

Bron foto: artfulblogger on Flickr.com (CC-licentie)

Informatiekwaliteit op het juiste niveau!

mei 2011

De tijd dat organisaties hun gegevens voor zich hielden zoals een kip op zijn eieren past, is gelukkig voorbij. Publieke organisaties wisselen veel gegevens uit, zoals adressen, gegevens over inkomsten en eigendommen en informatie over iemands werkervaring en opvoedrisico's. Zij gebruiken veel data die niet in eigen processen is ontstaan. Daardoor kan de betekenis van gegevens veranderen of verloren gaan, wat tot een lage gegevenskwaliteit leidt. Wij stellen een afwegingskader voor dat de processen en hun gegevens dichter bij elkaar brengt. Dat strookt vaak wel, maar niet altijd met de gedachte van basisregistraties...

Door Ton Monasso, Tom Kronenburg en Ko Mies

In dit artikel geven wij een visie op de vraag hoe uitvoerende organisaties hun effectiviteit kunnen vergroten door betere gegevens te gebruiken en gegevens beter te gebruiken. Het 'schaalniveau' waarop de gegevens worden beheerd staat daarin centraal. We schetsen eerst kort een historisch perspectief en gaan dan in op de voor- en nadelen van centralisatie van gegevens. Vervolgens presenteren we een afwegingsprincipe voor het bepalen van het juiste niveau voor het beheren van gegevens. Tot slot beschrijven we de stappen die nodig zijn om vanuit de huidige situatie tot verbeteringen te komen.

Centralisatie en decentralisatie wisselen elkaar af

'In den beginne', voor de tijd van de grootschalige informatietechnologie, bestond geen keuze: gegevens werden verzameld en beheerd binnen het proces waar ze werden gebruikt. In het geval van een bijstandsaanvraag, bijvoorbeeld, was het de sociale dienst die alle aanvraagformulieren, de

uitkeringshistorie en de voor controle benodigde gegevens opsloeg. Dit is een voorbeeld van gedecentraliseerd gegevensbeheer. In deze 'fase 1' hadden organisaties hun eigen registraties, aansluitend op hun eigen taken en processen.

Sindsdien is de complexiteit van overheidsprocessen stevig toegenomen en zijn veel meer mogelijkheden ontstaan om informatietechnologie grootschalig in te zetten. Er zijn meer processen ontstaan, die onderling afhankelijk zijn, verschillende variaties kennen en steeds meer gegevens gebruiken om de afhankelijkheid van dure medewerkers te verminderen. Om de gegevens tussen processen te kunnen delen, ontstonden in de tweede fase inzichtvoorzieningen. Organisaties behielden hun eigen registraties, maar konden onderling wel informatie opvragen, die zij vervolgens weer in hun eigen registraties verwerkten. Voorbeelden daarvan zijn de inzage in het kentekenregister van de RDW door de politie, Suwinet voor gemeenten en UWW en het raadplegen van de Gemeentelijke Basisadministratie door andere overheidsorganisaties.

Het denken en handelen in de Nederlandse publieke sector bevindt zich nu in een volgende fase. Gegevens die in veel processen worden gebruikt, zijn ondergebracht in basisregistraties, zoals de Gemeentelijke Basisadministratie (GBA) en de Basisadministratie Adressen en Gebouwen (BAG). Het gebruik van die registraties is voor overheden verplicht. Organisaties proberen eigen administraties van dezelfde gegevens uit te bannen door alle kennis die overheden hebben over de basisgegevens centraal te delen. Niet langer wordt geprobeerd om 'alles' te delen, maar de gedeelde gegevens worden wel gecentraliseerd beheerd.

Centralisatie leidt tot schaalvoordelen en meer mogelijkheden voor uitwisseling...

