

Datagedreven werken in relatie tot gegevensmanagement

Datagedreven werken en gegevensmanagement zijn termen die binnen organisaties steeds vaker in één adem worden genoemd.

Gegevensmanagement levert een onmisbare bijdrage aan datagedreven werken.

Gegevensmanagement zorgt er nl. voor dat data goed zijn en goed beschikbaar zijn. Het Tactisch Katern Gegevensmanagement van VNG-Realisatie biedt gemeenten een praktische handreiking om gestructureerd te bouwen aan de beschikbaarheid van goede data.

Datagedreven werken

Datagedreven werken is het op basis van onbewerkte feiten en data een beweging op gang krijgen naar meer inzicht, kennis en gefundeerde beslissingen. In de volle breedte van het bedrijf. Teneinde een datagedreven organisatie te realiseren waar iedereen op basis van betrouwbare data beslissingen neemt en niet meer op basis van aannames." *Bron: Gemeente Zwolle / Passioned Group*

Het gebruik van data neemt een enorme vlucht

Zowel in de stad als binnen de muren van het stadskantoor/gemeentehuis.

Het gebruik van data neemt toe, en daarmee het belang van goede, en beschikbare data.

Denk aan:

- Beslissingsondersteuning
Beschrijvende, diagnostiserende, voorspellende, voorschrijvende analyses
- Met proces-analyse tot effectiviteits- en efficiencywinst komen (b.v. process mining, Six Sigma, Lean);
- Geautomatiseerde uitvoering van onze processen:
geautomatiseerde besluitvorming (algoritmen), IoT-aansturing, robotic process automation;
- De maatschappij digitaliseert:
met data-toepassingen ontstaat er een andere vorm van interactie (participatie) met onze stakeholders.
- De maatschappij digitaliseert:
een autonome ontwikkeling waar de gemeentelijke organisatie bij wil en moet blijven

Data-gebruik onder architectuur

Om gegevens uit veel bronnen te kunnen ontsluiten is een duidelijke ICT- of gegevens-architectuur essentieel. Ook is het belangrijk een bedrijfsbreed gegevenswoordenboek te hebben zodat je zeker weet dat je het over hetzelfde hebt. In de vakliteratuur praten we dan over master data management en data governance. Binnen GEMMA hebben we dit samengevoegd onder de kop gegevensmanagement.

Door hier bij elk project (datagedreven werken, ICT-project, etc.) vanaf het begin aandacht aan te besteden en zo de datakwaliteit, privacy, archivering en beveiligingsaspecten te garanderen, wordt er vertrouwen in de data opgebouwd. Daarnaast heeft het een positief effect op de samenwerking, op de ICT-infrastructuur en op het 'in control' zijn over de data. Ook is het belangrijk om ervoor te zorgen dat de benodigde kennis en vaardigheden binnen de organisatie ontwikkeld worden.

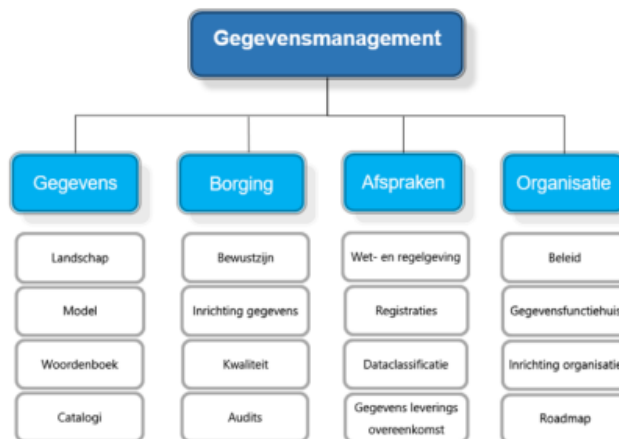
Databewustzijn moet groeien binnen de informatievoorzienings- en ICT-afdelingen, maar ook binnen de lijnorganisatie. Het degelijk organiseren van gegevensmanagement betekent dat je verantwoordelijkheid moet en mag durven nemen (ook als het gaat om eigenaarschap van systeem en data).

De basis van Datagedreven Werken

Daarmee is gegevensmanagement één van de heipalen van Datagedreven Werken. Het hele fundament bestaat uit:

- Kennis, kunde, capaciteit
Heb je de mensen in huis om datagedreven te kunnen sturen? Of je dat nu geheel zelf uitvoert of geheel uitbesteed hebt: ook in het laatste geval zul je de kennis in huis moeten hebben om de opdracht goed te formuleren en om erop toe te zien dat volgens opdracht geleverd wordt;
- Data-faciliteiten
Datawarehousing, datalakes, Urban Data Platforms¹, gegevensmakelaar, etc.
De techniek om gegevens beschikbaar te stellen, onder de juiste voorwaarden en met -waar benodigd- logging van gegevensgebruik
- IT-governance
Wie beslist, en hoe, over data-toepassingen, over datagebruik?
- Gegevensmanagement
Laatstgenoemd maar zeker niet 'least'.
Welke gegevens hebben we, voldoen die aan de kwaliteit, welke afspraken rondom gebruik zijn er gemaakt?
En ook: welke gegevens hebben we niet en hoe komen we daar dan aan?. Hoe weten we hoe betrouwbaar deze gegevens van 'buiten' zijn, hoe bewaken we of regisseren we dit? Basis- en kernregistraties zijn nog relatief eenvoudig, maar hoe gaan we om met het beheersen van data uit IOT bijvoorbeeld.

