

# **Data management**

**Wat u moet en kunt doen voor het optimaal  
beheren en gebruiken van data in uw organisatie**

**VOOR BETERE ICT BESLISSINGEN**



# Data management

**Wat u moet en kunt doen voor het optimaal  
beheren en gebruiken van data in uw organisatie**



**Uitgegeven door het ICT informatiecentrum, Houten**

**© 2020 ICT informatiecentrum**

Alle rechten voorbehouden. Het is de ontvanger van deze publicatie verboden de inhoud ervan geheel of gedeeltelijk te verveelvoudigen, openbaar te maken, digitaal te verspreiden of op welke wijze dan ook te distribueren, zonder voorafgaande toestemming van de uitgever. Hoewel deze uitgave met zorg is samengesteld, aanvaardt de uitgever geen enkele aansprakelijkheid voor schade ontstaan door het gebruik ervan en fouten of onvolkomenheden in de gepubliceerde teksten.

# Inleiding

U heeft steeds meer data tot uw beschikking. Operationele systemen, zoals uw ERP of CRM software, hebben hun eigen databases met commerciële en financiële data. Daaromheen komt veel data beschikbaar via andere systemen, websites en digitale communicatie. Ook voor deze data zult u over databases moeten beschikken. U kunt uw data gebruiken voor analytische doeleinden, zodat u bedrijfsprocessen kunt verbeteren en beter onderbouwde beslissingen kunt nemen. Maar ook als u niet direct een analytische toepassing heeft, moet u een oplossing hebben om uw data op te slaan, te verwerken, terug te vinden én te beveiligen. Cybercriminelen zitten niet stil en wettelijk gelden er steeds meer verplichtingen om de veiligheid van met name persoonsgegevens te waarborgen. Al deze aspecten liggen op het terrein van data management.

Beschikt u over data, dan heeft u dus altijd te maken met het beheer daarvan. Het is uw verantwoordelijkheid dat goed te regelen. Wat daarbij komt kijken, leest u in dit boekje van het ICT informatiecentrum. U leest over wat data management is, over master data, opslag, databases, data lakes en de integratie van data uit verschillende bronnen. Datakwaliteit komt uitgebreid aan de orde, evenals data security. Kennis en informatie dus, die nodig is als u projecten rondom data management start, het beheer van data in uw organisatie wilt verbeteren of als u zich wilt voorbereiden op de volgende stap: het toepassen van data analytics en business intelligence.

Heeft u hulp nodig, omdat u met data management aan de slag wilt, heeft u vragen over dit thema, zoekt u een specialist of wilt u meer kennis en informatie, laat het ons weten. Wij helpen u graag!

# Met dank aan ...

De inhoud van dit boekje is tot stand gekomen met medewerking van diverse specialisten voor dit thema. Wij noemen ze 'kennispartners'. Staat een kennispartner als bron bij een artikel vermeld, dan kunt u het artikel plaatsen in de context van die kennispartner en zijn of haar activiteiten.

Het is goed om te weten dat niemand ons betaalt voor publicatie van zijn of haar bijdrage aan deze publicatie. Ook betaalt het ICT informatiecentrum niemand voor een tekstuele bijdrage. Zo zijn en blijven wij volledig onafhankelijk in de samenstelling van onze publicaties.

Uiteraard danken wij onze kennispartners voor hun medewerking. Mede dankzij hun ervaringen, visies, tips en adviezen worden ICT beslissingen iedere dag weer een beetje beter.

# Inhoud

Inleiding	3
Met dank aan ...	4
Introductie	7
Wat is data management?	10
Waarom data management?	12
Vijf elementen van een data strategie	14
Master data management	17
Voordelen van een multi-domein MDM systeem	20
Data governance	25
De zin en onzin van data management en data governance	28
Rollen en verantwoordelijkheden in de data governance organisatie	31
Drie actuele data governance trends	35
Vijf lessen bij de implementatie van data governance	37
Data opslag	41
Databases	45
Data warehouse	51
In 5 stappen naar een succesvol datawarehouse	58
Data lake	62
Data migratie	65
Data integratie	68
Data kwaliteit	71
Datakwaliteit; geen project maar een proces	74
Tien basistips voor het verhogen en waarborgen van datakwaliteit	78
Wanneer is data van goede kwaliteit?	82
Belang van datakwaliteit	84
Kosten van slechte data	91
Wat is datapreparatie en waarom is het zo belangrijk?	95
Afwijkende transactiecijfers: wat zijn dat en waar komen ze vandaan?	99

Management van datakwaliteit: zo pak je dat aan	103
Het belang van een data dictionary	106
De waarde van data kwaliteit tooling	112
Wat is data security?	115
Data security tips en adviezen	120
5 tips voor het opstellen van informatiebeveiligingsbeleid	126
Data management platform	130
5 voordelen van een data management platform	132
Voordelen van een customer data platform	138
Data hub: informatie, inzicht, actie!	141
De juiste dataoplossing: PIM of PDM	144
Waarmee kunnen wij u helpen?	146
Kennispartners	147

# Introductie

Bron: ICT informatiecentrum | [ictinformatiecentrum.nl](https://www.ictinformatiecentrum.nl)

De noodzaak van goed data management is de afgelopen jaren groter geworden, vooral omdat in bijna iedere organisatie de hoeveelheid data enorm is toegenomen. Zo zijn er steeds meer databronnen in uw bedrijfssoftware en IT systemen. Daarnaast verlopen interne en externe communicatie inmiddels hoofdzakelijk via digitale kanalen, zoals netwerken en internet. Ook is het belang van databeveiliging toegenomen door stijgende cybercriminaliteit en maakt wet- en regelgeving, zoals de AVG, het noodzakelijk dat uw databeheer aan strengere eisen voldoet.

## **Data management**

Voordat u uw data zinvol kunt gebruiken, is het structureren, valideren, samenvoegen, eenduidig maken, aanvullen en opslaan ervan noodzakelijk. Deze taken en begrippen vallen onder data management. Andere begrippen en functionaliteiten die te maken hebben met data management zijn dataopslag (databases, datawarehouse, data lake, datacenters), data kwaliteit, data governance, data migratie, data security, data virtualisatie en data integratie.

## **Data analytics**

Heeft u uw data management eenmaal op orde, dan kunt u gebruik maken van data analytics of business intelligence om 'waarde' uit uw data te halen. Door het analyseren van data ontdekt u trends, krijgt u nieuwe inzichten of legt u verbanden die belangrijk voor u kunnen zijn. Denk hierbij aan het optimaliseren van uw bedrijfsprocessen, het analyseren van de markt waarin u actief bent of het nemen van strategische en tactische beslissingen nemen.

## **Master data management**

Met master data of stamdata worden de belangrijkste gegevens bedoeld

over personen, zaken en locaties, die essentieel zijn voor een goede bedrijfsvoering. Master data management (MDM) is gericht op het beheer van deze belangrijkste gegevens. De moeilijkheid van master data management is dat de stamdata voorkomt in verschillende systemen. MDM zorgt voor consistente data en maakt deze bruikbaar voor de verschillende systemen. Meer hierover leest u in het boekje [Master data management](#) van het ICT informatiecentrum.

### **Document management**

Document management richt zich op het beheer van de in uw organisatie aanwezige documenten. De IT oplossing die u hiervoor gebruikt, is een document management systeem (DMS). Naast de verwerking en opslag van digitale documenten, zorgt een DMS voor de digitale opslag van papieren documenten en video- en audiobestanden. Ook document management heeft als doel de efficiency van uw organisatie te verhogen. Goed documentbeheer zorgt ervoor dat informatie in documenten snel vindbaar en beschikbaar is. Meer informatie vindt u in een aparte sectie over document management of de speciale website over [DMS software](#).

### **Content management**

Data die in de vorm van tekst of beeld gebruikt wordt op websites en in documenten, wordt aangeduid met content. Het beheer hiervan is content management. Alles wat op websites online gepubliceerd wordt, valt onder web content management. Voor het beheer daarvan gebruikt u een content management systeem (CMS). Gaat het om content management binnen het bedrijfsproces van uw organisatie, dan spreken we over enterprise content management (ECM).

### **Opslag en beheer**

Data management richt zich op het beheer van uw data en daarmee ook de opslag ervan. Voor dataopslag kunnen relatief eenvoudige databases toegepast worden. Maar komt data uit verschillende bronnen, dan zal deze centraal verzameld moeten worden. Hiervoor maakt u gebruik van een



datawarehouse, verzamelt u alle data in een data lake of richt u een data management platform in.

### **Dataopslag**

Dataopslag gaat over de wijze waarop digitale data wordt bewaard of weggeschreven. Dit kan op verschillende manieren. Data kan lokaal worden opgeslagen op een interne of externe harde schijf, zolang het gaat om een enkele gebruiker. In bedrijfsnetwerken kan data op een lokale server bewaard worden. Tegenwoordig vindt voor een steeds groter deel dataopslag plaats in de cloud. Hiermee worden oplossingen via internet bedoeld die gebruik maken van opslag van data bij datacenters.

### **Datacenters**

Data management gaat over telkens toenemende hoeveelheden data. Om diverse redenen kan het wenselijk of nodig zijn het beheer van data en de bijbehorende applicaties extern onder te brengen. U kunt daarvoor gebruik maken van een datacenter. Internet verbindt u met het datacenter. Dit betekent dat u altijd en overal via internet toegang heeft tot uw data en de bijbehorende applicaties.

### **Hoe begint u met data management?**

Data management betreft zowel eenvoudige als zeer complexe oplossingen en alles daartussenin. Door de veelheid aan mogelijkheden, keuzes en oplossingen komt u al snel op een terrein waar uw eigen kennis en ervaring niet toereikend zijn. In projecten rond data management is dan ook bijna altijd in een vroeg stadium externe hulp nodig. Laat u tijdig informeren en begeleiden. Veel informatie vindt u in de boeken, whitepapers en andere media van het ICT informatiecentrum. Onze DATA box biedt u toegang tot alle kennis en informatie die u nodig heeft om uw dataproject te kunnen starten en uit te voeren. Maak er optimaal gebruik van!