Het beheer van gegevens op een niveau 'boven' de uitvoerende processen heeft ontegenzeggelijk voordelen. Als hetzelfde gegeven in meerdere processen gebruikt wordt, is er iets voor te zeggen om dat centraal te beheren. Ten eerste leidt centraal beheer vaak tot schaalvoordelen. De beheerkosten van de data gaan omlaag en het is mogelijk specialistischer (beheer)expertise in te schakelen. Ten tweede is het bevorderlijk voor standaardisatie en daarmee samenwerking. Omdat de adresgegevens van burgers feitelijk landelijk beheerd worden - de gemeenten voeren het uit, maar de meeste afspraken zijn landelijk geüniformeerd - kan de Sociale Verzekeringsbank met een enkel 'datalijntje' verhuizingen van zowel Middelburgse als Maastrichtse inwoners verwerken. Omdat voertuiggegevens landelijk worden beheerd, is het mogelijk om Groningse snelheidsdruivers op de Brabantse wegen te beboeten. Ook hoeven burgers en bedrijven een verhuizing nog maar bij één instantie door te geven, in plaats van bij veel aparte loketten.

...maar ook tot vervreemding van de uitvoerende processen

Centralisatie heeft ook nadelen, die overigens vaak worden onderschat. Deze ontstaan omdat de gegevens buiten de context van het uitvoerende proces worden gebracht, en daarmee een deel van

hun (contextuele) betekenis verliezen of een andere betekenis krijgen.

De betekenis kan veranderen door gebruik in een andere context dan waarin ze zijn ontstaan. De Wetenschappelijke Raad voor het Regeringsbeleid noemt dat de de- en vervolgens hercontextualisering van informatie. Denk aan de uitwisseling van informatie over jongeren die risico's lopen in hun ontwikkeling. Stel dat informatie uit een Elektronisch Kinddossier en het Elektronisch Leerdossier met elkaar verbonden wordt. De notitie die een leerkracht maakt over een 'vermoeden van kindermishandeling' kan bij de jeugdarts alle bellen doen rinkelen. Beiden bedoelen met dezelfde term echter iets heel anders. Een leerkracht die bedreven is in les geven, maar geen expert is rondom fysieke aandoeningen, kan in een tijdens de gymles geconstateerde blauwe plek al aanleiding zien om een 'vermoeden van mishandeling' uit te spreken. Een jeugdarts zal daar veel minder snel over spreken. Hij of zij ziet veel blauwe plekken en kan deze veel beter verklaren. Pas na een waarschijnlijkheidsdiagnose zal in een dossier een 'vermoeden van mishandeling' worden opgetekend. Het is lastig om informatie los van de context te interpreteren. Het gebeurt echter wel doordat men zich niet bewust is van verschillen in referentiekaders.

Als gegevens te ver af komen te staan van de processen, organisaties of medewerkers die ze hebben opgetekend, gaat betekenis verloren. Degene die de gegevens beheert, bijvoorbeeld een gemeente als het om adresgegevens gaat, heeft geen zicht meer op de relatie van de gegevens met de doelen waarvoor zij gebruikt worden. Stel dat de GBA 0 januari registreert als geboortedatum, als alleen het geboortjaar van een asielzoeker zonder papieren bekend is. Voor de gemeentelijke diensten hoeft dat geen probleem te zijn, als hier een gemeentebrede afspraak over is gemaakt. Iedereen begrijpt dan dat eigenlijk 'onbekend' bedoeld wordt. Wanneer nu de Sociale Verzekeringsbank diezelfde gegevens gaat gebruiken om AOW-rechten uit te keren, kan echter een probleem ontstaan. Niet elke afdeling burgerzaken gaat immers op dezelfde wijze met onbekende geboortedata om. Sommige gebruiken 1 januari, andere 1 juli en weer anderen vullen 0-0 of 0 januari in. De Sociale Verzekeringsbank gaat er wellicht vanuit dat de geboortedatum

van de asielzoeker in kwestie in januari ligt, dus zal zij vanaf januari uitkeren. Deze burger ontvangt daarmee mogelijk (gemiddeld) een half jaar te vroeg een AOW-uitkering.

Als iemand meerdere malen niet reageert op post van de gemeente, kan diegene worden geregistreerd als 'vertrokken, onbekend waarheen'. De burger kan echter heel goed een studie- of bijstandsschuld hebben. Stel dat de sociale dienst de adressen zou beheren, dan zou iemand veel minder snel op 'vertrokken onbekend waarheen' worden gezet. Het belang van correcte adresgegevens is immers veel zichtbaarder, en het is voor de dienst van groter belang om de persoon in kwestie op te sporen.