Gegevensmanagement is het geheel van activiteiten om in een organisatie ervoor te zorgen dat de juiste gegevens van de juiste kwaliteit op het juiste moment beschikbaar zijn.



¹ Zie 'Regie op data uit de stad', VNG Realisatie, juni 2019

Ze kunnen niet zonder elkaar

Datagedreven werken en gegevensmanagement zijn niet hetzelfde. Gegevensmanagement is één van de randvoorwaarden om datagedreven te kunnen werken. Hieronder volgt een aantal voorbeelden die aangeven waarom. Onderstaande onderwerpen zijn ook terug te lezen in de bijlagen over CRISP-DM en DELTA. Dit zijn methodes hoe om te gaan met het uitvoeren van een DGW project en om te analyseren hoe ver een organisatie is met DGW.

Consequent in kwaliteit

Gegevensmanagement organiseert de verantwoordelijkheid en bewustwording voor de gegevenskwaliteit: *garbage in, garbage out*. Gegevensmanagement voert ook audits hierop uit, zodat bij de start van een datagedreven werken project duidelijk is wat de datakwaliteit is. Een voorbeeld is 'volledigheid'. Een gemeente wil een voorspelling doen op basis van een dataset met locaties, tijden, omvang e.d. van inbraken in de binnenstad. Dit vereist uiteraard dat alle gegevens consequent zijn ingevoerd in het systeem. Als het veld 'locatie' niet consequent is ingevuld, dan wordt de voorspelling waar de komende tijd regelmatig inbraken zullen plaatsvinden erg onbetrouwbaar.

Wie, wat, waar

"Gegevensmanagement heeft als taak om in beeld te hebben welke gegevens er beschikbaar zijn, wat de kwaliteit ervan is en wie er voor deze gegevens verantwoordelijk is. Deze gegevensvragen worden door gegevensmanagement beantwoord. Bij een datagedreven werken project is één van de fases 'het in beeld brengen van de benodigde gegevens'. Op het moment dat dit bekend is, is er niets zo vervelend dat niet duidelijk is waar die benodigde data te vinden zijn."

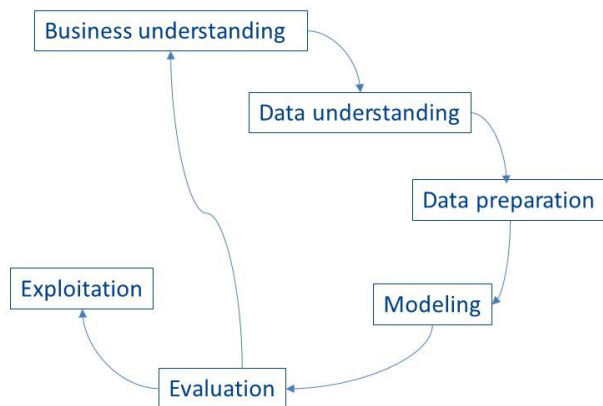
Gegevensmanagementproducten als gegevenslandschap en gegevenswoordenboek zijn hierbij onmisbaar.

Wie is de baas over de gegevens??

Er zijn veel gegevens beschikbaar. Kan elk datagedreven werken project dan zo maar worden uitgevoerd met deze data? Is er iemand de 'baas' over deze gegevens? Wie is er én voelt zich er verantwoordelijk voor dat de juiste data beschikbaar zijn, beschikbaar gesteld worden en gebruikt worden? Gegevensmanagement bepaalt en organiseert wie welke taken, verantwoordelijkheden en bevoegdheden heeft. Producten als het functiehuis en de organisatie rond gegevensmanagement helpen hierbij.

