



BTO 2017.041 | Mei 2017

BTO rapport

BTO Trends

Diepte-artikel

Data

BTO

Diepte-artikel data

BTO 2017.041 | Mei 2017

Opdrachtnummer

400554/208/001

Projectmanager

Jos Frijns

Opdrachtgever

BTO - Thematisch onderzoek - Trends

Kwaliteitsborger

Andrew Segrave

Auteur

Henk-Jan van Alphen

Verzonden aan

Dit rapport is verspreid onder BTO-participanten en is openbaar.

Jaar van publicatie
2017

Meer informatie
MSc, Henk-Jan van Alphen
T +31 30 606 9626
E Henk-Jan.van.Alphen@kwrwater.nl

Keywords
Digitalisering; Data; Trends;
Toekomstverkenning

PO Box 1072
3430 BB Nieuwegein
The Netherlands

T +31 (0)30 60 69 511
F +31 (0)30 60 61 165
E info@kwrwater.nl
I www.kwrwater.nl



Watercycle
Research
Institute

BTO 2017.041 | Mei 2017 © KWR

Alle rechten voorbehouden.

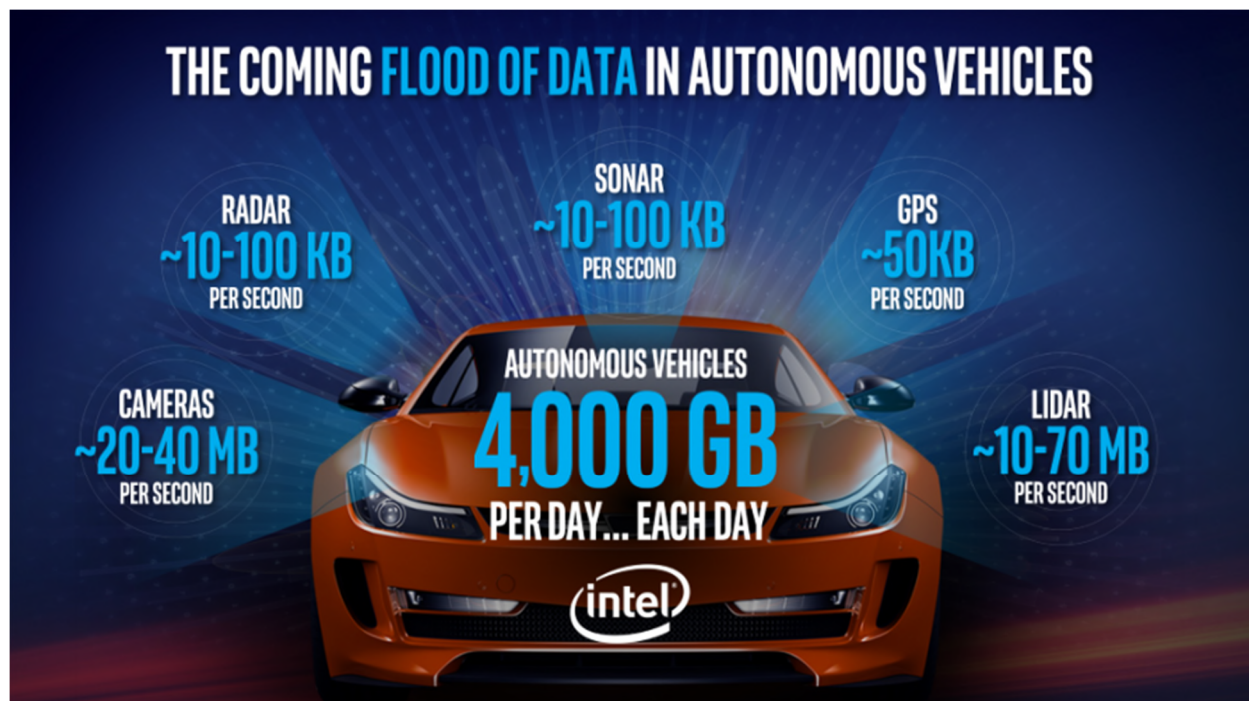
Niets uit deze uitgave mag worden vervoelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen, of enig andere manier, zonder voorafgaande schriftelijke toestemming van de uitgever.

