

Datagestuurd werken: een veranderkundige aanpak

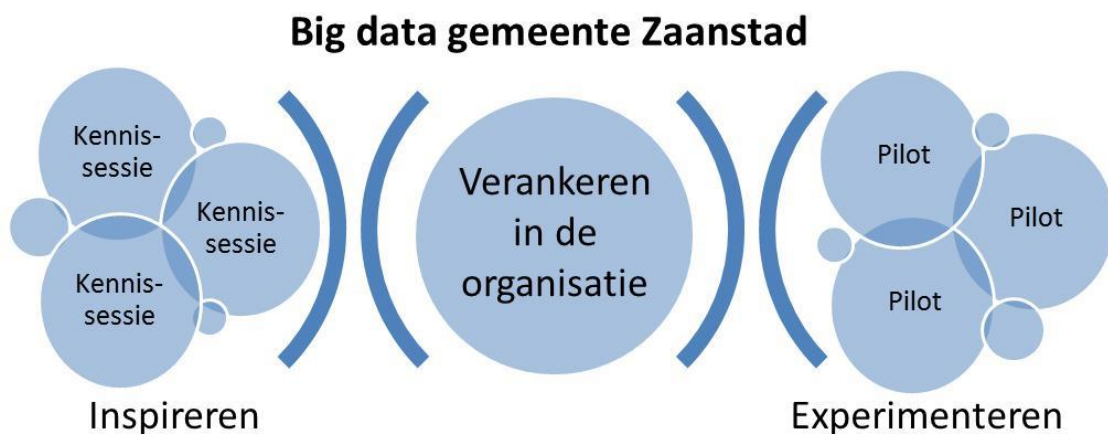
Door Tom Pots

Bij veel publieke organisaties bestaat het geloof dat slim omgaan met data leidt tot betere resultaten voor burgers, zoals schonere straten, veiliger buurten, betere voorzieningen, etc. Maar hoe kom je van dit geloof tot daadwerkelijk betere resultaten op basis van data? Dit vraagt meer dan een data-analist aannemen en die een paar opdrachten laten uitvoeren. Waar begin je? Wat zijn essentiële onderdelen van je aanpak? Hoe begeleid je deze verandering? Wat moet je organiseren? Wat zijn belangrijke valkuilen en succesfactoren? In dit verhaal neem ik je mee in wat datagestuurd werken vraagt van jou en jouw organisatie én hoe je de eerste stap kunt zetten om op basis van data maatschappelijke vraagstukken aan te pakken.

Haal inspiratie van buiten, begin klein en sluit aan op wat de organisatie aan kan

Binnen de gemeente Zaanstad kwam data een aantal jaar geleden op de agenda zoals bij veel organisaties. 'Big data' was op dat moment één van meest besproken onderwerpen op congressen. Een leidinggevende van de IT-afdeling heeft dit opgemerkt en besloten dat we hier iets mee moesten. Aan mij de vraag om dit op te pakken en te onderzoeken wat de kansen en mogelijkheden waren voor onze organisatie. Waar begin je dan?

In onze organisatie was er geen gemeenschappelijk beeld wat big data was, geen ervaring met big data projecten en geen eerdere resultaten. Kortom, we zijn gestart met inspiratie van buiten, kleine experimentele pilots en verankeren in de organisatie (zie onderstaande figuur). Het doel van de kennissessies was tweeledig: inzicht in de kansen en mogelijkheden van big data en de ideeënstroom op gang te brengen. Op het moment dat de energie goed is en een idee interessant genoeg kan er geëxperimenteerd gestart worden. We dachten groot, maar zijn klein begonnen én sloten aan op wat de organisatie aan kon.



