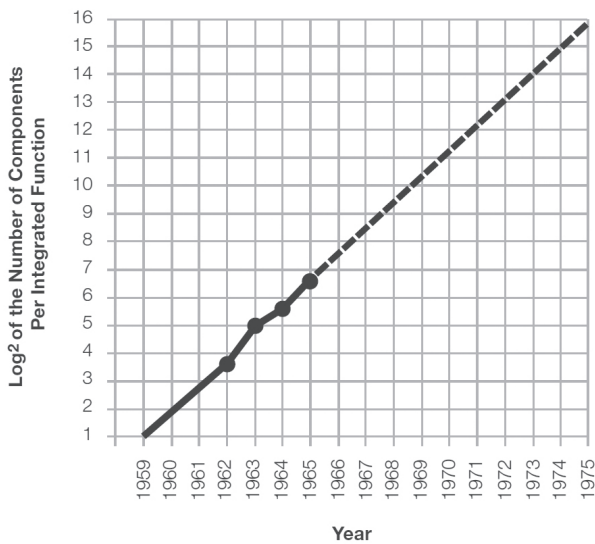


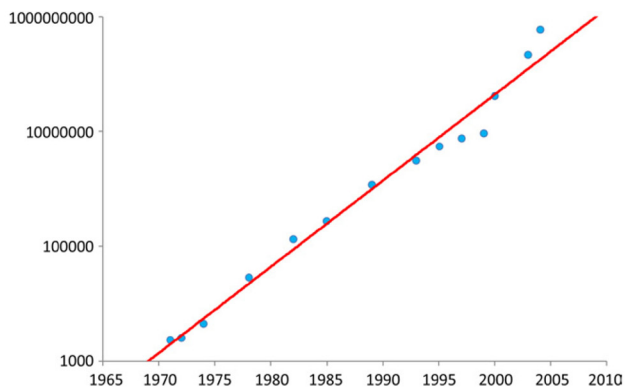
## Data + modellen = dynamische informatie (=sexy)

In 1965 voorspelde pionier Gordon Moore dat het aantal componenten, transistoren, op een geïntegreerd circuit, een computerchip jaarlijks zou verdubbelen, zie Figuur 1.



Figuur 1 - Voorspelling aantal transistoren op een chip in 1965 (Moore, 1965).

Zijn voorspelling is zeer houdbaar gebleken, het aantal transistoren op een chip is sindsdien jaarlijks met een factor 1,58 gegroeid (Robison, 2012), zie Figuur 2.



Figuur 2 - Analyse aantal transistoren op een chip 1970-2010 (Robison, 2012).

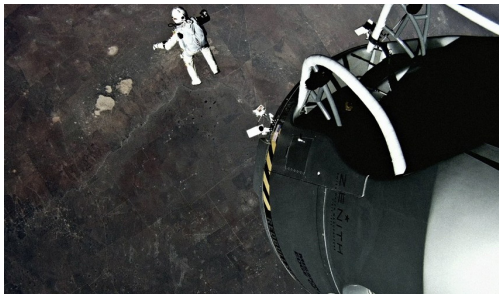


Ignaz Worm  
PWN  
TU Delft



## Data + modellen = dynamische informatie (=sexy)

Omdat de ontwikkeling van chips gepaard is gegaan met beheersbare kosten is de totale rekenkracht van computers op de wereld explosief gegroeid tot  $6 \times 10^{12}$  MIPS (millions instructions per second) in 2007 (Hilbert and López, 2011). In 1986 zat het grootste deel van deze rekenkracht in calculators, in 2000 in PC's. Zeven jaar later hadden game consoles en mobiele telefoons beide al meer rekenkracht dan PC's in 2000. De totale hoeveelheid opgeslagen niet volatiele digitale data groeide tot 295 exabyte, dus  $295 \times 10^9$  gigabyte in 2007 (Hilbert and López, 2011). Uitgedrukt in CD-ROMs in een hoesje, een stapel van 42m bij 42m met de hoogte waarvan Felix Baumgartner op 14 oktober 2012 naar de aarde sprong, 39 km.