BTO Managementsamenvatting

Ontwikkel voor je waterbedrijf een strategie rond het verzamelen, analyseren en toepassen van data

Auteur(s) Henk-Jan van Alphen, MSc

Vitale economische en maatschappelijke processen zijn tegenwoordig afhankelijk van ICT systemen die in essentie werken op basis van schier eindeloze reeksen nullen en enen. De digitale samenleving wordt gekenmerkt door de eenvoud waarmee grote hoeveelheid data gegenereerd, opgeslagen, gedeeld en geanalyseerd kan worden. Dit leidt in hoog tempo tot de ontwikkeling van nieuwe kennis, diensten en producten. KWR analyseerde vier trendalerts op het gebied van digitalisering: over big data, citizen science, robotisering & kunstmatige intelligentie en DNA-sequencing. Uit deze trends blijkt dat de hoeveelheid beschikbare data en het belang daarvan voor de bedrijfsvoering sterk zullen toenemen. Veel processen rondom data zullen bovendien door computers zelf kunnen worden uitgevoerd. Waterbedrijven doen er goed aan om een strategie te ontwikkelen rond het *verzamelen, opslaan, analyseren* en *toepassen* van data. Een belangrijk dilemma is of je daarbij moet werken vanuit een beoogde toepassing of dat je op basis van je data toepassingen kunt ontwikkelen. Voor de opslag van data gelden privacyregels die in 2018 aangepast gaan worden, maar naar verwachting ook daarna in ontwikkeling blijven. Ten slotte moeten data- en ICT-vaardigheden ook onderdeel worden van strategische personeelsbeleid om de mogelijkheden van ICT te kunnen benutten.



Voorbeeld hoe nieuwe technologie kan leiden tot een enorme toename van de hoeveelheid data. Iedere zelfsturende auto produceert 4000 Gb aan data.

Belang: integraal overzicht van de verschillende trends op het gebied van data

In de afgelopen jaren heeft KWR een aantal trends onderzocht die in verband staan met data en datamanagement: over big data, citizen science, robotisering & kunstmatige intelligentie en DNA-sequencing. Door de inzichten omtrent data uit deze vier onderzoeken te integreren ontstaat een duidelijker beeld van de kansen en uitdagingen die de trends bieden op het gebied van data en datamanagement.

Aanpak: analyse van trendalerts aangevuld met recente literatuur

Voor dit onderzoek hebben de onderzoekers van KWR de vier trendalerts geanalyseerd op hun impact op vraagstukken op het gebied van data en datamanagement. De analyse is aangevuld met een bureaustudie van recente relevante literatuur.

Resultaten: data nemen toe in hoeveelheid en in het belang voor (water)organisaties

De vier onderzochte trends wijzen op een toename van de hoeveelheid data en bieden strategische uitdagingen voor het verzamelen, opslaan, analyseren en toepassen van data in de eigen organisatie. Veel processen rondom data zullen bovendien door computers zelf kunnen worden uitgevoerd. Drinkwaterbedrijven doen er goed aan een data-management-strategie te ontwikkelen. Een goed startpunt kan zijn om in kaart te brengen welke processen door data zouden kunnen worden ondersteund. Ook dient datamanagement onderdeel te zijn van het strategisch personeelsbeleid, vanwege de specifieke vaardigheden die ICT en data management vragen.

Implementatie: gebruik inzichten voor een strategische discussie over data en data management

Dit artikel, samen met de onderliggende trendalerts, is bedoeld om onderwerpen te agenderen voor aanvullend onderzoek of strategie-ontwikkeling. Waterbedrijven kunnen de inzichten en dilemma's uit dit artikel gebruiken om kritisch naar de (potentiële) rol van data in hun eigen organisatie te kijken en een strategische discussie over dit onderwerp op te starten.