# Wat is data management?

Bron: ICT informatiecentrum | [ictinformatiecentrum.nl](http://ictinformatiecentrum.nl)

Data management staat meer dan ooit in de belangstelling. Het feit dat u dit leest, geeft aan dat ook u interesse heeft in oplossingen die kunnen helpen bij het beheer van uw gegevens. Dat is verstandig, omdat de hoeveelheid data in en rond uw bedrijfsvoering alleen maar toeneemt. En daarmee waarschijnlijk ook de behoefte om daar op een effectieve en efficiënte wijze gebruik van te maken. Hoe beter u in staat bent om data op een goede manier te gebruiken, des te groter zijn de voordelen die u ermee kunt behalen. Het resultaat kan zich vertalen in lagere kosten, hogere winsten en concurrentievoordelen. Belangrijk dus om u te verdiepen in de toepassing en mogelijkheden van data management.

## **Doel van data management**

Data management gaat over het onderhouden, actualiseren, beheren en beveiligen van data. Hiervoor worden gegevensbestanden gecontroleerd op mutaties en zo nodig aangepast. Met nieuwe en aanvullende data uit externe bronnen verrijkt u bestaande bestanden. Dankzij koppelingen met externe bestanden kan het actueel houden van data steeds vaker geautomatiseerd worden. Doel van data management is om ervoor te zorgen dat de aanwezige data compleet, betrouwbaar en op tijd beschikbaar is voor toepassingen die van die data gebruik maken. Dat kan uw bedrijfssoftware zijn of andere IT systemen. Op basis van goede en complete data kunt u - na analyse daarvan - de uitvoering van bedrijfsprocessen optimaliseren en wordt u geholpen bij het nemen van de juiste beslissingen. Door de toevoeging van de term 'enterprise' voor het begrip data management (enterprise data management, ofwel EDM) wordt benadrukt dat dit het beheer van gegevens van bedrijven en organisaties betreft. Omdat dat hier vanzelfsprekend is, beperken wij ons tot het begrip data management.

## **Belang van databeheer**

Betrouwbare data kan helpen bij het goed uitvoeren van uw bedrijfsprocessen en het nemen van goede managementbeslissingen. Steeds meer dwingen wettelijke bepalingen u tot het gestructureerd beheer van data, zoals rond persoonsgegevens het geval is. De betrouwbaarheid en kwaliteit van uw data kan echter niet gegarandeerd worden, als deze niet actief onderhouden wordt. Data verouderd immers snel en is dan verminderd of zelfs geheel niet bruikbaar. Naast actualiteit is de consistentie van data belangrijk. Als data in verschillende databases binnen uw organisatie betrekking heeft op hetzelfde item, is het van belang dat deze data in alle databases hetzelfde is. Zo wordt voorkomen dat fouten in de spelling van bijvoorbeeld bedrijfsnamen, persoonsnamen of adressen door de verschillende systemen gezien worden als respectievelijk twee bedrijven, twee contactpersonen en twee vestigingen. Uitgangspunt is dat organisaties die hun data management goed op orde hebben, daar veel strategische en operationele voordelen mee behalen.

## **Datawarehouse**

Het gegevensbeheer in uw organisatie kan worden vergemakkelijkt met een centrale database, waar alle bedrijfsdata op één plek bewaard en beheerd wordt. Het bewaken van de kwaliteit van de data en het voorkomen van fouten wordt hiermee eenvoudiger. Andere bedrijfsapplicaties kunnen gebruik maken van deze database, zodat iedere gebruiker over dezelfde kwalitatief hoogwaardige data beschikt. Het ontsluiten van data is van groot belang om de waarde ervan optimaal te benutten. Gebeurt dit niet, dan is er sprake van data silo's. Dit zijn van elkaar gescheiden en geïsoleerde datasets die niet of nauwelijks worden benut. De oplossing die u nodig heeft voor het centraal verzamelen en opslaan van data uit verschillende bronnen, wordt aangeduid als een datawarehouse.

# Waarom data management?

Bron: ICT informatiecentrum | [ictinformatiecentrum.nl](http://ictinformatiecentrum.nl)

Of u tijd, geld en aandacht gaat besteden aan het verbeteren van uw databeheer is afhankelijk van de waarde die data voor uw organisatie en bedrijfsvoering kan hebben. De onderstaande vragen kunnen u inspireren hierover een standpunt te formuleren.

- Neemt u al uw beslissing op basis van zoveel mogelijk beschikbare gegevens?
- Bent u altijd overtuigd van de actualiteit en betrouwbaarheid van uw data?
- Is data altijd beschikbaar op het moment dat deze binnen uw organisatie nodig is?
- Weet u welke informatie er in uw organisatie beschikbaar is?
- Zijn er klachten in uw organisatie over de beschikbaarheid van data?
- Zijn basisgegevens over klanten en leveranciers in alle systemen altijd hetzelfde?
- Voldoet u aan alle regelgeving met betrekking tot privacygevoelige data?
- Combineert u klantdata en marktgegevens bij het nemen van strategische beslissingen?
- Wat kunnen de gevolgen zijn als u conclusies trekt op basis van onbetrouwbare data?
- Weert u zeker dat de data die u gebruikt altijd juist is?
- Is de veiligheid van uw data gegarandeerd?
- Zijn er bedrijven in uw branche die concurrentievoordeel halen uit hun databeheer?
- Kunt u snel beschikken over financiële, commerciële of productiegerichte overzichten?

- Verwacht u een toename van beschikbare data in de komende jaren?
- Verwacht u dat beter inzicht in de prestaties van uw organisatie u helpt met het aansturen ervan?
- Verwacht u dat beter inzicht in klanten en de markt waarin u actief bent u helpt met betere besluiten daarover?

# Vijf elementen van een data strategie

Bron: Tentive | [tentive.nl](https://tentive.nl)

In de afgelopen 20 jaar is het belang van data sterk toegenomen. Vandaag de dag wordt data zelfs beschouwd als een 'asset' niet als een bijproduct van een applicatie. Organisaties moeten daarom data strategieën ontwikkelen die overeenkomen met de realiteit van vandaag. In dit artikel leest u wat we onder data strategie verstaan en uit welke vijf componenten een gedegen data strategie moet bestaan.

## **Wat is data strategie?**

Data strategie beschrijft een 'reeks keuzes en beslissingen die samen een actie op hoog niveau uitstippelen om doelen op hoog niveau te bereiken.' Deze strategie omvat bedrijfsplannen om informatie te gebruiken voor een concurrentievoordeel en om organisatiedoelen te ondersteunen (DAMA International, 2018). Een data strategie vereist een goed begrip van de data behoeften die inherent zijn aan de bedrijfsstrategie (Knight, 2017). Het is nodig om de manier te verbeteren waarop data wordt verzameld, bewaard, beheerd, gedeeld en gebruikt. Data strategie vereist gemeenschappelijke methoden en geformaliseerde processen die worden gedefinieerd als onmisbaar voor gedeeld beheer van bedrijfsinformatie.

## **Wat zijn de vijf componenten van een data strategie?**

Een gedegen data strategie bestaat uit onderstaande vijf componenten:

### **Identificatie**

Het identificeren van data en het erkennen van de betekenis en het bruikbaarheid ervan (ongeacht locatie, typologie en bron) is een onmisbare voorwaarde een goede data strategie. Het verwerken en benutten van data is immers niet mogelijk als de data geen naam heeft, niet in een specifiek format staat en er geen waarde aan is

toegekend. Als een organisatie (meta)data niet beschrijft, is het alsof informatie wordt genegeerd, omdat men het liever niet kent. Als data daadwerkelijk als een 'asset' wordt beschouwd, dan moet de data strategie ervoor garant staan dat alle data kan worden geïdentificeerd en beschreven.

### **Opslag**

Het veilig archiveren van data is zonder twijfel een cruciale activiteit. In de regel zijn technologische archiveringskeuzes meer gericht op het opslaan van de data als een doel op zichzelf. Deze aanpak gaat voorbij aan het eenvoudig kunnen delen en verplaatsen van data tussensystemen. Hiervoor is vaak geen plan. Een goede data strategie zorgt ervoor dat zodra data is gecreëerd, deze eenvoudig toegankelijk is en door iedereen kan worden gedeeld, zonder dat iemand anders zijn eigen kopieën hoeft te maken.

### **Distributie**

In tegenstelling tot het verleden toen 'solitaire' applicaties werden gebruikt voor het opslaan en gebruik van data, wordt data tegenwoordig door een groot aantal systemen gebruikt ter ondersteuning van bedrijfsprocessen en managementbeslissingen. Het delen van data is niet langer een gespecialiseerde, technische capaciteit die wordt overgelaten aan de grillen van applicatieontwikkelaars. Het is een 'business' behoefte geworden en moet dienovereenkomstig worden aangepakt.

### **Integratie**

Data integratie is veel meer dan de traditionele data warehouse aanpak. Het omvat alle data (gestructureerd en ongestructureerd) en alle databewegingen tussen de verschillende applicaties. Binnen de meeste organisaties is data integratie geen centrale functie. Elk ontwikkelteam gebruikt zijn eigen logica om data te linken aan verschillende applicaties. Er is alleen oog voor de behoefte van de

specifieke opdrachtgever. Omdat men langs elkaar heen werkt, worden er ook geen synergievoordelen behaald. Een goede data integratie strategie zorgt ervoor dat data wordt opgeschoond en samengevoegd in een finale dataset die op een consistente en herhaalbaar manier door iedereen kan worden gebruikt.

### **Governance**

Vaststellen, toepassen en communiceren van regels voor datagebruik zorgt voor beter beheer. Dit is vooral van belang wanneer de hoeveelheid informatie om te beheren en te delen groeit. Beleid en regels moeten worden geformaliseerd, rollen en taken moeten worden vastgesteld en moeten worden methoden geïmplementeerd, teneinde de kwaliteit en veiligheid van de data te garanderen.



# Master data management

Bron: ICT informatiecentrum | [ictinformatiecentrum.nl](https://www.ictinformatiecentrum.nl)

Master data management is gericht op het zo goed mogelijk beheeren van de belangrijkste data in uw bedrijf, bijvoorbeeld over uw klanten, leveranciers of producten. Deze master data of stamdata komt voor in verschillende systemen. Eenduidigheid van deze gegevens is belangrijk, omdat u alleen dan zeker weet dat het altijd over hetzelfde gaat.

