



Universiteit
Leiden
The Netherlands

Computer says no: understanding digital authority

Weggemans, D.J.

Citation

Weggemans, D. J. (2025, February 13). *Computer says no: understanding digital authority*. Retrieved from <https://hdl.handle.net/1887/4180592>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4180592>

Note: To cite this publication please use the final published version (if applicable).

S | SAMENVATTING

In de afgelopen decennia zijn we overspoeld met innovatieve, communicerende technologieën. Overal om ons heen vinden we digitale technologieën die informeren, adviseren, aanraden, beweren, instrueren en bevelen. Ze spelen een rol in allerlei omgevingen variërend van ons huis tot op de weg en van op werk tot in onze vrije tijd. In veel gevallen is het logisch om deze digitale communicaties te volgen. Ze zijn tenslotte ontworpen om uitdagingen aan te pakken en we hebben de waarde van computers al op veel gebieden gezien. Helaas is geen enkele technologie foutloos. Er kunnen altijd bugs, slechte algoritmen, biased data en invoerfouten zijn die ons op een dwaalspoor kunnen brengen. Maar hoewel de meesten van ons zich hier wel van bewust zijn, doen we geregeld maar al te graag wat ons digitaal wordt verteld.

Verhalen van mensen die in de problemen komen doordat zij automatisch een digitale technologie gehoorzamen komen veelvuldig in het nieuws – zo zijn er bijvoorbeeld talloze berichten over gevaarlijke situaties die ontstonden doordat mensen blindelings hun GPS-systeem volgden. Het vormt ook een populair thema voor satire. In de televisieshow *Little Britain* is er bijvoorbeeld een vrouwelijk karakter dat aan een bureau zit. Als haar klanten beschrijven wat ze willen, begint ze te typen met elke keer hetzelfde antwoord dat volgt: “computer zegt nee”. De uitdrukking is iconisch geworden voor het beschrijven van die situaties waarin mensen worden geleid door wat een computer hen ook maar vertelt.

Met de digitalisering van onze wereld zien we nu dat digitale systemen en omvangrijke databases betrokken raken bij steeds belangrijkere beslissingen. Verkeersovertredingen en fraude worden tegenwoordig op grote schaal digitaal geïdentificeerd en (semi-)automatisch afgehandeld. Bij het herkennen en voorkomen van allerlei risico's zijn algoritmen en data centraal komen te staan. Zo zien we binnen de politie steeds meer voorbeelden van digitale systemen om criminaliteit te voorspellen en worden rechters wereldwijd steeds vaker voorzien van digitale risico-inschattingen bij het straffen en vrijlaten van verdachten. Daarnaast worden digitale technologieën ook gebruikt om kwetsbare personen die mogelijk in gevaar zijn beter te kunnen signaleren - zoals minderjarigen die het risico lopen om verwaarloosd of mishandeld te worden. De term *digital security governance* verwijst naar deze nieuwe benadering van werken aan veiligheid en het gebruik van allerlei digitale

technologieën voor het verzamelen, verwerken, opslaan, delen en toepassen van informatie. Het resultaat is dat digitale technologieën direct betrokken zijn geraakt bij beslissingen die een substantiële impact hebben op het welzijn van individuen en groepen – beslissingen over arrestaties, straffen, vrijlatingen, toegang en uitsluiting. Een ontwikkeling die op verschillende plekken geregeld leidt tot grote zorgen.

Tegen een achtergrond waarin organisaties voortdurend op zoek zijn naar geavanceerde technologieën voor veiligheid en andere wezenlijke kwesties, is de bereidheid van mensen om naar een computer te luisteren zowel intrigerend als verontrustend. Hoe ver kunnen we gaan? En waarom? Het volgen van digitale technologieën om vervolgens het verkeerde te doen is paradoxaal. Ten eerste omdat deze technologieën doorgaans zijn geïntroduceerd met het idee dat zij onze cognitie en besluiten juist verbeteren. Ten tweede omdat we “dingen” intuïtief nog vaak zien als neutrale en passieve instrumenten die we kunnen gebruiken om onze doelen te bereiken.