Creëer een gemeenschappelijke taal

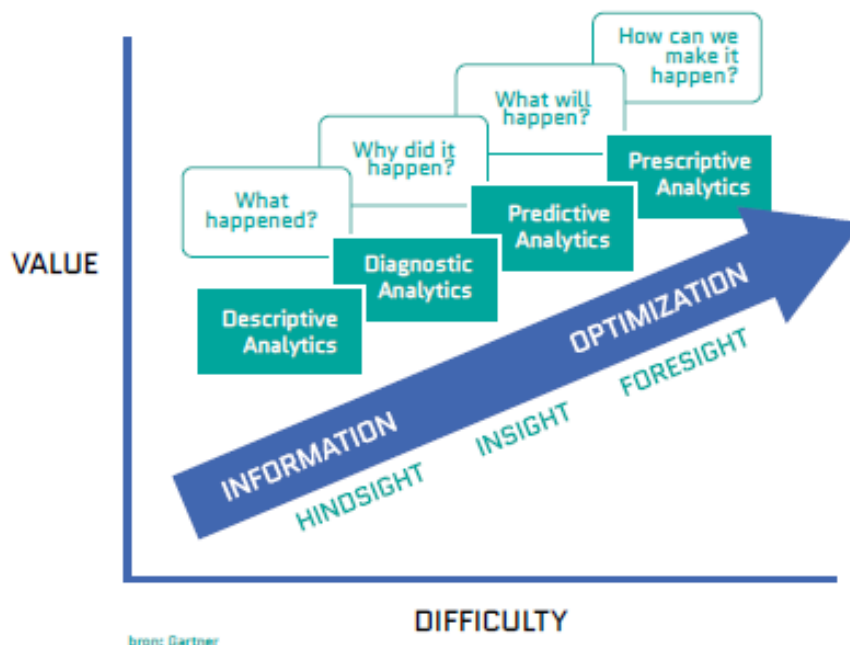
Tijdens het grote aantal pilots kwamen we er ook achter dat een gemeenschappelijke taal in onze organisatie over data ontbrak. Allerlei termen, zoals big data, datagestuurd werken, data science, machine learning, onderzoek & statistiek, business intelligence werden kriskras door elkaar gebruikt met wisselende betekenissen.

Een belangrijke eerste stap was om te stoppen met de term big data, omdat big data vooral gericht is op voorspellende modellen en het woordje big voor ons eigenlijk niet relevant is. Ook als iets niet als

big data getypeerd kan worden kun je er relevante inzichten uit halen. We hebben gekozen voor datagestuurd werken omdat dit meer gericht is op het gebruiken van praktische data-analyses in je dagelijkse werk. De volgende vraag is wat betekent datagestuurd werken voor ons? Binnen de gemeente Zaanstad gebruiken we onderstaand model van Gartner om een gemeenschappelijk beeld te creëren wat datagestuurd werken kan betekenen in het dagelijks handelen. Elk maatschappelijk vraagstuk kan hier geconcretiseerd worden. We doen het hieronder aan de hand van fraude:

- Beschrijvende analyse: Wie fraudeert, waar komt het voor en hoe groot is de omvang van fraude?
- Diagnosticerende analyse: Wat zijn de oorzaken van deze fraude, waarom doen mensen het en hoe doen ze het?
- Voorspellende analyse: Kunnen we voorspellen waar fraude gaat plaatsvinden en wat zijn voorspellende risicoprofielen van personen en situaties?
- Voorschrijvende analyse: Welke aanbevelingen kunnen we professionals concreet geven om fraude te voorkomen?

Datagestuurd werken in de echte wereld is dus data-analyses (beschrijvend, diagnosticerend en/of voorspellend) gebruiken bij jouw dagelijkse werkzaamheden. Bij het onderdeel 'organiseer capaciteit: datasetatelier en data-analyseteam' wordt nader toegelicht wie deze analyses kunnen uitvoeren.



Start met de juiste analyse

We zijn met datagestuurd werken dus gestart op de hype van big data. De ontwikkeling van big data gaat vooral over vooruitkijkend (voorspellend en voorschrijvend). Voorspellende modellen maken op basis van veel data wordt door veel organisaties gezien als een zeer interessante ontwikkeling. Veel organisaties waren aan het experimenteren in data science pilots. Wat is de waarde van voorspellende analyses voor onze organisatie? Ook wij wilden dit ervaren en zijn bij de start drie data-science pilots gestart in het maatschappelijk domein. We kwamen er al snel achter dat de eerste stap eerst goed terugkijken is. Als de beschrijvende en diagnosticerende analyses niet bestaan of gebruikt worden, dan ben je bij voorspellende analyses kansloos vanaf het begin.