## **Wat is master data management?**

Master data bestaat uit de belangrijkste gegevens over mensen, plaatsen en dingen, die voor een optimaal bedrijfsproces cruciaal zijn. Master data management (MDM) richt zich op het beheer van deze master data, die ook aangeduid wordt als brondata of stamdata. Master data is aanwezig in de verschillende applicaties die u gebruikt. Denk aan de software en databases voor CRM, administratie, logistiek of productie. Om de communicatie en samenwerking tussen deze applicaties mogelijk te maken, is het noodzakelijk dat de master data eenduidig is en met elkaar in verband kan worden gebracht. Met een goede oplossing voor master data management wordt dat doel bereikt. Zeker in grote en complexe organisaties is het noodzakelijk om te streven naar originele en eenduidige data. Master data management is dan ook praktisch onmisbaar in grote ondernemingen en in bedrijven die groeien, waardoor het aantal applicaties, gegevens en eindgebruikers toeneemt. Omdat MDM oplossingen steeds toegankelijker worden voor kleinere organisaties, profiteren ook zij in toenemende mate van de voordelen van master data management.

## **Toepassing master data management**

Wanneer organisaties groter worden, fuseren of overnames doen, dan heeft dat gevolgen voor het beheer van hun data, zeker als dat voorheen in gescheiden omgevingen gebeurde. Bij fusies en overnames moet data uit

verschillende bronnen daarom geïntegreerd worden. Het integreren is een complexe aangelegenheid, die niet ten koste mag gaan van dagelijkse bedrijfsprocessen. Geldt dit voor u, dan biedt master data management in deze situaties een oplossing. Met MDM creëert u dan één enkele bron voor alle essentiële bedrijfsdata. Alle systemen die deze data gebruiken, kunnen terugvallen op de database met eenduidige master data.

Als er binnen uw organisatie een hoop veranderingen plaatsvinden, heeft dat ook zijn weerslag op het data management en het gebruik van data. Denk hier aan meerdere versies van dezelfde data die naast elkaar bestaan of gesloten data silo's waar niemand in uw organisatie iets mee doet, maar die wel bedrijfskritieke informatie (kunnen) bevatten. Ook hier biedt master data management u een oplossing.

Als de datakwaliteit of databetrouwbaarheid in het geding is, kan master data management ook een goede oplossing zijn. Het komt voor dat de herkomst van bepaalde data niet vaststaat, waarmee de betrouwbaarheid niet kan worden vastgesteld. Om vervolgens toch met de meest betrouwbare data te kunnen werken, is het creëren van een leidend brondatabestand een oplossing. Hetzelfde geldt voor data die het resultaat is van handmatige invoer die foutgevoelig is.

### **Voordelen**

Master data management zorgt voor het filteren en transformeren van data, zodat deze toegepast kan worden in verschillende systemen en voor verschillende onderdelen van het bedrijfsproces. Dat levert u diverse voordelen op:

- Uw hele organisatie beschikt over correcte en actuele data (brondata).
- U bespaart kosten, omdat u meervoudige administratie van gegevens tegengaat.

- U heeft een beter inzicht in bedrijfsprestaties, omdat data uit verschillende bronnen gecombineerd wordt.
- U kunt betere beslissingen nemen, omdat u over betrouwbaardere data beschikt.

# Voordelen van een multi-domein MDM systeem

Bron: Beeyond | [beeyond.nl](https://beeyond.nl)

Een multi-domein master data management systeem heeft meerdere voordelen ten opzichte van een single-domein MDM systeem. Het is echter vaak niet duidelijk welke gevolgen de keuze tussen beiden heeft voor organisaties. Dit artikel gaat in op deze toegevoegde waarde van een multi-domein master data management systeem.

## Breed inzetbaar

Multi-domein MDM platforms zijn platforms waar meerdere typen data beheerd kunnen worden. Deze kunnen een wondermiddel zijn in vergelijking met de complexiteit van single-domein oplossingen, waar slechts één type data beheerd kan worden. Dit is vooral merkbaar als men kijkt naar de kosten, omdat organisaties één oplossing kunnen implementeren die niet alleen werkt met de traditionele domeinen van klant- en productgegevens, maar ook met de vele andere data, zoals locaties, activa, medewerkers en machines.

Voor elke organisatie, onafhankelijk van de grootte, is het complex om alle wensen van de organisatie en de diverse afdelingen in verschillende single-domeinen onder te brengen. Zoals bij de afdeling marketing, die zowel toegang tot de gegevens van klanten nodig hebben als productgegevens. Een single-domein MDM systeem beperkt in dat geval de efficiëntie en het innovatieve vermogen van de organisatie.

## Wensen vereisen meestal multi-domein

De initiële vraag van organisaties richt zich vaak op een oplossing voor een bepaald domein. Het gaat dan over één type data. Er ontstaan dilemma's

naarmate de organisatie groeit en de MDM behoeften uitbreiden buiten een afdeling en de gehele organisatie omvatten. Achteraf bezien blijkt dat de initiële wens eigenlijk meer aspecten omvatte, maar deze werden niet als zodanig herkend en uitgesproken.

Op het moment dat een organisatie, na een te beperkt gedefinieerde wens, eenmaal gekozen heeft voor een single-domein MDM platform en dat geïmplementeerd heeft, ontstaan de problemen. Er is dan vrijwel geen weg meer terug. Het krijgen van financiering en draagvlak voor implementatie van een organisatie-breed multi-domein MDM systeem is immers vrijwel ondoenlijk als je recent geïnvesteerd hebt in implementatie van een single-domein MDM platform.

Het verdient daarom grote aanbeveling om bij het starten van een zoektocht naar de juiste oplossing de wensen goed in beeld te brengen en medewerkers uit alle disciplines binnen de organisatie samen te brengen om de wensen te inventariseren.

### **Voordelen voor databeheer**

Een van de meest opvallende voordelen van multi-domein MDM-platforms is data governance, ofwel het beheren en toezien op de kwaliteit van data. Met single-domein oplossingen, die gericht zijn op slechts een deel van de organisatie, is het moeilijk om ervoor te zorgen dat aan de gestelde data-normen wordt voldaan.

#### **Alleen schone data, volgens gestelde eisen**

Eindgebruikers hebben de neiging om een grote mate van zelfstandigheid te nemen bij het updaten en opslaan van gegevens. Dit kan resulteren in meerdere definities van termen die eigenlijk voor de gehele organisatie gelijk zouden moeten zijn. Gebruikers zullen aanzienlijk minder zelfstandigheid krijgen bij het creëren van definities en regels voor de gegevens. Dit omdat de algemene architectuur wordt vereenvoudigd, wat resulteert in een grotere

transparantie. Als gevolg van betere uitgangspunten en randvoorwaarden voor databeheer, kunnen functionele samenwerkingen tussen afdelingen eenvoudiger ontstaan. Dit resulteert in een verhoogde efficiëntie.

### **Vereenvoudigd centraal databeheer**

Een van de belangrijkste redenen voor databeheer is het waarborgen van de kwaliteit van gegevens, het verrijken en zorgen voor betrouwbare data, die er anders niet zou zijn. De taak voor databeheerders wordt aanzienlijk gemakkelijker wanneer er slechts één multi-domein MDM-platform is, waarin alle data beheerd wordt. Data hoeft maar een keer opgeschoond te worden en er is minder kans op overbodige data en andere factoren die leiden tot wantrouwen. Multi-domein MDM systemen zorgen ervoor dat je verschillende groepen van data uit verschillende delen van de organisatie samen kunt brengen, waardoor je delen van de organisatie samenbrengt maar hierbij toch eenvoudig de regie houdt over het eindresultaat. Niet zomaar alle data op een hoop, maar bedrijfsprocessen, regels en beleid inzake het databeheer, dat wordt ondersteund door de gehele organisatie. Dat leidt tot een vertrouwde opslagplaats van alle data, voor alle medewerkers binnen de organisatie.

### **Big data**

Big data functioneert als versneller in de richting van multi-domein MDM. De verscheidenheid van de gegevens die big data biedt, kan toegepast worden in veel domeinen. Om meer waarde te halen uit big data is het van belang dat organisaties hun eigen gestructureerde gegevens voldoende beheersen. Het is moeilijker om big data te integreren in een single-domein MDM, door het simpele feit dat de variatie van de gegevens mogelijk niet van toepassing is op een specifieke domein, terwijl het dat wel is voor andere domeinen. Een multi-domein MDM systeem helpt om overzicht te krijgen over big data, waardoor je meer waarde uit je data kunt halen.

Steeds vaker wordt direct gekeken naar een multi-domein MDM oplossing als het gaat om big data en enterprise information management (EIM). Het heeft geen zin om hierbij te kijken naar een single-domein oplossing. Men zoekt naar oplossingen om te kunnen omgaan met big data uitdagingen, zoals sociale media en ongestructureerde data.

### **Multi-domein musthaves**

Er zijn tal van multi-domein MDM-platforms beschikbaar, die elk specifieke sterke en zwakte punten hebben, waarbij het niet in de laatste plaats draait om de hoeveelheid klantsupport die een leverancier bereid is te bieden. Onderstaande checklist geeft een reeks criteria, functies en eigenschappen waarop MDM oplossingen zich van elkaar kunnen onderscheiden.

#### **Datakwaliteit**

Systemen moeten geautomatiseerde dataprofielen kunnen vastleggen die doelmatig de datakwaliteit controleert. Inkomende records moeten worden gestandaardiseerd en opgeschoond volgens de gebruiker verstrekte business rules. Het systeem moet gegevens (zoals BTW en KvK) op juistheid controleren.

#### **Integratie**

Het product moet API's voor realtime, batch integratie en Data-as-a-Service omvatten.

#### **Governance support**

Bij flexibele platforms kan data goed beheerd worden. Afzonderlijke componenten worden verenigd in de gebruikersinterface om zo een geïntegreerd beeld te kunnen geven.

#### **Afstamming**

Multi-domein MDM's moeten de ruwe gegevens die eerst worden ontvangen, bewaren in de hub en moeten een pad bieden voor het

bewerken van golden records als ze worden veranderd.

### **Modellering**

Oplossingen moeten flexibiliteit bieden, zodat het gedrag van het platform verandert al naar gelang de datamodeltering verandert. De hub moet relaties en entiteiten in de data weten te vinden en presenteren.

### **Loading**

Platforms moeten datamigratie en updates bieden die duplicaten binnen de inkomende gegevens en de gegevens in de hub voorkomen.

### **Toekomstige adoptie**

Er zijn voordelen voor het gebruik van zowel single- als multi-domein MDM platforms. Het belangrijkste verschil is dat een toenemend aantal multi-domein MDM's veel van de functionaliteit van de single-domein MDM's en daarbij nog veel meer kunnen leveren. Een multi-domein MDM is hiermee meer noodzaak dan luxe.