Rapport

Dit onderzoek is beschreven in rapport *Diepteartikel Data* (BTO-2017.041). De vier onderliggende trendalerts zijn:

- BTO 2015.070 Trendalert 28: Gentechnologie: acceptatie en risico's
- BTO 2015.067 Trendalert 27: Slimme robots: waar blijven ze nou?
- BTO 2013.071 Big Data: the art of the possible
- BTO 2016.034 Trendalert: Citizen Science: burgers meten steeds meer zelf.

Meer informatie

Henk-Jan van Alphen, MSc

T 0655281776

E Henk-Jan.van.Alphen@kwrwater.nl

KWR

PO Box 1072

3430 BB Nieuwegein

The Netherlands

Inleiding: digitalisering

Digitalisering is een fascinerend fenomeen waarvan de impact op de samenleving nauwelijks te overschatten valt. Vitale economische en maatschappelijke processen zijn tegenwoordig afhankelijk van ICT systemen die in essentie werken op basis van schier eindeloze reeksen nullen en enen. De digitale samenleving wordt gekenmerkt door de eenvoud waarmee grote hoeveelheden data gegenereerd, opgeslagen, gedeeld en geanalyseerd kan worden. Dit leidt in hoog tempo tot de ontwikkeling van nieuwe kennis, diensten en producten.

In de afgelopen jaren heeft KWR trends onderzocht die op termijn zullen leiden tot hoge eisen aan de digitale infrastructuur van de drinkwaterbedrijven. Elke van deze trends vraagt om een strategische respons van de drinkwaterbedrijven op het gebied van data. Daarom heeft de Themagroep Trends KWR gevraagd een diepte-artikel te schrijven waarin de strategische vraagstukken rondom data nader worden besproken. Deze bestudeerde trends zijn:

- Big data
- Citizen Science
- Robotisering en Kunstmatige Intelligentie
- DNA Sequencing

In dit artikel bespreken we de implicaties van elk van de vier ontwikkelingen voor toekomstig databeheer en brengen we in kaart welke strategische uitdagingen dit voor de watersector biedt.

Data

De vier genoemde trends hebben alle betrekking op de manier waarop we in de toekomst met data omgaan. Big Data gaat over hoe grote hoeveelheden data ons meer kunnen leren over onze omgeving en ons misschien in staat stellen om gedrag en ontwikkelingen te voorspellen. Citizen science bestaat voor een deel uit data-verzameling (en soms analyse) door burgers en hoe professionele onderzoekers daar gebruik van kunnen maken. Robotisering en Kunstmatige Intelligentie gaat over hoe computers kunnen leren (door data-analyse) en zelfstandig beslissingen kunnen nemen. DNA sequencing, ten slotte, gaat over het verzamelen en analyseren van zeer grote hoeveelheden data van allerlei biologisch materiaal.

In dit artikel bespreken we de impact van de trends aan de hand van de vier stappen die centraal staan in ieder data-gerelateerd proces:

1. Verzamelen
2. Opslag
3. Analyse
4. Toepassing

Hoewel dit een logische volgorde is (geen toepassing zonder analyse, geen analyse zonder opslag, geen opslag zonder verzamelen), moet je strategisch gezien juist precies andersom kijken. Eerst moet je de toepassing bepalen: welke vragen wil je beantwoorden? Welke problemen wil je oplossen? Welke dienst wil je aanbieden? Vervolgens moet je de toepassing omzetten in een vraag die je met data-analyse kunt beantwoorden en bepalen welke analysemethoden je wilt toepassen. De methode bepaalt vervolgens de datavraag (bestaande data of nieuwe data?) en de wijze waarop die data moet worden opgeslagen. Ten slotte kan dan bepaald worden welke data moet worden verzameld en op welke manier.