Creëer een paar successen en bouw daar een krachtig verhaal omheen

Zoals gezegd zijn we gestart met een groot aantal pilots in verschillende domeinen. We waren op zoek naar een paar lokale successen die onomstotelijk de waarde van datagestuurd werken laten zien. Hieronder twee voorbeelden die veelvuldig verteld zijn:

- In Keulen zaten vluchtelingen aan vrouwen. In de Europese media werd al snel geroepen dat er alleen testosteronbommen binnen kwamen. Dit werd door de landelijke media overgenomen. We hebben toen uitgezocht hoe het voor Zaanstad zat en uit de beschrijvende analyse bleek dat het aantal mannen en vrouwen ongeveer gelijk was. Bovendien was het gedeelte dat als testosteronbom kon worden bestempeld, maar een klein gedeelte van het geheel. Uit de analyse bleek echter ook dat 75% van alle statushouders in het verleden niet uit de bijstand is gekomen. Deze feiten hebben we in het politieke debat gebracht, waardoor het niet over testosteronbommen ging, maar over hoe krijgen we ze aan het werk. Een hele simpele beschrijvende analyse die echt het verschil maakte.
- In Zaanstad staan veel panden op (korte) houten funderingspalen. Met name grenenhout is kwetsbaar voor bacteriële aantasting. Hierdoor neemt de draagkracht van de fundering af en ontstaat er schade aan een pand. Funderingsherstel is dan noodzakelijk. Voor dit vraagstuk geldt dat de beschrijvende en diagnostiserende analyse op orde zijn en daarom hebben we hier een voorspellend model ontwikkeld waarmee funderingsproblematiek in Zaanstad voorspeld kan worden. We kunnen nu een adres invoeren en met 95% zekerheid voorspellen of het huis funderingsproblematiek heeft.

Het doel van deze verhalen was om andere inhoudelijke teams te verleiden om ook een pilot te starten en de waarde van datagestuurd werken te tonen aan de organisatie.

Creëer ruimte door de hele organisatie te laten kiezen voor data

Ik heb verteld hoe big data (daarna datagestuurd werken) op de agenda van onze organisatie is gekomen en hoe ik er verantwoordelijk voor werd. Los van het feit dat ik het thema een heel interessant en relevant vond, was ik wel alleen en had ik geen idee hoe de rest van de organisatie tegen het thema aankeek. Daarom ben ik parallel aan de inspiratiesessies en pilots gestart met het proces om te komen tot een gedragen informatievoorzieningsstrategie.

In de IV-strategie dwingen we onszelf om jaarlijks naar nieuwe ontwikkelingen te kijken en samen te duiden: willen we er iets mee of niet? Hiervoor hadden we informeel ook een soort van IV-strategie. Onze leidinggevende van de IT-afdeling ging naar een aantal congressen, legde de belangrijkste thema's op tafel en vroeg aan de informatiemanagers wie welk thema wilde op pakken. Het nieuwe proces is gestart met het interviewen van college, raad, directie, strategen, informatiemanagers, IT-leverancier, bank, sociaal wijkteam, zorginstelling, woningbouwcorporatie met de vraag welke relevante ontwikkelingen komen op de gemeente af en waar moeten wij iets mee? Uit deze ronde zijn twaalf thema's naar voren gekomen, die vervolgens aan alle leidinggevendenden van de gemeente Zaanstad zijn voorgelegd met de vraag: welke drie thema's moeten opgepakt worden en welke negen gaan we niet doen? Bijna net zo belangrijk als de top 3 is de keuze om negen thema's niet te doen.

Er zijn drie thema's geprioriteerd, waarbij datagestuurd werken met stip op 1 stond. Het risico van dit proces is dat het thema ook niet geprioriteerd kan worden. Dat is goed om te weten, omdat je dan energie steekt in een thema dat niet relevant wordt gevonden door de macht van de organisatie. Het proces om thema's binnen en buiten de organisatie op te halen en daarna te laten prioriteren door alle leidinggevendenden had een aantal voordelen

1. Het was niet meer mijn thema, maar dat van de organisatie.
2. Er stond een hek om de verandering, want de macht had gesproken
3. Inzet van capaciteit (geld en mensen) werd veel eenvoudiger

Nu ruim twee jaar verder ben ik niet alleen verantwoordelijk om resultaten te boeken, maar zijn we met 10-15 personen onze ambities op het gebied van datagestuurd werken aan het realiseren.