Deze voorbeelden maken ook een klassiek ketenprobleem zichtbaar. De baten van goede gegevens slaan vaak op een andere plaats neer dan de kosten. De beheerder van de GBA zal niet zomaar dure huisbezoeken afleggen als de sociale dienst of een grote uitvoeringsinstelling als DUO of UWV daarvan profiteert. De recente bezuinigingsrondes en de focus op een 'compacte overheid' hebben de oriëntatie van organisaties op hun eigen belang alleen maar versterkt en zullen de welwillend van overheidsorganisaties om elkaar te ondersteunen niet vergroten.

Tot slot: gegevens kunnen een hele nieuwe betekenis krijgen, als zij op een 'te centraal' niveau worden beheerd. Dat gebeurt door de samenloop van twee fenomenen. Ten eerste gebruiken overheden die zelf ver van de uitvoering af staan graag gegevens om hun beleid op te baseren en successen aan af te meten. Ten tweede zijn gegevens altijd een abstractie van de werkelijkheid: in de opbouw van een dataverzameling worden bepaalde elementen van de werkelijkheid wel en andere niet weergegeven. De door de overheid gebruikte informatie geeft dus maar een deel van de werkelijkheid weer. Dat kan leiden tot sturing op gegevens in plaats van op de werkelijkheid.

Een klassiek voorbeeld is de belastingheffing op de houtproductie uit bossen. Om de hoogte van de verschuldigde belasting te kunnen vaststellen, zijn nauwkeurige schattingen van de productie nodig. Dat is lastig als plantages sterk verschillen in oppervlakte, vruchtbaarheid van de bodem, bebosningsdichtheid en type bomen. Als hulpmaatstaf is in

de tweede helft van de 18e eeuw in Pruisen en Saksen een 'eenheid' geïntroduceerd die opbrengsten vergelijkbaar moest maken: de Normalbaum. De verwachte productie kon uitgedrukt worden in aantallen gevelde Normalbäume per jaar. Dit leidde echter niet alleen tot een betere belastingheffing, gegeven de werkelijkheid zoals die was. De werkelijkheid zelf veranderde ook. De plantages gingen daadwerkelijk bestaan uit Normalbäume - de Noorse spar bleek buitengewoon geschikt. Echte bomen tellen is immers veel makkelijker, zowel voor de bosbouwer als voor de overheid. Het voorkomt slepende conflicten over de hoogte van de belasting.

Een voorbeeld uit onze tijd zijn de vele jongeren die door sociale diensten werden doorgeschoven van de bijstand naar de Wajong. Door personen als bijstandsstatistieken te zien was het mogelijk om daar specifiek beleid op te maken. Mensen in de bijstand moesten 'uitstromen' naar werk, maar bleken door gemeenten op grote schaal overgeheveld te worden naar de Wajong regeling. Dezelfde personen, andere statistiek. We hebben nu in Nederland daadwerkelijk een hoop meer mensen die zich gehandicapt voelen doordat de nieuwe beleidsdefinities hen dat stigma oplegden.

In alle bovenstaande gevallen zou decentralisatie van gegevensbeheer voordelen opleveren. De gegevensdefinities zijn beter afgestemd op de per definitie gedecentraliseerde werkprocessen, er gaat geen of in ieder geval veel minder informatie over gegevens verloren en de werkelijkheid past zich niet aan aan het meetinstrument. Het is dus noodzakelijk om een balans te vinden tussen centralisatie en decentralisatie.

Organiseer de gegevens zo decentraal als het kan en zo centraal als het moet

Uitvoerende processen in de publieke sector hebben groot belang bij juiste gegevens: de weergave van de werkelijkheid. Het bepalen van het niveau waarop het beheer van gegevens wordt georganiseerd - in een uitvoerend proces, in een enkele organisatie, binnen een sector, landelijk of misschien wel internationaal - is geen eenvoudig vraagstuk. Centralisatie en

decentralisatie hebben, zoals hierboven geschetst, beide voor- en nadelen.