1. Toepassing

De belangrijkste strategische vraag rondom databeheer, gaat eigenlijk helemaal niet over data. Het gaat erom welk probleem een organisatie wil oplossen of welk doel ze wil bereiken. Aan de hand daarvan kan dan bepaald worden welke rol data daarin kan spelen. Dat is een fundamenteel verschil met hoe Big Data vaak gepresenteerd wordt. De suggestie wordt gewekt dat je uit een hele berg data van allerlei toepassingen kunt ontwikkelen, maar dat is een al te rooskleurige voorstelling van zaken. Grote hoeveelheden data bevatten namelijk ook grote hoeveelheden statistische verbanden, we zullen daar in de paragraaf 'analyse' verder op in gaan.

Op dit moment wordt zowel bij de toepassing van Big Data als bij kunstmatige intelligentie de beste resultaten geboekt bij specifieke toepassingen waarvoor veel data aanwezig is. In de effectenstudie over robotisering en kunstmatige intelligentie maken we onderscheid tussen specialistische intelligentie (het uitvoeren van heel specifieke taken) en generalistische intelligentie (een soort algemene vaardigheid om veel verschillende taken uit te voeren).

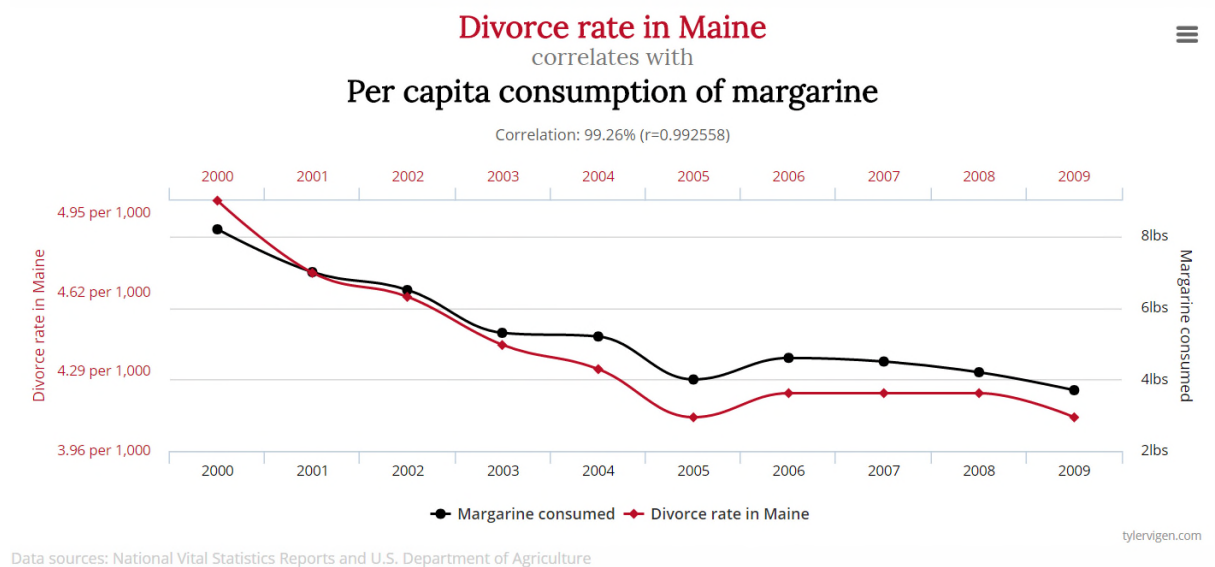
De voorbeelden die we tot nu toe hebben gezien zijn van de specialistische soort. De wereldberoemde schaakcomputer Deep Blue, kan bijvoorbeeld niet dammen. Het spraakherkenningsprogramma van Apple, Siri, kan geen kattenplaatjes herkennen. Wat de voorbeelden ook gemeen hebben, is de beschikbaarheid van onvoorstelbare hoeveelheden data (zoek maar eens naar kattenplaatjes op facebook). Als die data vervolgens zeer gestructureerd wordt aangeboden, dan is een computer instaat om daarvan te leren. Dat is machine learning, een 30 jaar oud principe waar alle kunstmatige intelligentie op gebaseerd is.