Ambities datagestuurd werken

1. Datalab



2. Datapakhuis



3. Data-academy



4. Dashboard burger & Leefomgeving



5. Toevoegen data-analyse aan teams



6. Verankeren in organisatie



Ontwikkel een middel om complexe vraagstukken aan te pakken op basis van data

Opgaven in de samenleving laten zich niet vangen door grenzen van een organisatie. Iedereen denkt na over datagestuurd werken. Kan data ons verbinden? Vanuit dat perspectief zijn we gestart met het Data Lab. Het doel van de eerste pilot was om de waarde hypothese te bewijzen, dus kunnen we een middel ontwikkelen dat helpt om maatschappelijke vraagstukken een beetje beter maken. We hebben Parteon, GGD, Politie, gemeente Purmerend, Agora, Eveen, UWV en VU/UvA gevraagd om mee te doen, samen de aanpak te ontwikkelen en de waarde voor andere maatschappelijke vraagstukken te beoordelen. Hieronder de aanpak op hoofdlijnen:

1. Maatschappelijke vraagstuk prioriteren

In deze stap wordt duidelijk wat de doelen van de deelnemende organisaties zijn, wat de kern van het vraagstuk is en wat we willen weten om het vraagstuk aan te pakken. Na deze fase kan een organisatie bepalen in of uit te stappen. Extramuralisering & overlast is het eerste maatschappelijke vraagstuk dat opgepakt is.

2. Samen feitelijke foto maken

In de tweede stap worden alle relevante informatie uit factsheets en onderzoeken gehaald die al in de verschillende organisaties aanwezig zijn over het voorliggende thema, zodat we in beeld krijgen wat we samen al weten en welke verdiepende analyses nog uitgevoerd moeten worden. De inhoudelijk experts extramuralisering & overlast van de deelnemende organisaties zeiden na de feitelijke foto: "wat weten we al veel zonder dat we het weten, wat ligt er al veel zonder dat we het gebruiken".

3. Veilig delen data

Bij elke data-analyse speelt privacy een belangrijke rol en deze opgave wilden voor alle toekomstige vraagstukken in het Data Lab slechten. Het is ons gelukt om een werkwijze te ontwikkelen waarin de privacy gewaarborgd is en het mogelijk maakt om data te gebruiken voor analyses. Het doorlopen van de aanpak veilig delen data zorgt ervoor dat de je erachter komt wat

wel en niet mag. Een deelnemer zei na deze stap dat hij nog niet eerder had meegemaakt dat zulke gevoelige data tussen organisaties daadwerkelijk gedeeld werd op een veilige manier.

4. Verdiepende data-analyses

In de vierde fase worden verdiepende analyses (beschrijvend, diagnosticerend en voorspellend) uitgevoerd.

5. Oplossingen bepalen vanuit eigen rol en positie deelnemers

In fase 5 is het beschouwen afgelopen. Er mogen geen vragen meer gesteld worden. Nu gaan we aan de slag met inzichten en oplossingen.

1. Wat zijn opvallende inzichten (op basis van stap 2 - feitelijke foto en stap 4 - verdiepende analyses)?
2. Wat betekenen deze inzichten voor de praktijk (per organisatie en/of tussen organisaties) en waarom?
3. Wat zijn mogelijke oplossingsrichtingen voor het vraagstuk?

Beperkt gebruik van WMO door mensen met verward gedrag, slechts 10 personen zorgen voor grootste gedeelte van overlastmeldingen, meldingen zijn gecentreerd, maar zorg verspreid, beperkte regie op de meldingen en veel meldingen over vervuiling zijn een paar voorbeelden van de gemeenschappelijke inzichten. Op basis van de inzichten zijn er mogelijke oplossingsrichtingen vormgegeven.