Naarmate organisaties groeien, is het vaak in hun voordeel om de traditionele silo-aanpak te verlaten en hun MDM volledig binnen de onderneming te integreren. Met de vraagstukken met betrekking tot beheer en de relevantie van big data als drivers, zal de trend naar multi-domein MDM's alleen maar toenemen.



# Data governance

Bron: ICT informatiecentrum | [ictinformatiecentrum.nl](https://www.ictinformatiecentrum.nl)

Data governance staat aan de basis van data management. Het betreft alle afspraken en spelregels waarmee u onder andere vastlegt hoe u van alle data gebruik wilt maken, wat uw doelen ermee zijn, hoe u de kwaliteit waarborgt en hoe de veiligheid van de data geregeld wordt. Met data governance legt u structuren vast en bepaalt u datastandaarden, uw visie en strategie. En u legt vast wie welke verantwoordelijkheden hiervoor heeft.

## **Wat is data governance?**

Data governance is een set aan afspraken en gekozen standaarden die gezamenlijk vastleggen op welke manier uw organisatie met uw data omgaat. Het heeft als primair doel om de consistentie, betrouwbaarheid en ongehinderde beschikbaarheid van uw data te garanderen, zodat de werkprocessen in uw organisatie optimaal ondersteund worden. Data governance betreft een aantal verschillende werkzaamheden rondom het gebruik en beveiligen van data binnen een organisatie, alsmede de controle over de processen die dit mogelijk maken. Het heeft niet alleen met de technische en logistieke kant van het opslaan van data te maken. Het betreft ook een noodzakelijke combinatie voor het naleven van wetten, regels en procedures. Daarnaast draait het om de mensen die met de data werken. Gegevensbeheerders, databewakers of data stewards vervullen een belangrijke rol bij het realiseren van data governance.

## **Waarom is data governance belangrijk?**

Als uw organisatie over onbetrouwbare informatie beschikt, kan u dat belemmeren in uw activiteiten, omdat er op basis van die onbetrouwbare informatie minder goede of verkeerde keuzes gemaakt worden. Ook omdat u beschikt over een groeiende hoeveelheid aan data, kan goed overzicht ontbreken. Dat kan er toe leiden dat u (onbewust) niet aan alle wettelijke

verplichtingen voldoet. Data governance helpt u hierbij door het introduceren van een set met afspraken en standaarden, zodat de data en datastromen binnen uw organisatie een stuk overzichtelijker worden. Zonder doeltreffende data governance zal uw concurrentiepositie op lange termijn kunnen afnemen. Ook is de kans kleiner dat u alle data conform geldende regelgeving – de zogeheten ‘compliance’ – behandelt.

### **Hoge datakwaliteit**

Als data niet onderhouden wordt, neemt de kwaliteit ervan af. Als gevolg daarvan kunnen gebruikers hun eigen bronnen gaan gebruiken. Dit betekent niet alleen meer werk, maar ook meer onduidelijkheid over de juistheid van de data voor andere gebruikers. Data governance stelt regels op die bijdragen aan de bewaking van de datakwaliteit.

### **Wettelijke eisen**

Het voldoen aan de wettelijke eisen en het controleren ervan brengt hoge kosten met zich mee. Wanneer u niet voldoet aan (nieuwe) regelgeving loopt u het risico op zowel reputatieschade als geldboetes. Door data governance neemt de datakwaliteit toe en daarmee de kans af op het per ongeluk niet naleven van geldende regelgeving.

### **Steeds meer data**

U heeft steeds meer data en informatiebronnen tot hun beschikking. Waarschijnlijk vergaart u niet enkel meer informatie uit uw eigen systemen, maar maakt u ook steeds vaker gebruik van algemeen beschikbare data. Het samenvoegen en verwerken van deze data uit verschillende bronnen kunnen moeilijkheden opleveren met betrekking tot datakwaliteit. Data governance zorgt voor een structuur en werkwijze waarop u op een goede manier met (nieuwe) data binnen uw organisatie kunt omgaan.

### **Beheersing van kosten**

Het bewaken van de datakwaliteit en het handmatig corrigeren van fouten in een database is een tijdrovende en kostbare activiteit. Wanneer data governance goed wordt toegepast, ligt de standaard datakwaliteit een stuk hoger. Hierdoor zijn minder correcties nodig en bespaart u tijd en kosten. Dit staat nog los van het feit dat kwalitatief betere informatie waardevollere inzichten kan opleveren voor uw organisatie. De data zal een grotere waarde gaan vertegenwoordigen.

### **Commercieel voordeel**

Er zijn organisaties die veel geld geïnvesteerd hebben in een nieuw CRM systeem, maar nog steeds geen perfect beeld hebben van hun klanten. Met data governance kunnen regels opgesteld worden waardoor een beter beeld ontstaat, omdat data beter en overzichtelijker wordt opgeslagen in de database die gekoppeld is aan de CRM software.

### **Hoe realiseert u data governance?**

Als u data governance wilt toepassen, moet u hiervoor eerst een overkoepelende strategie voor uw data management definiëren. Hierin legt u vast hoe data wordt opgeslagen en beschermd. Ook wat betreft datagebruik binnen uw organisatie moeten standaarden en werkprocedures ontwikkeld worden, die vervolgens gemonitord worden middels controles of audits. Binnen uw organisatie kunnen speciale data governance teams worden samengesteld, die procedures opstellen voor de omgang met data. Als iedereen in uw organisatie zich betrokken voelt bij het data management, zullen zij zich willen inspannen om de regels en procedures na te leven. De kwaliteit van het data management wordt hiermee geoptimaliseerd.

# De zin en onzin van data management en data governance

Bron: Hot ITem | [hotitem.nl](http://hotitem.nl)

Je hoort het tegenwoordig te pas en te onpas: als organisatie móet je aan de slag met data management en data governance. Je staat vast niet te springen van enthousiasme als een consultant je dit weer eens komt vertellen. Want dat gebeurt toch al? Er zijn beheerafdelingen, functioneel- en technisch beheerders, DBA's ..., ze beheren al van alles! Dit is een zeer logische gedachte. Echter, meestal is data niet datgene wat beheerd wordt.

## Hoe zit het dan met het beheer van data?

Om die vraag te beantwoorden wordt er online volop gestrooid met allerlei complexe verklaringen van datatermen en buzzwords. Om te begrijpen waar het echt om gaat, werkt het veel beter om jezelf een paar vragen te stellen:

- Wat zijn de meest gehoorde klachten over de informatievoorziening in jouw organisatie?
- Weet je wat voor informatie je allemaal produceert? Heb je daar een overzicht van?
- Klopt alle informatie die je produceert? Hoe weet je dat?
- Wat zijn de consequenties als de informatie niet zou kloppen?
- Waar is de informatie op gebaseerd? Waar komt het vandaan?
- Hoeveel tijd ben je kwijt aan rework en/of uitzoeken van verschillen tussen cijfers?
- Hoeveel tijd besteedt de analist daadwerkelijk aan data analyse? Hoeveel van zijn tijd is hij bezig met het verzamelen en prepareren van data zoals koppelen en cleansen?

## **Geheim van een goede data administratie**

De problemen die ten grondslag liggen aan (de antwoorden op) bovenstaande punten zitten vaak in de administratie van je data. Dat is helemaal niet vreemd, want zo werkt het ook met je privé- of bedrijfsadministratie. Als je die niet op orde hebt, weet je niet meer precies wat je hebt, heb je misschien wel dubbele verzekeringen, weet je niet goed wat de huidige stand van zaken is en kost het veel uitzoekwerk om die data boven water te krijgen. Op een gegeven moment is het misschien zo'n rommeltje, dat je alles in een schoenendoos wegstopt en vol goede moed opnieuw begint.

Zo werkt het ook met je data administratie. Je moet bedenken hóe en wát je gaat administreren. Je moet een proces inrichten, zodat je het structureel bijhoudt. En je moet de gereedschappen en mensen hebben om dat goed te doen. Pas dan ga je de vruchten plukken. Om een paar dingen te noemen: geen dubbel werk meer en kloppende cijfers. Geen verkeerde zakelijke beslissingen gebaseerd op verkeerde cijfers. Geen discussies over cijfers; weten wat cijfers betekenen en hoe ze berekend zijn. Vertrouwen in je BI- en datawarehouse-omgevingen. Echte BI selfservice. Geen ver- of herbouw van je BI- en datawarehouse omgevingen meer. En soepele migraties...

Organisaties drijven tegenwoordig evenzeer op data als op geld. Zonder data zouden de meeste organisaties niet eens kunnen bestaan. Is het dan niet zinvol (en hoog tijd) die data ook eens te gaan administreren en af te spreken hóe je dat gaat doen?

## **De crux: groot denken, klein beginnen**

Met de juiste afspraken (governance) begin je met iets kleins en het succes laat de olievlek groter worden. Geen succes? Dan ook geen grotere olievlek, en moet je misschien andere afspraken maken en opnieuw goed beginnen. Laten we het voorbeeld van de privé administratie er nog eens bij pakken. Als eerste spreken we met onze partners af dat we in ieder geval afstappen van de schoenendozen. Voor het opslaan van afgehandelde administratie schaffen we multomappen aan met daarin een paar tabbladen. Zaken die

nog afgehandeld moeten worden, gaan in de schaal op tafel en die maken we elke zondag leeg. Daarna stoppen we het meteen in de multomap onder het juiste tabblad.

Kijk eens: we zijn afspraken aan het maken over hóe we gaan administreren, we zijn aan het 'governen'. Die multomap en perforator, dat is onze tooling, onze randvoorwaarden. De wijze waarop we de tabbladen indelen, dat is onze metadata, ofwel data over data. Misschien doen we het nog niet meteen op de meest handige manier of gedetailleerd genoeg, maar dat leren we al doende. We zijn in ieder geval van de schoenendoos afgestapt en hebben besloten dat we het anders gaan doen!

# Rollen en verantwoordelijkheden in de data governance organisatie

Bron: Tentive | [tentive.nl](https://tentive.nl)

Als consultant word je al in een vroegtijdig stadium aangeleerd dat elke organisatie uniek is. Of dat daadwerkelijk het geval is, willen we graag in het midden laten. Een goede data governance organisatie bestaat in ieder geval uit een handvol specifieke rollen met duidelijk omschreven verantwoordelijkheden. Welke dat zijn, wordt hieronder beschreven.