Eerder lieten wetenschappers al zien dat technologie veel meer is dan een instrument en ons op allerlei manieren kan beïnvloeden. Zo werd bijvoorbeeld duidelijk dat technologieën ons kunnen dwingen (Latour, 1992), overtuigen (Fogg, 2003), subtiel sturen (*nudging*) (Thaler & Sunstein, 2008) of misleiden (Castelfranchi, 2000). Toch bieden deze inzichten onvoldoende handvatten om bovengenoemde momenten van verregaande volgzzaamheid te begrijpen. En ook de lens van menselijk vertrouwen in technologie lijkt een onvolledig beeld te geven van wat hier aan de hand is – zelfs als we de output van een computer niet geloven, negeren we soms toch ons eigen oordeel en doen we wat ons digitaal wordt gezegd. Het lijkt erop dat digitale technologieën hun rol als vertrouwde assistenten kunnen ontgroeien en ons op dominantere wijze kunnen sturen in wat we denken en doen. Dit brengt ons bij het idee van digitale autoriteit.

HET GEZAG DER DINGEN

Autoriteit (of *gezag*) is een omstreden begrip. Het wordt op veel verschillende manieren gebruikt en verwijst naar een diversiteit aan actoren – van experts en deskundigen tot degenen die formele ambten bekleden, zoals politieagenten en rechters. Wat zij gemeen hebben is dat hun communicaties mensen ertoe aan kunnen zetten te accepteren wat er ook tegen hen wordt gezegd. Autoriteiten beschikken over de capaciteit om dingen gedaan te krijgen. Het woord is afgeleid van het Latijnse *auctoritas*, dat in brede zin verwijst naar het vermogen om een dwingende kracht op mensen uit te oefenen “groter dan die van advies, maar minder dan die van bevel” (Mommsen, 1887 c.f. Arendt 1961: 123). Autoriteit is een speciaal soort invloed, verschillend van zowel dwang als overreding. Met autoriteit hoeft iemand niet te overtuigen, te dreigen of te dwingen om iets te bereiken. In plaats daarvan accepteren mensen wat autoriteiten zeggen uit vertrouwen en respect, of kunnen ze om strategische redenen doen alsof dit het geval is (Lincoln, 1994: 9). Autoriteit berust boven

alles op de vrijwillige erkenning van het subject dat naar de ander moet worden geluisterd. Het kan ook altijd worden betwist en ontkend. Maar zodra mensen autoriteit erkennen zullen zij geneigd zijn om alles dat tegen hen gezegd wordt te volgen, ongeacht hun eigen oordeel. Het lijkt dan net alsof we via een ander in de wereld denken en handelen.

In dit boek beschrijf ik autoriteit vanuit een interactionistisch perspectief, als iets dat voortkomt uit de uitwisselingen tussen ten minste twee actoren in een bepaalde setting. Iemand (of *iets*) heeft niet zomaar autoriteit. Het bestaat niet bij de gratie van individuele kwaliteiten zoals ambt, kennis, rijkdom, afkomst of charisma. Autoriteit is een gevolg, of een uitkomst van sociale interacties. Om dit idee verder te kunnen vatten onderscheid ik in hoofdstuk 3, een drietal conceptuele lagen. Ten eerste beschrijf ik autoriteit als een relationeel concept. Autoriteit impliceert een asymmetrische relatie waarin de een door de ander(en) niet als gelijke wordt beschouwd, maar in zekere zin als meerdere. Ten tweede is er een situationele laag. De erkenning van autoriteit ligt verankerd in de aard van de situatie. Er is een hele reeks signalen (items, gebaren, gedragingen, etc.) die autoriteit binnen een bepaalde setting overbrengt. Deze signalen kunnen mensen ertoe aanzetten een gehoorzame of accepterende rol aan te nemen. Ten derde bestaat er een culturele laag van autoriteit. Onze ideeën over autoriteit zijn cultureel gevormd. Gedurende ons leven, binnen het gezin, op school en later in de openbare en professionele ruimte, leren we wie we kunnen vertrouwen en dienen te respecteren. Op deze manier worden onze ideeën en verwachtingen van autoriteit geconditioneerd. Kortom, dit gelaagde perspectief op autoriteit benadrukt dat autoriteit wordt geconstrueerd in relaties, situaties en culturen. Hiermee verwordt autoriteit niet tot een statische of stabiele entiteit, maar staat de dynamiek en haar open essentie centraal. Dit doet recht aan het feit dat claims en erkenning van autoriteit in de loop van de tijd kunnen conflicteren en veranderen. En zo ontstaat ook ruimte voor reflectie op potentiële nieuwe autoriteiten.

