naar de digitale waarheid' (Henseler, 2010). Deze rede markeerde het begin van het Lectoraat E-Discovery, nadat ik in 2009 als lector was begonnen bij de Hogeschool van Amsterdam. In die rede, en in het bijbehorende boekje, keek ik vooral terug op twee decennia (1990-2010) waarin ik eerst het forensisch computeronderzoek en naderhand E-Discovery in Nederland heb zien groeien.

In deze lectorale rede wil ik de evolutie van digitaal bewijs in het forensisch onderzoek beschrijven, die mijns inziens op het punt staat om een revolutionaire ontwikkeling door te maken. Om dat te kunnen begrijpen is het goed om stil te staan bij het fenomeen digitaal bewijs en op welke wijze digitaal bewijs in het forensisch onderzoek en in de opsporing wordt gebruikt.

Figuur 1 geeft een overzicht van het aantal geregistreerde misdrijven in Nederland in 2016 opgesplitst in categorieën. Iedere categorie bestaat uit een onderverdeling. Zo valt bijvoorbeeld computervredebreuk onder categorie 2.2.4 en omkoping onder 2.4.6 (openbaar gezag misdrijf overig).



Figuur 1. Aantal geregistreerde misdrijven in 2016 per categorie[1].

[1] Volgens het CBS, zie http://statline.cbs.nl

Het aantal geregistreerde misdrijven is waarschijnlijk lager dan het totaal aantal misdrijven. Vooral van moderne vormen van identiteitsfraude en oplichting via internet wordt lang niet altijd aangifte gedaan. Zo waren er volgens het CBS in 2015 4,4 miljoen delicten in de traditionele criminaliteit. Hiervan is 36% gemeld bij de politie en van 27% is werkelijk aangifte gedaan[2].

Het forensisch computeronderzoek in het begin van de jaren '90 ging vooral over bewijs zoeken op computers en in elektronische zakagenda's die met een wachtwoord waren beveiligd. In die jaren was digitaal forensisch onderzoek ook vooral iets wat speelde binnen de opsporingsdiensten. Bedrijven hadden er nog weinig mee te maken. Veel communicatie verliep altijd nog via de reguliere post en de fax. Internetverbindingen waren in die tijd vooral in gebruik bij de academische wereld.

In het volgende decennium, in de jaren 2000-2010, ging het steeds vaker over e-mails, documenten en, afhankelijk van het soort verdenking, om foto's, video's en internet-geschiedenis. Digitaal bewijs in bedrijfsonderzoeken naarbijvoorbeeld mededinging, omkoping en interne fraude zijn toen sterk gegroeid. In de VS zijn de Federal Rules of Civil Procedure aangepast en was daarmee voor eens en voor altijd duidelijk dat elektronisch opgeslagen informatie (ESI) ook bewijs is en niet genegeerd mag worden.

In 2009 was E-Discovery vooral nog een Amerikaans-Engels verschijnsel en bedrijven in continentaal Europa die te maken hadden met een moeder- of dochterbedrijf in die markten, schakelden regelmatig de Big 4 (Deloitte, PWC, KPMG en EY) in om digitaal bewijs te verzamelen. In Nederland werd er weliswaar ook gezocht naar digitaal bewijs, bijvoorbeeld in het onderzoek naar de Vastgoedfraude[3], maar toch nooit op de schaal waarin men in de VS en de UK dit soort projecten uitvoerde.

Door de aanhoudende exponentiële groei van digitale informatie (verdubbeling iedere 18 maanden) was het dan ook geen verassing dat men in de VS op zoek ging naar een manier om de kosten van E-Discovery beheersbaar te maken. In 2012 schreef ik over predictive coding, een nieuwe ontwikkeling die wat mij betreft de geschiedenis in mag gaan als een mijlpaal van de inzet van kunstmatige intelligentie. Bij predictive coding[4] voorspelt de computer aan de hand van een model of een e-mail of document relevant is.

[2] https://www.cbs.nl/nl-nl/nieuws/2016/46/vandalisme-meestal-niet-gemeld-bij-politie
[3] Zie video-opname van presentatie door Willem Koops van Spigthoff Litigators over Digitaal Bewijs in de Vastgoedfraude.
Zie https://hva.mediamission.nl/Mediasite/Play/5106154ffbff45e698417d0f43c254111d
[4] Zie publicatie in de AutomatiseringGids, nu online te vinden op http://agconnect.nl/artikel/predictive-coding-glazen-bol-of-black-box

Het bijzondere is dat dit model automatisch is gemaakt door een computer aan de hand van voorbeelden die door een expert zijn beoordeeld. Deze machine learning-technieken bestonden al langer, maar de groeiende kosten en de (nieuwe) wetenschap dat mensen slecht zijn in het beoordelen van grote hoeveelheden documenten[5], zorgden ervoor dat een aantal rechters in de VS toestond dat deze vorm van kunstmatige intelligentie mocht worden toegepast.

De laatste tien jaar, van 2007-2017, worden gekenmerkt door de opkomst van sociale media, de smartphone en de cloud en vooral ook de combinatie van die drie. Nu zijn we aangekomen op een punt dat digitaal bewijs, dat hiervan afkomstig is, soms een bijna vergelijkbare rol aanneemt als die van het traditioneel bewijs. De reden daarvoor is dat smartphones door sociale media heel persoonlijk zijn geworden. Door het contact met de cloud wordt veel vaker dan voorheen locatie en tijdstip van de smartphone vastgelegd en dus van de gebruiker.

Traditioneel was E-Discovery vooral gericht op het analyseren van documenten om te achterhalen wie wat wist en wanneer. Nu is het ook mogelijk om te bepalen waar de gebruiker zich bevond. Daarmee komen de digitale sporen langzaam meer op het terrein van de klassieke forensische sporen zoals vingerafdrukken, voetsporen en DNA.

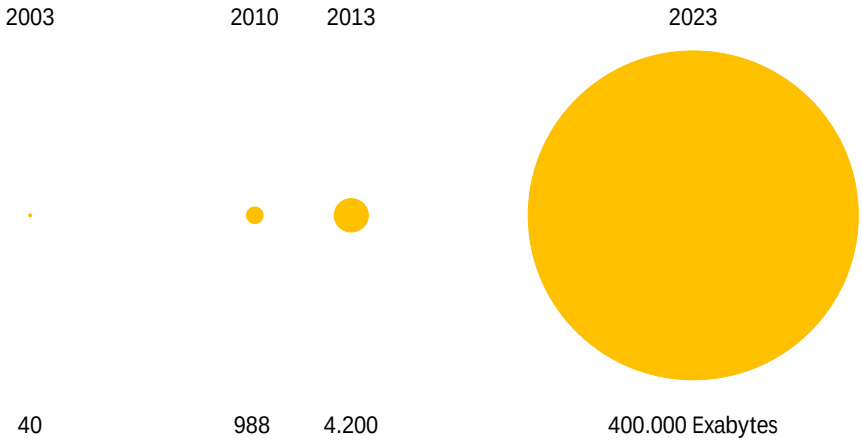
Met deze ontwikkeling is naar mijn mening de revolutie van digitaal bewijs reeds begonnen en nog zeker niet aan het einde. In het vervolg van deze lectorale rede wil ik u meenemen in verschillende trends in de ICT die onomkeerbaar zijn en waarvan ik verwacht dat ze een belangrijke rol zullen spelen in de revolutionaire ontwikkeling die digitaal bewijs ook de komende jaren zal doormaken. Het gaat om de volgende trends.

Het internet der dingen

In de komende jaren zullen niet alleen computers en telefoons met hun gebruikers de online wereld bevolken. Er is een internet der dingen 'internet of things' (IoT) aan het ontstaan waarbij al onze apparatuur wordt aangesloten op internet en informatie zal versturen en ontvangen. De kosten van het internetgebruik zullen dalen en de verbindingssnelheden zullen blijven stijgen. Het IoT zal de komende jaar explosief groeien waardoor de exponentiële groei van de hoeveelheid digitale informatie voorlopig nog niet voorbij is.

[5] Zie het artikel van Blair en Maron uit 1985 "An Evaluation of Retrieval Effectiveness for a Full-Text Document-Retrieval System" in Communications of the ACM 28(3):288-299. In 2011 publiceerde Grossman en Cormack een studie "Technology-Assisted Review in E-Discovery Can Be More Effective and More Efficient Than Exhaustive Manual Review" in het Richmond Journal of Law and Technology, Vol 17(3) waaruit bleek dat reviewers veel beter presteren met de hulp van een computer.

Deze exponentiële groei wordt door IDC geschat op een verdubbeling iedere 18 maanden, wat moeilijk te bevatten is. Onderstaande figuur probeert te illustreren wat het verschil is tussen de groei tussen 2003 en 2013 en de groei de daarop volgende 10 jaar van 2013 tot 2023. De hoeveelheid data in 2003 is letterlijk een stip in vergelijking met de enorme schijf die overeenkomt met de hoeveelheid informatie die in 2023 voorspeld wordt.

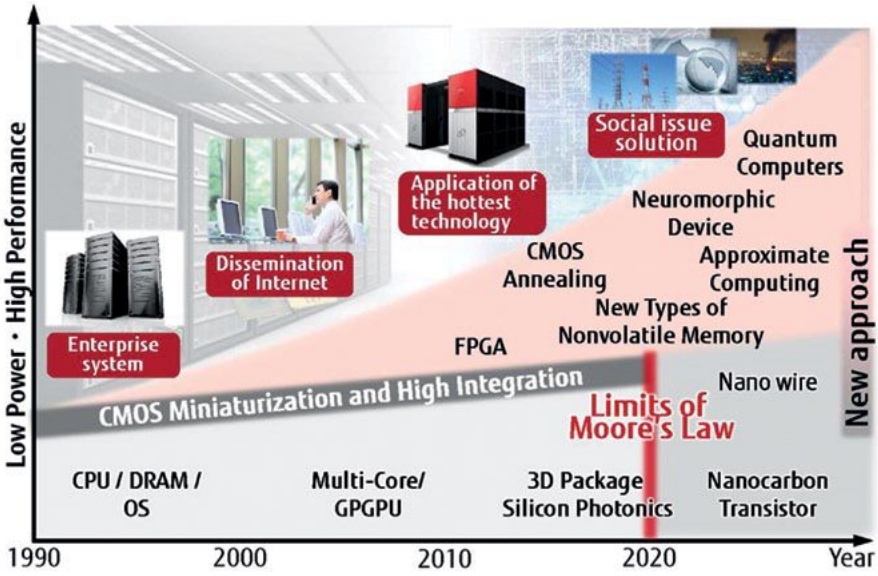


Figuur 2. Exponentiële groei van de hoeveelheid informatie.

De opkomst van kunstmatige intelligentie

De hierboven geschetste groei van de hoeveelheid informatie en de verdere digitalisering van onze complete samenleving worden nog verder versterkt doordat onze computers slimmer, sneller, kleiner en goedkoper worden. Volgens de Wet van Moore zal het aantal transistors op een chip iedere achttien maanden verdubbelen en daarmee ook de snelheid. De voorspellingen wisselen, maar de algemene verwachting is dat deze wet nog in ieder geval tot 2020-2025 van toepassing blijft, totdat de miniaturisering zo ver is doorgegaan dat de fysieke grenzen zijn bereikt.

De hardware wordt niet alleen steeds maar kleiner. Ook nieuwe computerarchitecturen, die geoptimaliseerd zijn voor deep learning-algoritmes, zullen ervoor zorgen dat computers enorme hoeveelheden data snel kunnen analyseren. In combinatie met het IoT worden nieuwe toepassingen mogelijk die veel verder gaan dan de automatisering die we tot nu toe kennen in onze omgeving. Het volgende figuur probeert een indruk te geven van deze ontwikkeling.



Figuur 3. Evolutie van nieuwe computer hardware[6].

Samensmelting van cyberspace en de werkelijke wereld

Tot nu toe was de computer vooral een instrument, een hulpmiddel om informatie automatisch mee te kunnen verwerken. Dat is langzaam aan het veranderen en de komende jaren zal de computer steeds meer een verlengstuk van ons worden.

Figuur 4 bevat een frame uit een TED prestentatie[7] in 2016 door Alex Kiplan van Microsoft Hololens. Hij demonstreert live hoe je een gesprek kunt hebben met een bewegend hologram van iemand die ergens anders is terwijl er ook nog wetenschappelijke data, in dit geval het oppervlakte van Mars, getoond wordt. Dankzij beeldherkenning kan de Hololens hologrammen in de omgeving plaatsen en gebaren van de drager herkennen.

[6] Overgenomen uit https://nworeport.me/2016/12/13/in-2016-supercomputers-merged-with-ai-and-by-2018-there-will-ai-machines-with-exaflop-power/
[7] https://www.ted.com/talks/alex_kipman_the_dawn_of_the_age_of_holograms#t-645065

De combinatie van kunstmatige intelligentie en het IoT zal uiteindelijk leiden tot een versmelting van cyberspace en de werkelijke wereld. We krijgen nieuwe userinterfaces waarmee we bijna ongemerkt onze zintuigen en ons geheugen laten ondersteunen door de computer.

Figuur 4. Alex Kipman van Microsoft Hololens communiceert met een hologram.

Deze ontwikkeling past in het bredere perspectief van de vierde industriële revolutie die volgens deskundigen[8] aanstaande is en die de manier waarop we leven, werken en met elkaar omgaan fundamenteel zal veranderen. Het zal geen verrassing zijn dat diezelfde revolutie ook zal leiden tot een revolutie van digitaal bewijs.

[8] Zie artikel van het World Economic Forum op https://www.weforum.org/agenda/archive/fourth-industrial-revolution

2. De smartphone als digitaal bewijs

Wie had in 1989 bij de eerste experimenten van Tim Berners-Lee[9] kunnen bedenken hoe het internet binnen 25 jaar ons leven zou veranderen? Wie had kunnen voorspellen dat na de eeuwwisseling niet e-mail maar sociale media voor een explosieve groei van informatie zouden zorgen? En wie had gedacht dat de smartphone vanaf 2009 binnen zeven jaar de desktopcomputer zou voorbijstreven als meest gebruikte device om online te zijn[10]?

De drie trends die in de inleiding zijn geïntroduceerd, zouden niet mogelijk zijn geweest zonder opkomst van sociale media, de smartphone en de cloud in de afgelopen tien jaar. Daarom wil ik in dit hoofdstuk uitleggen wat de impact van deze ontwikkelingen is geweest op de manier waarop nu tegen digitaal bewijs wordt aangekeken.

2.1 Smartphones

De smartphone heeft de aard van het digitaal bewijs flink gewijzigd. Vóór de smartphone bevatten telefoons telefoonnummers, gesprekslijsten en eventueel sms-berichten. Wat later konden gebruikers foto's maken met hun telefoon en werd e-mail toegevoegd zodat werknemers ook onderweg bereikbaar waren. De Blackberry was een van de eerste telefoons die veilig, betrouwbaar en gebruikersvriendelijk e-mail wist te integreren in een mobiele telefoon. Maar het bleef vooral bij een kopie van de e-mail die ook op de desktop en op de mailserver van het bedrijf was te vinden.

Met de komst van de iPhone in 2007 en de daarop volgende modellen, het app-store concept en de introductie van Android, is de smartphone niet alleen smart maar vooral ook persoonlijk geworden. Bovendien werd al snel het nut onderkend van de smartphone als een mobiel apparaat waar je niet alleen mee kunt bellen, maar waarop je ook kunt navigeren en waarmee je foto's en video's kunt delen. Het duurde niet lang of al deze eigenschappen maakten de smartphone het ideale apparaat om overal en nergens informatie te delen met vrienden via sociale media of publiek toegankelijk te maken.

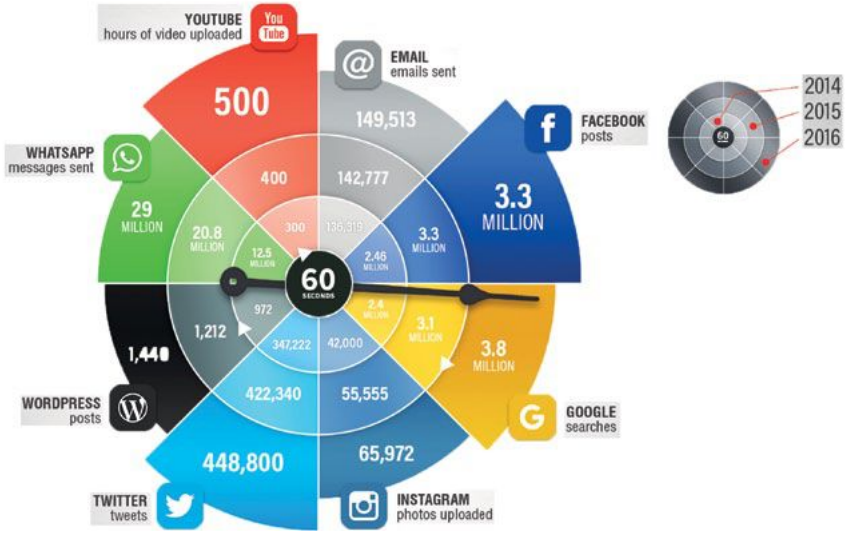
[9] Tim Berners-Lee stelde in maart 1989 een nieuw soort informatiemanagement-systeem voor. In datzelfde najaar implementeerde hij de eerste versie van het Hyper Text Transfer Procol (HTTP). Daarmee legde hij de basis voor ons huidige internet (https://en.wikipedia.org/wiki/Tim_Berners-Lee).

[10] Zie bijvoorbeeld http://bgr.com/2016/11/02/internet-usage-desktop-vs-mobile

In 2010 lag de nadruk in het forensisch onderzoek nog vooral op het onderzoeken van digitale opslagmedia, laptops en desktopcomputers. Mobiele telefoons werden nog maar sporadisch in beslag genomen. In de jaren daarna is dat drastisch veranderd. In 2010 bestond circa 20% van het digitaal bewijs dat door opsporingsdiensten werd onderzocht, uit telefoons en de overige 80% bestond uit laptops, computers en opslagmedia. Inmiddels is die verhouding compleet omgedraaid en maken smartphones 80% van het bewijs uit. Er zijn zelfs steden waar de politie alleen nog maar smartphones en geen enkele computer meer op digitaal bewijs onderzoekt.

2.2 Open source intelligence

Smartphones en sociale media hebben geleid tot een enorme explosie van informatie op het internet, die (tot op zekere hoogte) vrij toegankelijk is. Hoeveel informatie er wordt geproduceerd is duidelijk te zien in onderstaande figuur, die bovendien laat zien dat de hoeveelheid informatie gestaag doorgroeit.



Figuur 5. Wat gebeurt er in 60 seconden op internet?[11]

[11] Overgenomen uit http://www.smartinsights.com/internet-marketing-statistics/happens-online-60-seconds

In 2014 verscheen het boek “Social Media: Het nieuwe DNA”, waarin de auteurs uitleggen op welke wijze sociale media op internet kunnen helpen bij het verbeteren van de maatschappelijke veiligheid. Een van de auteurs, Arnout de Vries, gaf op het Symposium E-Discovery in Nederland in 2015 een tot de verbeelding sprekende presentatie met uitgebreide voorbeelden van de informatie die over privépersonen op sociale media zijn te vinden.