Het is te verwachten dat data-gebaseerde toepassingen in de toekomst een prominentere rol gaan spelen in de beslissingsondersteuning. Naarmate de intelligentie van systemen toeneemt, zal ook het vermogen beslissingen te ondersteunen (of zelfstandig beslissingen te nemen) met resultaten van data-analyse toenemen. Een andere interessant toepassingsgebied is het combineren van data van verschillende sectoren om bijvoorbeeld de inrichting en onderhoud van de stedelijke infrastructuur te sturen. Dit gebeurt op dit moment onder de vlag van de Smart City.

2. Analyse

2.1 Correlatie ≠ causatie

Het analyseren van data gebeurt traditioneel door computers opdracht te geven rekenkundige operaties uit laten voeren op een dataset. Deze opdrachten noemen we algoritmen. Een algoritme kan een computer bijvoorbeeld de opdracht geven te zoeken naar correlaties tussen verschillende variabelen. Een correlatie is een eerste stap in het bepalen of er wellicht een causaal verband is tussen twee variabelen. Het lastige is echter dat heel veel correlaties toevallig zijn. En dat probleem wordt groter als je meer analyses uitvoert op meer data. En dat is precies wat kunstmatige intelligentie en big data beogen. Waartoe dat kan leiden toont de onderstaande grafiek:



FIGUUR 1 : CORRELATIE TUSSEN TWEE ONGERELATEERDE VARIABLEN

Een bijna perfecte correlatie, maar waarschijnlijk geen causaal verband.

Het zo maar analyseren van grote hoeveelheden data kan dus problematisch zijn. Een analyse moet ook worden gedreven door een heldere vraagstelling. Mazzocchi (2015) betoogt dat een combinatie van ongericht zoeken en hypothesevorming het beste werkt. Hij laat ook zien dat echt ongericht zoeken helemaal niet kan, want al bij de definitie van wat als 'data' wordt gezien, spelen vooronderstellingen (hypotheses) over hoe de wereld in elkaar zit, een rol.

Een goede vraagstelling kan bijvoorbeeld voorkomen uit strategische vragen die spelen bij het management. Je hebt in je organisatie dus mensen nodig die strategische vragen uiteindelijk kunnen vertalen naar duidelijke instructies voor een computer.

Behalve dat data-analyse kan leiden tot allerlei onjuiste verbanden, kan het ook leiden tot verbanden die misschien wel juist zijn (statistisch), maar een onwenselijk effect hebben. Onder de noemer Bigot Data zijn allerlei voorbeelden bekend van ongewenste uitkomsten als gevolg van data-analyse. Google liet enige tijd hoogbetaalde functies vaker aan mannen zien dan aan vrouwen (op basis van effectiviteitsanalyse) en facebook varieerde advertenties op basis van etniciteit. In die gevallen hoeft er geen sprake te zijn van intentionaliteit bij de ontwikkelaars, een slimme computer kan ook zelf met die verbanden komen.

2.2 Machine learning

De ontwikkeling van kunstmatige intelligente laat zien dat computers soms ook zelf dingen kunnen leren zonder duidelijke algoritmen. Een computer kan bijvoorbeeld katten leren herkennen door heel veel kattenplaatjes te bekijken. Als een computer verteld krijgt op welke plaatjes wel en op welke plaatjes geen kat te zien is, kan ze het op een gegeven moment zelf herkennen. Hoe meer plaatjes, hoe beter het werkt. Maar hier zien we dat het gaat om een zeer specifieke vraag, met zeer duidelijk gestructureerde data in zeer grote hoeveelheden.

Machine learning lijkt dan wel vanzelf te gaan, er is toch heel specialistische kennis nodig om het proces in goede banen te leiden. Daarbij geldt: hoe meer 'intelligenter' de computer, hoe meer hoger de expertise van de begeleider moet zijn. Er wordt weleens de vergelijking gemaakt met een hondentrainer. De hond leert zelf, maar de trainer moet het leerproces inrichten en de hond waar nodig corrigeren.