*Figuur 3 - Felix Baumgartner sprong op 14 oktober 2012 van 39 km hoogte naar de aarde (foto [www.redbullstratos.com](http://www.redbullstratos.com)).*

De rol van ICT en datastromen in ons dagelijks leven mag duidelijk zijn. Als de datasets zo groot zijn dat de gebruikelijke software om datasets in te lezen, op te slaan, te beheren en analyseren ze niet kan bevatten, spreken we van 'big data'. Strategie consultant McKinsey beschrijft big data als de nieuwe uitdaging in innovatie, competitie en productiviteit (Manyika et al., 2011). Kunnen we daar in de drinkwatersector iets mee? Ik denk van wel, alhoewel de schaal zal passen bij wie we zijn en wat we doen. PWN slaat jaarlijks 15 gigabyte aan procesdata op en beheert 3000 gigabyte aan klantgerelateerde data. We zullen door het koppelen van interne en externe databronnen, veelal gekoppeld aan een geografische locatie, en virtuele data gegenereerd met modellen, meer en meer informatie genereren die voor een individu van waarde is. Dat individu kan een klant zijn, een manager, een technoloog, een bedrijfsvoerder, een storingsmonteur, enz.

Hierna een terugblik op de afgelopen jaren, de integratie van procesmodellen met geëmuleerde procesautomatisering, het komende jaar, de human sensor, en de komende jaren, een breder onderzoek om van data informatie te maken.

### **Integratie procesmodellen geëmuleerde procesautomatisering software**

Drinkwaterbedrijven zijn overgegaan op volautomatische bedrijfsvoering om een hogere en constantere waterkwaliteit te bereiken, minder afhankelijk te zijn van menselijke beperkingen (zoals de maximale werkduur en persoonlijke voorkeuren in de bedrijfsvoering), een hogere leveringszekerheid en lagere kosten. Als bedrijfsvoerders door de toegenomen afstand tot het zuiveringsproces niet meer adequaat kunnen handelen tijdens calamiteiten kan dat leiden tot verstoring of onderbreking van de drinkwaterlevering. Bovendien kan de procesautomatisering (PA) zelf falen. In de afgelopen jaren zijn emulatie platforms van PA-systemen geïntroduceerd in de drinkwatersector. Op deze platforms kan software worden uitgewisseld tussen een PLC (programmable logic controller) in de zuivering en een virtuele controller op een PC, vice versa, zonder iets te wijzigen in de software, zie Figuur 4.

Dit systeem is waarheidsgetrouw in bediening en gedrag van de software, maar er is een belangrijk verschil met de werkelijkheid, het gedrag van het proces ontbreekt. Door een hydraulisch model en een waterkwaliteitsmodel te koppelen aan het geëmuleerde PA-systeem wordt dit ondervangen, zie Figuur 5.

Er zijn drie toepassingen verkend van dit systeem, i) technologen de mogelijkheid bieden regelingen te ontwerpen en te evalueren, ii) het trainen van bedrijfsvoerders in een nagenoeg waarheidsgetrouwe 'human-in-the-loop' simulator, en iii) het functioneel testen van PA-software (Worm, 2012).

De overdracht van kennis in een simulator is bij PWN bepaald in de Waterspot simulator met vier testgroepen. Drie daarvan, ervaren bedrijfsvoerders (EB), minder ervaren bedrijfsvoerders (MEB) en



*Figuur 4 - Emulatie van procesautomatisering software. Een PC gaat zich gedragen als één of meerdere virtuele procescontrollers (PLC's).*



*Figuur 5 - Door een procesmodel te koppelen wordt een nog meer waarheidsgetrouwe simulator gerealiseerd.*

leken (L60x) trainden met een versnelde simulatie, de vierde groep, leken, trainde op normale snelheid (L1x). De trainees leerde hoe ze de waterkwaliteit van het effluent van de ontharding reactoren konden regelen. Tijdens het testen van de overdracht van de kennis, wanneer ze geconfronteerd werden met andere procesverstoringen dan tijdens de training, presteerde L60x significant slechter dan de bedrijfsvoerders (EB en MEB). Er werd geen onderscheid gemeten tussen de prestaties van de EB en de MEB en L60x presteerde beter dan L1x. Deze resultaten lijken het idee te bevestigen dat het leren regelen van een langzaam en complex proces kan verbeteren door op een herkenbare manier te trainen met versnelde simulatie.

Met de ontwerpmethodode voor regelingen voor drinkwaterzuiveringen (Van Schagen et al., 2010) is een nieuwe regeling voor de pelletontharders op PWN zuivering Wim Mensink ontworpen. Een Stimela waterkwaliteitsmodel is opgesteld en gevalideerd. Met dit procesmodel zijn de effecten van de bestaande regeling van de ontharding vergeleken met een nieuwe regeling en werd het effect van

een lagere toevoer van RO-water onderzocht. Aangevoerd kon worden dat door de nieuwe regeling de capaciteit van de zuivering vergroot kon worden, en dat het gebruik van energie en chemicaliën verminderd kon worden. Tegelijkertijd kon de totale hardheid van het drinkwater dichter bij de gewenste waarde gebracht worden.