Tijdens zijn presentatie liet Arnout een filmpje uit 2012 zien van de Vlaamse Dave Guillaume, die als helderziende in een oosters uitziende tent op een markt in Brussel aan willekeurige voorbijgangers zijn gaven toonde[12]. De bezoekers waren allen erg onder de indruk als Dave de meest intieme details van hun leven onthulde. Wat ze niet wisten was dat Dave een speciaal team in een ander deel van de tent had geïnstalleerd dat in hoog tempo allerlei informatie over de betreffende bezoeker naar boven wist te halen door te zoeken in sociale media.

Het zoeken in open bronnen op internet wordt ook wel open source intelligence (OSINT) genoemd. Deze term is afkomstig van de inlichtingendiensten die al decennia lang informatie verzamelen uit openbare informatiebronnen, zowel offline als online. Andere bekende intelligence-vormen zijn HUMINT (human intelligence) en SIGINT (signal intelligence).

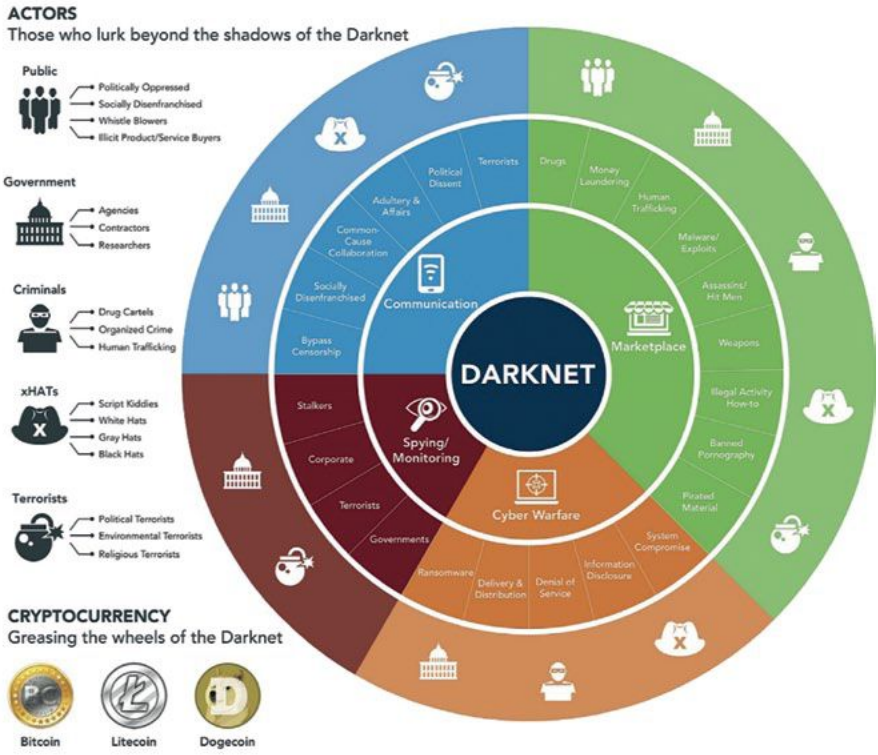
Naast sociale media zijn er talloze andere openbare informatiebronnen te vinden op internet zoals marktplaatsen, websites en chatrooms. Al deze bronnen zijn vervolgens weer doorzoekbaar gemaakt door algemene zoekengines zoals Google en Bing, maar ook door specifieke zoekmachines die zijn gespecialiseerd in uiteenlopende onderwerpen zoals personen, bedrijven, vacatures en boeken.

Lang niet alle informatie op internet kan worden gevonden door de bekende zoekmachines. Het internetgedeelte dat verborgen blijft wordt ook wel aangeduid als Deep Web. Een groot deel van het Deep Web is afgesloten omdat bedrijven en personen hun privégegevens en gevoelige bedrijfsinformatie niet met iedereen willen delen.

Criminelen blijven ook graag uit het zicht van opsporingsdiensten. Zij verschuilen zich op het Dark Net. Het Dark Net is een deel van het Deep Web dat wel voor iedereen toegankelijk is (met de juiste software) maar waarbij de identiteit van de aanbieders van informatie niet makkelijk te achterhalen is. Onderstaande figuur bevat een globaal overzicht van het ondergrondse netwerk dat door criminelen in het Dark Net is aangelegd.

[12] Het bewuste filmpje is op YouTube te vinden via https://www.youtube.com/watch?v=F7pYHN9iC9I

Deze figuur illustreert in feite de grote mate van verwevenheid tussen de fysieke (criminele) wereld en het internet.



Figuur 6. Criminele activiteiten, actoren en betaalmiddelen op het Dark Net[13].

[13] Overgenomen uit http://bitwitt.com/welcome-Dark Net-underground-underground/

2.3 Uitdagingen

Softwarefabrikanten van forensische tools zijn de afgelopen jaren druk aan de slag gegaan met het toevoegen of uitbreiden van functionaliteit om smartphones te kunnen onderzoeken. Het op forensische manier veiligstellen van informatie in een smartphone is geen sinecure. Gelukkig worden de tools steeds eenvoudiger, maar de beveiliging is wel een groeiend probleem.

Begin 2016 werd de wereld opgeschrikt door nieuws dat de FBI de firma Apple via de rechter wilde dwingen om mee te werken aan het ontgrendelen van een beveiligde iPhone. De iPhone was van een terrorist die in 2015 samen met zijn vrouw veertien mensen had doodgeschoten. Apple bleek echter niet in staat te zijn om de beveiliging te doorbreken en de FBI heeft uiteindelijk professionele hackers ingehuurd die erin slaagden om toegang tot de telefoon te krijgen.

Het doorbreken van de beveiliging is zeker niet de enige uitdaging. Tegenwoordig hebben telefoons een geheugen van soms wel 64Gb of meer. Een telefoon kan duizenden of wel tienduizenden foto's en filmpjes bevatten en het aantal berichten in bijvoorbeeld WhatsApp loopt makkelijk op tot in de tienduizenden. Het doorzoeken van al die informatie op relevant bewijs kan een enorme klus zijn.

Verder is er nog de grote variëteit aan geïnstalleerde apps. Gebruikers van Android konden in maart 2017 kiezen uit maar liefst 2,8 miljoen apps en gebruikers van Apple konden kiezen uit 2,2 miljoen apps[14]. Veel apps slaan gegevens lokaal op in de telefoon in een eigen database. Het mag duidelijk zijn dat de beschikbare tools niet alle apps ondersteunen. Alleen de meest gebruikte apps worden geanalyseerd. Dat zijn in veel gevallen internationaal bekende apps en niet de typisch lokale apps in een land of regio waar gebruikers vaak gebruik van maken. Onderstaande figuur geeft het topje van de ijsberg weer van populaire apps in de Nederlandse Android app store.

[14] Zie https://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/

Figuur 7. Overzicht van meest populaire apps in de Nederlandse Android app store[15].

Naar aanleiding van een stageopdracht bij de Nationale Politie heeft Vince Noort, student forensische ICT van de Hogeschool Leiden, een script ontwikkeld waarmee in databases van apps automatisch wordt gezocht naar patronen die lijken op datum tijd, e-mailadressen, telefoonnummers en GPS-coördinaten. Omdat het script onderdeel is van een bestaande commerciële tool is het gebruik ervan redelijk eenvoudig. Onlangs heb ik over dit onderzoek een presentatie gegeven op de DFRWS US 2017 conferentie; de presentatie is beschikbaar op de website van de conferentie[16].

Gebruikers kunnen van hun smartphone een back-up maken op hun computer. In sommige gevallen is dat heel handig voor onderzoekers van digitaal bewijs. Als deze back-up niet beveiligd is, is het heel eenvoudig om de informatie in die back-up forensisch te onderzoeken. Echter, niet alle gebruikers maken een back-up omdat zij het te lastig vinden of omdat zij simpelweg geen computer (meer) hebben.

De laatste jaren wordt opslag van data (en dus back-ups) in de cloud steeds goedkoper. Apple biedt haar gebruikers opslag in de iCloud aan. De eerste gigabytes zijn gratis, maar wie zijn goedgevulde 64Gb iPhone 6 wil opslaan in de iCloud moet een betaald abonnement afsluiten. Microsoft heeft OneDrive en via Google kunnen gebruikers werken met Google Drive. Ook Dropbox is heel populair.

[15] Statistieken afkomstig uit https://www.applyzer.com/?mmenu=worldcharts
[16] Presentatie "Finding Digital Evidence in Mobile Devices". PDF-versie beschikbaar via https://www.dfrws.org/conferences/dfrws-usa-2017/sessions/finding-digital-evidence-mobile-devices

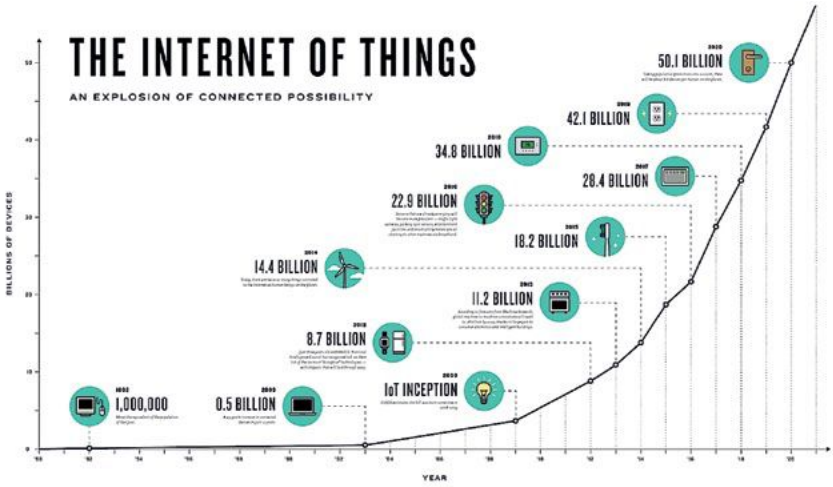
Naast apps voor opslag zijn er veel apps op de smartphone die rechtstreeks gekoppeld zijn aan een (eigen) clouddienst. Denk bijvoorbeeld aan e-mail, WhatsApp, Facebook, Twitter, Snapchat en Instagram. Gebruikers die hun smartphone beveiligd hebben met een vingerafdruk of toegangscode, slaan hun gebruikersnaam- en wachtwoordcombinaties van deze apps op in de telefoon. Dat is makkelijk omdat ze niet telkens hun wachtwoord hoeven in te vullen.

Inmiddels zijn er forensische computertools beschikbaar die precies weten waar deze gegevens in de smartphone zijn opgeslagen. Ze beschikken over programma's die met behulp van deze inloggegevens de data van de gebruiker vanuit de cloud kunnen kopiëren. Daarmee kunnen ook gegevens die niet meer op de smartphone staan, bijvoorbeeld omdat het geheugen vol was, toch onderzocht worden als digitaal bewijs.

Aan uitdagingen voor digitaal forensische onderzoekers is dus zeker geen gebrek. Door bedrijven en kennisinstellingen wordt meer dan ooit gewerkt aan nieuwe oplossingen en uitbreiding van de bestaande tools. Het lijkt bijna onmogelijk om nog sneller te ontwikkelen, maar toch zal dat moeten als het digitaal bewijs zich inderdaad revolutionair zal gaan ontwikkelen zoals ik in de volgende hoofdtukken zal gaan beschrijven.

3. Het internet der dingen

Het IoT is zeker niet nieuw, maar de verwachte groei ervan de komende jaren is op zijn minst spectaculair, zoals in onderstaande figuur is te zien. De groei is nu nog vooral afkomstig uit de maakindustrie, de utilities- en de transportsector. Consumenten kwamen in 2014 op de vierde plaats, maar na 2020 wordt ook in de consumentensector een sterke groei verwacht.



Figuur 8. Explosieve groei van het IoT[17].

Ons dagelijkse leven zal steeds meer beheerst gaan worden door slimme apparaten die ons gaan verlossen van eenvoudige taken. Deze apparaten zijn aangesloten op het internet en kunnen met elkaar en met ons communiceren. In dit hoofdstuk wil ik uitleggen wat de gevolgen zijn van deze ontwikkeling voor digitaal bewijs.

De ontwikkeling van het IoT is begrijpelijk in het perspectief van de ontwikkeling van de smartphone, die in het vorige hoofdstuk is beschreven. Door de smartphone zijn mensen onderling veel meer verbonden met elkaar en met het internet. Vervolgens zien we dat ook auto's, huizen, gebouwen en steden verbonden worden door de inzet van online sensoren. Dankzij de apps op smartphones kunnen mensen met zulke online sensoren communiceren en gaan auto's, huizen, gebouwen en steden rekening met ons houden.

[17] Overgenomen van https://www.i-scoop.eu/internet-of-things-guide

De transportindustrie is al jarenlang bezig met het toepassen van sensoren die gebruikt worden om de bedrijfszekerheid te vergroten en de exploitatiekosten te verlagen. Het onderhoud van vliegtuigen wordt al grotendeels geregeld via online sensoren die gebruikt worden om onderhoudswerkzaamheden in te plannen. Gegevens uit motor-managementsystemen in onze auto's werden al gebruikt als de auto in de garage was voor onderhoud, maar langzamerhand worden deze systemen ook online gekoppeld zodat de gegevens continu beschikbaar zijn.

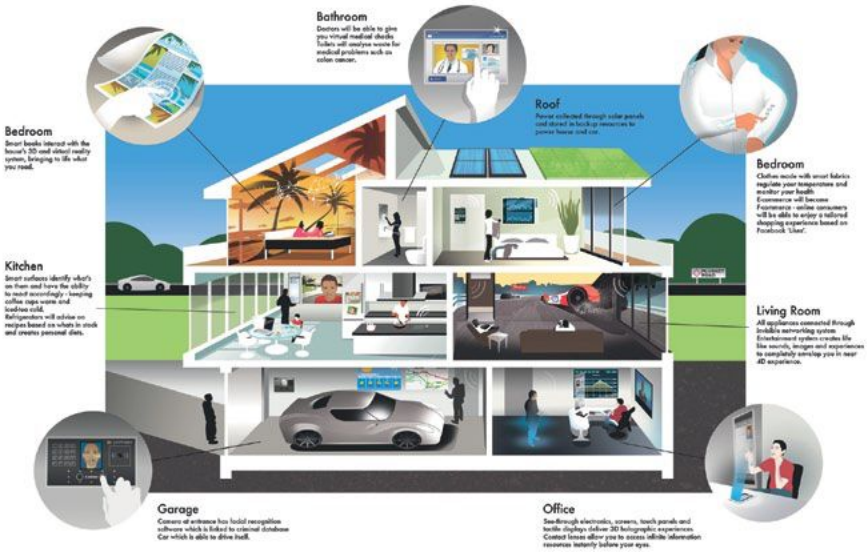
De ontwikkeling zoals de transportindustrie nu doormaakt, verspreidt zich verder naar andere domeinen zoals de gezondheidszorg. Naarmate mensen steeds meer van zichzelf gaan meten, zal er ook een moment ontstaan dat de gezondheidszorg gaat profiteren als deze data ingezet kan worden voor preventie of vroegtijdige detectie van ziektes op een moment dat behandelen nog zeer effectief kan zijn. Met die belofte is het haast niet voor te stellen dat deze ontwikkeling niet zal plaatsvinden.

3.1 Smart homes, gebouwen, havens en steden

In huis krijgt het IoT gestaag vorm. Tegenwoordig heeft ieder huis een eigen router met draadloos internet. Smart home gadgets bestonden al veel langer, maar nu krijgen we gadgets die aan het internet gekoppeld worden. Gadget klinkt nog als een speeltje, maar steeds vaker zijn het apparaten met een serieuze functie. De zelfdenkende programmeer-bare thermostaat bestaat al lang, maar sinds kort kennen we Nest en Toon. Die houden contact met de buitenwereld en via onze smartphone kunnen we ze op afstand uitlezen en zelfs bedienen.

Onze televisie is gekoppeld aan het internet en houdt bij welke programma's er te zien zijn en waar we naar gekeken hebben. Het alarmsysteem waarschuwt ons via een app en met camera's kunnen we op afstand zien wat er in huis gebeurt. Dit zijn de meest normale dingen met internet in ons moderne huis, maar voor de early adopters zijn er talloze apparaten te koop die via het internet uitgelezen en bediend kunnen worden.

Terwijl wij in de privésfeer thuis rustig experimenteren, wordt het IoT ook op grotere schaal zichtbaarder. Complete kantoorgebouwen worden bestuurd door online systemen: toegangssystemen, liften, luchtbehandeling, alarmsystemen, bewakingscamera's enzovoorts. Moderne kantoren met flexplekken geven aan welke werkplekken nog vrij zijn. Printen gaat via een follow me-systeem, waarbij een printopdracht op iedere willekeurige printer in het kantoor opgehaald en geprint kan worden.



Figuur 9. Illustratie van een smart home[18].

Cybercriminaliteit werd vooral gezien als criminaliteit die zich in cyberspace afspeelt. Naarmate onze samenleving verder digitaliseert, zien we echter steeds vaker dat cybercriminaliteit wordt ingezet om traditionele criminaliteit te plegen. Criminelen krijgen door dat het veel eenvoudiger, goedkoper en minder gevaarlijk is om hun misdaden te plegen met behulp van digitale hulpmiddelen.

In juli van dit jaar werd Nederland opgeschrikt door het nieuws dat drugscriminelen in de haven van Rotterdam opereren met behulp van cybercriminelen. Eind vorig jaar sloegen bedrijven ook al alarm toen bleek dat de douane als gevolg van computerstoringen in de Rotterdamse haven containers zonder controle doorlaat. Dat geeft te denken over de virusuitbraak die een aantal containerterminals in Rotterdam eind juni dit jaar heeft getroffen. Was dat een (mislukte) poging om losgeld te innen of was de echte reden dat men de controle wilde saboteren?

[18] Overgenomen van http://smarthomepensacola.com/

Dat lijkt misschien ver gezocht, maar het kan nog gekker. Een paar jaar geleden, in 2013, bleek al dat de Haven van Antwerpen zo lek was als een mandje en dat criminelen containers met drugsladingen eenvoudig konden ophalen door de computers van het havenbedrijf te hacken. Zelfs toen de gaten in de firewalls waren gedicht, hebben de criminelen spionage-apparatuur aangebracht in de havenkantoren en konden ze hun criminele activiteiten voortzetten. Het mag duidelijk zijn dat criminele organisaties in toenemende mate digitaliseren om hun criminele activiteiten te ondersteunen.