Bijlage 1: CRISP-DM

Datagedreven werken:



Eén van de methodes die je kunt toepassen bij datagedreven werken is de CRISP-DM methode. Het startpunt van de methode is Business Understanding. Welke doelstellingen moeten worden ondersteund? Welke vragen moeten beantwoord worden? Welke middelen zijn beschikbaar? **Dit zijn vragen die een relatie hebben met bewustwording, beleid, wet- en regelgeving,**

De volgende stap is Data Understanding. Welke data is noodzakelijk? Waar is deze data te vinden? Hoe bruikbaar en betrouwbaar is deze data? **De laatste twee vragen kunnen snel en degelijk beantwoord worden als gegevensmanagementproducten (zoals: gegevenslandschap, model, woordenboek, kwaliteit, dataclassificatie) op orde zijn.**

Als de data beschikbaar is volgt de stap Data Preparation. Moet de data opgeschoond worden? Moeten er verschillende databronnen gecombineerd worden? Eventueel moet de data in een voor de gekozen analysemodellen in het juiste formaat geschikt gemaakt worden. **Deze vragen hebben een relatie met de producten kwaliteit, gegevenslandschap, woordenboek, roadmap, GLO's,**

De volgende stap is Modeling. Hierin worden de meeste geschikte modellen en methoden voor het betreffende vraagstuk geselecteerd en worden de modellen gebouwd en getest. **Hiervoor kunnen de beschikbare woordenboeken en catalogi worden gebruikt om in ieder geval dezelfde naamgeving te hanteren. Daarnaast zorgt gegevensmanagement voor het borgen van deze modellen zodat deze bij volgende data-gedreven werken initiatieven hergebruikt kunnen worden.**

In de stap Evaluation worden de resultaten van de ontworpen modellen beoordeeld. Indien de resultaten voldoen aan de gestelde eisen kan er overgegaan worden naar de laatste stap, Deployment. In deze stap worden de dataproducten daadwerkelijk in gebruik genomen in de processen van de organisatie. **Deze stap kan ook gebruikt worden om de stappen/ontwikkelingen die gezet zijn op het gebied van gegevensmanagement te evalueren bij te stellen waar nodig en te borgen.**

Om de startpositie en de groeimogelijkheden van een organisatie op het gebied van datagedreven werken in kaart te brengen, kun je gebruik maken van het “DELTA-model of analytical capabilities” van Davenport and Harris.²

² 2 Competing on analytics, updated edition Thomas H. Davenport, Jeanine G. Harris.

Bijlage 2: DELTA Analytic Maturity model

Analytical Maturity Model (DELTA)

	DATA	ENTERPRISE	LEADERSHIP	TARGETS	ANALYSTS
STAGE 5 Analytical Competitors	Relentless search for new data and metrics	All key analytical resources centrally managed	Strong leadership passion for analytical competition	Analytics support the firm's distinctive capability and strategy	World-class professional analysts and attention to analytical amateurs
STAGE 4 Analytical Companies	Integrated, accurate, common data in central warehouse	Key data, technology and analysts are centralized or networked	Leadership support for analytical competence	Analytical activity centered on a few key domains	Highly capable analysts in central or networked organization
STAGE 3 Analytical Aspirations	Organization beginning to create centralized data repository	Early stages of an enterprise-wide approach	Leaders beginning to recognize importance of analytics	Analytical efforts coalescing behind a small set of targets	Influx of analysts in key target areas
STAGE 2 Localized Analytics	Data useable, but in functional or process silos	Islands of data, technology, and expertise	Only at the function or process level	Multiple disconnected targets that may not be strategically important	Isolated pockets of analysts with no communication
STAGE 1 Analytically Impaired	Inconsistent, poor quality, poorly organized	n/a	No awareness or interest	n/a	Few skills, and these attached to specific functions



In welke groeifase een organisatie zich bevindt wordt beïnvloed door de volgende focusgebieden van datagedreven werken:

- Data:

Is gericht op de beschikbaarheid en de kwaliteit van de aanwezige data binnen een organisatie.

Wordt georganiseerd binnen gegevensmanagement met de hoofdproducten gegevens en borging

- Organisatie:

Gaat over aspecten zoals de inrichting van data governance binnen een organisatie, de bedrijfsstrategie, de strategische visie op datagedreven werken, inzicht in de benodigde skills en de inrichting van de organisatie en inzicht in belangrijkste prestatie-indicatoren.

Wordt georganiseerd binnen gegevensmanagement met het hoofdproduct organisatie.

- Leiderschap:

Staat met reden in het midden van het model. Het uitdragen van de visie en mogelijkheden van datagedreven werken door het management, maar ook medewerkers die als sponsor fungeren is een kritische factor in het daadwerkelijk transformeren naar een datagedreven organisatie. Leiderschap betekent ook het faciliteren van datagedreven werken en het creëren van de mogelijkheden en omstandigheden voor datagedreven initiatieven.

Wordt deels ook binnen gegevensmanagement georganiseerd met het deelproduct bewustwording

- Doelen:

Geeft inzicht in de belangrijkste kritische prestatie-indicatoren (KPI's) van een organisatie. Op basis van dit inzicht is het mogelijk datagedreven initiatieven te prioriteren, keuzes te maken, kansen te ontdekken en laaghangend fruit te identificeren.

- Vaardigheden:

Gaat over vaardigheden en competenties van de medewerkers. Welke analytische skills zijn er aanwezig en welke data science skills zijn er noodzakelijk om tot een datagedreven organisatie te komen?

Zit ook deels in de gegevensmanagementproducten functiehuis en organisatie