Tenslotte is de toegevoegde waarde onderzocht van proces simulatie modellen in de 'factory acceptance test' (FAT) van PA software. Twee test teams testten hetzelfde stuk aangepaste software. Eén team gebruikte hiervoor een testsysteem dat bestond uit een emulatie van de software geïntegreerd met een proces simulatie model, het AVC systeem, zoals te zien in Figuur 5. Het andere team gebruikte hetzelfde systeem met in plaats van het proces simulatie model, basale parameter relaties, het VC systeem. Elk team speurde één (verschillende) fout op van de dertien fouten die voor het experiment in de software waren gestopt, de meeste fouten waren tijdens de

Data + modellen = dynamische informatie (=sexy)

unit test door het ontwerpteam gevonden. Tijdens de functionele test vond het AVC team drie fouten in de 78 geteste items, het VC team vier. Maar het AVC team beoordeelde 1% van de test items 'niet van toepassing', het VC team 17%, een indicatie dat met AVC meer getest kan worden. Met AVC kunnen niet overtuigend meer fouten worden gevonden dan met VC. Niet het proces gedrag is relevant voor de FAT van PA software, maar het over- of onderschrijden van proces grenzen.

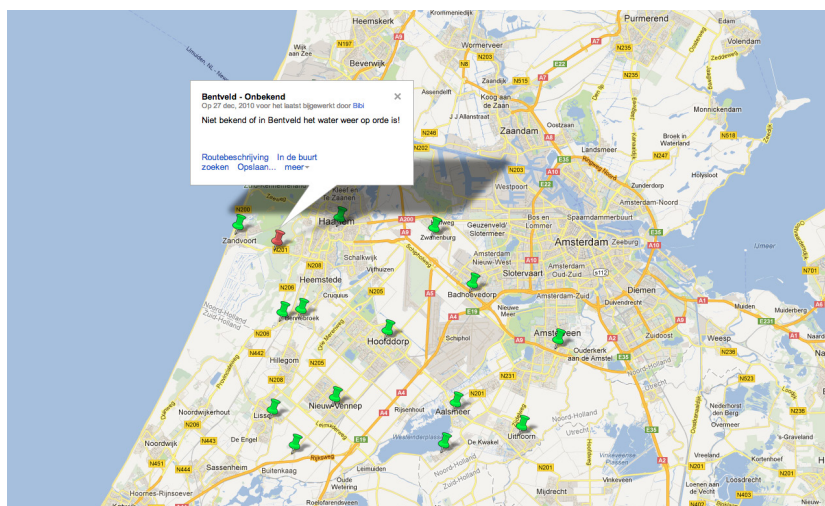
**Human sensor**

Besmettingen in het distributienet, grote lekken, kleine lekken... veroorzakers van ondermaatse leveringsminuten zijn het! En als ze optreden duurt het soms dagen voordat de bron gevonden is, zoals in 2009 in Rijsenhout toen een boer regenwater in het net pompte. Acht dagen zaten er tussen de eerste melding en het bepalen van de vermoedelijke locatie. Niet voor niets richten leveranciers en onderzoekers zich op het ontwikkelen van sensoren die in het distributienet geplaatst kunnen worden. Sensoren die lage concentraties moeten detecteren. Die weinig onderhoud mogen vergen en waarvan de data ontsluiting niet te veel mag kosten. Van deze sensoren zijn er heel veel nodig om een dekkend systeem te hebben. Is dat mogelijk? Het is moeilijk dit goed voor elkaar te krijgen ... en vooral duur...

De uitval van Hoofddorp tijdens de Kerst van 2010 heeft ons geleerd dat onze eigen klant ook sensoren zijn. Binnen een uur circuleerde een kaartje op internet waar de klachten over waterdruk te zien waren, zie Figuur 6. En neemt het aantal telefoontjes naar het klant contact centrum niet toe als er een lek of besmetting is?

Doelstelling van dit project is om klacht-data, tweets en gemiste telefonische oproepen (allemaal geografisch gelabeld en voorzien van tijdstip) te gebruiken om de locatie van een probleem in het distributienet te bepalen. Het presenteren van deze informatie op een kaartje is al waardevol voor klanten, bedrijfsvoerders en storingsmonteurs, maar in de human sensor wordt de klachtdata ook gebruikt als input voor een leidingnetmodel met 'backtrace (terugreken)' functie.