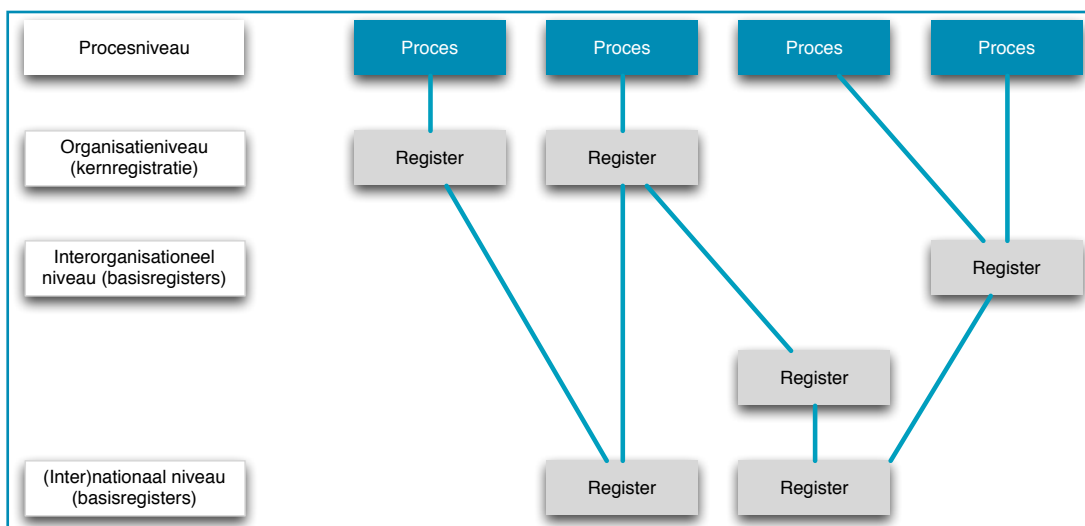
Om te bepalen op welk niveau gegevens moeten worden vastgelegd en beheerd, dient allereerst gekeken te worden naar de wijze waarop men het gegeven in de uitvoerende processen gebruikt. Kijken we naar bijvoorbeeld het adresgegeven dan zien we dat de jeugdzorg dit gebruikt om zo snel mogelijk contact te kunnen leggen met een persoon. Het maakt minder uit als dat betekent dat er meerdere adressen zijn, als wat onduidelijkheid bestaat over de schrijfwijze van de straat of als de postcode mist. Als het adresgegeven echter gebruikt wordt om een stembiljet of belastingaanslag te versturen of de vestiging van een bedrijf te regelen, dan is de correctheid, uniciteit en verifieerbaarheid van veel groter belang. Van minder belang is of die op het adres geregistreerde persoon ook daadwerkelijk een groot deel van de nachten op dit adres overnacht, zoals in de jeugdzorg belangrijk is. We zien dat men in dit geval zou moeten besluiten om twee adressen te hanteren: het 'snelle contactadres' of 'waarschijnlijke verblijfadres' en het 'officiële overheidspost- en vestigingsadres'.

Aan de andere kant kunnen we ons ook voorstellen dat scholen, jeugdzorg, huisarts en sociaal werkers wel het gegeven 'waarschijnlijke verblijfadres' delen, omdat hier grote voordelen te behalen zijn bij het gebruik maken van elkaars gegevens. Wanneer een hulpverlener een nieuw adres constateert, dan is dit ook relevant voor anderen in het sociale domein. Mocht het onverhoopt toch niet kloppen, dan is er ook geen man overboord, omdat de definitie van het 'snelle contactadres' ook enige onzekerheid in zich

heeft en men hier dus al van tevoren rekening mee kan houden.

Dit is een voorbeeld van het principe "Zo decentraal als het kan, zo centraal als het moet". We moeten informatie delen tussen zorgverleners omdat hun gezamenlijke effectiviteit anders sterk afneemt maar we kunnen en willen wel differentiëren tussen zorg en administratie. We zouden hier willen spreken van een 'subsidiariteitsprincipe van gegevensuitwisseling' omdat gezocht moet worden naar een balans tussen decentraal en centraal gegevensbeheer. Het zoekproces naar het juiste niveau moet van onder naar boven lopen en dus uitgaan van de processen en het gegevensgebruik op de werkvloer van de diverse sectoren of instanties. Door per gegeven (of gegevensgroep) te bepalen of gebruik in andere processen van toepassing is, kan tot gezamenlijk beheer worden besloten. Niet centraler, omdat dan vervreemding optreedt, en niet decentraler, omdat samenwerken dan wordt bemoeilijkt. Dit is het 'subsidiariteitsprincipe van gegevensuitwisseling'.