Het concept van autoriteit is lange tijd perifeer gebleven binnen ons denken over technologie. Hoewel onze levens vol zitten met autoriteiten en de autoriteitsrelatie zo oud is als de menselijke beschaving zelf, is de term zelden gebruikt in relatie tot technologieën. In de hoofden van velen waren autoriteiten mensen zoals politieagenten, leraren, ouders of experts op een bepaald onderwerp – en niet dingen. Hoogstens werd gesproken over technologieën als instrument van menselijke autoriteiten of als factor die nieuwe autoriteitsrelaties tussen mensen tot stand kan brengen. Maar de autoriteit van technologieën zelf is echter vaak over het hoofd gezien of afgewezen. Toen nog nauwelijks sprake was van interactie tussen mensen en technologieën, was dit overigens heel logisch: er viel niets te gehoorzamen of accepteren.

EEN STUDIE NAAR DIGITALE AUTORITEIT

Nu digitale technologieën steeds meer kunnen, onderling verbonden zijn en toegang hebben tot een wereld aan data, zien we dat er nu wel af en toe gerichter gesproken wordt over de autoriteit van digitale technologieën. Ondanks dat dit vaak nog de vorm heeft van vluchtige verwijzingen, laat dit een voorzichtige verschuiving zien in het denken over de aard van autoriteit. Kennelijk is het niet meer zo vreemd om de term expliciet te koppelen aan technologische artefacten. Maar hoewel het toepassen van het idee van autoriteit op technologieën interessante wegen kan bieden voor het analyseren van hun sterke invloed op wat mensen denken en doen, brengt het ook nieuwe uitdagingen met zich mee en vereist het meer conceptuele helderheid. Want hoe moeten we deze digitale vorm van autoriteit begrijpen? Waar komt het vandaan en hoe is het opgebouwd? Hoewel technologieën nu misschien kunnen communiceren, ontberen ze nog steeds verstand en hebben ze geen intentie zoals mensen, laat staan enige vorm van vrijheid (Verbeek, 2011: 13).

De centrale vraag van dit proefschrift is:

In hoeverre kunnen digitale technologieën als autoriteiten worden beschouwd?

Het doel is om een bepaalde vorm van dwingende technologische invloed op menselijke beslissingen te conceptualiseren, om het fenomeen van mensen die voorbij hun eigen oordeel naar digitale technologieën luisteren beter te begrijpen. Een dergelijk nieuw begrip helpt ook om potentiële risico's zichtbaar te maken en deze te beperken. Hiervoor introduceer ik het idee van *digitale autoriteit*. Hiermee verwijs ik niet naar de aanwezigheid van menselijke autoriteit in de digitale wereld of gezaghebbende communicatie van mensen *via* computers. Ook gaat het hier niet over digitale technologieën die geprogrammeerd zijn om zelf autonome beslissingen te nemen in situaties waarin mensen afwezig zijn - denk bijvoorbeeld aan een computer die de werking van machines in een fabriek regelt of een volledig autonoom wapensysteem dat doelen aanvalt zonder enige menselijke tussenkomst. In plaats daarvan draait het bij digitale autoriteit om de technologieën zelf en de sterke invloed van hun communicaties op het menselijk handelen en denken. Specifieker verwijst het naar digitale technologieën en hun capaciteit om output te generen die door mensen als vanzelfsprekend wordt geaccepteerd uit vertrouwen, respect of hun bereidheid te doen alsof dit zo is. Zo kan digitale autoriteit zowel gaan over een bepaald soort invloed van digitale technologieën op mensen, als over de technologieën waarvan algemeen wordt aangenomen dat ze een dergelijke invloed kunnen uitoefenen. Dit omvat een breed scala aan digitale technologieën en richt niet uitsluitend op robots en geavanceerde vormen van artificiële intelligentie (AI). Daarnaast geldt dat digitale autoriteit niet noodzakelijkerwijs opzettelijk ontworpen hoeft te zijn. En hoewel ik in deze dissertatie primair verwijs naar het veiligheidsdomein vanwege de potentieel significante impact op de levens van mensen, kan digitale autoriteit overal voorkomen. Empirische onderzoeken en observaties vormen een

belangrijk fundament voor deze filosofische verkenning van digitale autoriteit en helpen om dit fenomeen verder te begrijpen.