2.3 Real time en voorspellende analyse

De meeste traditionele data-analyse is gericht op het verleden. Big data en kunstmatige intelligente maken het mogelijk om real time data-analyse te doen, die bijvoorbeeld gebruikt kan worden voor het signaleren van calamiteiten of het direct aanpassen van een systeem aan veranderingen in de omgeving. Dat vereist wel een zeer betrouwbare analyse: er is geen tijd om de resultaten te toetsen of te evalueren. Een andere, nog waardevollere vorm van analyse is de voorspellende analyse. Deze vorm van analyse kan worden gebruikt om onderhoud te plannen, productievolumes aan te passen of calamiteiten te voorkomen. Het inherente probleem van het voorspellen van calamiteiten is, dat als de respons succesvol is, de calamiteit niet optreedt. Er kan vervolgens niet bepaald worden of de voorspelling correct is. Dit lijkt triviaal, maar in veel sectoren wordt een adequate reactie op een calamiteit beter beloond, dan het voorkomen ervan.

2.4 Praktische aspecten

Het analyseren van data kost rekenkracht en dat is duur. Daarom kunnen organisaties ervoor kiezen om dit uit te besteden. Dat kan op allerlei manieren. Organisaties kunnen bijvoorbeeld een prijsvraag uitschrijven en hun data beschikbaar stellen aan de deelnemers, die dan zelf de vraag proberen op te lossen. Dit zien we terug in de ontwikkeling van citizen science: je gebruikt de creativiteit en intelligentie van burgers om vraagstukken op te lossen. Dat betekent natuurlijk wel dat je ze (in ieder geval tijdelijk) toegang moet geven tot je data.

Sommige organisaties 'lenen' kleine beetje rekenkracht van particulieren om grote datasets te analyseren. Een van de bekendste en langlopende voorbeelden is SETI@home. SETI staat voor Search for ExtraTerrestrial Intelligence en is een initiatief van wetenschappers van de University of California at Berkeley. Deelnemers krijgen kleine pakketjes gegevens van een radiotelescoop toegestuurd om te analyseren op hun PC, als ze die even niet gebruiken. Zo kunnen de wetenschappers tegen lage kosten zeer veel data analyseren, in hun zoektocht naar buitenaards leven.

Het is ook mogelijk om rekenkracht in te huren. Clouddiensten als Amazon, Google en Microsoft bieden rekencapaciteit te huur aan. Het voordeel hiervan is dat organisaties zelf geen IT systemen hoeven aan te schaffen of, minstens zo belangrijk, up-to-date te houden. Het nadeel is dat de data, net als bij de andere voorbeelden, wel naar een externe partij verstuurd moet worden. Ook kunnen de kosten van dit soort diensten behoorlijk oplopen als de analyses intensiever worden.

De meest verstrekende oplossing is om als organisatie zelf de benodigde infrastructuur aan te schaffen. De leeftijd van computers en servers zijn echter een stuk korter dan die van de waterinfrastructuur. Om actueel te blijven moeten mogelijk iedere vijf jaar alle systemen worden vervangen. Ook heb je als organisatie experts nodig om de systemen te laten draaien, te beschermen tegen hackers, kortom een flinke uitbreiding van de afdeling systeembeheer.

3. Opslag

Dezelfde praktische overwegingen gelden voor de opslag van data in databases. Databases liggen namelijk ten grondslag aan alle data-gestuurde toepassingen. Google, Facebook, Twitter, allemaal zijn ze gebaseerd op databases. De databases van de toekomst worden nu ontwikkeld op de technische universiteiten. De TU Delft heeft de 'kleinste harde schijf ooit' ontwikkeld, waarbij iedere bit bestaat uit slechts één chlooratoom. In theorie kan je alle boeken die de mensheid ooit heeft geschreven, bewaren op een postzegel. Hij functioneert echter alleen nog in een vacuümomgeving bij -196 graden Celsius. Verder wordt er geëxperimenteerd met opslag in DNA, harde schijven gevuld met helium, magneettoepassingen en opslag in hologrammen.