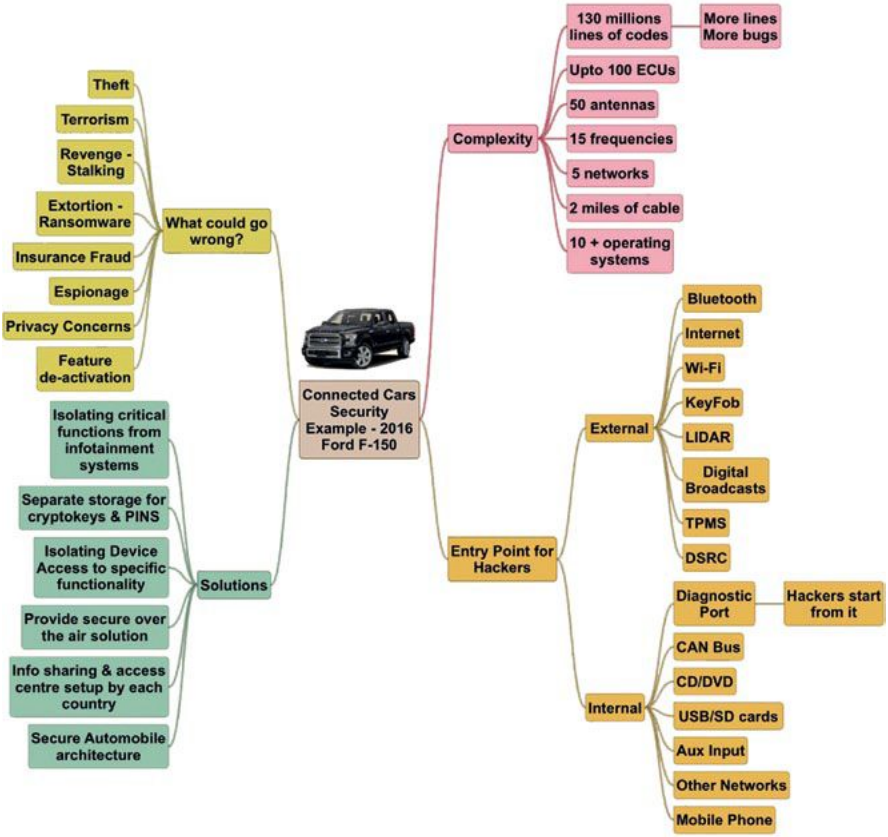
Vijf jaar geleden in Antwerpen moesten criminelen nog hun eigen apparatuur binnensmokkelen om systemen te manipuleren. Door in te breken op het IoT kunnen criminelen ongemerkt meekijken en gegevensstromen in de haven manipuleren.

3.2 Connected cars

Nog voor dat de smartphone zijn intrede deed, was de auto-industrie al bezig met het idee dat ieder vervoersmiddel uitgerust zou moeten worden met een computer. Navigatiesystemen hebben de afgelopen tien jaar een enorme ontwikkeling doorgemaakt en inmiddels verbazen we ons niet meer over de technische hoogstandjes die in onze auto's zijn verwerkt.

Nieuw is dat onze auto's ook verbonden worden met het internet. Tesla werkt de software in de auto bij zonder dat de eigenaar daaraan te pas komt en verzamelt constant informatie. Autobezitters kunnen Tesla verzoeken om deze functie uit te schakelen, maar dat betekent ook dat je bepaalde waarschuwingen van Tesla niet meer zult ontvangen. In de vliegtuigindustrie is het al gewoon dat fabrikanten tijdens de vlucht alles monitoren zodat voorspeld kan worden wanneer een onderdeel vervangen moet worden, voordat het kapot gaat. Deze technieken zullen ook tot de auto doordringen.

Fabrikanten bedenken nieuwe toepassingen voor de bestaande sensoren. Zo worden steeds meer auto's uitgerust met camera's, die bijvoorbeeld ook gebruikt kunnen worden om vrije parkeerplaatsen in de stad te signaleren. Die informatie kan gedeeld worden tussen auto's zodat ook anderen die op zoek zijn naar een parkeerplaats van die informatie profiteren. Maar uiteindelijk zullen diezelfde auto's ook zelfstandig kunnen rijden en een parkeerplaats buiten de stad kunnen opzoeken.



Figuur 10. Mindmap over de beveiliging van de connected car[19].

Bovenstaande figuur schetst de verschillende aspecten die een rol spelen bij de beveiliging van de connected car en geeft daarmee ook gelijk inzicht in de verschillende onderdelen en complexiteit die een rol spelen bij het zoeken naar digitaal bewijs.

[19] Overgenomen van Connected Car Security uit learnings van RSA Singapore 2017 van @mayanklau op Twitter, zie https://twitter.com/mayanklau/status/890593885461794816

3.3 Quantified self

Als we vliegtuigen en auto's kunnen monitoren en vroegtijdig reparaties kunnen uitvoeren, kunnen we dat dan ook niet bij mensen? Dat ligt natuurlijk erg gevoelig, maar eigenlijk is het antwoord volmondig ja. Alleen zijn er (nog) geen fabrikanten die ons kunnen dwingen om onszelf continu te laten monitoren. Medische data is een zeer gevoelig onderwerp en het automatisch uitwisselen van informatie over onze gesteldheid stuit vooralsnog op verzet.

Wat nu als de komende jaren het probleem van privacy wordt opgelost en mensen, net als apparaten, ook preventief gemonitord worden? Wat nu als blijkt dat met deze maatregelen wij ons leven significant kunnen verlengen? Het zou waarschijnlijk veel geld en ellende besparen en ongetwijfeld zouden we er veel van leren. Dit lijkt misschien ver weg, maar op dit moment zijn er talloze initiatieven gaande onder de noemer die bekend staat als de *quantified self*[20].

Wie heeft er tegenwoordig niet een stappenteller op zijn mobiele telefoon? Wie meet zijn hartslag en voert deze in de telefoon of legt zijn telefoon naast zijn bed om zijn slaapritme te controleren? Er zijn genoeg mensen die dit doen en de apps maken het mogelijk om je prestaties te vergelijken met andere gebruikers. Weleens van een slim bed gehoord? De sensoren in het matras meten hoe goed je hebt geslapen. En als je wakker wordt zet je bed automatisch je slimme koffiezetapparaat aan.

Onderstaande figuur dateert uit 2014 en geeft een uitgebreid overzicht van allerlei vormen van draagbare (wearable) technologie. De informatie die deze apparaten vastleggen zijn niet alleen bruikbaar vanuit een quantified self-perspectief, maar kan ook dienen als bewijs. In september 2016 vond de Officier van Justitie in de zaak van de moord op Koen Everink dat de gegevens van de stappenteller van de verdachte niet overeenkomen met het verhaal van diezelfde verdachte[21].

Op het eerste gezicht lijkt het niet onlogisch om gebruik te maken van deze informatie. Maar zonder gedegen onderzoek is het gevaarlijk om aan te nemen dat er een verband is.

[20] Zie o.a. The Quantified Self door M. Swan. Verschenen in Big Data, Vol 1, No. 2, Juni 2013. http://online.liebertpub.com/doi/pdfplus/10.1089/big.2012.0002
[21] Zie o.a. https://www.rtlnieuws.nl/nederland/om-stappenteller-mark-de-j-bewijst-dat-hij-liegt-in-zaak-everink
[22] Zie https://www.rtvutrecht.nl/nieuws/1607338/apple-vindt-stappenteller-geen-bewijs-in-moordzaak-koen-everink.html

In april van dit jaar betoogde de advocaat van de verdachte dat de stappenteller op de iPhone geen bruikbaar bewijsmiddel is volgens een verklaring van het technologiebedrijf Apple[22]. Apple stelt dat de 'stappen' ook geteld kunnen zijn door schokkende bewegingen in de auto. Uit een opmerking van de raadsman is af te leiden dat het NFI na onderzoek toch tot een andere conclusie is gekomen[23], alhoewel dat alleen op basis van deze bron moeilijk is vast te stellen.

Figuur 11. Afbeelding uit 2014 met een overzicht van wearable technologie.

Wel of geen bewijs, het mag duidelijk zijn dat dit soort draagbare sensoren als bewijs kunnen dienen als de Officier van Justitie of de advocaat van de verdachte van mening is dat het de zaak dient. Interessant is vervolgens om te zien dat deskundigen in de rechtszaal blijkbaar verschillende interpretaties geven aan het bewijs. Dit bevestigt nogmaals de noodzaak voor meer deskundigen die in staat zijn om digitaal bewijs te kunnen verzamelen en in de context te kunnen plaatsen.

[23] Zie https://www.rtvutrecht.nl/nieuws/1607338/apple-vindt-stappenteller-geen-bewijs-in-moordzaak-koen-everink.html

4. De opkomst van kunstmatige intelligentie

Alles om ons heen wordt slim. Auto's, huizen, gebouwen en steden worden verweven met het internet en vormen het IoT. Maar hoe slim is dat IoT nu eigenlijk? Wordt het slimmer dan wijzelf of worden we zelf ook slimmer?

Twintig jaar geleden wist IBM met Deep Blue Kasparov te verslaan. Toen werd voor iedereen duidelijk dat door heel veel te rekenen met een slim programma de computer in ieder geval slim genoeg is om goed te kunnen schaken. De laatste paar jaar gaan de ontwikkelingen erg hard en blijkt dat de computer slimmer is dan we dachten.

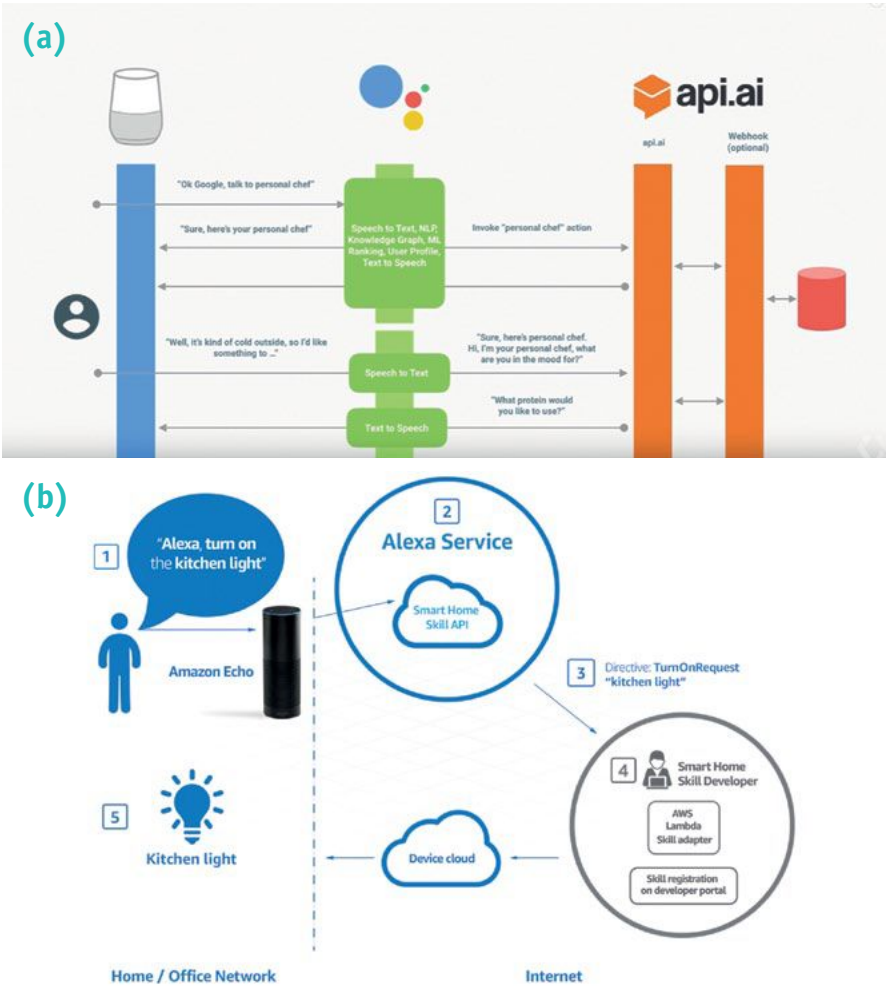
4.1 Kunstmatige intelligentie en natuurlijke taalverwerking

De algoritmen, het grote, ijzersterke geheugen met daarin alle mogelijke spelvariaties en de brute rekenkracht, en de beperktheid van de toepassing beloofden op dat moment nog weinig voor andere toepassingen. Maar IBM heeft niet stil gezeten en heeft Watson gebouwd. Nu niet met speciale, maar met standaard hardware. In 2010 heeft IBM met Watson de beste spelers in de VS van het televisiespel Jeopardy verslagen.

Na Jeopardy heeft Watson inmiddels vele toepassingen, zoals het verbeteren van medische diagnoses. Na een jaar training is Watson in staat om symptomen van een patiënt te vergelijken met 600.000 medische rapporten, 2 miljoen pagina's uit 42 medische tijdschriften en 1,5 miljoen patiëntverslagen en klinische proeven. Op dit moment is Watson een aparte business unit van IBM met omstreeks 10.000 werknemers. Watson wordt nu niet alleen ingezet om artsen te helpen bij de analyse van röntgen- en MRI-beelden, maar ook bij genomics-onderzoek in de bestrijding tegen kanker.

In tegenstelling tot Deep Blue spreekt Watson wel tot de verbeelding van mensen als een kunstmatige, intelligente computer. De zogenaamde Deep Q&A-technieken komen steeds meer terug in andere toepassingen op internet, die steeds beter weten waar we naar op zoek zijn zodat ze gerichtere antwoorden geven. Chatbots hebben de afgelopen twee jaar een enorme ontwikkeling doorgemaakt. Alle grote spelers zoals IBM, Micsosoft, Facebook, Amazon en Google zijn met platformen gekomen om snel chatbots te ontwikkelen waarbij ontwikkelaars gebruik kunnen maken van geavanceerde technieken.

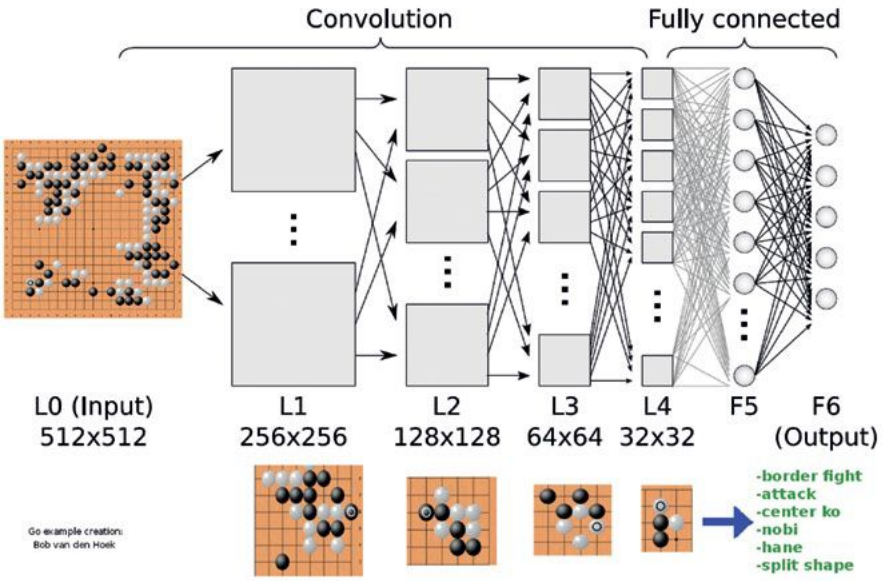
Onderstaande figuur illustreert twee verschillende architecturen. Figuur (a) is van het Google Api.Ai platform en figuur (b) is van de Amazon Alexa architectuur. Aan het einde van dit hoofdstuk gaan we dieper in op de spraakinterfaces van deze chatbots, maar ook zonder spraakherkenning bieden ze veel geavanceerdere en vooral meer gebruikersvriendelijke dialoogsystemen dan we tot voor kort kenden.



Figuur 12. Chatbot-ontwikkeling in (a) het API.AI platform van Google, (b) het Amazon Alexa framework.

4.2 De computer krijgt gevoel

Vorig jaar wist DeepMind (een dochterbedrijf van Google) met het programma AlphaGo geheel onverwacht de wereldkampioen Go te verslaan. Experts hadden voorspeld (op basis van de bekende technieken) dat het nog zeker tien jaar zou duren voordat kunstmatige intelligentie tot zoiets in staat zou zijn. Begin dit jaar is de wedstrijd nogmaals gespeeld. In 2016 wist de menselijke kampioen nog wel een partij te winnen, maar dit jaar was hij compleet kansloos[24]. Leek de computer in 2016 nog enigszins menselijk, in 2017 leek de computer meer op een Go-god.



Figuur 13. Vereenvoudigd overzicht van het deep neural network dat is toegepast in de architectuur van AlphaGo[25].

[24] Zie https://www.nytimes.com/2017/05/23/business/google-deepmind-alphago-go-champion-defeat.htm

[25] Overgenomen van http://deeplearningskysthelimit.blogspot.nl/2016/04/part-2-alphago-under-magnifying-glass.html

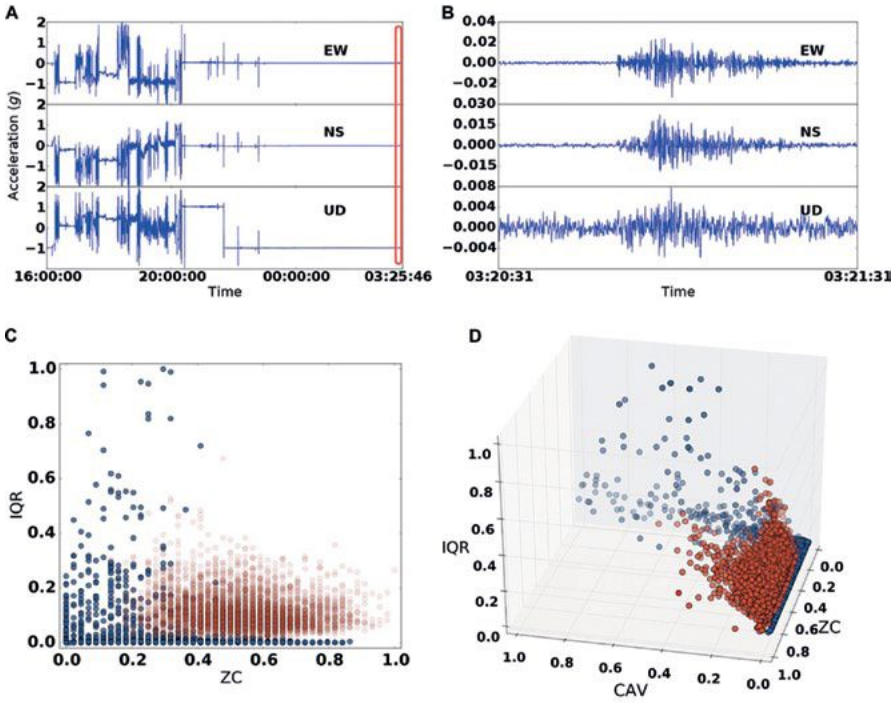
Het meest bijzondere van deze overwinning in vergelijking met voorgaande spel-spelende computers is dat de computer zelf Go heeft leren spelen. Voor Deep Blue was in 1997 nog speciale hardware gebouwd en werd gebruik gemaakt van de meest geavanceerde schaakalgoritmes en openingsbibliotheken. Nu is het blijkbaar voldoende om de computer eerst te trainen op voorbeeldpartijen van Go-grootmeesters en vervolgens te verbeteren door hem tegen zichzelf te laten spelen.

De successen met deep learning werden tot nu toe nog geboekt met data die door mensen op het internet is gezet. Denk aan de foto's op Facebook, teksten en vertalingen op Wikipedia en de kennis die is vastgelegd in het semantische web. Dankzij het IoT kunnen computers nu rechtstreeks data uit sensoren ontvangen. Met het IoT krijgt de computer eindelijk gevoel in de meest letterlijke zin van de betekenis. Dit gevoel, in combinatie met deep learning en internet, gaat veel verder dan het automatiseren van alledaagse activiteiten. Automatische systemen zullen steeds meer autonomie krijgen, waarvan de zelfrijdende auto misschien wel het beste voorbeeld is.

DeepSense[26] is een deep learning framework dat werkt op mobiele telefoons en dat gebruikt kan worden om data van mobiele sensoren te classificeren. Denk bijvoorbeeld aan het herkennen van de activiteit die de bezitter van de smartphone op een bepaald moment uitvoert (lopen, staan, fietsen enzovoorts) gebaseerd op bewegingssensoren. Een andere toepassing van deze technologie is de identificatie van personen gebaseerd op hun eigen manier van lopen.