Om de ontwikkeling en werking van digitale autoriteit te kunnen begrijpen ga ik in het tweede deel van het boek (hoofdstuk 4, 5 en 6) dieper in op de drie lagen van autoriteit: *relatie*, *situatie* en *cultuur*. Zij vormen de analytische kern van deze filosofische verkenning naar digitale autoriteit.

AUTORITEITSRELATIES MET TECHNOLOGIEËN

Hoofdstuk 4 richt zich op de relatie tussen mensen en technologieën. Centraal staat de vraag of hier ook autoriteitsrelaties zouden kunnen ontstaan, en hoe deze er dan uit zouden zien. Technologie laat zich niet goed begrijpen wanneer we deze benaderen als slechts een middel of instrument om onze doelen mee te bereiken. In plaats daarvan zette Don Ihde (1990) de relaties tussen mens en technologie centraal. Zo wordt zichtbaar hoe technologie mede vormgeeft aan onze ervaringen van de werkelijkheid en aan hoe wij hierbinnen aanwezig zijn. Ihde beschrijft diverse relaties tussen mens en technologie. Zo laat hij onder andere zien hoe we de wereld kunnen ervaren *via* een technologie - zoals bijvoorbeeld het geval is met een bril die door ons wordt “ingelijfd” en zo beïnvloedt hoe we alles om ons heen ervaren, of een thermometer die ons een bepaalde weergave van de wereld laat zien die wij vervolgens moeten interpreteren. Daarnaast beschrijft Ihde nog de zogeheten *alteriteitsrelatie* waarin technologische artefacten verworden tot quasi-anderen. Dit is het geval met veel digitale technologieën waar we nu mee *interacteren*.

Autoriteitsrelaties met technologieën gaan nog een stap verder. Hierbinnen worden menselijke percepties en handelingen sterk gevormd door een technologie die niet alleen een bron van interactie vormt, maar daarbij ook centraal staat in onze relatie met de werkelijkheid. Dit leert ons dat lang niet alle technologische artefacten potentiële autoriteiten zijn. We zullen een bril, een pen of een bijl niet als autoriteiten herkennen, omdat ze geen bron van interactie zijn. Maar zoals gezegd kunnen we met digitale technologieën een gevoel van “anderheid” ervaren, vanwege hun schijnbare autonomie en de interacties die ze veroorzaken. Meer specifiek zijn technologieën met hun digitalisering gaan communiceren. Lange tijd werd autoriteit gezien als een puur sociale relatie, juist omdat het alleen mensen waren die mededelingen deden die geloofd of gevolgd moesten worden. Maar met de komst van digitale technologie wordt de communicatie van dingen ook actief herkend door mensen – of het nu een pieptoon, een kunstmatige stem of een tekst op een scherm is. Digitale technologieën zijn een expliciete bron van communicatie en interactie geworden. Het bijzondere aan autoriteitsrelaties is echter dat zij fundamenteel asymmetrisch van aard zijn. Er zal dus een zekere mate van ongelijkheid tussen de mens en de technologie moeten worden ervaren die ons motiveert om te doen of te denken wat ons digitaal wordt verteld. In een autoriteitsrelatie is de invloed die technologieën uitoefenen

op menselijke percepties en gedrag zo dwingend dat deze kenmerkend is voor de relatie tussen mensen en hun wereld. We doen en denken dingen simpelweg omdat de technologie ons dat vertelt. Je kunt zeggen dat we in deze configuraties de wereld ervaren *door of volgens* een technologie.