## **Stuurgroep**

De 'steering committee' of stuurgroep bestaat uit de uitvoerende leiders van de organisatie. Dit kunnen personen op C-niveau zijn, maar meestal zijn het managers die verantwoordelijk zijn voor specifieke bedrijfsactiviteiten die tezamen een data domein vormen, zoals finance, klanten en materialen. De stuurgroep heeft tot doel om functionele beslissingen te vergemakkelijken. Kenmerken van een effectieve stuurgroep zijn:

### **Juiste omvang**

De stuurgroep moet groot genoeg zijn om alle stakeholders te kunnen vertegenwoordigen, maar klein genoeg zijn om snel belangrijke informatie te analyseren en beslissingen te nemen.

### **Strategie**

De stuurgroep bepaalt de algemene strategie met specifieke resultaten en houdt de governance organisatie verantwoordelijk voor tijdslijnen en uitkomsten.

### **Wegnemen van barrières**

De stuurgroep moet een middel zijn om organisatorische belemmeringen weg te nemen, niet een gewone vergadering om te luisteren naar de voortgang rapportages van project teamleden.

### **Geen substituut**

De stuurgroep moet geen substituut zijn voor hands-on sponsorschap.

### **Data eigenaar**

De 'data owner' of data eigenaar is verantwoordelijk voor de data binnen een specifiek data domein. Een data eigenaar moet ervoor zorgen dat de informatie binnen zijn domein correct wordt beheerd door verschillende systemen en bedrijfsactiviteiten. Data eigenaren maken vaak deel uit van de stuurgroep, hetzij als stemhebbend lid of als lid zonder stemrecht. Specifieke verantwoordelijkheden voor de data eigenaar zijn:

- Goedkeuren van data glossaries\* en data definities.
- Zorgen voor de juistheid van informatie, zoals gebruikt binnen en door de gehele organisatie.
- Directe datakwaliteit activiteiten.
- Beoordelen en goedkeuren van de master data management (MDM) aanpak, resultaten en activiteiten.
- Samenwerken met andere data eigenaren om dataproblemen en onbegrip tussen de verschillende bedrijfseenheden op te lossen.
- Tweedelijnsbeoordeling van dataproblemen die zijn geconstateerd door data stewards.
- Input geven aan de stuurgroep over softwareoplossingen, beleid of wettelijke vereisten die van invloed zijn op het data domein van de data eigenaar.

De rol van de data eigenaar wordt vaak vervuld door senior medewerkers. Zo kan de financieel directeur de data eigenaar zijn van de financiële gegevens



in de organisatie. Dit niveau van senioriteit houdt echter in dat een data eigenaar vaak niet in de gelegenheid is dagelijks betrokken te zijn bij activiteiten die zijn gericht op het beheren van de datakwaliteit. Omdat 'data ownership' vaak geen voltijdbaan is, wordt de data eigenaar in de regel ondersteund door een of meerdere data stewards.

### **Data steward**

Het belangrijkste verschil tussen een data eigenaar en een data steward is dat de laatstgenoemde verantwoordelijk is voor het dagelijks beheer van de kwaliteit van de gedefinieerde datasets. De data steward is de 'subject matter expert' (SME) die de betekenis en het gebruik van de informatie begrijpt en communiceert. Hij geeft de data eigenaar inzicht in de algemene doelstellingen van de data, maar zal ook sterk betrokken zijn bij de details over hoe dit kan worden bereikt. Organisatiebreed werkt een data steward vaak samen met andere stewards in een 'data steward council'. In dit besluitorgaan worden beslissingen genomen over en oplossingen bedacht voor mogelijke dataproblemen. De data steward vertegenwoordigt de data eigenaar in de meeste discussies. Hij zal zich richten tot de data eigenaar en/of de stuurgroep als in de council niet tot overeenstemming kan worden gekomen over hoe een dataprobleem opgelost zou moeten worden. Andere taken voor de data steward zijn onder meer:

- Opstellen van datadefinities en documenteren van toegestane waardes.
- Definiëren van regels voor het creëren van data, het gebruik van data of het afleiden van data.
- Identificeren en documenteren van huidige en gewenste datasystemen.
- Stellen van datakwaliteit doelen.

Sommige organisaties hebben formele data steward rollen gedefinieerd, vaak uitgevoerd door mensen uit de business die hiervoor zijn vrijgemaakt. Andere organisaties wijzen 'data stewardship' verantwoordelijkheden toe

aan werknemers die ook andere verantwoordelijkheden hebben. Onafhankelijk van hoe de positie is vormgegeven, een effectieve data steward volgt de van tevoren afgesproken data definities, identificeert datakwaliteit problemen en zorgt ervoor dat de Business de gespecificeerde standaarden volgt (IBC, 2018).

### **Tot slot**

Elk van bovengenoemde rollen is een belangrijk onderdeel van een goed geleide data governance organisatie. Wie deze rollen het beste binnen een organisatie kunnen vervullen en hoe deze rollen elkaar onderling beïnvloeden, hangt echter af van de MDM volwassenheid van de betreffende organisatie. Zou elke organisatie dan uiteindelijk toch uniek zijn?

*\* Data glossaries zijn beschrijvingen van de kenmerken van de bedrijfskritische data elementen in de verschillende systemen. Deze beschrijving is van belang om ervoor te zorgen dat iedereen weet hoe data elementen gedefinieerd zijn, waar de data elementen opgeslagen zijn en wat de kwaliteit ervan is. Bij het uitvoeren van projecten zal vaak teruggegrepen worden naar deze data glossary om te zien hoe data elementen geraakt worden door bepaalde veranderingen (Feltkamp, 2018).*

# Drie actuele data governance trends

Bron: Tentive | [tentive.nl](https://tentive.nl)

In dit artikel worden 3 actuele trends besproken rond het onderwerp data governance.

## **Datakwaliteit nog meer borgen binnen data governance**

Nog altijd is datakwaliteit de meest hardnekkige uitdaging binnen datamanagement. Organisaties hadden tot voor kort slechts aandacht voor datakwaliteit op het moment dat omzet of reputatie gevaar liepen. Deze stellingname is echter snel aan het veranderen. Complexe naleving van wet- en regelgeving en de almaar toenemende snelheid en schaal van data, leiden ertoe dat organisaties prioriteit geven aan datakwaliteit. Sterker, organisaties beschouwen datakwaliteit als een uiterst belangrijk component van data governance. Door een op datakwaliteit gebaseerd raamwerk voor data governance op te zetten, geven organisaties allereerst een impuls aan de waarde van hun bedrijfskritische data. Daarnaast zorgen zij er hiermee voor dat datakwaliteit issues vroegtijdig worden gesignaleerd en opgelost voordat deze over verschillende systemen kunnen verspreiden. Het besef groeit dat het uiterst onverstandig is om een afwachtende houding aan te nemen om naderhand geconfronteerd te worden met het bewijs dat slechte data inderdaad een negatief effect heeft op de business en de continuïteit daarvan. Tegen die tijd is de schade echter al aangericht.

## **Privacywetgeving vereist krachtige data governance**

De Algemene Verordening Gegevensbescherming (AVG) van de Europese Unie is in 2018 geïmplementeerd. Inmiddels dient de AVG als een wereldwijde katalysator voor data privacywetgeving. Organisaties die de nieuwe wetgeving onvoldoende serieus nemen, riskeren naast hoge geldboetes ook reputatieschade. Organisaties worden dus gedwongen mogelijke gaten in de naleving te evalueren en waar nodig aan te pakken.

Met het oog op het naleven van de privacywetgeving is krachtige data governance daarom onontbeerlijk. Niet alleen om persoonsgegevens te identificeren en te beschermen, maar ook om de afkomst van de data te volgen, als deze worden verplaatst van bronnen naar systemen en processen. Daarnaast speelt datakwaliteit een cruciale rol bij het verkleinen van nalevingsrisico's. Slecht onderhouden data en data van slechte kwaliteit kunnen beide eenvoudig leiden tot niet-naleving van de privacywetgeving, met alle negatieve gevolgen van dien op het merk en de bedrijfsresultaten van een organisatie.

### **Zonder goede data governance geen selfservice analytics**

BI tools met machine learning en artificial intelligence die selfservice analytics mogelijk maken, zijn al in 2019 van start gegaan en het gebruik ervan neemt alleen nog maar toe. Eindgebruikers krijgen door selfservice analytics de mogelijkheid om met een gemakkelijk te hanteren BI oplossing snel eigen rapportages en analyses te maken. Ze zijn niet meer afhankelijk van IT om rapporten voor ze te maken. Bovendien kunnen ze op deze manier veel sneller en flexibeler voorzien in de eigen informatiebehoefte. Dit lijkt op het eerste gezicht een positieve ontwikkeling. Maar let op: zonder goede data governance in de BI omgeving kunnen meerdere versies van de waarheid ontstaan (bijvoorbeeld door inconsistente formules en definities) en de datakwaliteit verminderen. Daarnaast is het ook geen goed idee om gebruikers onbeperkt toegang te geven tot allerlei data en iedereen toe te staan rapportages te maken die binnen en buiten de organisatie kunnen worden gebruikt. Kortom, BI en data governance gaan daarom hand in hand, meer dan ooit tevoren.

# Vijf lessen bij de implementatie van data governance

Bron: Cmotions | [cmotions.nl](https://cmotions.nl)

Advanced analytics, realtime reporting, big data, smart data en data virtualisatie, van datacenters naar centers of data; zomaar een paar begrippen en zinsneden die ons de laatste jaren om de oren vliegen. Echter, wat zijn deze begrippen waard wanneer je helemaal niet weet of je de bron ervan, de onderliggende data, wel of niet kan vertrouwen? Dit vertrouwen is te organiseren. Dit artikel gaat in op het begrip data governance en geeft u vijf geleerde lessen uit de praktijk.

## Data als fundament

Veel organisaties gebruiken in toenemende mate data als basis of uitgangspunt voor hun beslissingen. Marketing voert klantanalyses uit, finance creëert het inzicht in de resultaten en de risico afdeling denkt in control te zijn over de processen. Allemaal voorbeelden van zaken die gedreven zijn door het vertrouwen in onderliggende data. We kunnen data daarom zien als het fundament onder bedrijfsprocessen, rapportages en analyses binnen een organisatie. Vanuit deze conclusie is het vreemd dat er in veel organisaties maar weinig aandacht wordt besteed aan de invulling en organisatie van dit fundament. Ofwel, aan het begrip data governance.

## Data governance

Wat verstaan wij dan onder data governance? Wij hanteren de volgende definitie: 'data governance verwijst naar alle beheersmaatregelen die de beschikbaarheid, bruikbaarheid, integriteit en de veiligheid van de data binnen een organisatie moeten waarborgen'.