De ingrediënten zijn grotendeels beproefd en beschikbaar, de combinatie is vernieuwend en kent een zeer gunstige verhouding tussen investering en opbrengst. Vermoedelijk om die reden is dit idee door experts van het Koninklijk Nederlands Waternetwerk overtuigend gekozen als 'het beste IT-idee in de watersector 2012'. PWN, Vitens en Brabant water dragen bij aan de ontwikkeling van dit concept.



Figuur 6 - Een klant van PWN publiceerde een overzicht van tweets over de onderbreking van de levering vanuit pompstation Hoofddorp.

### Systemanalyse drinkwaterprijs en risico

De huidige benchmark wordt door aandeelhouders vooral gebruikt om druk te zetten op beheersing van kosten, en minder op het verbeteren van de volksgezondheid of het vergroten van het klantvertrouwen. Het is dan ook goed nieuws dat de VEWIN zich heeft zich voorgenomen een verbeterde benchmark te ontwikkelen met de TU Delft. Want er lijkt ruimte om waterkwaliteit, leveringszekerheid en kosten weer meer in verband te zien. Eén methode om deze systemanalyse te doen is het kwantitatief koppelen van de drinkwaterprijs aan het gewenste risico niveau. Maar een laag risiconiveau alleen is geen garantie voor klantvertrouwen, daar is ook een optimaal klantcontact voor nodig. Het is een uitdaging dat optimale klantcontact te uniformeren vanwege efficiëntie, maar tegelijk voldoende specifiek te maken om zo veel mogelijk aan te sluiten op individuele wensen. In meer technische termen is het boeiend te onderzoeken hoe we de klant 'in-the-loop' van de bedrijfsprocessen kunnen brengen om zo het drinkwatersysteem te optimaliseren.

Ik verwacht dat de invulling van optimaal de komende jaren zal veranderen. Door het steeds verder dalen van de detectiegrenzen van waterkwaliteitsmetingen wordt het uitgangspunt 'we verwijderen tot onder de detectiegrens' te kostbaar. Voor waterkwaliteit groeien we naar een nieuwe definitie toe, bijvoorbeeld 'de waterkwaliteit staat boven elke verdenking'. Zolang zich geen nieuwe bedreigende stoffen aandienen lijkt het daarmee verantwoord het streven naar een steeds betere drinkwaterkwaliteit af te remmen. Op het gebied van leveringszekerheid, of preciezer beschikbaarheid, lijkt het er op dat we minder storingen hebben dan de klant accepteert. Uit klantonderzoek bij PWN blijkt dat één storing per jaar die 4 tot 6 minuten duurt zonder meer geaccepteerd wordt (Sijpersma, 2012), maar ook bij meer en langere storingen is de klant niet bereid meer geld te betalen voor zijn water. Is het niet optimaal om wat meer storingen te hebben en daardoor meer klantcontacten en zo de kans het klantvertrouwen te vergroten?

Voor een kwantitatieve systemanalyse met de klant 'in-the-loop' is het integreren van interne en externe databronnen en virtuele data van modellen een geweldige uitdaging met grote potentie. En dan nu net als Felix Baumgartner weer landen op aarde.

### Referenties

1. Hilbert, M., López, P., 2011. The world's technological capacity to store, communicate, and compute information. *Science* 332(6025), 60-65.
2. Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., Hung Byers, A., 2011. Big data: the next frontier for innovation, competition, and productivity. Rapport: Seoul, San Francisco, London, Washington.
3. Moore, G.E., 1965. Cramping more components onto integrated circuits. *Electronics* 38(8), 114-117.
4. Robison, R.A., 2012. Moore's Law: Predictor and Driver of the Silicon Era. *World Neurosurgery* 78(5), 399-403.
5. Sijpersma, R., 2012. Goed drinkwater, tegen elke prijs? Rapport: Velsersbroek.
6. Van Schagen, K.M., Rietveld, L.C., Veersma, A., Babuska, R., 2010. Control-design methodology for drinking-water treatment processes. *Water Science and Technology: Water Supply* 10(2), 121-127.
7. Worm, G.I.M., 2012. Integration of process models of drinking water treatment plants and emulated process automation software, PhD thesis, Delft University of Technology. Water Management Academic Press: Delft.