Onderstaande figuur illustreert hoe een smartphone gebruikt kan worden om een aardbeving te detecteren. Eén smartphone is misschien niet nauwkeurig, maar als grote aantallen gebruikers dezelfde app gebruiken, dan kunnen al die smartphones gezamenlijk vanwege hun spreiding heel accuraat niet alleen de sterkte van een aardbeving ter plaatse, maar ook het centrum van de aardbeving bepalen.

[26] Blog "MyShake: A smartphone seismic network for earthquake early warning and beyond" op https://blog.acolyer.org/2017/05/17/deepsense-a-unified-deep-learning-framework-for-time-series-mobile-sensing-data-processing/



Figuur 14. Aardbeving waargenomen door een smartphone: (a) tijdspanne van 12 uur met beweging in de X, Y en Z as, (b) uitvergroting van de rode rechthoek in a op het moment van een aardbeving, (c) en (d) alledaagse bewegingen hebben grote amplitude en lage frequentie (blauwe stippen) en de aardbeving veroorzaakt hogere frequenties met een meer gematigde amplitude[27].

De voorspelde groei in omvang en fijnmazigheid van het IoT lijkt misschien een hype. Maar schijn bedriegt. Smartdust was vroeger een term uit sciencefiction waarbij een wolk van microscopische machines voor allerlei taken kon worden ingezet. Inmiddels bestaan er echte smartdust-toepassingen[28]. De energievoorziening van dat soort systemen was een probleem, maar zonnecellen worden steeds sterker en kleine en ook andere innovaties helpen. Zo kwam het Delftse bedrijf NOWI met een innovatie waarmee sensoren in hun stroombehoefte voorzien door energie af te tappen uit standaard aanwezige Wifi-signalen.

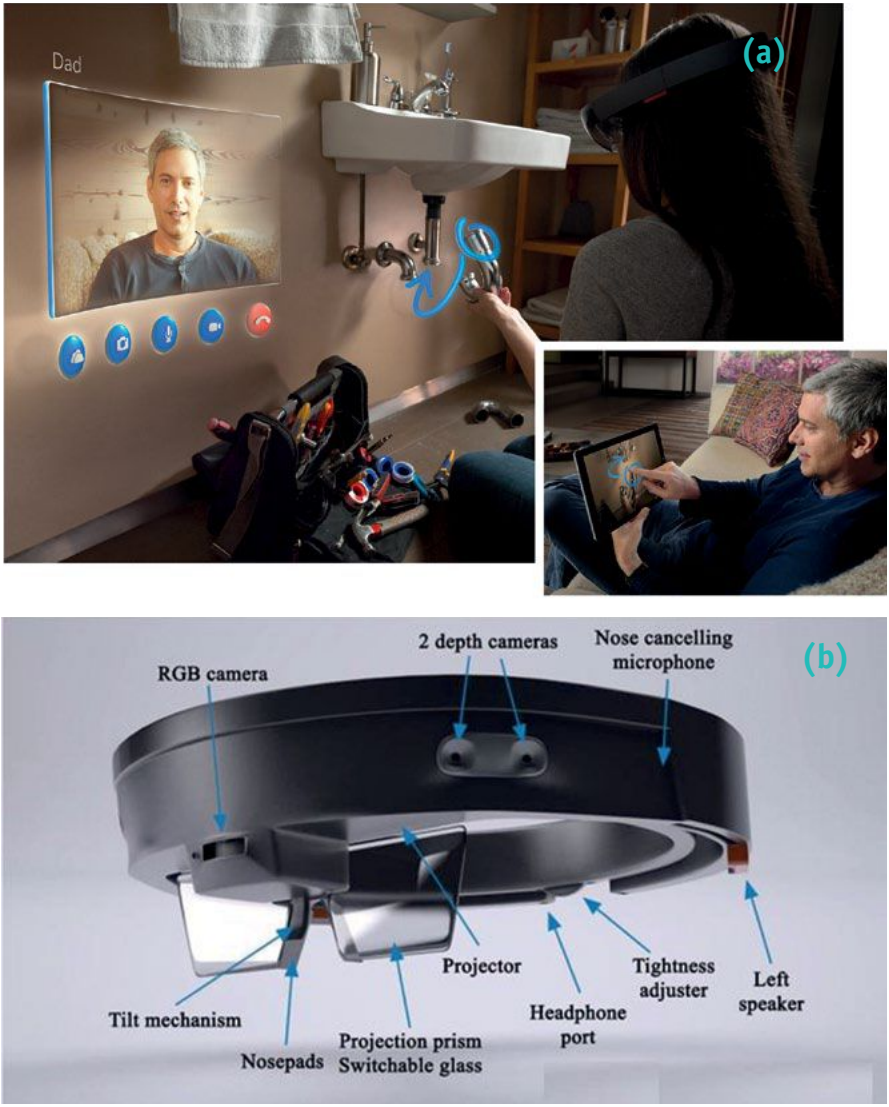
[27] Overgenomen van http://advances.sciencemag.org/content/2/2/e1501055.full
[28] Zie bijvoorbeeld http://computerworld.nl/markttrends/94263-wat-is-smartdust

Dat betekent dat het onderhoud van zulke sensoren flink goedkoper wordt, waardoor business cases om massaal online sensoren te installeren interessant worden. Zulke sensoren hoeven niet slim te zijn als het er maar veel zijn (zoals in het voorbeeld van de smartphones die aardbevingen kunnen detecteren). Zodra de big data die dan ontstaat met deep learning-algoritmen geanalyseerd worden, kunnen patronen herkend worden en verbanden worden gelegd.

4.3 Versmelting fysieke wereld en cyberspace

Onze fysieke wereld en cyberspace zijn geleidelijk aan het samensmelten. Nu al door spraak-gestuurde assistenten die onopvallend met ons meeluisteren en vragen kunnen beantwoorden. Augmented reality (AR)-brillen voegen daar nog een visuele component aan toe door hologrammen om ons heen te projecteren en gebaren te herkennen.

Layar is een Nederlands bedrijf dat is opgericht in 2009. Het bedrijf lanceerde een browser voor mobiele platformen waarin AR werd toegepast waarmee informatie aan de fysieke wereld kon worden toegevoegd. Die techniek wordt overigens al veel langer toegepast bij Defensie, bijvoorbeeld in de heads-up displays van straaljagerpiloten. Pas in 2015 kwam Microsoft als eerste grote bedrijf met een AR-bril gebaseerd op Windows 10, zonder de noodzaak om de bril met een PC te verbinden. Dit jaar volgden Facebook AR, Apple ARKit en Google ARCore met software waarmee onze smartphones bruikbaar worden als AR devices. In de tussentijd zijn er nog talloze andere AR-initiatieven van de grond gekomen van startups die ook werken aan vergelijkbare hardware.



Figuur 15. Augmented reality: (a) Dochter krijgt van haar vader instructies via de hololens hoe ze de afvoer van de wastafel moet aansluiten, (b) samenstelling van de Hololens augmented reality-bril van Microsoft[29].

[29] Overgenomen uit http://technewsexpres.blogspot.nl/2016/03/microsoft-hololens.html

Zero learning-curve user interfaces

AR-deskundigen[30] voorspellen het einde van beeldschermen en smartphones zoals we die kennen. Belangrijke reden daarvoor is dat tot nu de computers hun gebruikers altijd dwingen om via een abstract model bepaalde functies te kunnen bedienen. Dit voert terug naar het zogenaamde WIMP-model[31]. De letters WIMP staan voor Windows Icons Menus en Pointer. Dit model voor user-interactie is al in 1973 in het befaamde Xerox PARC in Palo Alto ontwikkeld en staat nog steeds model voor onze huidige Windows- en iOS-gebaseerde user interfaces.

Apple heeft met iOS en de iPad de bediening van de computer een stuk eenvoudiger gemaakt. Er waren wel eerdere pogingen, maar door de nog gebrekkige hardware faalde bijvoorbeeld de Windows Tablet PC die Microsoft in 2002 probeerde te lanceren. Met de huidige tablets kunnen ouderen zonder enige computerervaring eenvoudig door de foto's van hun kleinkinderen bladeren of met hun vingers een spelletje patience spelen. Maar op het moment dat ze het scherm wat lichter of donkerder moeten instellen, verdwalen ze in menu's en submenu's.

Met AR, touch interfaces en spraak-gestuurde assistenten kunnen we voorgoed afscheid nemen van het toetsenboard en ingewikkelde menustructuren. De computer leert communiceren met ons via onze zintuigen. We hoeven niet meer een cursus te volgen, een handleiding te lezen of een YouTube-filmpje te bekijken om te leren hoe we de computer of een apparaat moeten gebruiken. Met kunstmatige intelligentie zal het apparaat zelf begrijpen wat we proberen te doen, of het ons vragen om ons vervolgens te helpen dat te bereiken.

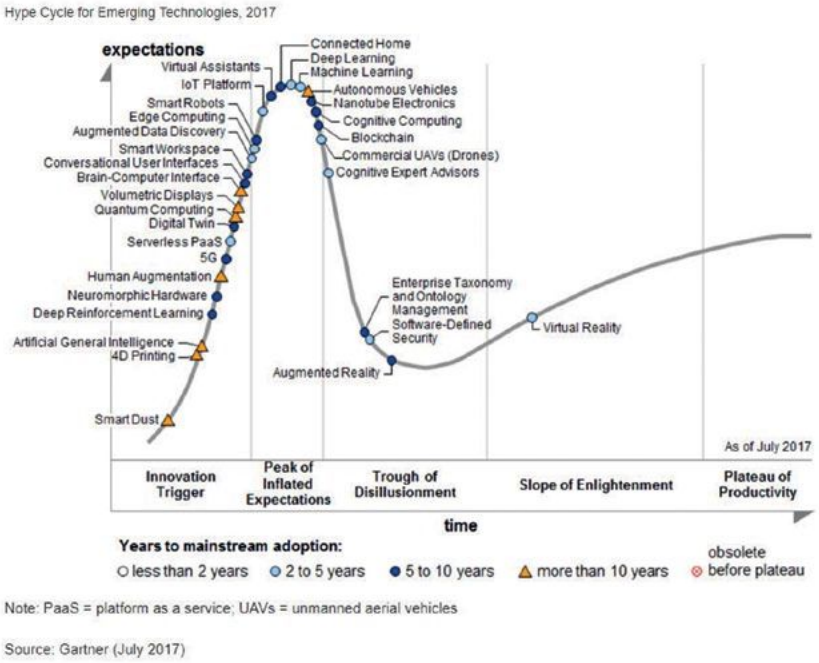
Head sockets: hoezo maar vijf zintuigen?

Met behulp van augmented reality, haptische interfaces en geluid kunnen we extra informatie toevoegen aan ten minste drie van onze vijf zintuigen, namelijk zien, horen en voelen. Met behulp van de computer kunnen we daar op allerlei manieren informatie aan toevoegen die onze zintuigen normaal gesproken niet kunnen waarnemen. Bijvoorbeeld het zichtbaar maken van onhoorbaar geluid of onzichtbaar licht. Of een sensor op onze huid die gaat trillen als die naar het noorden wijst. Ook het labelen van personen en gebouwen die zich in ons gezichtsveld bevinden is in feite een verrijking van onze normale zintuigen.

[30] Zie bijvoorbeeld het interview met Alex Kipman van Microsoft in een interview met Business Insider in maart http://www.businessinsider.com/microsoft-alex-kipman-hololens-2017-2 of een verslag van de keynote van Mark Zuckerberg op de jaarlijkse Facebook F8 conferentie https://www.businessinsider.nl/facebook-f8-mark-zuckerberg-on-augmented-reality-2017-4

[31] https://en.wikipedia.org/wiki/WIMP_(computing)

Maar het blijft niet bij “augmentatie” van onze bestaande zintuigen. Elon Musk is vastberaden om met het bedrijfje Neuralink een interface te bouwen die het menselijke brein in staat zal stellen om nog veel sneller informatie met het IoT uit te wisselen en uiteindelijk ook de communicatie tussen mensen onderling te verbeteren. Volgens Musk is een nieuwe manier van communiceren tussen mens en machine noodzakelijk om het menselijke brein in staat te stellen om bij te blijven met de snel verbeterende kunstmatige intelligentie. Musk is niet de enige overigens. Facebook presenteerde op de F8- conferentie dit jaar het typen van woorden door te denken en het horen van geluid via de huid.



Figuur 16. Gartner hype cycle for emerging technologies, 2017.

Bovenstaande figuur toont de hype cycle voor emerging technologies van Gartner in 2017. Volgens Gartner moeten we nog meer dan tien jaar wachten op brain-computer interfaces, maar zal virtual reality binnen twee tot vijf jaar en augmented reality binnen vijf tot tien jaar volwassen worden. Interessant is overigens om te zien dat volgens deze hype cycle in 2016 virtual reality ook nog vijf tot tien jaar nodig zou hebben, maar blijkbaar zijn de ontwikkelingen zo hard gegaan dat deze verwachting is bijgesteld. Als bedrijven zoals van Elon Musk en Facebook serieus werk maken van brain-computer interfaces, dan zou Gartner in 2018 misschien ook wel de verwachting op dit punt kunnen bijstellen.

5. Een revolutie

De titel van deze lectorale rede is “De (r)evolutie van digitaal bewijs” en in de vorige twee hoofdstukken heb ik de ontwikkeling van het IoT, de opkomst van kunstmatige intelligentie en de versmelting van de fysieke wereld en cyberspace beschreven. Dat digitaal bewijs evalueert is evident, maar waarom zou er sprake zijn van een revolutie?

5.1 Niet gewoon méér data maar big data

De explosieve groei van digitale informatie kwam in de inleiding al aan bod. Dat er meer data is, is op zich geen reden voor een evolutie. Het feit dat deze data op steeds meer verschillende plaatsen wordt gemaakt en verwerkt is wel bijzonder. Welbeschouwd had digitaal forensisch onderzoek in pakweg de periode 1990-2000 vooral te maken met gegevens op pc’s, elektronische zakagenda’s, diskettes, cd’s, computertapes en de fax.

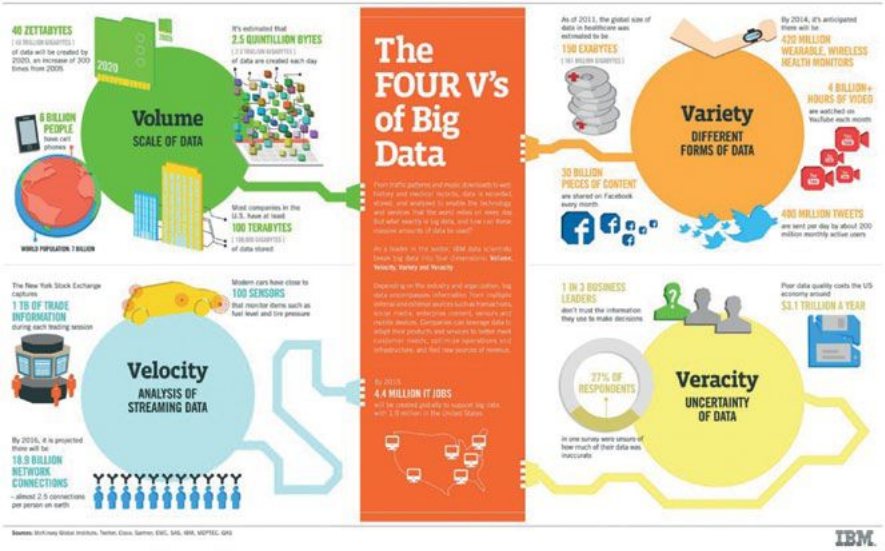
Aan het einde van de jaren negentig ontstond er een internet bubble[32] die in 2000 uit elkaar zou spatten. Het geloof in de onbegrensde mogelijkheden van het internet had even een schok gekregen en veel dot com bedrijven hebben de crash niet overleefd. Maar de economische recessie die daarop volgde duurde vrij kort en de technologische ontwikkelingen kwamen snel op gang. De digitale camera’s voor consumenten werden steeds beter. De cd werd opgevolgd door de dvd en de diskette werd vervangen door usb-sticks en harde schijven met een usb-koppeling. Naast pc’s kwamen er servers. Eerst fileservers, toen e-mailservers en daarna databaseservers, webservers en steeds meer applicatieservers. Nokia kwam met nieuwe telefoons die meer functionaliteit kregen. De Blackberry, Palm Pilot en Psion organizers werden populairder en verdrongen de elektronische zakagenda’s van Sharp en Casio.

De introductie van de iPhone in 2007 en even later van de iPad markeerde het begin van het samengaan van computer en telefoon. Onze smartphone kan tegenwoordig bijna net zoveel als onze computer, alleen werkt het wat onhandiger en is het scherm een stuk kleiner. Veel mensen bezitten een smartphone en hebben geen pc meer of zelfs nooit gehad. Het virtuele toetsenbord op onze telefoon evolueert snel en maakt het mogelijk om berichten te verrijken met emoji’s of woorden te swipen in plaats van te typen. Gebruikers die communiceren in een taal met ingewikkelde tekens zijn al helemaal overgestapt op spraakherkenning.

[32] Zie bijvoorbeeld https://en.wikipedia.org/wiki/Dot-com_bubble

Niet alleen de telefoon heeft zich in die tijd snel ontwikkeld. Ook de auto's kwamen in de periode 2000-2010 steeds voller met elektronica te zitten. Inmiddels beschikt iedere nieuwe auto over elektronische navigatie. Daarnaast zijn allerlei zaken in de auto vervangen door de computer. De snelheidsmeter, klimaatcontrole, motormanagement, start/stop, open/dicht: functies die we al kenden maar die nu door een computer bestuurd worden.

Data die zo anders is, kan niet meer op de traditionele manier opgeslagen, verwerkt en gepresenteerd worden. De manier waarop we deze data moeten gaan analyseren en presenteren zal fundamenteel anders worden. De term die iedereen inmiddels kent is big data. De forensische kenmerken van big data en big data-systemen zijn zo anders dat we op een heel nieuwe manier ervoor zullen moeten zorgen dat onderzoekers de gegevens in hun context kunnen plaatsen.



Figuur 17. De vier V's van big data: Volume, Variety, Velocity en Veracity[33].

[33] Overgenomen uit http://www.ibmbigdatahub.com/infographic/four-vs-big-data

Bovenstaande figuur illustreert de vier V's van big data: volume, variety, velocity en veracity:

- *Volume* geeft aan dat het om grote hoeveelheden gaat. Zoveel dat opslag op een enkel systeem niet meer mogelijk is, waardoor het traditionele paradigma waarin data naar de cpu brengen niet meer werkt. Bij big data wordt de cpu naar de data gebracht met bijvoorbeeld het MapReduce-algoritme.
- *Variety* verwijst naar de variëteit in de data. Tot voor kort werden gestructureerde data in relationele databases opgeslagen met voorgedefinieerde recordschema's die beheerd worden door Relational Database Management Systemen. Met big data is dit echter niet meer vol te houden. Middels NoSQL-databases is het mogelijk om data zonder vooraf gedefinieerde records-structuur op te slaan en terug te zoeken.
- *Velocity* verwijst naar de snelheid waarmee de data verwerkt moet worden. Sommige big data, zeker ook uit het IoT, komt in zulke grote hoeveelheden en in real-time beschikbaar dat deze streaming verwerkt moet worden.
- *Veracity* ten slotte verwijst naar de betrouwbaarheid van de data. Met zoveel data die ook nog gedistribueerd (en soms redundant) is opgeslagen, is het essentieel om de betrouwbaarheid van die data te kunnen garanderen.