Die ontwikkelingen zijn hard nodig, want alle trends wijzen op een enorme toename van data. De nieuwste DNA sequencing technieken produceren nu tot 1,8 Terabyte (1000 Gigabyte) in een analyse. De verwachting rond Big Data is nu dat er in 2040 wereldwijd 150 miljard sensoren zijn, die continu data produceren. Ook de data afkomstig van mobieltjes (citizen science) en smart meters zal toenemen. Al heeft die data vervolgens weer nodig voor machine learning. Hoe sla je zulke enorme hoeveelheden data op? Dat roept vergelijkbare vragen op als bij rekenkracht. Houd je dat 'dichtbij' of kan je dat ook aan andere partijen overlaten.

Daarbij moet een afweging gemaakt worden tussen kosten, veiligheid van de data, toegankelijkheid voor derden en wetgeving met betrekking tot data-privacy. Die afweging zal anders zijn voor verschillende typen data. Voor persoonsgegevens gelden andere overwegingen dan voor data uit de eigen infrastructuur of metingen uit de omgeving. Dat maakt de kwestie waar de data de bewaren en op welke manier de data samen te laten komen nog complexer.

Een eigen datacentrum biedt waarschijnlijk de beste veiligheid en privacy (als de beveiliging goed geregeld is), maar is pas bij grote hoeveelheden data kostenefficiënt. Een gedeeld datacentrum van alle waterbedrijven heeft kostenvoordelen en zorgt dat de data binnen de sector blijft.

Als de opslag uitbesteed wordt is de locatie van de server ook nog van belang. De EU heeft regels voor clouddiensten binnen de Europese Economische Ruimte (EER) en heeft een lijst van 'witte landen' die ook voldoende veiligheid bieden. Voor opslag in andere landen (waaronder de VS) moeten aanpassingen gedaan worden in de beveiliging van persoonsgegevens. Dat geldt ook als de data in Europa is opgeslagen, maar beschikbaar moet zijn voor een helpdesk in India.

De EU heeft in 2016 strengere privacywetgeving aangenomen, die het verkeerd of onterecht opslaan van persoonsgegevens met hoge boetes bestraft. Dit is zeer relevant als de watersector in de toekomst meer gebruik willen maken van klantdata, bijvoorbeeld afkomstig van smart meters, of citizen-science-toepassingen. De wetgeving zal op 25 mei 2018 in werking treden. Op http://ec.europa.eu/justice/newsroom/data-protection/infographic/2017/index_en.htm is in een infographic weergegeven wat dit voor organisaties kan betekenen. Volledige informatie over de hervorming is hier te vinden: http://ec.europa.eu/justice/data-protection/reform/index_en.htm.

3.1 Blockchain

De ontwikkeling van Blockchain-technologie is een bijzondere vorm van data-opslag. Aan een Blockchain kan alleen nieuwe data worden toegevoegd, maar geen data verwijderd. Het is vergelijkbaar met een dagboek waar steeds meer pagina's bijkomen. Wat ook bijzonder is aan de blockchain is dat de data is verspreid over een heel netwerk. Bij iedere toevoeging, moet de toegevoegde data gevalideerd worden door het hele netwerk. Blockchain wordt vooral gebruikt voor het opslaan van transactie-data zoals financiële en juridische transacties. Door de voortdurende validatie is een blockchain-database bijna immuun voor fraude en inbraken.

4. Verzamelen

Anders dan experimentele metingen of vragenlijsten is het verzamelen van sensordata en andere vormen van Big Data een continu proces, waarbij het ook nog eens om zeer grote hoeveelheden gaat. De data kan afkomstig zijn van zeer verschillende bronnen, zoals social media, logbestanden van machines, web analytics, locatie-gegevens etc. Maar sensoren vormen veruit de grootste bron van data. Wetenschappers van de TU Delft schatten dat er over tien jaar 150 miljard sensoren via netwerken met elkaar zijn verbonden. De hoeveelheid data die we produceren verdubbelt dan ieder etmaal. We produceren dan in een dag tijd net zoveel data als we tot dan toe in de hele geschiedenis hebben gedaan.