Maar hoe zouden we het bestaan van deze asymmetrische relatie tussen mens en technologie kunnen begrijpen? Hoewel dit een vraag is die het hele boek beslaat, bespreek ik in hoofdstuk 4 enkele initiële elementen die kunnen worden beschouwd als constituerende onderdelen van digitale autoriteitsrelaties. Centraal staat hierbij het idee van *gedifferentieerde toegang* – autoriteit draait vaak om het idee dat iemand een bijzondere toegang heeft tot belangrijke kennis of tot specifieke machtsinstrumenten. Net als bij menselijke epistemische autoriteiten (zoals wetenschappers, experts en ervaringsdeskundigen) ervaren we bij digitale technologieën vaak dat ze toegang hebben tot een wereld aan unieke inzichten en informatie waar wij zelf niet bij kunnen. Ook kunnen we een bepaalde machtsongelijkheid ervaren in onze interacties met een digitale technologie – bijvoorbeeld zodra digitale systemen binnen een organisatie verplicht dienen te worden geraadpleegd bij het nemen van een beslissing of een technologie zo is geprogrammeerd dat het afwijkend gedrag simpelweg niet toelaat – ook wel *technoregulering* genoemd. Iemand kan het dan misschien niet eens zijn met de technologie, maar er uiteindelijk toch voor kiezen om te doen wat wordt opgedragen, omdat de technologie anders gehoorzaamheid wel zal afdwingen: de poort gaat niet open, het document wordt niet verzonden of het systeem reageert gewoon niet. Omdat technologieën geen emotionele wezens zijn die medelijden hebben met een ander, kan dit idee van mogelijke dwang extra krachtig zijn: technologieën maken namelijk geen uitzonderingen.

Deze verschillen in toegang tot kennis en formele macht maken dat ook binnen digitale autoriteitsrelaties overtuiging en dwang op de achtergrond een belangrijke rol spelen. De latente aanwezigheid van deze andere manieren van invloed geeft de asymmetrie van autoriteitsrelaties tussen mensen en technologieën vorm en zorgt voor acceptatie, vertrouwen, respect of de bereidheid te doen alsof dit zo zou zijn. Het leidt tot de veronderstelling dat in gevallen van twijfel of onwil de technologie in staat is om ons te overtuigen door middel van de juiste argumenten of simpelweg kan dwingen om mee te werken. In de praktijk zijn technologieën natuurlijk vaak niet geprogrammeerd om zichzelf uit te leggen en zijn ze nog vaak afhankelijk van anderen om overtuigende argumenten aan te dragen voor hun output of om dwang uit te oefenen in tijden van verzet. Maar waar het om draait is het idee dat mensen denken dat er logischerwijs bewijs bestaat voor digitale claims of dat de technologie in staat is om medewerking eventueel af te dwingen. Het is deze veronderstelling die mensen vooraf overtuigt dat naar de computer geluisterd dient te worden. Echter, als belangrijke herinnering, wanneer dwang of overreding daadwerkelijk wordt toegepast, zou er geen sprake zijn van digitale autoriteit maar slechts van digitale dwang of overreding.

DIGITALE AUTORITEITSSITUATIES

Hoofdstuk 5 gaat vervolgens in op mogelijke autoriteitssituaties met digitale technologieën. Hannah Arendt (1963) en Stanley Milgram (1974) betoogden decennia geleden al dat gezag berust op de aard van de situatie. In het kort: als mensen een autoriteitssituatie herkennen zullen zij een gehoorzame rol aannemen en doen of denken wat hen wordt gezegd. Om dit beter te begrijpen beschreef ik de noties van definitie van de situatie en scripts. Wanneer mensen in een nieuwe situatie terechtkomen, vragen ze zich af: wat is er aan de hand? En hoe moet ik me hier gedragen? (Goffman, 1986; Van den Berg, 2009). Om deze vragen te beantwoorden, vertrouwen mensen veelal op aangeleerde scripts. Scripts zijn kennispakketten waarmee mensen door hun dagelijks leven kunnen navigeren. Ze specificeren de rollen van personen en objecten binnen een bepaalde situatie en bevatten bepaalde actiepatronen (Schank & Abelson, 1977). Een omgeving met de juiste signalen zal vervolgens een specifiek script activeren en bepaalde ervaringen en gedragingen stimuleren. Zo kunnen we ook begrijpen hoe mensen autoriteit herkennen en hierop reageren: met de juiste aanwijzingen kunnen scripts worden geactiveerd die mensen ertoe aanzetten iets als een autoriteitssituatie te definiëren om vervolgens een accepterende, eerbiedige of zelfs onderdanige rol aan te nemen. Er bestaan tal van gezaghebbende items, plaatsen, symbolen of gedragingen die als situationele signalen fungeren die individuen helpen autoriteit te herkennen en daarmee een mentale shortcut bieden om gepast te reageren. Zo wordt duidelijk dat degenen die autoriteit claimen en zich op een bepaalde manier aan ons presenteren, een situatie effectief trachten te vormen voor de ander: met behulp van de juiste signalen stimuleren ze iemand om de situatie te definiëren als autoriteitssituatie.