5.2 Disruptie maakt alles anders

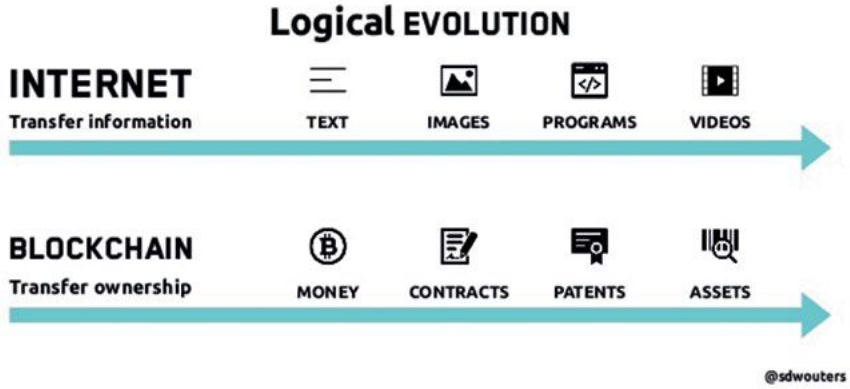
Tot zover dus gewoon (heel veel) meer data en meer soorten, wat nieuwe eisen stelt aan de manier waar we data opslaan, verwerken en presenteren. Op zichzelf heeft big data al voor een revolutie gezorgd, maar er is nog meer aan de hand. We krijgen niet alleen meer data, maar we gaan ook steeds anders met elektronische data om.

In de jaren tachtig werd de personal computer geïntroduceerd, terwijl grote bedrijven gebruik maakten van mainframes. Het forensisch computeronderzoek in het begin van de jaren negentig was sterk gericht op de personal computer en alle opslagmedia die men kon tegenkomen. In sommige gevallen moesten digitale gegevens uit een geautomatiseerde administratie worden veiliggesteld. Het ging vooral om data die op één plaats werd gemaakt, verwerkt en bewaard.

In de loop van de jaren negentig werd e-mail geïntroduceerd op de werkvloer en kwam het internet op. Binnen de academische wereld was het al gewoon om met collega's op andere universiteiten te communiceren. Tussen bedrijven onderling is nog lange tijd via fax gecommuniceerd, terwijl internet e-mailsystemen snel aan populariteit wonnen. E-mail nam de functie van de fax over en is vervolgens ook vanaf de jaren 2000 een grote rol gaan spelen in E-Discovery-onderzoeken binnen bedrijven.

Als we nu terugkijken, kunnen we vaststellen dat we vooral traditionele taken zijn gaan automatiseren. Dat wil zeggen, de werkwijze veranderde niet echt, maar met behulp van de computer konden we dingen sneller en op veel grotere schaal uitvoeren. Met de komst van de smartphone en de ontwikkeling van nieuwe diensten en apps wordt de manier waarop we werken en communiceren nu wel getransformeerd. AirBnB, NetFlix, Ebay, Alibaba, KickStarter, Spotify, Amazon en Uber zijn daar goede voorbeelden van. Zij zijn er ingeslaagd om een business model te introduceren waarbij zij de infrastructuur aanbieden maar waarbij ze zelf niet over content beschikken. AirBnB heeft geen hotels. NetFlix heeft geen films, Uber heeft geen taxi's enzovoorts. De algemene verwachting is dat we nu nog maar aan het begin staan van wat nu al de "sharing economy" wordt genoemd.

Blockchain (zie figuur 18) is een voorbeeld van een disruptieve technologie. Het principe van de blockchain is bekend geworden in de periode 2010-2015 door de introductie van de Bitcoin, een digitaal betaalmiddel. Vanaf 2014 zijn er ook andere toepassingen bedacht voor de blockchain. Denk daarbij aan slimme contracten, slim onroerend goed, registraties enz. Sinds 2015 groeit het aantal toepassingen van blockchain explosief. Bijzonder daarbij is dat blockchain-toepassingen ontstaan in compleet verschillende markten. Na de financiële sector gaan nu ook de gezondheidssector, telecommunicatie, reisindustrie, overheid, het IoT en de maakindustrie aan de slag met blockchain.



Figuur 18. Vergelijking tussen het internet en blockchain, waarbij internet gezien wordt als techniek om informatie uit te wisselen en blockchain als techniek om eigendom uit te wisselen. In beide gevallen zonder tussenkomst van een derde partij[34].

[34] Overgenomen uit https://www.slideshare.net/SamWouters/why-how-what-the-blockchain

Bovenstaande figuur illustreert dat blockchain op dit moment in het stadium verkeert waarin het internet zich bevond in 1992. Het zag er niet mooi uit, bedrijven hadden het nog niet omarmd en werkten zelfs tegen als ze de indruk hadden dat het een bedreiging voor hun businessmodel was. Dat betekent niet dat het weer twintig jaar gaat duren voordat blockchain eenzelfde ontwikkeling heeft doorgemaakt als het internet. Zeer waarschijnlijk gaat dat een stuk sneller.

5.3 Augmented intelligence

Tot nu toe heb ik uitgelegd dat digitaal bewijs zal veranderen doordat de digitale informatie verandert met de komst van nieuwe apparatuur en nieuwe toepassingen. Daarbij komt dat we tot nu toe computers en smartphones hebben gebruikt om onze traditionele werkzaamheden te automatiseren. We gaan nu een periode tegemoet waarin we ons werk op een heel andere manier zullen gaan uitvoeren en onze werkzaamheden zullen veranderen. Daardoor zullen we opnieuw moeten bedenken hoe we digitaal bewijs kunnen vinden en hoe we dat nieuwe bewijs moeten analyseren.

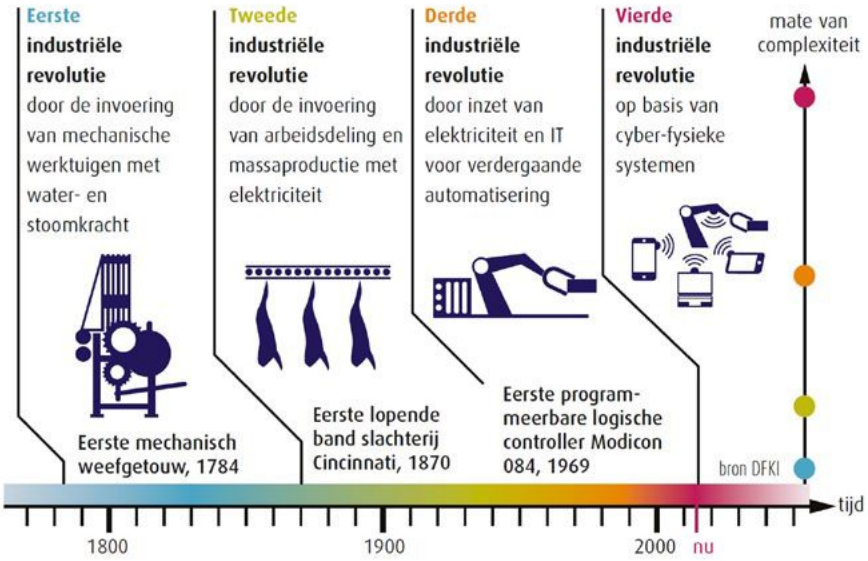
In het vorige hoofdstuk legde ik uit dat de computer door de combinatie van deep learning met het IoT gevoel krijgt. Computers kunnen daarmee nog niet zelf denken, maar we staan nu wel aan het begin van een tijdperk van niet-denkende computers die veel slimmer worden dan we tot voor kort voor mogelijk hielden. Uiteindelijk worden onze natuurlijke intelligentie en waarnemingsvermogen daardoor groter en is er sprake van *augmented intelligence*[35].

Het begrip augmented intelligence heb ik voor het eerst in 2013 bedacht toen ik een hoofdstuk moest schrijven voor het boek "De Informatiemaatschappij in 2023"[36]. Door sommigen wordt dit begrip (inmiddels) ook aangeduid *intelligence augmentation*[37]. Gartner heeft voor het eerst in 2017 in de hype cycle for emerging technologies (zie figuur 16) de term *Human Augmentation* (ook wel aangeduid als "Human 2.0") geïntroduceerd.

[35] Al in 1962 werd het onderwerp Augmenting Human Intellect uitgebreid besproken door Douglas Engelbert http://www.dougengelbart.org/pubs/augment-3906.html
[36] Hoofdstuk "Op zoek naar de digitale waarheid in 2023" gepubliceerd in het boek " De informatiemaatschappij in 2023" verschenen onder redactie van Dr. G.J. van Bussel.
[37] The Future isn't HAL – it's Intelligence Augmentation, op https://qz.com/921099/the-future-of-ai-isnt-hal-its-intelligence-augmentation/

Gartner verstaat daaronder de verbetering van cognitieve en fysieke capaciteiten van de mens. Het gaat daarbij dus niet alleen om augmented intelligence, maar ook om slimme protheses die betere prestaties leveren dan tot nu toe menselijk gezien mogelijk is. Human augmentation wordt door Gartner geschaard onder een van de drie toptrends in 2017, namelijk Transparently Immersive Experiences samen met onder andere virtual reality, augmented reality en de connected home[38].

Dankzij augmented intelligence wordt de interactie tussen mens en machine fundamenteel anders. Mensen hoeven niet langer meer in abstracties te kunnen denken. De computer zal sneller begrijpen wat de intentie van de gebruiker is en daarop anticiperen. Daardoor zullen meer mensen steeds meer uren per dag, bewust of onbewust met hun persoonlijke assistent, online zijn, waardoor zij moeilijk anoniem kunnen blijven.



Figuur 19. De vier industriële revoluties tot nu toe[39]

[38] http://www.gartner.com/smarterwithgartner/top-trends-in-the-gartner-hype-cycle-for-emerging-technologies-2017/
[39] https://www.deingenieur.nl/artikel/fabrieksrevolutie

5.4 Vierde industriële revolutie

Het IoT, de opkomst van kunstmatige intelligentie en uiteindelijk de versmelting van cyberspace en de fysieke wereld passen in het bredere perspectief van de vierde industriële revolutie. De vierde industriële revolutie bouwt voort op de derde revolutie, de digitale revolutie, en wordt gekenmerkt door samensmelting van verschillende technologieën waardoor de lijnen vervagen tussen de fysieke, digitale en biologische domeinen.

Tot voor kort lieten we sporen achter doordat we bewust gebruik maken van computer of smartphone. Als cyberspace en fysieke wereld samensmelten, is de computer niet meer een hulpmiddel maar wordt die een verlengstuk van onszelf. Hoe de vierde industriële revolutie er precies uit gaat zien, is niet duidelijk. Wel lijkt me duidelijk dat onze menselijke "footprint" steeds meer sporen zal nalaten in cyberspace, en digitaal bewijs zal in alle vormen van opsporing en forensisch onderzoek een cruciale rol gaan spelen.

Een voorbeeld van zulk bewijs is bijvoorbeeld een router waarin sporen zijn terug te vinden van apparaten die in het bezit zijn van een verdachte. Of de gegevens uit een stappenteller die in strijd zijn met de verklaring van de verdachte. We kunnen talloze voorbeelden bedenken van apparaten die mogelijk digitale sporen bevatten die bewijs kunnen leveren of die kunnen helpen bij de opsporing.

Door die versmelting van onze fysieke ruimte en cyberspace zal niet alleen het bewijs maar ook de misdaad zelf een andere vorm aannemen. Het begrip computervredebreuk is inmiddels redelijk bekend; dit is in feite een zeer gespecialiseerde vorm van computercriminaliteit. Computervredebreuk is ingewikkeld en vergt speciale kennis en ervaring. Maar wat te denken van vermogensdelicten, heling, vernieling, identiteitsfraude en diefstal?

Veel van deze misdrijven kunnen ook in cyberspace gepleegd worden door social engineering, chantage, kopiëren van een digitale sleutel enzovoorts. Criminelen zijn heel vindingrijk als het gaat om het vinden van zwakke plekken, met of zonder technische hulpmiddelen. Slachtoffers zullen aangifte doen en rechercheurs zullen in staat moeten zijn om aangiftes op te nemen en sporen veilig te stellen op een plaats delict in cyberspace.

Bestaande criminaliteit in de fysieke wereld zal in cyberspace nieuwe vormen aannemen. Digitale sporen die daarbij achterblijven en de analyse van die sporen zullen de revolutie van het digitale bewijs nog verder aanwakkeren.

6. Onderzoek van het lectoraat

Zodra mensen informatie zijn gaan verwerken met computers, is digitaal bewijs een rol gaan spelen in het forensisch onderzoek. Tot nu toe bestond dat bewijs vooral uit digitale gegevens afkomstig uit computers, verwisselbare opslagmedia en smartphones. Bij dit soort gegevens moet gedacht worden aan e-mails, documenten, multimedia-bestanden, communicatiegegevens, locatiegegevens en internetzoekopdrachten die aan de hand van datum en tijd gebruikt kunnen worden om bewijs te leveren.

In de voorgaande hoofdstukken heb ik uiteengezet hoe ik denk dat het IoT, de opkomst van kunstmatige intelligentie en uiteindelijke de versmelting van cyberspace en de fysieke wereld, ons leven zullen veranderen de komende vijf tot tien jaar. Het gevolg is ook dat digitaal bewijs een prominente rol zal gaan spelen in alle vormen van geregistreerde misdaad (zie figuur 1).

Tot nu toe ging het om digitale sporen die wij rechtstreeks achterlaten door bewust gedrag op momenten dat we van een computer of smartphone gebruik maken. Wat gaat er gebeuren naarmate cyberspace en fysieke wereld samensmelten en wij onbewust op veel grotere schaal dan voorheen digitale sporen achterlaten van al ons doen en laten? Wat heeft deze verandering voor invloed op het digitaal forensisch onderzoek?

Die vraag wil ik in dit hoofdstuk verder onderzoeken en tegelijk zal ik een onderzoeks-agenda opstellen voor het lectoraat, die anticipeert op nieuwe toepassingen van Forensische ICT die noodzakelijk zijn om digitaal bewijs te kunnen blijven onderzoeken. Voordat ik toekom aan die agenda zijn er twee onderwerpen die mede bepalend zijn voor het perspectief waarmee het onderzoek dat wordt geagendeerd, zal worden uitgevoerd.

6.1 Redenen van wetenschap

Voordat ik in ga op de technische (onderzoeks)uitdagingen, is het goed om eerst stil te staan bij het recherchewerk van de opsporingsambtenaar en de "redenen van wetenschap". Deze term hoorde ik voor het eerst in het "Digitale Evidence Dashboard"-project[40] waarin TNO, In-pact en Tracks Inspector samen onderzoek deden naar de wijze waarop opsporingsinstanties met digitaal bewijs zouden moeten omgaan. Vóór dat project wist ik niet dat dit een basisbegrip is in de opleiding van een (bijzondere) opsporingsambtenaar.

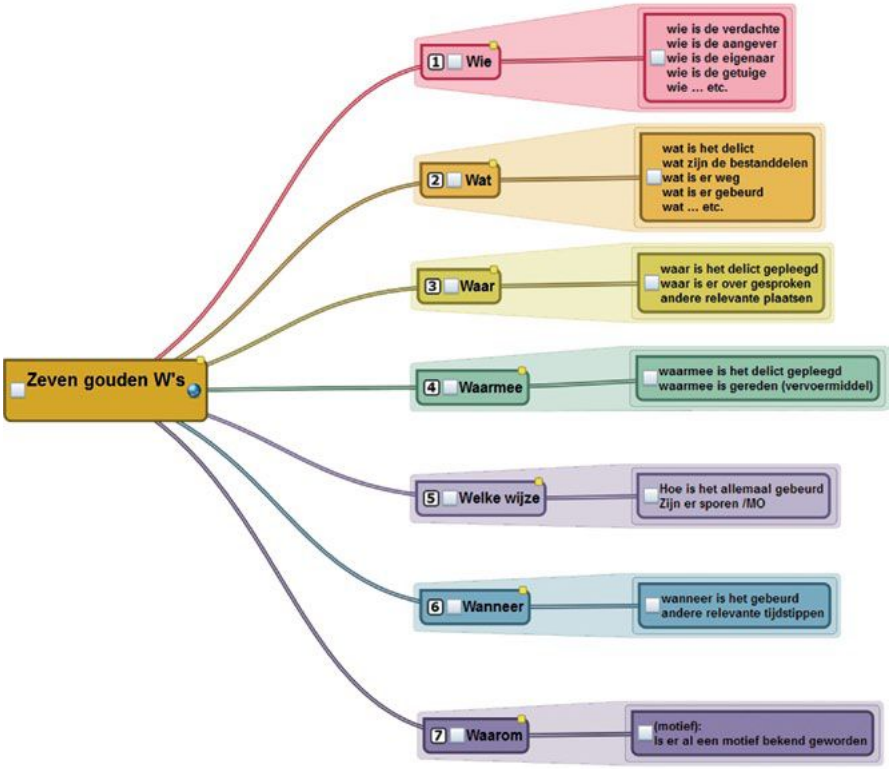
Volgens het document "Een goed PV, daar vang je boeven mee"[41] doet de opsporingsambtenaar verslag van de feiten en omstandigheden, door hem zelf waargenomen, en maakt hiervan een proces verbaal (PV) op. Belangrijk is te vermelden wat de opsporingsambtenaar heeft waargenomen. Wat heeft hij gezien, gehoord, geroken, geproefd, gevoeld? Wie heeft hem verteld wie de dader zou zijn of wat de dader gedaan zou hebben?

De opsporingsambtenaar wordt geacht in zijn PV te vermelden hoe hij aan zijn kennis is gekomen, ook wel "redenen van wetenschap" genoemd. Het gaat om de eigen waarneming van feiten en omstandigheden. Meningen, gissingen of conclusies zijn géén feiten en omstandigheden en vormen daarom geen bewijs.

De Officier van Justitie en de rechter verwachten dat iedere opsporingsambtenaar objectief de waarheid aan het licht brengt. De feiten waarnaar gezocht wordt, en daarbij de vragen die helpen bij de opsporing van een strafbaar feit, zijn de zeven gouden vragen, ook wel de zeven W-vragen genoemd:

1. Wie kan in verband worden gebracht met het misdrijf?
2. Wat is er gebeurd?
3. Waar is het misdrijf gepleegd en waar kunnen mogelijk sporen gevonden worden?
4. Met welke middelen is het misdrijf gepleegd?
5. Op welke wijze is het misdrijf gepleegd?
6. Wanneer is het misdrijf gepleegd?
7. Waarom is het misdrijf gepleegd?

[40] Zie http://digitalevidencedashboard.nl/
[41] Uitgave van het kwaliteitsprogramma Politie-OM Den Haag
https://www.politieacademie.nl/kennisenonderzoek/kennis/mediatheek/PDF/91730.PDF



Figuur 20. De zeven gouden W's[42].

Uit interviews met tactische rechercheurs in het eerdere genoemde DED- project (zie voetnoot 40) is gebleken dat zij worstelen met het formuleren van redenen van wetenschap als het gaat om feiten die uit digitaal bewijs volgen. Vermoedelijk wordt dit veroorzaakt door het abstractieniveau dat doorgaans nodig is om digitale sporen te begrijpen. Denk bijvoorbeeld aan een "user account", documenten die zijn geordend in een "directory structuur", de verdachte bevond zich op een ander "IP-adres op het internet", "de beveiliging van de firewall is doorbroken".