Om data governance goed te implementeren, maken we gebruik van een

data governance framework. Dit framework bevat, naast de inrichting en organisatie van processen en procedures, ook de daadwerkelijke uitvoering van het ingerichte beleid. Belangrijk uitgangspunt is hierbij dat het ingerichte framework bijdraagt aan het realiseren van de data strategie binnen de business strategie.

### **Data governance en de relatie met data management**

Vaak wordt data governance verward met data management. Echter, wij zien data management als één van de essentiële onderdelen van data governance. Om deze onderlinge samenhang te duiden, is het wenselijk om te kijken naar het model van de DAMA (data management association international); het data management body of knowledge (DMBoK). In dit model is snel zichtbaar welke verbindende en sturende rol data governance heeft en welke raakvlakken er zijn op het gebied van andere aandachtsgebieden, zoals verschillende onderdelen van data management. En hoezeer deze sturende rol ook van belang is, deze kan niet goed uitgevoerd worden zonder de invulling of in ieder geval aandacht voor de verschillende deelgebieden. Ofwel, om de relatie tussen data governance en data management te duiden: data governance verhoudt zich tot data management zoals boekhoudregels zich verhouden tot financiële administratie.

### **Vijf geleerde lessen**

#### **Laat regelgeving niet het vertrekpunt zijn**

Vaak wordt data governance pas ingericht, omdat de regelgever/toezichthouder bepaalde eisen stelt. Het is bijvoorbeeld plots belangrijk te kunnen duiden waar bepaalde informatie in een (al dan niet wettelijke) rapportage op gestoeld zijn. Vaak is het invoeren van data governance op een dergelijk moment te laat. Buiten het feit dat men eigenlijk sowieso continu vraagtekens moet zetten bij de bron van rapportages en de betrouwbaarheid hiervan, is het invoeren van data governance geen lapmiddel om snel inzicht te

kunnen verschaffen. Ons advies is dus ook: organiseer data governance sowieso, ook de interne organisatie heeft hier recht op en vooral baat bij!

### **Datagovernance is geen 'IT-feestje'**

Vaak wordt data governance geïnitieerd door de IT organisatie. Data wordt immers vaak als iets technisch en dus als IT beschouwd. Hier zijn wij het niet mee eens. Data governance is onzes inziens informatiemanagement over de hoogste ordening van de organisatie heen en ligt, net als informatiemanagement, zeker niet alleen bij IT. Ons advies is om data governance te organiseren bij een afdeling die neutraal instaat tussen IT en de business, inclusief stafafdelingen als finance, risk, HR). Te denken valt aan een informatiemanagement afdeling of bij compliance.

### **Start met governance, IT invulling volgt later**

Vaak wordt de oplossing voor het ontbreken van een goede data governance gezocht in het beschikbaar stellen van tooling. Wij geloven hier niet in. Zorg eerst voor goed beschreven processen en heldere taken, bevoegdheden en verantwoordelijkheden. Kies voor een snelle manier om metadata in reporting repositories, business glossaries en data dictionaries vast te leggen. Maar zorg wel dat deze later te importeren zijn wanneer er wel tooling in het spel komt. En kijk bij tooling ook naar integratie met reeds gebruikte applicaties. Reeds gebruikte reporting software is een prima plaats om de koppeling te maken met een business glossary of data dictionary. Door deze aanpak is in ieder geval op dat moment de waarde van het data governance traject al bewezen, los van de software.

### **Continue commitment is key**

Zonder draagvlak van het (top-)management zal de implementatie van data governance stroef verlopen. Vooral wanneer de benefits

niet direct zichtbaar zijn, kost het veel overredingskracht om de mensen die het werk moeten verzetten (lees: zaken (beter) moeten gaan documenteren) in de meewerkende modus te krijgen. Daarom staat of valt een data governance traject met het hebben van commitment van het management. Pas wanneer het bij het management een terugkerend onderwerp van gesprek is en wordt uitgedragen in de organisatie, zal data governance een succes worden.

### **Denk groot, start klein**

De implementatie van data governance kan een groots en meeslepend traject worden. Wanneer je echter klein begint, maar wel het grote plaatje in het achterhoofd houdt, is het wel mogelijk om de juiste snelheid te maken en te leren van de verschillende stappen. Bij het afronden van elk klein stukje is het mogelijk te laten zien wat er bereikt is; de business glossary voor de afdeling marketing en een overzicht van rapportages voor finance. Allemaal kleine stukjes van de puzzel die het demonstreren meer dan waard zijn.

### **Conclusie: gewoon doen!**

Data governance klinkt misschien als complex, groots, duur en wellicht zelfs onnodig. Maar we willen toch altijd met zekerheid kunnen zeggen dat de data waarop we rapportages maken, inzichten verwerven en beslissingen nemen zo goed mogelijk vertrouwen? En in ieder geval begrijpen waar deze data als bron vandaan komt? Maak het daarom klein. Laat zien dat met een kleine inspanning de eerste stappen snel gezet zijn en dat de resultaten de gehele organisatie verder kunnen brengen.



# Data opslag

Bron: ICT informatiecentrum | [ictinformatiecentrum.nl](https://www.ictinformatiecentrum.nl)

Bij veel IT toepassingen die u gebruikt heeft u geen keuze waar uw data opgeslagen wordt. De aanbieder van de toepassing bepaalt dat voor u, zoals bij banktransacties, social media of webshops. Bij uw eigen data heeft u meestal wel de keuze waar uw data bewaard wordt. Voor data opslag of data storage heeft u keuze uit drie mogelijkheden: lokaal, in de cloud en een mengvorm dan beide, de multi cloud.

## **Wat maakt dataopslag belangrijk?**

De data waarover u beschikt, beschouwt u als uw eigendom. Het bevat informatie over uw bedrijfsprocessen, klanten, leveranciers, resultaten en andere zaken waarvan u niet wilt dat anderen daar inzage in hebben. Bescherming en veiligheid van data, het terrein van data security, spelen dan ook een steeds belangrijker rol. Dit vooral ook omdat de risico's van diefstal, gijzeling en vermindering van data de laatste jaren fors zijn toegenomen. Vanuit de wetgeving liggen er duidelijke eisen waaraan uw data management en data storage moeten voldoen. Daarnaast stelt u zelf eisen aan beschikbaarheid van uw data en de snelheid van de systemen waarmee u werkt, toegang hebben tot die data. Bovendien worden de technische oplossingen voor dataopslag steeds complexer, zodat u zich ook kunt afvragen welke kennis u daarover zelf in huis wilt hebben. Ook spelen kosten een rol. Al deze aspecten maken de keuze van de manier waarop u uw data opslag organiseert belangrijk.

## **Welke dataopslag vormen zijn er?**

Als het soort dataopslag niet gekoppeld is aan de IT oplossing waarvan u gebruik maakt, kunt u zelf de keuze maken waar u uw data opslaat. U heeft de keuze uit drie mogelijkheden:

- Lokale data opslag (on premise data storage)
- Cloud opslag (cloud storage)
- Hybride opslag: een mengvorm van lokale en cloud opslag (multi storage)

### **Lokale dataopslag**

Bij lokale opslag (on premise) staat alle data op uw eigen servers in uw eigen omgeving. Dat geeft veel organisaties nog een veilig gevoel en u heeft de volledige controle op wie van de data gebruik kan maken. De vraag is echter of het gevoel van veiligheid terecht is. Uw organisatie is op veel manieren via internet verbonden met de buitenwereld. Dat betekent evenzoveel mogelijkheden om van buitenaf uw systemen binnen te dringen. Veiligheid is daarmee een relatief begrip en niet gegarandeerd. Als u data security niet goed geregeld heeft, is uw data kwetsbaar voor diefstal, gijzeling of verminking.

Om het risico daarop te verkleinen moet u veiligheidsmaatregelen treffen. Dat bent u sowieso verplicht om te kunnen voldoen aan de eisen die de wetgeving stelt aan het beheer van sommige soorten data, zoals persoonsgegevens (AVG). De kans is groot dat u deze kennis over dataveiligheid niet in huis heeft en gebruik zult moeten maken van externe expertise. Daarnaast zult u ook de opslagmedia zoals harde schijven zelf moeten beheren, zelf back-ups moeten maken, zelf de software moeten onderhouden die daarvoor nodig is en telkens zelf de updates daarvan moeten installeren. Ook hierbij is de vraag wat u zelf kunt, hoe afhankelijk u bent van anderen en welke risico's u loopt dat bij het beheer daarvan ook fouten worden gemaakt.

### **Cloudopslag**

Bij cloudopslag staat uw data niet op uw eigen servers, maar wordt uw data opgeslagen bij uw cloudprovider of datacenter. Via internet heeft u toegang tot uw data. Bij veel dataverkeer is een snelle en stabiele internetverbinding

daarom een vereiste. Alle aspecten rondom dataveiligheid worden door de provider geregeld. Dat geldt ook voor het beheer van de opslagmedia, de software die daarvoor nodig is en de updates daarvan. Omdat uw data via internet toegankelijk is, is de locatie waarvandaan u uw data raadpleegt niet van belang. Overal ter wereld waar internet is, heeft u toegang tot uw data. Thuiswerken is eenvoudig en medewerkers die vaak onderweg zijn, hebben de data doorlopend ter beschikking. Komt u opslagcapaciteit te kort, dan kunt u die eenvoudig uitbreiden. De kosten bestaan uit een vast maandbedrag en zijn afhankelijk van de opslagcapaciteit die u nodig heeft.

### **Dataopslag lokaal en in de cloud**

Als u ERP software in de cloud\_ of CRM software in de cloud gebruikt, heeft u geen keuze waar uw data staat. De provider van uw cloudoplossing regelt dat voor u. Gebruikt u on premise ERP software (lokaal) of andere lokale operationele systemen, dan kan het vanwege snelheid praktisch zijn om de actuele data daarvan op uw eigen servers te hebben staan. Data van andere systemen of historische data van operationele systemen kunt u dan opslaan in de cloud. Hiermee ontstaan mengvormen van beide oplossingen die aangeduid worden met de begrippen hybride cloud of multi cloud. Deze oplossingen bieden u het beste van twee opslagmogelijkheden en maken u minder afhankelijk van specifieke cloudproviders. Ook faciliteren multi cloud oplossingen mogelijkheden voor data migratie tussen de lokale opslag en de cloudopslag. Raakt data onverhoopt beschadigd, dan zorgt deze technologie ook voor het herstel (datarecovery) ervan.