Maar wat zegt dit over digitale autoriteit? Vooralsnog ging het immers enkel over autoriteitssituaties met andere mensen. Toch lijken ook digitale technologieën gebruik te maken van dezelfde sociale scripts en op vergelijkbare wijze autoriteitssituaties voor ons te kunnen creëren. Uit onderzoek komt naar voren dat digitale technologieën sterke sociale reacties bij ons kunnen oproepen – het soort reacties dat we traditioneel reserveren voor interacties met andere mensen (Reeves & Nass, 1996; Turkle, 2005). En hiervoor hoeven de sociale signalen van technologieën niet bijzonder geavanceerd te zijn en hoeven mensen zich niet bewust te zijn van hun sociale gedragingen. Het is daarom misschien niet langer nodig om een sterk onderscheid te maken tussen menselijke en technologische autoriteitssituaties: door de juiste presentatie kunnen technologieën op een zeer vergelijkbare wijze centraal komen te staan in autoriteitssituaties als menselijke autoriteiten dat doen. Zo zien we nu dat digitale technologieën traditionele autoriteitsrollen vervullen die voorheen werden vervuld door menselijke gezagsdragers, zoals artsen, leraren, deskundigen of politiefunctionarissen. Ook worden er allerlei titels gegeven die dit nog eens benadrukken: technologieën die bijvoorbeeld als “expert” of “specialist” worden bestempeld. En soms zijn er meer verfijnde signalen, denk aan een bepaalde taalstijl of autoriteitssymbolen zoals certificaten of logo's.

Al deze signalen die een rol spelen bij de constructie van menselijke autoriteitssituaties kunnen in interacties met digitale technologieën vergelijkbare reacties oproepen.

Kortom, digitale technologieën hoeven er niet precies hetzelfde uit te zien en te handelen als mensen, om subjecten ertoe te brengen autoriteitssituaties te herkennen en er dienovereenkomstig op te reageren. En het maakt niet uit of hun kennis of bevoegdheden echt zijn – of technologieën nu legitieme autoriteiten zijn of niet. Waar het om gaat is de gezaghebbende indruk die ze kunnen maken die een autoriteitsscript kan triggeren. In deze situaties, een belangrijke toevoeging, functioneren digitale technologieën niet als proxy voor menselijke autoriteit. Uit verschillende experimenten is gebleken dat individuen zich niet richten op degenen achter de computer - de programmeurs of hun bazen. Integendeel, teruggrijpend op de eerdere discussie over de relaties tussen mens en technologie: het is de digitale technologie waarop we onze sociale reacties richten.

EEN CULTUUR VOOR DIGITALE AUTORITEIT

Hoofdstuk 6 richt zich op de vraag in hoeverre er plaats is voor digitale autoriteit in hedendaagse culturen. Er bestaan aanzienlijke verschillen tussen culturen in de mate waarin mensen hiërarchieën en ongelijkheid in sociale relaties leren accepteren (Hofstede et al, 2010). Sommige culturen zijn meer op autoriteit georiënteerd dan anderen. Ook groeien mensen op in verschillende omgevingen met eigen ideeën over technologische innovatie. Toch zijn de voorbeelden van mensen die blind hun digitale technologieën volgen wereldwijd verspreid. Als we vervolgens inzoomen op hoe *binnen* samenlevingen ideeën van digitale autoriteit zouden kunnen postvatten ontdekken we twee conflicterende technologische narratieven. Enerzijds leren mensen over hoe technologie elk probleem kan oplossen en alles makkelijker kan maken. Anderzijds is er doorlopend aandacht voor de gevaren en risico's van technologische innovatie. Binnen het veiligheidsdomein zien we bijvoorbeeld hoe digitale technologieën vaak worden beschouwd als waardevolle oplossing voor misdaad en allerhande dreigingen, maar ook als een potentieel gevaar voor het individu. Onze moderne culturen zitten vol met elementen van deze twee fundamenteel verschillende verhalen van acceptatie en afwijzing, van moderne technologie als iets goeds of slechts. Wat zij gemeen hebben is een deterministische kijk op de invloed van nieuwe technologieën op samenlevingen – er wordt aangenomen dat zij de wereld kunnen redden of ruïneren. Het benadrukken van de vormende of transformerende capaciteiten van technologieën kan zorgen voor ontzag en de constructie van digitale autoriteitsrelaties stimuleren. Deze positieve en negatieve ideeën zullen er echter ook toe leiden dat mensen op tegenstrijdige manieren reageren op digitale technologieën. Over het algemeen is de moderne samenleving dus een omgeving van ambivalentie die digitale autoriteit zowel kan stimuleren als beperken.