Dit subsidiariteitsprincipe toegepast op de huidige ideeën omtrent een stelsel van basisregistraties levert een getrappt model op. In dat model staan de identificerende basisregisters bovenaan een stelsel van meer sectorale of zelfs individuele gegevensregisters die informatie kunnen gebruiken uit hoger liggende niveaus, of deze naar gelang de definitie kunnen aanpassen. Het NHR voor bedrijven, het BSN-register dat nu onderdeel is van het GBA en het voertuigenregister van de RDW zijn voorbeelden van identificerende registers. Hieronder is een mogelijke configuratie van zulke registers weergegeven.



Op het hoogste niveau zien we de uitvoerende processen. Een proces gebruikt gegevens uit een bedrijfsregister of soms uit gemeenschappelijk registraties. Sommige bedrijfsregisters gebruiken gegevens uit gemeenschappelijke registraties als aanvulling op de eigen gegevens. Bedrijfsregisters en gemeenschappelijke registraties kunnen gegevens uit de nationale basisregisters gebruiken. Het aantal gegevens in een register zal in de bedrijfsregisters hoger zijn dan in de gemeenschappelijke registers, terwijl deze weer meer gegevens bevatten dan de nationale basisregistraties.

Van analyse naar actie

Als we het principe van “Zo decentraal als kan, zo centraal als moet” toepassen op de huidige situatie in het Nederlandse openbaar bestuur, leidt dat tot vijf aanbevelingen voor verandering:

1. Betrek vertegenwoordigers van de uitvoerende processen intensiever bij de doorontwikkeling van centrale registraties. De processen zouden de ‘vraag’ moeten formuleren, die door de beheerder van de registratie kan worden verwezenlijkt.
2. Voorkom dat het stelsel van basisregistraties teveel gegevenselementen bevat. Beperk de basisregistraties tot die gegevens die een groot bereik hebben, of beperk de toepassing van de basisregistraties tot die processen die met het gegeven overweg kunnen. Concreet betekent dit bijvoorbeeld dat de GBA een minder prominente rol zou moeten vervullen in de zorg- en welzijnssector.

3. Werk toe naar een getrappt stelsel van basis- en kernregistraties, waarbij ook op internationaal, sectoraal en organisatieniveau registraties bestaan waarmee gegevens worden gedeeld.
4. Waardeer succesvolle gegevensuitwisselingen en registraties op sectoraal niveau. Probeer van bijvoorbeeld Suwinet geen ‘one size fits all’-voorziening voor steeds meer organisaties te maken, maar houdt dit beperkt tot een dienend instrument voor partijen in het domein werk & inkomen.
5. Houd meer rekening met de verschillende contexten van gebruik. Presenteer de gegevens op een wijze die past bij de context. Geef bijvoorbeeld aan wanneer gegevens niet beschikbaar zijn, wanneer ze niet meer actueel zijn of wanneer de validiteit door iemand in twijfel is getrokken.

Ir. Ton Monasso en drs. Tom Kronenburg zijn senior adviseur bij onderzoeks- en adviesbureau Zenc. Ir. Ko Mies is partner van Zenc.

Over Zenc

Zenc is een maatschappelijk betrokken, innovatief adviesbureau. Wij zijn gericht op maatschappelijke organisaties in het publieke domein en de non-profitsector. Onze praktijk bestaat uit organisatieadvies, onderzoek, opleiding, en project-management op innovaties.

Wij helpen opdrachtgevers met vraagstukken rond samenwerking en gegevensuitwisseling tussen overheden, reorganisatie, procesherontwerp en (e-)dienstverlening. In interactieve processen komen wij met onze opdrachtgevers tot gedragen probleemanalyses en draagvlak voor creatieve oplossingen.