### **Technologie**

De vertrouwde vorm van dataopslag is door gebruik te maken van harde schijven. Steeds vaker wordt echter ook gebruik gemaakt van flash dataopslag technologie. Data wordt hierbij niet meer weggeschreven op een schijf, waarbij het uitlezen en wegschrijven van de data relatief veel tijd kost, maar in een geheugenchip. Deze werkt veel sneller, maar is op dit moment nog duurder. Omdat de onderhoudskosten van storagetechnologie die gebruik maakt van harde schijven eerder toenemen dan afnemen en

flashstorage minder onderhoud nodig heeft en langer meegaat, is de verwachting dat het gebruik van conventionele data storage met harde schijven in de toekomst vaker vervangen zal worden door flash storage. In toenemende mate gaat ook artificial intelligence (AI) een rol spelen bij het beheer van data en de opslag, vindbaarheid en kwaliteit ervan.

### **Welke dataopslag heeft u nodig?**

Het soort dataopslag dat u nodig heeft, hangt af van veel factoren. Veiligheid en eigen kennis zijn belangrijk. Ook het soort data en de hoeveelheid ervan spelen een rol. Maar ook de flexibiliteit, technische aspecten, voorkeuren voor technologie, kosten en de zekerheid daarover. Voor een verstandige keuze is het raadzaam u zich hierover goed te laten informeren.

# Databases

Bron: ICT informatiecentrum | [ictinformatiecentrum.nl](https://www.ictinformatiecentrum.nl)

Databases komen voor rondom alle soorten bedrijfssoftware en een groot deel van de bestaande IT oplossingen. Zodra er met software of IT oplossingen gegevens worden verwerkt en bewaard, is een database noodzakelijk. Deze bevat namelijk de informatie waar de systemen gebruik van maken.

## **Wat is een database?**

Een database is een systeem waar data geordend en overzichtelijk verzameld wordt. Dat gebeurt meestal in omvangrijke tabellen, waardoor informatie snel en via een vaste structuur van kolommen en rijen opgeslagen en teruggevonden kan worden. Door de sterke groei van de hoeveelheid beschikbare data stijgt de vraag naar flexibelere opslagtechnieken die organisaties ook in staat stellen grote hoeveelheden ongestructureerde data op te slaan. Er bestaan daarom meerdere soorten databases. Vanuit een database is informatie snel bruikbaar door het eraan gekoppelde computersysteem.

Voor het definiëren van een database zijn vier aspecten van belang. Zo moet nieuwe informatie makkelijk bewaard kunnen worden (create), de beheerder in staat gesteld worden al bestaande informatie makkelijk te doorzoeken (read) en te onderhouden (update). Tenslotte moet informatie kunnen worden verwijderd (delete), zonder negatieve operationele gevolgen voor het systeem. Voor het hebben van een betrouwbare en solide database is het noodzakelijk dat alle gegevens correct en eenduidig zijn opgeslagen. Daarnaast mag de inhoud van de database uitsluitend door bevoegden worden uitgebreid, onderhouden of verwijderd.

## **Database ontwikkelingen**

De database en de bijbehorende technieken hebben grote ontwikkelingen doorgemaakt sinds de introductie van databases in de jaren zestig. Waar de eerste databases nog erg eenvoudig en statisch waren, zijn deze door de jaren heen steeds groter en flexibeler geworden. Door de snelle ontwikkeling van het internet en de daarbij behorende datastroom, steeg de vraag naar snelle verwerking van grote hoeveelheden, veelal ongestructureerde data. Moderne databases voorzien in die mogelijkheid en onderhouden zichzelf grotendeels autonoom door gebruik te maken van machine learning.

## **Databases versus datawarehouses en data lakes**

Binnen het thema data management zijn er veel termen die op elkaar lijken. De begrippen database, datawarehouse en data lake zijn daarvan voorbeelden. Het betreft alle drie technieken om data op te slaan, maar toch zijn er wezenlijke verschillen. Zo is een database een plaats om geordend informatie op te slaan, meestal gekoppeld aan een separaat IT systeem. Data van databases van verschillende systemen kunnen worden samengevoegd in een datawarehouse. Met een datawarehouse krijgt u een volledig en eenduidig beeld van uw data. De structuur en omvang van een datawarehouse biedt voordelen bij data analyses, dat u antwoord geeft op businessvraagstukken. Voor de analyses maakt u gebruik van business intelligence tools.

Een data lake werkt anders. Waar databases en datawarehouses gebaseerd zijn op structuur en gestructureerde data, is dat bij een data lake niet het geval. In een data lake worden grote hoeveelheden informatie in ruwe en onbewerkte vorm opgeslagen. De ongestructureerde vorm maakt een data lake zeer geschikt voor datamining en machine learning. Veel organisaties willen de enorme hoeveelheid data die ze genereren en tot hun beschikking hebben ergens opslaan, zonder dat daar direct een doel voor geldt. Over het algemeen geldt hoe meer data, hoe betrouwbaarder analyses worden. In theorie kunt u dus betere managementbeslissingen nemen als u beschikt

over meer informatie. Een data lake gebruikt u als opslagmedium voor een enorme hoeveelheid aan data die ogenschijnlijk niet directe waarde heeft.

### Soorten databases

Er is een grote variëteit aan databases. Wat de beste keuze voor u is hangt af welk doel uw organisatie ermee heeft. Databases kunnen op verschillende aspecten worden gecategoriseerd. Dat kan op basis de opslagtechniek die ze gebruiken of andere onderscheidende eigenschappen. Enkele van de meest gangbare databasetechnieken worden hieronder kort besproken.

- Een **relationele database** (SQL) is opgebouwd uit rijen en kolommen. Het is de meest effectieve vorm voor het opslaan en aanspreken van (grotere hoeveelheden) geordende informatie.
- In een **flat file database** worden gegevens in pure tekst opgeslagen, in een soort versimpelde vorm van een relationele database. Omdat het zo simpel is, is het voor geringe uitwisseling van informatie erg doeltreffend. Bij grote hoeveelheden data wordt het al snel onoverzichtelijk.
- Een **objectgeoriënteerde database** bestaat uit vaste pakketjes data (objecten). Denk bij een object aan een schematische opsomming van informatie over een onderwerp, zoals accountinformatie van een enkele gebruiker. In dit geval zal het object bijvoorbeeld informatie bevatten, zoals de gebruikersnaam, het e-mailadres en het klantnummer.
- Een **JSON database of documentdatabase** slaat in verschillende documenten data in een standaardformaat op. Bij elk document hoort een unieke sleutel, waarmee de data opgeroepen kan worden.

- Een **NoSQL database** is een niet-rationele database die ervoor zorgt dat niet- en semi-geordende data samen opgeslagen en verwerkt kunnen worden.
- Een **OLTP database** is een database waaruit in een hoog tempo tussen meerdere gebruikers veel informatie-uitwisselingen gedaan kunnen worden. Voorbeelden van zulke uitwisselingen zijn het plaatsen van een bestelling en het versturen van een bericht.
- Een **OLAP database** is bedoeld voor analytische doeleinden. Het verzorgt snelle informatie-uitwisselingen tussen verschillende databases en de analysesoftware.
- Een **multi-model database** koppelt verschillende databasetechnieken aan elkaar tot één geïntegreerde backend, zodat één informatiebron verschillende datasoorten kan ondersteunen.
- Een **datawarehouse** is een centrale verzamel- en verwerkplek voor informatie uit verschillende databases. Het heeft als taak om op een overzichtelijke manier te voorzien in de vraag naar data binnen een organisatie. Een datawarehouse biedt grote voordelen bij het gebruik van business intelligence (BI).

Ook zitten er verschillen in bepaalde vormen van een database, zonder te kijken naar welke opslagmethode er wordt gebruikt. Het gaat dan bijvoorbeeld over de manier waarop de database bereikbaar is of hoe het systeem wordt onderhouden.

- In een **gedistribueerde database** staan verschillende bestanden op diverse opslaglocaties. Door via een netwerk computers te verbinden kunnen bestanden uitgewisseld worden. Er is dus niet



één centrale plek waar alle data opgeslagen is, maar alle data kan wel vanuit verschillende plekken worden aangesproken.

- Een **cloud database** is een database op een externe locatie die via het internet wordt bereikt. De database bevat uw data, maar wordt onderhouden door de cloudaanbieder (DataBase-as-a-Service).
- Een **open source database** is een database waarvan de onderliggende broncodes openbaar zijn. Dit geeft als voordeel dat de databasesoftware naar wens van de gebruiker kan worden aangepast.

Databases kunnen meerdere van bovenstaande benamingen hebben. Zo bestaan er relationele cloud databases (een database opgebouwd uit rijen en kolommen die door een cloudaanbieder gehost wordt) en open source object georiënteerde databases (databases die werken op basis van pakketjes schematische opsommingen en waarvan de broncode openbaar is).

### **Wat zijn zelfsturende databases?**

Een zelfsturende database beheert zich grotendeels zelf met behulp van machine learning. Het tijdrovende en kostbare routinewerk dat voorheen werd uitgevoerd door een databasebeheerder, wordt hierbij overgenomen door computers. Taken zoals beveiligen, updaten en backups maken zijn geautomatiseerd. Zelfsturende databases gaan naar verwachting grote invloed krijgen op de manier waarop organisaties hun informatie exploiteren. Terwijl de kosten afnemen worden de beveiliging en algemene prestaties sterk verbeterd.

### **Hoe beheert u een database?**

Wanneer een database in omvang toeneemt, wordt het steeds moeilijker en duurder om het systeem te beheren. Onderhoud is enorm van belang om

prestaties en de betrouwbaarheid van de database te waarborgen. Daarom wordt het beheer en onderhoud van een database gedaan met een database management system (DBMS). Deze geeft de beheerder inzicht in de status van de database. Vanuit een DBMS kunnen onderwerpen zoals beveiliging (security), schadeherstel (recovery), prestatie monitoring en updates geregeld worden.

# Data warehouse

Bron: ICT informatiecentrum | [ictinformatiecentrum.nl](https://www.ictinformatiecentrum.nl)

Data moet veilig worden opgeslagen en direct beschikbaar zijn als u of de applicatie die er gebruik van maakt, deze nodig heeft. Voor dit doel bestaan databases. Verschillende applicaties, zoals ERP, CRM, DMS of HRM systemen, maken gebruik van hun eigen databases. Voor eenduidigheid van gegevens die in verschillende systemen voorkomen en om gegevens uit verschillende systemen met elkaar in verband te kunnen brengen, is een centrale database nodig: het datawarehouse. Deze bevat gestructureerde data uit de afzonderlijke databases.