Na het eerste Data Lab waren alle deelnemers positief over de waarde hypothese van het Data Lab voor andere maatschappelijke vraagstukken. Het doel was vervolgens om het Data Lab door te ontwikkelen van pilot naar ecosysteem waar maatschappelijke organisaties (zorginstellingen, politie, GGD, woningcorporaties, scholen, etc.), wetenschap (VU/UvA) en gemeenten gezamenlijk aan opgaven in de samenleving werken op basis van data. In de eerste pilot hebben we de werkwijze voor toekomstige Data Labs ontwikkeld en inmiddels hebben 35 gemeenten en maatschappelijke partners de intentie uitgesproken om onderdeel te zijn van het Data Lab en bij voor de organisatie relevante vraagstukken deel te nemen. Vraagstukken die in de opstart, uitvoering of afgerond zijn: extramuralisering & overlast, Jongens/meisjes in het onderwijs, ontwikkeling bedrijventerrein Molletjesveer Noorderveld, versnellen woningbouwopgave, vroegsignalering, schuldenproblematiek, operationele zorgvragen, energie transitie en brandveiligheid. Het is onze ambitie om tien tot vijftien Data Labs in 2019 uit te voeren voor maatschappelijke vraagstukken die zich niet laten vangen door de grenzen van een organisatie.

Organiseer capaciteit: datasetatelier en data-analyseteam

Als je aan de slag wil met maatschappelijke vraagstukken op basis van data dan zijn twee soorten capaciteit noodzakelijk. Wie gaat data ontsluiten? Wie gaat data analyseren? Beiden vragen moeten verankerd worden in je organisatie:

1. Datasetatelier ontsluit data

Het datasetatelier zijn de informatie-architecten en gegevensbeheerders die datasets prepareren vanuit de bron en deze publiceren in het datapakhuis in drie verschijningsvormen (zie volgende onderwerp - ontsluit in samenhang alle beschikbare data en visualiseer die in dashboards).

2. Data-analyseteam analyseert data

De volgende vraag is wie maken de beschrijvende, diagnosticerende en voorspellende analyses? In Zaanstad gebruiken we onderstaand model om deze vraag te beantwoorden.

- Business Intelligence ontwikkelaar voor beschrijvende analyses

Beschrijvende analyses zijn bij uitstek geschikt om te automatiseren in dashboards. Binnen de gemeente Zaanstad gebruiken we Qlikview voor de beschrijvende analyses.

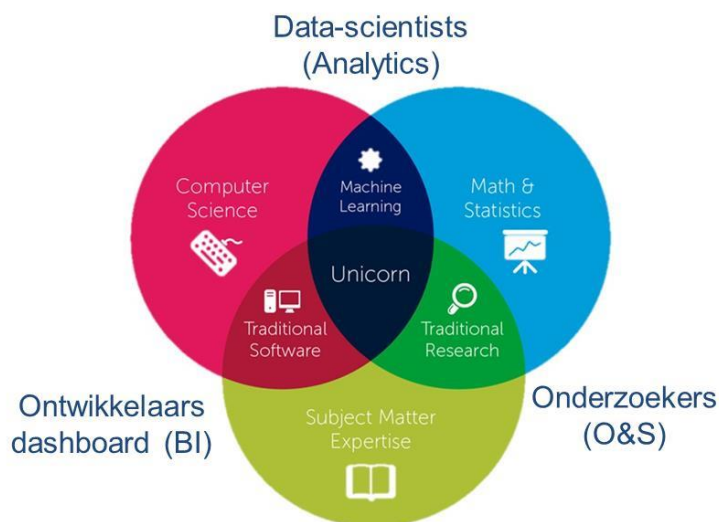
- Onderzoekers voor diagnostiserende analyses

Combineer je wiskunde & statistiek met inhoud dan gaat het over traditioneel onderzoek. Relevant om de diagnostiserende analyse te maken.

- Data-scientists voor voorspellende analyses

Ten slotte is er een nieuwe discipline bijgekomen. Data-scientists die machine learning toepassen om voorspellende analyses te maken.