[42] Overgenomen uit http://vraaghetjevakdocent.blogspot.nl/p/het-maken-van-een-goed-rapport-of.html

In de voorgaande hoofdstukken heb ik uitgebreid betoogd dat door de versmelting van cyberspace en de fysieke wereld er meer en andersoortig digitaal bewijs zal komen. Misdrijven zullen zich ook steeds meer voor een deel in cyberspace gaan afspelen. Opsporingsambtenaren zullen dus moeten leren welke feiten ze kunnen vinden in digitaal bewijs en ze zullen ook moeten leren wat de redenen zijn van die wetenschap.

Gelukkig bieden ook bij digitaal bewijs de zeven W-vragen houvast. De wie-vraag kan vaak beantwoord worden door te achterhalen welke gebruiker schuil gaat achter een e-mailadres, user account of telefoonnummer. Datum en tijd van een bestand of spoor zeggen iets over wanneer iets is aangemaakt, gewijzigd of gezien. GPS-locaties of Wifi-namen kunnen iets over de locatie zeggen. De zoekgeschiedenis uit de browser kan inzicht geven in het motief, op welke wijze en met welke middelen een misdrijf is gepleegd.

Voor een digitaal expert is dit logisch, maar voor een opsporingsambtenaar zonder digitale expertise kan dit in de praktijk een (groot) obstakel zijn. Het is daarom van belang dat het lectoraat uitleg geeft aan het werkveld dat dit een probleem is en hulp aanbiedt hoe dergelijke vaardigheden bijgebracht kunnen worden.

Naast het schrijven van een goed PV is de interpretatie van zo'n PV en van rapportages door digitaal forensisch gerechtelijk deskundigen minstens zo belangrijk. Dat betekent dat niet alleen aandacht besteed moet worden aan de opleiding van opsporingsambtenaren, maar ook aan de gebruikers in de rechtsketen (rechters, Officieren van Justitie en advocaten) die meer te maken zullen krijgen met PV's en rapporten waarin feiten en conclusies over digitaal bewijs worden gepresenteerd.

Het Nederlands Register voor Gerechtelijk Deskundigen (NRGD)[43] is dé forensische kwaliteitsorganisatie in Nederland, die een constante kwaliteit van de inbreng van deskundigen in de rechtsgang waarborgt en bevordert. Het NRGD is opgericht naar aanleiding van de Wet deskundige in strafzaken, die in werking is getreden in 2010 en die eisen stelt aan de kwaliteit, betrouwbaarheid en bekwaamheid van deskundigen.

Aan het hoofd van het NRGD staat het onafhankelijke College gerechtelijk deskundigen (College[44]), waar ik sinds 2014 lid van ben. Inmiddels zijn er negen deskundigheids-gebieden geregistreerd[45].

[43] Zie de website van het NRGD op https://www.nrgd.nl
[44] Het College van het NRGD https://www.nrgd.nl/over-het-nrgd/organisatie/college/index.aspx
[45] Zie het jaarverslag NRGD 2016 https://www.nrgd.nl/binaries/NRGD%20jaarverslag%20 2016_tcm39-278419.pdf

De afgelopen jaren ben ik betrokken geweest bij de normering van het deskundigheidsgebied voor Digitaal Forensisch Onderzoek (DFO)[46]. Inmiddels zijn ook de eerste deskundigen geregistreerd. Een beschrijving van dit proces en een samenvatting van de afbakening en normen zijn te vinden in een publicatie in het tijdschrift Expertise en Recht[47].

6.2 High-tech crime versus low-tech crime

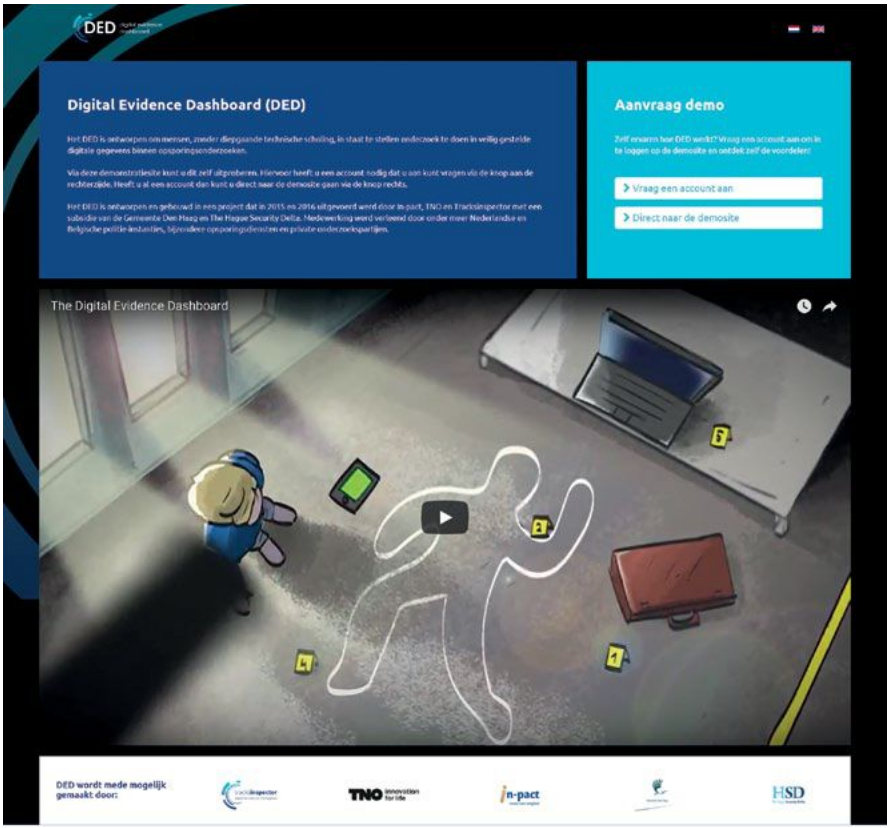
Momenteel is het IoT niet erg veilig. Dat was al geruime tijd zo, maar inmiddels leidt dat ook tot maatschappelijke onrust. In de VS heeft een aantal senatoren een nieuwe wet voorgesteld waarin zij pleiten voor het verbeteren van de cybersecurity van het IoT[48]. Ook in Nederland maakt de politiek zich zorgen. Vlak voor de zomer van 2017 werd door de Tweede Kamer met grote meerderheid een motie aangenomen waarin zij ervoor pleiten dat het kabinet onderzoekt of IoT-apparaten beter beschermd kunnen worden tegen computercriminaliteit[49].

Wat is er aan de hand? Fabrikanten maken zich op dit moment vooral druk over functionaliteit en nieuwe businessmodellen en nog weinig over veiligheid. Maar zelfs als fabrikanten zich daarover druk gaan maken is dat nog geen garantie voor een veilig IoT. In onze fysieke wereld lukt het ook niet om 100% veiligheid te realiseren. Streven naar veiligheid is goed, maar het is ook noodzakelijk om de juiste wetgeving te hebben. Om die wetgeving te handhaven, overtredingen te kunnen onderzoeken en verdachten te kunnen veroordelen, is opsporing en forensisch onderzoek nodig.

Incident response bij cybersecurity-incidenten richt zich nu vooral op *high-tech crime*. Bijvoorbeeld: iemand heeft onbevoegd toegang gekregen tot een computer en maakt daar misbruik van. Hoe heeft die persoon dat gedaan, wat heeft hij gedaan, wat is de schade, weten we zeker dat de inbreker uit het systeem is? Dat zijn de vragen die beantwoord moeten worden aan de hand van digitale sporen. Opsporing en forensisch onderzoek worden bemoeilijkt door de snelle ontwikkeling van nieuwe slimme apparaten en communicatieprotocollen waarvan de interne werking nauwelijks bekend is.

[46] Zie de sectie over DFO op de NRGD website Zie het onderdeel Digitaal Forensisch Onderzoek op de website van het NRGD https://www.nrgd.nl/registreren/digitaal-forensisch-onderzoek.aspx
[47] Henseler, H. en van Loenhout, S. (2016). "Registratie gerechtelijk deskundigen Digitaal Forensisch Onderzoek". Verschenen in Expertise en Recht 2016/5 (oktober)
[48] https://www.warner.senate.gov/public/_cache/files/8/6/861d66b8-93bf-4c93-84d0-6bea67235047/8061BCEEBF4300EC702B4E894247D0E0.iot-cybesecurity-improvement-act---fact-sheet.pdf
[49] https://www.nrc.nl/nieuws/2017/06/15/tweede-kamer-wil-veiligheidseisen-voor-internet-of-things-apparaten-a1563102

Publicaties in de media worden gedomineerd door voorbeelden van high-tech crime. Maar lang niet alle criminaliteit met digitale sporen is high-tech crime. De digitale sporen van "gewone" criminaliteit zijn een veel groter probleem. Niet omdat ze ingewikkelder zijn, maar omdat deze vormen van criminaliteit veel vaker voorkomen. Denk aan een fraude bij een bedrijf, die alleen onderzocht kan worden aan de hand van de financiële administratie die onbereikbaar is ergens in de cloud. Of de verdachte die mogelijk belastend materiaal op zijn smartphone heeft ,die niet uitgelezen kan worden door de beveiliging. Dit soort criminaliteit duid ik voor het gemak hier aan als *low-tech crime*.



Figuur 21. Digitale sporen in low-tech crime (uit het DED-project[40]).

De illustratie over criminaliteit en actoren in het Dark Net (figuur 6) illustreert in feite dat het huidige internet nu al veel wordt gebruikt voor gewone criminaliteit. Door de versmelting van cyberspace met onze fysieke wereld zal het digitaal bewijs in low-tech crime nog sneller toenemen. Opsporing en forensisch onderzoek worden gebruikt om te bewijzen dat een delict is gepleegd door de verdachte aan de hand van sporen die op de plaats van het delict zijn gevonden. Digitaal forensisch onderzoek zal een noodzakelijke aanvulling worden op klassiek forensisch onderzoek (vingerafdrukken, DNA enzovoorts) om te bepalen of een verdachte zich op een bepaalde plaats en bepaalde tijd bevond, zowel in de echte wereld als in cyberspace.

6.3 Onderzoeksagenda

Van een lectoraat bij een opleiding die gespecialiseerd is in Forensische ICT (FICT) en gelet op het werkveld van de afgestudeerde studenten, mag verwacht worden dat de nadruk van het onderzoek vooral ligt op technische aspecten van digitaal forensisch onderzoek. Dat is een zeer breed terrein. Op basis van de bestaande expertise zijn drie thema's gekozen die centraal staan in het onderzoek: IoT, open source intelligence en E-Discovery. In de volgende paragrafen zal ik deze drie agendapunten verder uitwerken en toelichten.

6.3.1 Het IoT Forensics Lab

Voor digitaal forensisch onderzoek aan het IoT is gespecialiseerde apparatuur, kennis en software nodig. Denk hierbij aan chip-off, een speciale hardware-techniek om geheugens uit te lezen als een apparaat beveiligd of zwaar beschadigd is, bijvoorbeeld door brand of water. In Nederland is dit soort apparatuur en kennis alleen beschikbaar bij especialiseerde labs van opsporingsdiensten, bij het NFI en bij bedrijven die elektronische beveiligingen evalueren, bijvoorbeeld voor elektronische betaalsystemen van banken.

Dit soort organisaties en hun laboratoria zijn erg gesloten en alleen toegankelijk voor de eigen medewerkers en enkele stagiaires. Soms is ook een screening nodig voor stagiaires, waardoor de instroom lastig is, zoals op dit moment bij de Nationale Politie. Daarom heeft Hogeschool Leiden het IoT Forensics Lab geopend op de campus van The Hague Security Delta (HSD). De totstandkoming van dit lab volgt op de opening vorig jaar van een onderzoeksruimte van het lectoraat Digital Forensics & E-Discovery van de specialisatie Forensische ICT. Doel van het nieuwe lab is om nauwer te kunnen samenwerken met het werkveld van digitaal forensisch onderzoekers.

Figuur 22. Artistieke impressie van het IoT Forensic lab op de HSD-campus.

Derde- en vierdejaarsstudenten kunnen dankzij het lab nu ook tijdens hun stage ervaring opdoen met geavanceerde digitale forensische technieken in opdracht van bedrijven die zelf niet over de kennis en geavanceerde apparatuur beschikken die nodig zijn voor forensisch onderzoek aan IoT-apparaten. Ook tweedejaarsstudenten zullen ervaring opdoen in het lab door praktijkonderwijs. Docenten FICT en specialisten van de afdeling Digitale Technologie van het NFI, Politie, FIOD, Defensie en van het bedrijfsleven zullen adviseren bij de inrichting en zullen op de HSD instructies en gastcolleges geven aan studenten, docenten en professionals over het uitvoeren van IoT forensics. Door deze samenwerking wil de Hogeschool haar onderwijs vernieuwen en de samenwerking met het werkveld intensiveren.

Nieuwe augmented reality-technieken zoals Microsoft Hololens die in de inleiding (zie figuur 4) en in hoofdstuk 4 (zie figuur 15) staan beschreven, bieden overigens uitstekende mogelijkheden om in het vervolg op het oude CSI The Hague project[50] in het IoT Forensic lab te gaan experimenteren. Professionals zouden bijvoorbeeld met een Hololens instructies kunnen krijgen hoe ze een forensische kopie van een telefoon moeten maken. Hier lijkt samenwerking mogelijk met de specialisatie Mediatechnologie[51] van dezelfde opleiding Informatica aan de Hogeschool Leiden, waarbij stagiaires van beide specialisaties kunnen samenwerken.

Ook virtual reality biedt interessante e-learning-mogelijkheden voor forensische onderzoekers. Op de DFRWS 2017 conferentie in Duitsland[52] werd een demonstratie gegeven van een plaats-delictsimulatie waarin bewijsstukken met digitaal bewijs te zien zijn in virtual reality en waarin de gebruikers moeten proberen om een kabel aan een computer aan te sluiten of zelf toegang te krijgen tot een computer in die ruimte. De simulator staat beschreven in een artikel "Development and Initial User Evaluation of a Virtual Crime Scene Simulator Including Digital Evidence"[53]. Overigens is het lectoraat nauw betrokken bij de organisatie van de DFRWS EU conferenties en ben ik lid van de board of directors[54] van de DFRWS-organisatie die sinds 2001 in de VS bestaat.

De focus van de HSD op Security en Forensics heeft op de campus een verscheidenheid aan bedrijven, overheidsorganisaties en kennisinstellingen opgeleverd. Daarmee is een goede basis gelegd voor het aanvragen van subsidies voor toegepast onderzoek. Zo werkt de Hogeschool Leiden momenteel samen met het Cyber Threat Intell-lab van TNO en het Cyber Expertise Centrum van de Haagse Hogeschool aan de oprichting van een Nationaal IoT Security & Forensics Lab om startups en MKB-bedrijven te adviseren bij innovaties op het gebied IoT Security en Forensics.

[50] Zie https://www.thehaguesecuritydelta.com/projects/project/36 en het filmpje https://www.youtube.com/watch?v=5HrxQ3ZsPm8
[51] https://www.hsleiden.nl/informatica/opbouw-studie/media-technologie.html
[52] Zie de DFRWS-website https://www.dfrws.org/conferences/dfrws-eu-2017/schedule/program
[53] Conwey, A., James, J. en Gladyshev, P. (2015). "Development and Initial User Evaluation of a Virtual Crime Scene Simulator Including Digital Evidence". Verschenen in de proceedings van de 7th Inernational Conference on Digital Forensics en Cybercrime. Zie ook https://doi.org/10.1007/978-3-319-25512-5_2
[54] Board of Directors van DFRWS http://dfrws.org/board

6.3.2 Open source intelligence

Op dit moment is het internet het meest zichtbare stukje cyberspace dat iedereen kent. Op het internet gebeurt van alles en is ook allerlei informatie te vinden in open bronnen. Denk aan websites van bedrijven en organisaties, online marktplaatsen, wiki's, weblogs, github, chatrooms en sociale media. Zoekmachines zoals Google en Bing helpen ons al die informatie te doorzoeken, maar in de praktijk blijkt dat lastig, zeker als je niet precies weet waar je naar op zoek bent. Google en Bing zijn goed in het zoeken in ongestructureerde informatie. Aan de hand van een paar steekwoorden geven ze pagina's terug die gerangschikt worden op basis van relevantie. Bij Google wordt die relevantie bepaald door het PageRank-algoritme. Dat is een slim algoritme dat populaire pagina's, dat wil zeggen pagina's waarnaar veel verwijzingen bestaan, hoger op de resultatenlijst zet dan andere pagina's. Die rangschikking is belangrijk want gebruikers kijken meestal niet verder dan de eerste pagina met zoekresultaten. 'The best place to hide a dead body is page 2 of Google search results'[55].

Volgens information retrieval-begrippen zijn zoekmachines zoals Google wel precies maar leveren ze tegelijkertijd lang niet alle relevante resultaten. De normale gebruikers van Google vinden dat geen probleem. Als je een restaurant zoekt, heb je voldoende aan een paar goede restaurants en is het niet relevant om alle goede restaurants te bekijken. Voor digitale sporen ligt dat iets anders: het PageRank-criterium is voor open source intelligence niet echt bruikbaar. Een onderzoeker wil juist alle sporen verzamelen om vervolgens te analyseren wie de dader is, om een patroon te ontdekken of om een eigen relevance ranking-criterium toe te passen op de resultaten.

Een zoekterm levert al snel honderdduizenden resultaten op. Veel te veel om allemaal te kunnen lezen. Daarom worden er speciale OSINT-tools ontwikkeld die in staat zijn om grote hoeveelheden ongestructureerde informatie van het internet te verzamelen en automatisch te structureren. Dat klinkt mooi, maar de praktijk is zeer weerbarstig en het automatisch analyseren van ongestructureerde webpagina's is niet eenvoudig. Vandaar dat onderzoek nodig is om slimme oplossingen te verzinnen voor het verzamelen van informatie. Zie bijvoorbeeld het kader Orchideeën, waarin de samenvatting is opgenomen van een afstudeerstage door Rick Verdoes die is uitgevoerd in samenwerking met het Naturalis Biodiversity Center en het lectoraat Biodiversiteit van collega-lector Barbara Gravendeel[56].

[55] Deze quote is afkomstig uit een blog post op inbound.org van drie jaar geleden: https://inbound.org/discuss/the-best-place-to-hide-a-dead-body-is-page-2-of-the-google-search-results

[56] Zie ook de websites van het lectoraat Biodiversiteit https://www.hsleiden.nl/biodiversiteit/lectoren van het Naturalis Biodiversity Center https://www.naturalis.nl/nl/

Zoeken naar orchideeën op internet

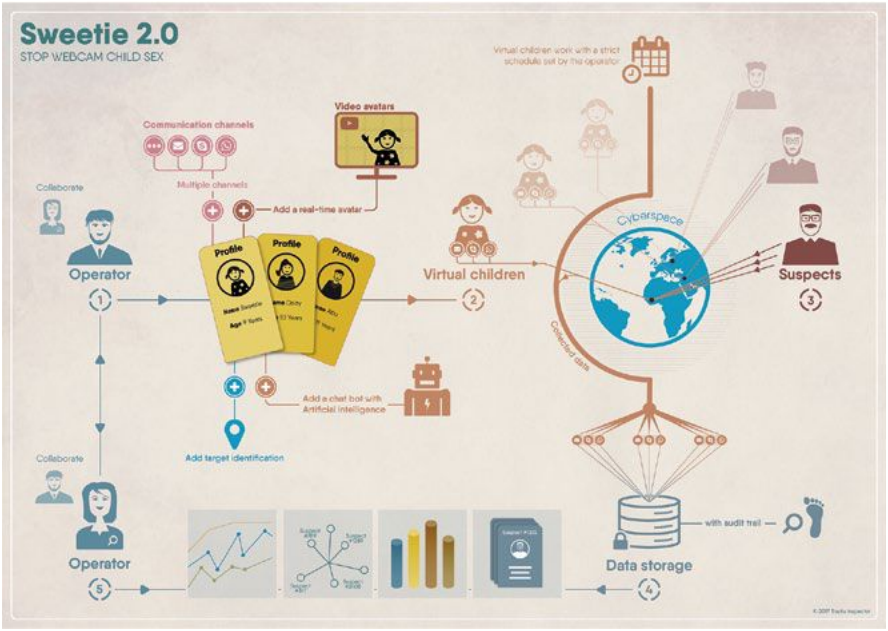
Wilde orchideeën, geoogst als salep, en de illegale export van deze orchideeën is toegenomen[57]. Deze export vindt steeds vaker plaats op online markten zoals eBay. De illegale oogst en export zorgen ervoor dat de populaties van wilde orchideeën afnemen en lokaal uitsterven. Tot op heden zijn er geen gebruiksvriendelijke applicaties voor het in kaart brengen van salep-advertenties op eBay.