Maar er is nog een diepere culturele laag. Er is al veel geschreven over dat de moderne tijd wordt gekenmerkt door het toegenomen bewustzijn van mensen over de onzekerheid en complexiteit van hun wereld. Dit heeft onder andere te maken met de erosie van traditie en religie als ook met het feit dat allerhande nieuwe gevaren steeds meer het gevolg blijken van menselijk handelen. Mensen maken zich hierdoor meer zorgen en voelen een groeiende verantwoordelijkheid voor wat de toekomst brengt. Aanvankelijk gaf de imposante ontwikkeling van wetenschap en technologie in moderne samenlevingen mensen de indruk van meer zekerheid en controle. Maar al snel leerde men dat binnen de wetenschap vaak onenigheid bestaat en kennis voortdurend wordt herzien. Ook dienden technologische catastrofes zoals kernrampen en grootschalige vervuiling zich aan. Zo creëerde de groeiende maatschappelijke rol van wetenschap en technologie uiteindelijk vooral een reeks nieuwe onzekerheden en angsten. Ulrich Beck (1992a) en Anthony Giddens (1991) stelden dat zo een risicocultuur of -maatschappij ontwikkelde, waarin een algemeen klimaat van onzekerheid heerst en een obsessie met toekomstige schade die ons zou kunnen overkomen.

Onder deze moderne omstandigheden zijn mensen steeds meer bezig met zoeken naar en creëren van zekerheid. Individuen nemen een *reflexieve* houding aan bij het maken van keuzes en handelen in hun wereld. Een onderdeel hiervan is dat men voortdurend actief is met het monitoren wie er kan worden vertrouwd en gevolgd. In een samenleving na religie en traditie, waar risico's nu worden ervaren als menselijke verantwoordelijkheden, moeten individuen op zoek naar nieuwe relevante expertise en betrouwbare anderen om naar te luisteren. Zo heeft het bewustzijn van de beperkingen van wetenschappelijke kennis en twijfel aan traditioneel gezag ruimte gemaakt voor de opkomst van andere autoriteiten. Individuen staan in de moderne tijd misschien wel meer dan ooit open voor nieuwe autoriteiten die hen de wereld kunnen helpen leren kennen en controleren - die ongemakkelijke gevoelens van onzekerheid en twijfel kunnen helpen tegengaan. Voor de meesten zal dit betekenen dat zij uiteindelijk hun weg zullen vinden door te vertrouwen op een reeks verschillende autoriteiten. En zodra digitale technologieën worden gepresenteerd als effectieve oplossingen tegen onzekerheid en complexiteit verworden moderne samenlevingen dan ook tot een vruchtbare omgeving voor het ontstaan van digitale autoriteit.

GRENZEN

Digitale autoriteit is echter geen gegeven. In het laatste deel van deze dissertatie beschrijf ik dat er veel zaken zijn die de constructie van digitale autoriteit kunnen belemmeren. Zo bekleden deze technologieën bijvoorbeeld geen officiële ambten en hebben zij niet dezelfde formele bevoegdheden die de positie van veel menselijke autoriteiten versterken. Daarnaast zijn er ook nog steeds tal van gebieden waar individuen over voldoende kennis

en ervaring beschikken om op hun eigen oordeel te vertrouwen. En helpen fouten van een technologie ook niet bij de ontwikkeling van digitale autoriteit. Digitale autoriteit is ook niet vrijgesteld van het reflexieve proces waarvan zij zelf ook het product is - mensen kunnen altijd op zoek gaan naar nieuwe bronnen van zekerheid.