In Zaanstad zijn alle mensen (Qlikviewontwikkelaar/onderzoekers/data-scientists) die beschrijvende, diagnostiserende en voorspellende analyses uitvoeren samengebracht in het data-analyseteam. Het data-analyseteam is in staat om data-analyses in de volle breedte van onze organisatie in gezamenlijkheid uit te voeren. De ene keer is een Qlikview dashboard nodig, de andere keer kwalitatief of kwantitatief onderzoek en dan is een voorspellend model de juiste oplossing.



Ontsluit in samenhang alle beschikbare data en visualiseer die in dashboards

We beschikken als gemeente over een groot aantal gegevensbronnen, die tot vorig jaar beperkt, dan wel slecht ontsloten waren. Dit maakt het delen van data tussen medewerkers, laat staan met maatschappelijke partners en burgers een tijdrovende en complexe bezigheid. Bovendien is niets zo frustrerend als een data-opdracht te willen starten, maar de datasets zijn niet beschikbaar. Onze informatie-architecten hebben toen het concept Datapakhuis bedacht. Het Datapakhuis faciliteert expliciet het delen van data met de binnen- en buitenwereld, waarbij de privacy gewaarborgd is.

Bij de start van het Datapakhuis hebben we een overzicht gemaakt van alle datasets en zagen we dat alle datasets te ordenen zijn langs twee thema's: burger en leefomgeving. Inmiddels hebben we alle datasets van een burger onder de knop zitten in drie verschijningsvormen: open, gesloten en private. Daarnaast is elke dataset in samenhang ontsloten, dus bsn (versleuteld), maand (tijd), wijk en buurt (plaats). Doordat alle datasets in samenhang zijn ontsloten is een dashboardontwikkelaar, onderzoeker en data scientist niet meer vijf dagen bezig met datapreparatie, maar heeft een dashboardontwikkelaar binnen een paar uur beeld en kunnen onderzoeker en data scientist veel sneller aan de slag.

Uit het data-analyseteam blijkt dat van alle data-analyses die we in de gemeente uitvoeren ongeveer 85% beschrijvend (wie, wat, waar en hoeveel) van aard is. Het proces van zo'n opdracht gaat nu

ongeveer als volgt: vraag medewerker, dataset(s) leveren door gegevensbeheerder(s), analyseren door onderzoekers of data-analist, antwoord op vraag. De doorlooptijd is snel 4-8 weken.

Hieronder een super interessant filmpje dat de kracht van beschrijvende analyses laat zien. Hans Rosling laat zien hoe snel hij feitelijke foto's kan maken en visualiseren. De wijze waarop hij inzichten genereert die je veronderstellingen ter discussie stelt is super interessant!

Let my dataset change your mindset

https://www.ted.com/talks/hans_rosling_at_state?utm_source=tcdcomshare&utm_medium=email&utm_campaign=tcdspread

Stel je voor dat wij iets vergelijkbaars voor onze datasets kunnen doen en we dit kunnen gebruiken in de maatschappelijke vraagstukken, zoals: vroegtijdig schoolverlaten, arbeidsmarkt, armoede, woningbouw, etc. We hebben inmiddels het merendeel van de datasets van de gemeente in samenhang ontsloten in het Datapakhuis en grofweg zijn ze te ordenen in twee categorieën:

1. Burger: bijv. alle producten in het maatschappelijke domein, kenmerken personen, etc.
2. Leefomgeving: bijv. woningen, leegstand, objecten openbare ruimte, etc.

Voor beide categorieën zijn we een dashboard aan het maken met beschrijvende analyses (wie, wat, waar en hoeveel), waarbij het dashboard burger inmiddels in gebruik is en het dashboard leefomgeving begin 2019 gereed is. Het doel is dat de doorlooptijd van opdrachten wordt gereduceerd van 4-8 weken, naar 15 minuten. Een strateeg, visie-specialist, etc. stelt de vraag niet aan een onderzoeker of data-analist, maar kijkt zelf.