Om vervolgstappen te nemen en deze illegale oogst en export tegen te gaan, is de online markt in kaart gebracht met behulp van een applicatie die automatisch zoekt naar advertenties in eBay. Met deze applicatie kunnen etnobotanisten met weinig IT-vaardigheden zelf onderzoek doen naar salep-advertenties op eBay. De resultaten zullen gemeld worden bij organisaties zoals het World Wildlife Fund (WWF), zodat zij vervolgstappen kunnen nemen om de populatie van wilde orchideeën te behouden.

Het zoeken naar advertenties op eBay met betrekking tot illegale handel in salep illustreert welke problemen en oplossingen er zijn. In feite zijn de gekozen aanpak, technologie en ontwikkelde oplossing een blauwdruk voor allerlei OSINT-projecten. Aangezien ieder onderwerp weer andere bronnen heeft en andere kenmerken, is telkens een nieuwe benadering nodig.

In het onderwijs van de specialisatie Forensische ICT worden studenten bewust gemaakt van de problemen die er zijn en worden ze gestimuleerd om oplossingen te vinden. De projectteams in het onderwijs beperken zich meestal tot handmatige zoektochten, waarbij ze leren om gestructureerd te werk te gaan en alle stappen in het onderzoek te documenteren. Tijdens een stage hebben studenten uitgebreid de tijd om de diepte in te gaan en kunnen ze speciale software ontwikkelen, zodat ze meer informatie automatisch kunnen analyseren.

[57] Jalal, J.S., Kumar, P. and Pangtey P.Y.S. (2008). *Ethnomedicinal Orchids of Uttarakhand. Western Himalaya.* Ethnobotanical Leaflets. Vol. 2008 : Iss. 1 , Article 164. Available at: http://opensiuc.lib.siu.edu/ebl/vol2008/iss1/164



Figuur 23. Infographic over de software in het Sweetie 2.0-project[58].

Met mijn bedrijf Tracks Inspector ben ik in 2015 in aanraking gekomen met het Sweetie 2.0-project, dat door de organisatie Terre des Hommes is geïnitieerd[59]. In het Sweetie 2.0-project wordt een heel geavanceerde vorm van open source intelligence toegepast. Traditionele OSINT-tools verzamelen informatie uit webpagina's en online databases, maar in Sweetie 2.0 is een automatische methode ontwikkeld om informatie in te winnen van andere internetgebruikers.

Die automatische methode in het Sweetie 2.0-project bestaat uit een chatbot die is ontwikkeld met behulp van technieken uit de kunstmatige intelligentie. De chatbot simuleert het meisje dat in het Sweetie 1.0-project aan de hand van scripts door medewerkers van Terre des Hommes was gesimuleerd. De chatbot geeft antwoord aan personen in een chatroom die een gesprek met haar beginnen. De chatrobot is geprogrammeerd met het doel om personen te ontmaskeren die tegen betaling Sweetie bloot willen zien terwijl ze weten dat ze minderjarig is.

[58] Voor meer informatie over de software in het Sweetie 2.0 zie de website van Tracks Inspector https://tracksinspector.com/blog/ti-software-sweetie-2-0.html

[56] Zie de website van Terre des Hommes https://www.terredeshommes.nl/programmas/sweetie-20-webcamseks-met-kinderen-de-wereld-uit

De werking van de software, die door Tracks Inspector voor het Sweetie 2.0-project is ontwikkeld, wordt geïllustreerd door de infographic in bovenstaande figuur.

De strategie die in Sweetie wordt toegepast, is veel breder toepasbaar. Zo wordt inmiddels door Terre des Hommes in het WATCH NL-project (zie kader), gebruik gemaakt van lokadvertenties op sekswebsites om personen te ontmaskeren die reageren op advertenties van minderjarige kinderen die seksuele diensten aanbieden (zie kader).

WATCH Nederland
WATCH Nederland stopt seksuitbuiting van minderjarigen door de markt van loverboys in Nederland kapot te maken. WATCH Nederland frustreert de vraagkant, want zolang er vraag is, is er aanbod. Watch Nederland spoort loverboys en hun klanten op door gebruik te maken van een meldpunt, web crawlers, nepadvertenties, lokprofielen en privé-detectives.

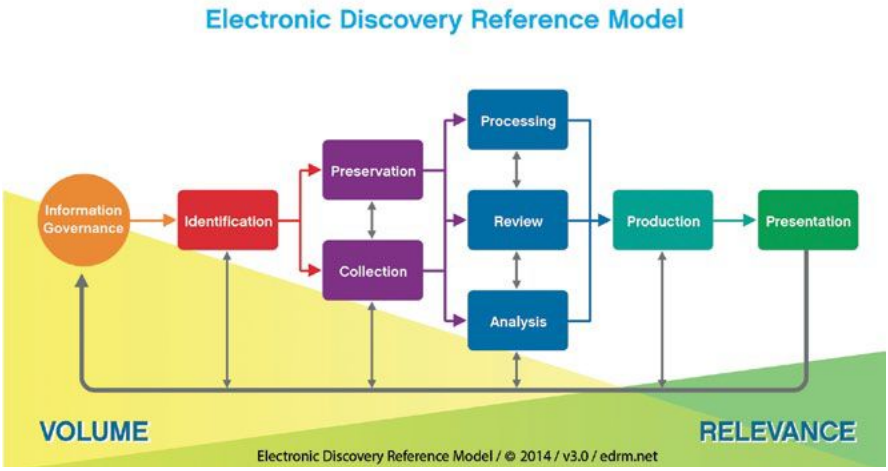
Mensenhandelaren gebruiken internet
Een loverboy is een bikkelharde mensenhandelaar die geweld noch chantage schuwt. Door handig gebruik te maken van internet is zijn handel geëxplodeerd. Hij hoeft niet meer fysiek bij zijn slachtoffer in de buurt te komen, hij spreekt haar of hem (ook jongens worden slachtoffer) direct aan op social media, webfora en chatsites. Zodra hij eenmaal één seksfoto of seksfilmpje heeft van het slachtoffer, dreigt hij het online te plaatsen als de minderjarige niet doet wat hij wil. Vervolgens maakt hij een advertentie aan op een populaire website en trekt hij tien tot twintig klanten per dag. Een loverboy verdient al snel zo'n 500 tot 1000 euro per dag per slachtoffer. Dat is meer dan hij kan verdienen met drugshandel, berovingen of inbraken. Bovendien is de pakkans nihil. Online is de loverboy namelijk vrijwel anoniem.

Het WATCH-project werkt (nog) niet met een chatbot, maar er wordt wel gekeken of technieken uit het Sweetie-project kunnen worden ingezet. De chatbot-benadering lijkt vooral geschikt voor het inwinnen van informatie voordat duidelijk is of een bepaalde gebruiker verkeerde intenties heeft. Denk bijvoorbeeld aan het bevragen van verkopers op Marktplaats die verdachte goederen aanbieden of het ontlokken van een reactie van een gebruiker op Twitter die een bedreiging heeft geuit aan het adres van een politicus.

De ontwikkeling van slimme webcrawlers, het slim analyseren van web-content en het inzetten van chatbots illustreren het toegepaste onderzoek waarin studenten een waardevolle rol kunnen vervullen. In sommige gevallen kunnen ze bedrijven of organisaties zoals Terre des Hommes en opsporingsinstanties van de overheid een stap verder helpen en doen ze praktijkervaring op die later in hun werk relevant kan blijken. Of het nu gaat om het ontwikkelen van een eigen methode of om het beoordelen of gebruiken van een commercieel product.

6.3.3 E-Discovery

De behoefte aan E-Discovery-technieken en -training neemt gestaag toe en het onderwerp wordt ook in Nederland steeds bekender. Het EDRM-model[60] is in 2006 voor het eerst in de VS geïntroduceerd; versie 2.0 heb ik in 2010 laten zien in mijn openbare les. Inmiddels zijn we toe aan versie 3.0 (zie onderstaande figuur). Het verschil met versie 2.0 is dat Information Governance (helemaal links) nu staat aangeduid met een cirkel.



Figuur 24. Het E-Discovery Reference Model.

[60] Het EDRM-model heeft een eigen website die nogal eens verandert. De beschrijving van het EDRM-model is te vinden op https://www.edrm.net/frameworks-and-standards/edrm-model

Met dit apart benoemen van Information Governance wordt onderkend dat dit deel van E-Discovery een zodanig belangrijk onderdeel/aspect is voor iedere organisatie dat het een eigen model verdient, namelijk het Information Governance Reference-model[61]. Bedrijven zijn gaan inzien dat information governance een belangrijk onderdeel is om hun organisatie voor te bereiden op een E-Discovery-onderzoek. Als in mei 2018 de nieuwe Europese richtlijn voor datalekken van kracht wordt, zal dit aspect nog belangrijker worden voor bedrijven[62]. Volgens de nieuwe General Data Protection Regulation kunnen bedrijven voor het lekken van persoonsgegevens bestraft worden met een maximale boete van 4% van de jaarlijkse wereldwijde omzet of 20 miljoen euro.

Ondanks de toegenomen aandacht in Nederland voor het onderwerp E-Discovery, is er nog steeds een groot verschil met E-Discovery-projecten in de VS en in de UK, waar legal discovery een wezenlijk onderdeel is van de rechtsgang. Het lectoraat wil zich vooral richten op de toepassing van E-Discovery-technieken in de Nederlandse context en met name de vertaling van de E-Discovery-lessen die de advocatuur en forensische accountants hebben geleerd naar de praktijk van opsporingsambtenaren.

In de praktijk van politie en bijzondere opsporingsdiensten (denk aan FIOD, ISZW, ACM) wordt het onderzoek nog steeds grotendeels uitgevoerd door digitale experts die gespecialiseerd zijn in forensisch ICT-onderzoek. Hun collega's (de tactische rechercheurs) geven aan naar wat voor soort bewijs wordt gezocht en laten in veel gevallen het zoekwerk over aan de digitale expert. Deze werkwijze verschilt flink van de werkwijze die forensische accountants en advocaten hanteren in het bedrijfsleven. Daar wordt al ruim tien jaar gebruik gemaakt van web-based review-platformen zoals Relativity, ZyLAB, Clearwell, iConect-Xera en Summation.

Dergelijke review-platformen worden zelden tot nooit ingezet door opsporingsinstanties in Nederland. Enerzijds wordt dit veroorzaakt door het kostbare prijsmodel waarbij per Gb betaald moet worden. Anderzijds is een forensisch onderzoek van een opsporingsambtenaar vaak veel meer dan alleen een document-review. Opsporings-ambtenaren hebben veel meer behoefte aan forensische digitale sporen.

[61] Het Information Governance Reference Model (IGRM) is ook te vinden op de EDRM-website, zie https://www.edrm.net/frameworks-and-standards/information-governance-reference-model
[62] Zie bijvoorbeeld http://www.eugdpr.org/

Het idee van web-based review en de mogelijkheid om tactische rechercheurs zelf eenvoudig digitaal bewijs te laten onderzoeken, wint aan populariteit. Leveranciers van expert-tools komen langzaamaan met een web-interface als alternatief voor hun Windows-programma. In veel gevallen echter probeert men met die web-interface nog vooral de ingewikkelde Windows-interface tot in detail na te bootsen, waardoor de web-interface wel beschikbaar maar eigenlijk nog te ingewikkeld is voor de tactische rechercheur.

In 2011 ben ik bij Fox-IT begonnen met de productontwikkeling van het concept Fox Tracks Inspector, dat de jaren daarvoor als idee was ontstaan naar aanleiding van de toenemende achterstanden in het digitale forensische lab van de politie in Den Haag. Het idee was om tactische rechercheurs met behulp van Tracks Inspector zelf in staat te stellen om digitaal bewijs eenvoudig te bekijken en zo het "laaghangende fruit" te vinden.

In de praktijk bleek het lastig om opsporingsinstanties te overtuigen dat er meer digitaal bewijs door tactische rechercheurs onderzocht moet worden. Zoals wel vaker blijkt het transformeren van bestaande werkprocessen moeilijk en tijdrovend. Toen Fox-IT zich steeds meer ging richten op cybersecurity en high-tech crime, heb ik in 2014 met een aantal collega's Tracks Inspector via een management buy-out als zelfstandig bedrijf op de HSD-campus voortgezet. Weliswaar hebben we nu klanten in België, Nederland, Engeland en Zuid-Afrika, maar vooralsnog zijn het vooral early adopters die besluiten om de achterstanden in de digitaal forensische labs door de manier van werken te transformeren.

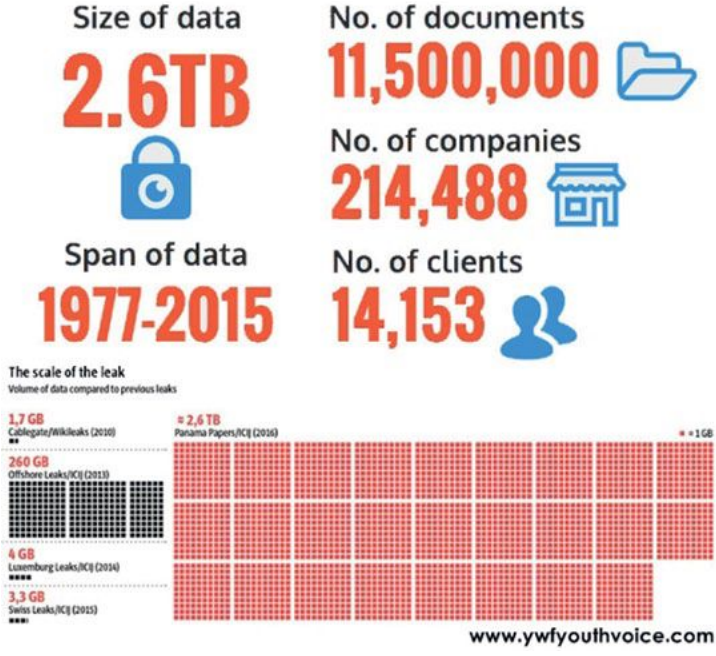
Toch zijn de voordelen overduidelijk. Digitale experts moeten zich bezig houden met datgene waarvoor ze zijn opgeleid, namelijk het onderzoeken van digitaal forensisch complex bewijs en tactisch onderzoek in high-tech crime. Door tactische rechercheurs zelf eerst digitaal bewijs te laten onderzoeken, kunnen ze niet alleen sneller een beeld krijgen van de bewijskracht maar ook veel specifieker aangeven welke bewijsstukken echt interessant zijn. Door digitale experts alleen geselecteerde bewijsstukken te laten onderzoeken, zal de hun werklast afnemen.

Daarbij komt dat de vraag naar digitale experts groot is en veel vacatures op dit moment niet vervuld kunnen worden. Bedrijven zijn bereid om hoge salarissen te betalen waardoor het voor overheidsinstanties lastig is om experts te vinden. Overigens betekent dat wel dat het arbeidsmarktperspectief voor de studenten van de specialisatie Forensische ICT zeer goed is.

De afgelopen jaren bleek het moeilijk om tactische rechercheurs meer zelf onderzoek te laten doen, maar het laatste jaar lijkt er een omslag gaande te zijn. Dit blijkt onder andere in de UK, waar eind vorig jaar de dienst die verantwoordelijk is voor de inspectie van de politie een duidelijk signaal heeft afgegeven[63]. Een van de knelpunten daarbij is dat rechercheurs moeten leren hoe ze in grote hoeveelheden informatie moeten zoeken. Ze moeten bewust worden gemaakt van de mogelijkheden van digitaal bewijs.

Een ander voorbeeld waaruit blijkt dat men oog krijgt voor de rol van tactische rechercheurs blijkt uit een opdracht van de FIOD die het lectoraat in augustus heeft verworven. Het lectoraat gaat samen met specialisten van de FIOD een masterclass E-Discovery ontwikkelen voor in totaal 300 FIOD- rechercheurs en -projectleiders. Dit nascholingstraject zal 2,5 jaar in beslag nemen en de kans is aanwezig dat ook andere bijzondere opsporingsdiensten zich zullen aansluiten.

Dat de FIOD deze stap neemt, is niet verrassend gelet op de steeds grotere hoeveelheden digitaal bewijs die ze in beslag nemen. Zie ook het kader over E-Discovery en de Panama Papers. De omvang van de Panama Papers wordt geïllustreerd in onderstaande figuur waarin een vergelijking wordt gemaakt met documentcollecties van andere grote datalekken zoals Wikileaks.



Figuur 25. Omvang van de Panama Papers in vergelijking met andere lekken[64].

E-Discovery en de Panama Papers[65]

De Panama Papers zijn een aansprekend voorbeeld van de inzet van E-Discovery. Journalisten van de Süddeutsche Zeitung brachten de zaak aan het rollen, nadat ze met een E-Discovery-tool 'gelekte' informatie doorzochten. Ze kregen vertrouwelijke gegevens met een omvang van 2,6 terabyte, 11,6 miljoen bestanden: Word, PowerPoint, spreadsheets, e-mails, pdf's. Te veel om allemaal door te lezen. Een E-Discovery-tool werd erop losgelaten en na twee weken leverde dit iets op wat de journalisten konden doorzoeken.

Belastingdienst gaat onderzoek doen

Staatssecretaris van Financiën Wiebes complimenteerde de journalisten en zei in 2016: 'Met deze informatie kan de Belastingdienst iets.' Hij was voornemens om een speciaal team van de Belastingdienst de zaak te laten onderzoeken. Is dat een goed idee? Jazeker, want de Belastingdienst in Nederland beschikt over dezelfde tool als de journalisten in Duitsland gebruikten. Dat klinkt goed, maar ondanks de inzet van deze tool waren 400 onderzoeksjournalisten een jaar bezig om de eerste nieuwsverhalen in 2016 te presenteren. Dat moet (en kan) beter.

Juiste tools, ook voor rechercheurs

Met E-Discovery kunnen hoeveelheden informatie worden doorzocht, die de menselijke vermogens te boven gaan. Zelfs als je niet precies weet wat je zoekt, kunnen deze tools opvallende uitschieters of verbanden opsporen. Maar de inzet van E-Discovery vraagt om kennis van de actuele technologie en ontwikkeling. De tool waar Wiebes aan refereert is vrij technisch en kan alleen gebruikt worden door getrainde experts. De Belastingdienst zou veel meer onderzoek kunnen doen als ze de allernieuwste E-Discovery- technieken inzetten, die eenvoudiger zijn en gebaseerd op machine-leren en semantisch zoeken. Als gewone rechercheurs ermee kunnen werken, vergroot je je onderzoekscapaciteit enorm.