Daarnaast geldt dat digitale autoriteit wordt begrensd doordat bij technologieën vooralsnog het vermogen ontbreekt om momenten van potentiële crisis waar te nemen en hier effectief op te reageren. Dit zijn de momenten waarop mensen, expliciet of impliciet, twijfelen of nee zeggen tegen een technologie. Technologieën ontberen een geest en hebben geen intentie of autonomie zoals mensen dat hebben, waardoor ze slechts een beperkt actierepertoire hebben om met weerstand en uitdaging om te gaan. Vooralsnog reageren de meeste technologieën slechts op de manier waarop ze zijn geprogrammeerd, waardoor ze vaak alleen maar in staat zijn hun output te herhalen aan een onwillig menselijk subject. Digitale technologieën zijn over het algemeen afhankelijk van anderen om een publiek weer voor zich te winnen – bijvoorbeeld van mensen die potentiële voordelen van naleving kunnen aangeven of door de negatieve gevolgen van weigering onder de aandacht te brengen. Dit maakt digitale autoriteit over het algemeen kwetsbaar en beperkt. Want wat als die anderen afwezig zijn?

Nu digitale innovatie steeds meer samenvalt met de voortdurende ontwikkeling van artificiële intelligentie, kunnen de grenzen van digitale autoriteit de komende decennia makkelijk verschuiven. Digitale technologieën zullen zich steeds gezaghebbender presenteren en steeds veerkrachtiger worden bij menselijk verzet. Tegelijkertijd laten de eerste digitale autoriteiten zich nu al duidelijk zien. Hoewel digitale technologieën ons de mogelijkheid bieden om gerichte beslissingen te nemen waar we anders niet toe in staat waren geweest, hebben ze ook tot nieuwe incidenten en risico's geleid. Een rode draad in deze dissertatie zijn de verhalen waarin mensen duidelijk te ver gaan in het volgen van digitale autoriteiten. Door individuele gevoelens van doelmatigheid (*efficacy*) en bevoegdheid (*empowerment*) te vergroten kunnen sommige potentiële valkuilen van digitale autoriteit worden verzacht. Zo kunnen bijvoorbeeld onderwijs, training en het uitdragen van duidelijke richtlijnen individuele oordeel- en besluitvorming stimuleren en mogelijk maken. Ten slotte blijft het belangrijk om te erkennen dat er altijd dingen kunnen gebeuren in omgevingen waarin er soms eenvoudigweg geen rationeel alternatief is voor het volgen van een digitale technologie. Individuen, organisaties en samenlevingen moeten zich hierop inrichten en structureel aandacht hebben voor de eventuele schadelijke impact van digitale autoriteit.

CONCLUSIE

In het laatste hoofdstuk concludeer ik dat digitale technologieën ons dwingen om veel opvattingen over de aard van artefacten en mensen te heroverwegen. Mensen zijn niet uitsluitend “gebruikers” en dingen niet slechts instrumenten. Dat blijkt ook uit die gevallen waarin we technologie respecteren en volgen. Het concept autoriteit helpt om de verregaande bereidheid van mensen om te luisteren naar een digitale technologie beter te begrijpen. En een analyse van de verschillende autoriteitslagen kan het gezaghebbende potentieel van specifieke technologieën verder verduidelijken. Digitale autoriteit is dan het resultaat van het samenvallen van een digitale autoriteitsrelatie, situatie en cultureel specifieke geconditioneerde verwachtingen en ideeën. Dit zorgt ervoor dat interactie met een digitale technologie zorgt voor vertrouwen, respect, acceptatie of gehoorzaamheid onder mensen.

Autoriteit is dus niet dood, zoals we soms lezen, maar lijkt zich te verplaatsen naar nieuwe plekken, zoals de wereld van de technologie. Dit biedt kansen voor betere besluiten maar zorgt dus ook voor wezenlijke nieuwe gevaren. Zeker op plaatsen waar de uitwisseling tussen mens en technologie de basis vormt voor belangrijke beslissingen die het welzijn van het individu direct raken - denk aan het veiligheidsdomein – is voorzichtigheid geboden. We zullen daarom na moeten blijven denken over verantwoorde ontwikkeling van digitale autoriteit in onze wereld. De laatste paragrafen zijn hier een eerste aanzet toe. Zo bespreek ik de waarde van verder empirisch onderzoek en het belang van de ontwikkeling van een ethiek van digitale autoriteit. Ook sta ik stil bij de complexiteit van verantwoordelijkheid, aansprakelijkheid en schuld in een context van digitale autoriteit.