In de komende maanden worden de datasets van de leefomgeving ontsloten. We hebben de ambitie om begin 2019 alle datasets van de gemeente onder de knop te hebben in de drie verschijningsvormen en ontsloten in twee dashboards: burger en leefomgeving, zodat enerzijds de doorlooptijd van beschrijvende analyses flink wordt gereduceerd en anderzijds het data-analyseteam tijd heeft voor de complexere analyses (diagnosticerend en voorspellend)

Doe het leren & inspireren niet alleen, maar altijd over grenzen heen

Bij de start zijn we begonnen met kennissessies om inspiratie van buiten te halen om de mogelijkheden te ontdekken van werken met data. De sessies waren alleen voor de ambtenaren van Zaanstad. Op een gegeven moment stelden we onszelf de vraag: waarom nodigen we de omliggende gemeenten en regionale partners eigenlijk niet uit? We kunnen dan het leren en inspireren samendoen. Er worden in de Data Academie trainingen, masterclasses en inspiratiesessies georganiseerd voor de gemeenten in Zaanstreek Waterland en maatschappelijke partners in de regio.

In de Data Academie staan de volgende uitgangspunten centraal: (1) Inspiratie & leren. Lokale successen van binnen en buiten de organisaties en mensen leren werken met data. (2) Kansen & mogelijkheden. Op basis van de verhalen van de lokale successen proberen we interesse te kweken om te gaan experimenteren met datagestuurde werken. Het doel is een leergemeenschap die die leert en ontwikkelt over de grenzen van de organisaties heen.

Stapelen van lokale verbeteringen

Datagestuurde werken betekent in onze organisatie het gebruiken van beschrijvende, diagnosticerende, voorspellende analyses per inhoudelijk thema. De hype zit op het voorspellen, én de gemeente Zaanstad, andere gemeenten en de publieke sector in zijn geheel hebben een wereld te winnen op het goed gebruiken van beschrijvende en diagnosticerende analyses. Als je naar een gemeente kijkt dan bestaat deze uit grofweg 50 inhoudelijke thema's van vroegtijdig schoolverlaten, funderingen, armoede, werkgelegenheid, jeugdzorg, woningbouw, etc. Het is de uitdaging om per inhoudelijk thema scherp te krijgen welke analyse door het team nu gebruikt wordt en dan een analyse toe te voegen. Bij woningbouw bleek de beschrijvende analyse niet op orde, dus zijn we bij deze analyse gestart. Bij funderingen waren de beschrijvende en diagnosticerende analyse op orde, dus was het startpunt de voorspellende analyse. Het idee hierachter is dat door de stapeling van lokale verbeteringen, waarbij inhoudelijke teams data-analyses beter toepassen, we als organisatie in zijn geheel meer datagestuurde zijn gaan werken.

Datagestuurde werken van idee tot concrete resultaten

Hieronder een overzicht van de belangrijkste lessen die wij hebben geleerd op onze weg naar het verbeteren van maatschappelijke vraagstukken op basis van data:

1. Haal inspiratie van buiten, begin klein en sluit aan op wat de organisatie aan kan
2. Creëer een gemeenschappelijke taal
3. Start bij de juiste analyse (beschrijvend, diagnosticerend of voorspellend)
4. Creëer een paar successen en bouw daar een krachtig verhaal omheen
5. Creëer ruimte door de hele organisatie te laten kiezen voor data
6. Ontwikkel een middel (bijv. Data Lab) om complexe vraagstukken aan te pakken op basis van data
7. Organiseer capaciteit: datasetatelier en data-analyseteam
8. Ontsluit in samenhang alle beschikbare data (bijv. Datapakhuis) en visualiseer die in dashboards
9. Doe het leren & inspireren niet alleen, maar altijd over grenzen heen (bijv. Data Academy)
10. Stapelen van lokale verbeteringen om als organisatie in zijn geheel meer datagestuurde te worden

Een mooi compliment aan ons team en de behaalde resultaten is dat er de afgelopen jaar veel interesse is vanuit de publieke sector voor de wijze waarop wij de verandering van datagestuurde werken aanvielen in Zaanstad. Ik ben daar trots op én we staan nog maar aan het begin. Er ligt een forse uitdaging om in de komende jaren om alle inhoudelijke teams die een gemeente rijk is, van jeugdzorg tot woningbouw, te leren werken met data.