[63] Zie het artikel "Police forces 'overwhelmed' by digital evidence, watchdog finds" op de website van de BBC http://www.bbc.com/news/uk-37846705 en https://www.justiceinspectorates.gov.uk/hmicfrs/news/news-feed/police-forces-urged-to-get-back-to-the-future

[64] Overgenomen uit http://www.ywfyouthvoice.com/2016/04/explaining-concept-of-panama-papers-and.html

[65] Overgenomen uit het persbericht waarin het E-Discovery Symposium 2016 wordt aangekondigd, zie http://www.hva.nl/content/nieuws/persberichten/2016/04/e-discovery.html

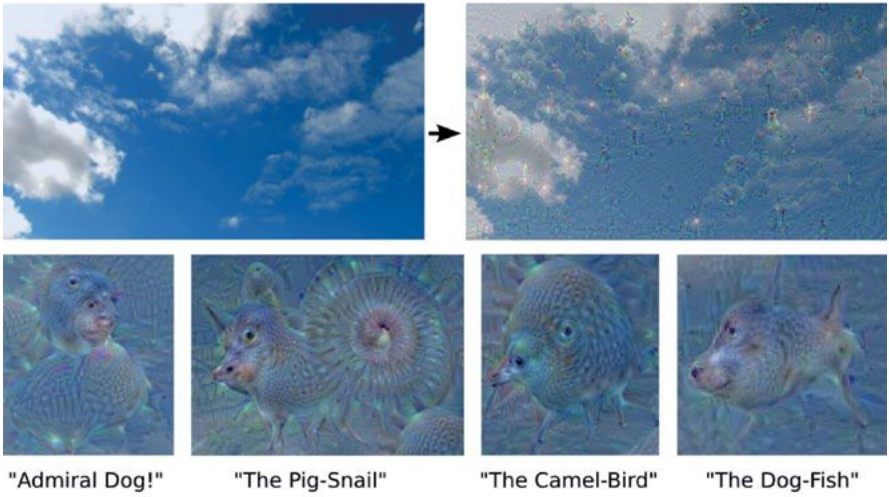
Onderdeel van de overeenkomst is dat Hogeschool Leiden en de FIOD samen het lesmateriaal blijven actualiseren naar de stand van de techniek. Dat is goed voor het onderwijs en goed voor de Forensisch IT (FIT)-experts, projectleiders en rechercheurs van de FIOD. Het is volgens mij een schoolvoorbeeld van hoe hoger beroepsonderwijs en het werkveld elkaar kunnen versterken en hoe een lectoraat daarbij een faciliterende rol kan spelen.

6.4 Blik op de toekomst

In hoofdstuk 4 en 5 heb ik geschreven over de opkomst van slimme computers die misschien nog niet intelligent zijn, maar wel veel slimmer zijn dan we tot voor kort voor mogelijk hielden. In de onderzoeksagenda die in de vorige paragraaf is uitgewerkt, heb ik drie onderzoeksthema's benoemd die inspelen op verschillende facetten van de revolutie die, denk ik, zal plaatsvinden in het digitaal bewijs. Naast die drie thema's zijn er nog vele andere en één daarvan wil ik ten slotte nog benoemen. Dat thema is deep learning forensics.

Als het gaat om autonome systemen die een hoge graad van automatisering hebben, zal het lastig worden om manipulatie van gegevens te bewijzen dan wel om aanwezige digitale sporen in zulke systemen op een forensische wijze te interpreteren. De huidige deep learning-algoritmen hebben kennis verworven aan de hand van trainingsvoorbeelden en het is lastig om met de hand na te gaan hoe deze kennis in de systemen is gerepresenteerd en hoe sensordata worden verwerkt. Daarbij komt dat deze algoritmen juist zo krachtig zijn door hun capaciteit om onvolledige informatie te kunnen verwerken.

Google heeft inmiddels spraakmakende voorbeelden op haar research blog gepubliceerd, die deze eigenschappen illustreren. Bijvoorbeeld een programma dat ontbrekende details op foto's kan reconstrueren (of beter gezegd fantaseren). Of een programma dat gebouwen, dieren en objecten op foto's kan herkennen maar deze beelden ook herkent in een wolkenlucht. Ook traditionele taalverwerking moet eraan geloven. In september 2016 heeft Google het "oude" translate-algoritme vervangen door een neuraal netwerk. Interessante bevinding daarbij is dat deep learning de verborgen lagen in het netwerk zo gevormd heeft dat vertalingen mogelijk zijn zonder dat het daar voorbeelden van heeft gezien.



Figuur 26. Wanneer deep learning-netwerken gaan hallucineren[66].

Wie denkt dat dit vooral academische toepassingen zijn, heeft het mis. Deze technologie wordt nu klaargestoomd voor de zelfrijdende auto. Nvidia, de maker van de graphics-processoren, liet begin 2017 op haar jaarlijkse developers-conferentie zien welke enorme ontwikkelingen er op dit moment gaande zijn om deep learning op grote schaal naar de consument te brengen. Daarbij is de zelfrijdende auto niet alleen een inspirerend voorbeeld, maar tegelijkertijd ook een belangrijke stimulans om de technologie mobiel en betaalbaar te maken.

Met dit soort systemen wordt het onderscheid tussen onvolledige informatie en gemanipuleerde informatie steeds moeilijker te maken. Het lijkt erop dat voor forensisch onderzoek aan slimme autonome apparaten naast IoT forensics ook onderzoek nodig is naar de representatie en verwerking van informatie in kunstmatige neurale netwerken, oftewel deep learning forensics.

Het gaat hierbij dus niet over de inzet van deep learning om forensisch onderzoek te verbeteren. Het gaat over digitaal bewijs dat verborgen zit in de gedistribueerde patronen van kunstmatige neurale netwerken. Mijn proefschrift in 1993 was getiteld “Neurons, Connections and Activation. The organization of representation in Artificial Neural Networks”. Met de verwachte opkomst van deep learning zou dit zomaar eens het nieuwe onderzoeksthema van het lectoraat kunnen worden over vier jaar.
Maar daar zal ik dan graag een nieuwe lectorale rede aan wijden.

[66] Afbeelding overgenomen uit Mordvintsev, A., Olah, C., Tyke, M. (2015) “Inceptionism: Going Deeper into Neural Networks” https://research.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html

Dankwoord

Deze lectorale rede zou er niet geweest zijn zonder Jos Griffioen. Jos kwam al een geruim aantal jaren naar het symposium E-Discovery in Nederland. We hadden al snel een klik en het was voor mij een aanleiding om diverse gastcolleges in Leiden te geven. Bij Fox-IT maakten we gretig gebruik van de FICT-studenten en zodoende ontstond er een voortdurende samenwerking. Begin 2016 vertelde Jos mij dat hem was gevraagd om een lectoraat op te gaan zetten. Hij wilde daarover graag met mij afstemmen om overlap te voorkomen met mijn lectoraat bij de Hogeschool van Amsterdam. Het gesprek nam al snel een andere wending toen ik opperde om zelf lector te worden in Leiden.

In Amsterdam ging mijn E-Discovery-onderzoek steeds meer in de richting van de specialisatie Informatiemanagement en het werkveld van de informatieprofessional. De gedachte om meer naar de technische Digital Forensics-kant van E-Discovery op te schuiven, sprak me direct aan vanwege mijn eigen achtergrond en mijn werk bij Tracks Inspector. De uitstekende koppeling met het werkveld die Jos en zijn collega's hebben ontwikkeld, paste bovendien veel beter bij mijn eigen netwerk dan bij het werkveld van de informatieprofessional. Daarnaast had ik inmiddels zelf ondervonden dat het cluster van The Hague Security Delta een unieke positie heeft verworven rondom het thema cybersecurity en forensics. Zo gezegd zo gedaan, en binnen een aantal maanden zou mijn overstap definitief worden[67].

Tijdens de overgangsperiode maakte ik al snel kennis met Patrick Pijnenburg. Van Jos had ik begrepen dat hij dankzij de steun van Patrick groen licht had gekregen om op zoek te gaan naar een lector. Al snel zou ik ondervinden dat Patrick een belangrijke motor is binnen de Hogeschool Leiden: dankzij zijn ondernemende instelling en relativeringsvermogen weet hij veel voor elkaar te krijgen en hij is een grote steun en toeverlaat voor de mensen die voor en met hem werken. Patrick, je bent een inspirerend voorbeeld voor de mensen om je heen, ook voor mij. Bedankt voor je niet aflatende inzet en het vertrouwen dat je aan ons schenkt!

Iedere lector heeft een Kenniskring. In de regel is dat een aantal docenten, zoals Jos, die een deel van hun tijd kunnen besteden aan werkzaamheden die verband houden met het lectoraat. Eigenlijk is mijn Kenniskring een team dat samen met mij het reilen en zeilen van het lectoraat regelt. Peter van der Wijden, docent FICT, geeft onder andere het vak E-Discovery, waardoor hij uitstekend past in mijn Kenniskring.

[67] Zie ook het interview "De Smartphone is het nieuwe DNA" door Martijn van Calmthout, dat in 2016 is gepubliceerd in een uitgave van de Hogeschool van Amsterdam met de titel "Forensics en Fashion. De gevolgen van digitalisering voor forensisch onderzoek en de mode-industrie" ter gelegenheid van het afscheid van lectoren Hein Daanen en Hans Henseler.

Omdat de meeste E-Discovery-boeken geschreven zijn vanuit een Amerikaans perspectief en gericht zijn op rechtenstudenten, moedigt Peter mij nog altijd aan om mijn lesboek over E-Discovery af te maken. Daar ben ik in 2014 ongeveer halverwege mee gestopt omdat er toen te veel andere zaken speelden. Maar dat boek gaat er komen. Al is het alleen maar omdat Peter, Jos en ik de komende 2,5 jaar gaan helpen om circa 300 medewerkers van de FIOD bij te scholen op het onderwerp E-Discovery. Peter, bedankt voor je aanmoediging, je hulp en voor je inzet in het avontuur dat we samen gaan beleven bij de FIOD.

Naast Jos en Peter wil ik Karin en Saskia bedanken. Karin Baak is mijn secretaresse, die, ondanks het beperkt aantal uren dat ze voor mij beschikbaar heeft, mij toch uitstekend weet te ondersteunen. Saskia Kanij en het web-team van Marketing & Communicatie verrichten ongelofelijk veel werk in verband met de PR van het lectoraat.
Saskia ondersteunt bij de website, de banners, de persberichten over het IoT-lab, interviews, grafische vormgeving en het drukwerk voor deze lectorale rede. Dankzij de hulp van Karin en Saskia, met af te toe de magische aanraking van Patrick, kan mijn lectoraat optimaal gebruik maken van alle diensten die de hogeschool te bieden heeft.

Het docententeam van de specialisatie FICT is natuurlijk groter dan alleen mijn Kenniskring. Waar mogelijk proberen we met elkaar activiteiten te ondernemen. Of het nu gaat om een gastcollege, contacten met het werkveld of deelname aan nationale en internationale conferenties. Er zijn genoeg gelegenheden waarop het lectoraat van het FICT-onderwijs profiteert en ik zal mijn best (blijven) doen om ervoor te zorgen dat het FICT-onderwijs en de docenten blijven profiteren van het lectoraat. Overigens weet ik uit ervaring dat de synergie tussen onderwijs en lectoraat niet vanzelfsprekend is.
Ik prijs me gelukkig met een specialisatie die er zo goed in geslaagd is om verbinding te maken met het werkveld. Hans, Jaap, Edward, Henk en Bas, bedankt voor jullie hulp bij de samenwerking tussen het onderwijs en het lectoraat. Met elkaar vormen we een ijzersterk team. Ook dank aan Vincent, het hoofd van de opleiding Informatica, voor de prettige samenwerking. In deze lectorale rede heb ik voldoende onderzoeksthema's geagendeerd om met de andere informaticaspecialisaties Software Engineering, Business Data Management en Mediatechnologie te gaan samenwerken.

Minstens zo belangrijk als de samenwerking met docenten en het werkveld is de samenwerking met de studenten. Dankzij de specialisatie is er grote interesse van studenten voor het lectoraat en voor de HSD-campus. Lisa, Josefien, Rick, Stefan, Martin, Nick, Jasper, Stephan, André, Bob, Danny, Vince en Anton, bedankt voor jullie inzet en enthousiasme! Dankzij studenten zoals jullie kunnen we professionals en docenten samenbrengen in innovatieve projecten. Hopelijk is jullie enthousiasme inspirerend voor al die studenten die nog gaan volgen. Ongetwijfeld zullen jullie een mooie baan gaan vinden in het FICT-werkveld en blijven we samenwerken.

Die inspiratie komt niet in de laatste plaats door het nieuwe IoT Forensic Lab, dat onlangs door ons is geopend op de HSD-Campus. Alweer een voorbeeld van teamwerk en visie, dat zonder de steun van het College van Bestuur van de hogeschool en de enthousiaste ondersteuning van de HSD (zowel de campus als de stichting) nooit was gerealiseerd. De strategische steun van de Gemeente Den Haag sterkt ons bij de verdere ontwikkeling; mede dankzij de door hen verstrekte subsidie zijn we in staat om te investeren in de modernste apparatuur. Bedankt ook Madison Gurkha, Data Expert, en het Nederlands Forensisch Instituut voor jullie steun en vertrouwen, in-cash en in-kind! Hopelijk dat Politie, Defensie, FIOD en nog vele andere bedrijven en organisaties jullie voorbeeld zullen volgen als sponsor en opdrachtgever voor stageopdrachten in het lab.

TNO en de Haagse Hogeschool wil ik ook bedanken voor het samen zoeken naar subsidies en bedrijven die met ons onderzoek willen doen in het kader van IoT Security & Forensics. Al die ballen zou ik niet in de lucht kunnen houden zonder de hulp van Wiebe Wamelink en Walter Zuijderduin. Inmiddels werk ik al tien jaar samen met Wiebe in allerlei initiatieven rondom de inzet van data science voor safety en security, en ik ben er trots op dat hij mij nu wil helpen bij het opzetten van nieuwe onderzoeksprojecten. Walter is subsidieadviseur bij de Hogeschool Leiden en toont zich zeer bekwaam in het doorgronden van subsidieregelingen. Niet alleen als het gaat om de subsidieregels, maar ook als het gaat om het analyseren van drijfveren die aan een bepaalde subsidieregeling ten grondslag liggen. De samenwerking is nog pril, maar voelt nu al goed aan.

Er zijn nog veel meer mensen die op een of andere manier hebben bijgedragen aan de activiteiten van het lectoraat. Het zijn er te veel om ze allemaal op te noemen, maar een aantal wil ik hier toch graag noemen. Om te beginnen wil ik Rens de Wolf en Marco de Moulin, mijn medeoprichters van Tracks Inspector, bedanken voor hun flexibiliteit. Twee dagen lector en drie dagen CEO zijn valt niet altijd even goed te combineren en de scheidslijn is niet altijd even makkelijk te maken. “Thought Leadership” is mooi voor bedrijven als PWC en Fox-IT, maar het is niet altijd duidelijk of en hoe een klein bedrijf daar de vruchten van kan plukken. Gelukkig is er nog steeds voldoende synergie en ik ben ervan overtuigd dat we samen, ieder met onze eigen kwaliteiten, Tracks Inspector linksom of rechtsom tot een succes kunnen maken. Mijn collega’s van het college en het bureau van het Nederlands Register Gerechtelijk Deskundigen (NRGD) wil ik bedanken voor de samenwerking en in het bijzonder voor de kans die ze mij hebben gegeven om mee te werken aan de normering van het deskundigheidsgebied Digitaal Forensisch Onderzoek. Het is een waardevolle toevoeging aan mijn contacten met het werkveld in zowel nationaal als internationaal verband.

Ten slotte wil ik Jolanda bedanken. Mijn vorige lectorale rede heb ik in de kerstvakantie geschreven en in het dankwoord heb ik toen beloofd dat dat niet meer zou gebeuren. Dat is redelijk gelukt, maar er blijven altijd nog voldoende andere werkgerelateerde afleidingen over... Gelukkig ook nog voldoende vakanties. Bedankt voor je geduld en weet dat ik zonder jouw steun dit niet zou kunnen combineren met al het andere wat in een mensenleven gebeurt. Droevige momenten maar ook hoogtepunten zoals het recente afstuderen van onze geweldige zonen Rutger en Matthijs. Jongens, we zijn trots op jullie en we hopen dat jullie samen met jullie vriendinnen net zo veel plezier in je werk en je leven zullen hebben als wij.

Curriculum vitae

Hans Henseler is sinds augustus 2016 aangesteld als lector bij het lectoraat Digital Forensics & E-Discovery van Hogeschool Leiden. Daarnaast is hij algemeen directeur en medeoprichter van het bedrijf Tracks Inspector, lid van het College van het Nederlands Register van Gerechtelijk Deskundigen en board member van de Board of Directors van DFRWS, een internationale conferentie op het gebied van digitaal forensisch onderzoek. Hij studeerde Informatica aan de TU Delft en is aan de Universiteit van Maastricht gepromoveerd op het onderwerp Artificiële Neurale Netwerken.
In 1992 is hij gaan pionieren op forensisch ICT-gebied en heeft hij de afdeling Digitale Techniek bij het NFI opgezet. Hij was onder andere succesvol bij het ontwikkelen van methoden om elektronische zakagenda's uit te lezen, die toen al veel door criminelen werden gebruikt. Sindsdien heeft hij voor verschillende bedrijven gewerkt, waar hij ervaring heeft opgedaan met het ontwikkelen van software en het leiden van digitaal forensische laboratoria. Van 1999-2000 was hij leider van de divisie Informatiesystemen bij TNO-TPD in Delft. Deze divisie was destijds haar tijd ver vooruit met geavanceerde oplossingen voor beeldherkenning en robotisering, parallelle computers en kennis-systemen, taaltechnologie en ketenlogistiek. Van 2000-2006 was hij als technisch directeur bij ZyLAB verantwoordelijk voor ontwikkeling, testen en de supportafdeling van ZyLAB's E-Discovery softwareoplossing. Van 2006-2010 was hij director Forensisch Technology Solutions bij PWC Eurofirms waar hij (internationale) E-Discovery-projecten heeft geleid en bedrijven heeft geadviseerd op het gebied van digitaal bewijs in fraude-onderzoeken, waaronder het Bouwfonds-onderzoek in Nederland en het onderzoek naar corruptiepraktijken bij Siemens in Duitsland. Van 2010-2014 was hij als partner bij Fox-IT verantwoordelijk voor de business unit Forensics. Hij was onder andere betrokken bij het forensisch onderzoek in het Diginotar-incident en het faillissement van Econcern en hij gaf met zijn team ondersteuning aan de Freeh Group die toezicht moet houden op de invoering van een compliance programma bij de firma Daimler in Stuttgart.

hogeschool Leiden

Hogeschool Leiden

Zernikedreef 11
2333 CK Leiden
Postbus 382
2300 AJ Leiden

071 - 518 88 00

info@hsleiden.nl

hsleiden.nl

facebook.com/HSLeidenNL

twitter.com/HSLeidenNL

linkedin.com/company/hogeschool-leiden

ISBN 978-90-821382-2-1

klimaatneutraal
natureOffice.com | NL-077-872221
gedrukt