Omdat Big Data vaak over 'real time' data gaat (bijvoorbeeld meteorologische gegevens) moeten de communicatielijnen voortdurend open staan. Afhankelijk van de hoeveelheid data, vereist dat verbindingen van een bepaalde breedte en moeten de verbindingen veilig zijn. Het zijn juist dit soort open dataverbindingen die een makkelijke prooi vormen voor hackers. Blockchain-technologie kan een behulpzame technologie zijn bij het verzamelen citizen science data, maar ook voor het uitlezen van smart meters.

De Smart City zal in de toekomst een belangrijke bron van data zijn. Hier wordt zeer veel verschillende data geproduceerd en gecombineerd. Voor een deel is dat data afkomstig uit sensoren, voor een deel persoonlijke data, transactiedata. De grote vraag is wie in de toekomst toegang heeft tot die data en onder welke voorwaarden en wie de eigenaar is van al die data. In "Who Owns the Future" pleit Jason Lanier voor het koppelen van eigendomsrecht aan wie de data produceert. Iedereen die vervolgens gebruik maakt van die data moet daarvoor een minimale vergoeding betalen. De data blijft te allen tijde eigendom van de producent. Ook hier kan Blockchain een rol spelen in het bijhouden van alle transacties. Op deze manier zouden de waterbedrijven de door hen zelf verzamelde data ook beschikbaar kunnen stellen voor analyse en toepassingen door anderen.

Conclusie

Data biedt veel strategische uitdagingen. De onderliggende trends wijzen erop dat de hoeveelheid data in de toekomst zeer sterk zal toenemen en dat we die data in toenemende mate kunnen gebruiken om de wereld om ons heen te begrijpen, te voorspellen en vorm te geven door onze beslissingen. Die processen zullen in de toekomst steeds vaker door computers autonoom worden uitgevoerd, slechts begeleid door mensen.

Waterbedrijven doen er goed aan om een strategie te ontwikkelen rond het verzamelen, opslaan, analyseren en toepassen van data. De toepassing moet daarbij leidend zijn, aangezien die voor de andere onderdelen bepaalt wat er nodig is. De ontwikkelingen staan echter niet stil. Het kan daarom toch verstandig zijn om ook zonder een directe toepassing op het oog te hebben al te beginnen met het systematisch verzamelen van data, bijvoorbeeld uit de eigen infrastructuur en bedrijfsvoering, die dan later voor toepassingen gebruikt kunnen worden.

Het is verstandig de toenemende rol van data en ICT ook in het HR-beleid te laten terugkomen. Enerzijds door steeds jonge ICT'ers in dienst te nemen die op de hoogte zijn van de laatste ontwikkelingen, anderzijds door ICT'ers ook in het hoger management een positie te geven. Dit om te voorkomen dat een gebrek aan kennis van (vaak complexe) ICT-systemen, goede beslissingen op dat gebied in de weg staan. Ten slotte zijn er ook mensen nodig die verbinding kunnen leggen tussen de echte wereld en de wereld van data en ICT toepassingen.

Voor data-analyse en -opslag geldt dat er zeer veel mogelijkheden zijn dat in te richten. Naast kosten en doelmatigheid zijn privacy en veiligheid de belangrijkste aspecten om rekening mee te houden. Wetgeving op dat gebied is complex en verandert voortdurend. In 2016 is voor het laatst nieuwe privacywetgeving in Europa aangenomen, die treedt in 2018 in werking.

Bronnen

Snijder, D (2017), *Data is Macht*, Stichting Toekomstbeelden der Techniek

Lanier, J. (2013) *Who Owns the Future*, Simon & Schuster