## **Wat is een datawarehouse?**

Een datawarehouse is een database voor het centraal opslaan en verwerken van gestructureerde data uit verschillende onderliggende databases. Deze onderliggende databases horen bij meerdere operationele systemen, zoals ERP software, CRM software, HRM software of andere soorten bedrijfssoftware. Bij het dagelijkse gebruik van de operationele systemen, benutten deze alleen hun eigen databases. Als er tussen verschillende bedrijfsprocessen met behulp van data verbanden gelegd of conclusies getrokken moeten worden, dan lukt dat niet meer door gebruik te maken van de afzonderlijke databases. De oplossing is dan om alle data samen te voegen in een centrale database: het datawarehouse.

De databases van operationele systemen bevatten gestructureerde data. Kenmerk daarvan is dat deze data in tabelachtige structuren opgeslagen kan worden. Ook het resultaat van de samenvoeging, het datawarehouse, is weer een database met gestructureerde data. Met hulp van business intelligence en data analytics kunt u waardevolle inzichten verwerven uit deze data en op basis daarvan beslissingen nemen. Een goed datawarehouse

draagt hierdoor in grote mate bij aan het succes van een datagedreven organisatie.

### **Datawarehouses en business intelligence**

De kern van business intelligence (BI) is dat de in de organisatie aanwezige data wordt gebruikt voor het nemen van managementbeslissingen. Het datawarehouse bevat de data waarop de BI dashboards hun visualisaties, meetgegevens en conclusies baseren. Datawarehouses en business intelligence zijn daarom aan elkaar verbonden. Door uw data gestructureerd en eenduidig in een datawarehouse op te slaan, kunt u het grootste rendement uit een business Intelligence oplossing halen.

### **Datawarehouse versus data lake**

Voor het opslaan van grote hoeveelheden data kan, naast een datawarehouse, ook gekozen worden voor een data lake. Het onderscheid tussen deze twee wordt bepaald door de manier waarop data verwerkt en opgeslagen wordt. In een datawarehouse wordt alle data gestructureerd, gecontroleerd en zonder fouten (opgeschoond) bewaard. U kunt daardoor gemakkelijker gebruik maken van business intelligence oplossingen. Bij een data lake worden alle soorten gestructureerde en ongestructureerde data in allerlei vormen (ook bestanden, beelden of berichten) en in originele vorm opgeslagen. Een data lake biedt voordelen bij de analyse van complexere data, omdat er geen rekening gehouden hoeft te worden met de beperkingen die de vaste structuur van een datawarehouse kan opleggen. Beide opslagtechnologieën hebben dus specifieke eigenschappen, waardoor de keuze bepaald wordt door welke data u heeft, welke u wilt gebruiken voor analyse en naar welke resultaten u zoekt. Een datawarehouse en data lake hoeven elkaar niet uit te sluiten. U kunt beide ook gelijktijdig naast elkaar gebruiken.

## Waarom een datawarehouse toepassen?

Om diverse redenen kunt u een datawarehouse toepassen:

### **Integratie**

U slaat data op verschillende plekken en in verschillende systemen op. Om een volledig overzicht te krijgen van alle bedrijfsprocessen, klanten en aanbieders is het van belang om alle data eenduidig te bundelen in een centrale database. Op die manier ontstaat een eenduidig beeld over uw gehele organisatie en uw activiteiten. Het streven is dat uw medewerkers data en databases makkelijker kunnen interpreteren door deze aan elkaar te koppelen en met elkaar in verband te brengen.

### **Snelle reactie**

De operationele systemen waarin de data gegenereerd wordt (ERP, CRM, enz.), zijn ontworpen om in een hoog tempo talloze kleine data-uitwisselingen uit te voeren. Voor het analyseren van uw bedrijfsdata, waarbij tegelijkertijd miljoenen data-elementen verzameld, samengevoegd, verwerkt en afgebeeld moeten worden, zijn de databases van deze systemen minder geschikt. De structuur van een datawarehouse leent zich daar beter voor. Deze is geoptimaliseerd om snel in één keer flinke hoeveelheden informatie te verstrekken.

### **Beschikbaarheid**

Als u over een datawarehouse beschikt, kunnen uw medewerkers eerder en beter de betekenis van data begrijpen. Met losse databases is het vaak lastiger om dit doel te bereiken, zeker als deze databases van verschillende aanbieders zijn, geen standaard structuren kennen of (bijvoorbeeld bij cloudtoepassingen) moeilijk bereikbaar kunnen zijn. Een datawarehouse zorgt voor een goede en snelle beschikbaarheid van alle data, waardoor uw medewerkers eerder tot waardevolle inzichten kunnen komen.

### **Rapportages**

Dankzij de gestructureerde opbouw van een datawarehouse kunnen uw medewerkers snel en eenvoudig analyses en rapportages maken, die zich bovendien makkelijk aan laten passen. Zo wordt bijvoorbeeld bij het analyseren van de financiën gemakkelijk geschakeld tussen maanden en kwartalen. Als u op die manier 'speelt' met het verwerken en visualiseren van data, kunt u waardevolle en unieke inzichten verwerven.

### **Self service**

Een datawarehouse in combinatie met toegankelijke business intelligence oplossing, zorgt ervoor dat uw medewerkers autonoom analyses en rapportages kunnen maken. Om niet-IT medewerkers zo min mogelijk afhankelijk te laten zijn van de BI- en IT-experts binnen of buiten uw organisatie, is het belangrijk te investeren in een goede infrastructuur tussen het datawarehouse en de BI tools. Door medewerkers goed te trainen en te coachen zorgt u ervoor dat ook niet-technische medewerkers op een hoog niveau analyses kunnen uitvoeren. Zij hoeven daarbij niet terug te vallen op de kennis van de IT afdeling.

### **Historische data**

Een losse, aan operationele systemen gekoppelde database registreert vaak niet de historie van de data. Er is een grote kans dat de oude waarde direct wordt vervangen door de nieuwe, zonder de oude waarde ergens op te slaan. Hierdoor gaat veel waardevolle informatie verloren, die juist kan zorgen voor een compleet beeld van gegevens en prestaties door de tijd heen. Door ook historische data in een analyse te verwerken, ontstaat een veel betrouwbaarder beeld, dan wanneer u alleen actuele informatie gebruikt. Een datawarehouse bewaart ook de historische data, door alle mutaties in de oorspronkelijke databases op te slaan.

### **Datakwaliteit**

De kwaliteit van data in de databases van operationele systemen kan niet altijd worden gegarandeerd. Er vinden standaard geen controles op de datakwaliteit plaats. Gebruikers hebben de mogelijkheid zelfstandig data aan het systeem toe te voegen of te wijzigen. Veelal blijken gegevens dan ook niet juist of onvolledig te zijn opgeslagen. Met een datawarehouse kan de datakwaliteit worden verbeterd door automatisch fouten te herkennen en te laten verbeteren.

### **Operationeel**

Door alle data te verzamelen in een datawarehouse en van daaruit analyses uit te voeren, wordt de druk op de onderliggende databases ontlast. Op die manier kunt u grote hoeveelheden data analyseren, zonder dat de bronsystemen operationele problemen ondervinden door overbelasting. Uw bedrijfsprocessen kunnen onbelemmerd voortgezet worden, terwijl via het datawarehouse complexe en veel data gebruikende analyses worden uitgevoerd.

### **Nadelen en beperkingen**

Een datawarehouse kent ook nadelen en beperkingen. Zo is het ontwerpen van een datawarehouse een tijdrovende en kostbare taak, die specialistische kennis vereist. Het is veel werk om alle data bij elkaar te zoeken, te verwerken en beschikbaar te stellen. Klein beginnen is belangrijk. Start met de bedrijfsdata die de hoogste potentiële waarde vertegenwoordigt. Voorkom dat verschillende bedrijfsafdelingen niet direct begrijpen waarom gegevens met elkaar moeten worden gedeeld. Leg daarom het nut en belang van een datawarehouse uit aan de verschillende afdelingen. Door op kleine schaal te starten met de implementatie van een datawarehouse kan de toegevoegde waarde ervan stap voor stap worden bewezen. Door met een simpele vorm te beginnen en deze telkens uit te breiden, wordt het draagvlak onder medewerkers langzamerhand groter.

## **Hoe implementeert u een datawarehouse?**

Voor het implementeren van een datawarehouse kunt u vijf stappen onderscheiden:

### **Informatiebehoefte in kaart brengen**

Bepaal welke managementbeslissingen u in de nabije toekomst wilt nemen, welke inzichten u wilt hebben, op welke vragen u antwoord wenst en met wat voor soort werkwijzen en gegevens uw organisatie te maken heeft. Door de informatiebehoefte in kaart te brengen, kunt u op een zo efficiënt mogelijke manier overstappen naar een datawarehouse.

### **Inventariseer databronnen**

Kijk welke databases de informatie bieden waaraan binnen uw organisatie behoefte is. Onderzoek of aan elke datavraag voldaan wordt, hoe de data opgeslagen is en van welk kwaliteitsniveau de data is. Bepaal hoe u omgaat met privacygevoelige informatie, omdat daar vanuit wettelijke verplichtingen extra restricties op zitten. Zo mogen alleen onder strikte voorwaarden bepaalde persoonsgegevens, zoals etnische afkomst of geloofsovertuiging, verwerkt worden in een datawarehouse.

### **Maak een planning**

Door naar de informatiebehoefte te kijken kunt u bepalen waar uw prioriteiten liggen. In plaats van het eerst verwerken van alle data, wordt aangeraden om te beginnen met in verhouding kleine en gemakkelijke projecten die niet veel data vereisen. Op die manier wordt tijd bespaard en zijn de eerste (positieve) resultaten van het datawarehouse snel zichtbaar. Bij acute informatievraagstukken, waarbij de benodigde informatie nog niet vanuit het datawarehouse beschikbaar is, kan gekozen worden voor een tijdelijke oplossing: door alleen actuele informatie te laden en te gebruiken en historische data achterwege te laten, kan in korte tijd een relatief



minder betrouwbaar, maar alsnog waardevol rapport gegenereerd worden.

### **Bepaal de techniek**

Er bestaan verschillende technieken voor een datawarehouse. Ontwerp het datawarehouse zorgvuldig, zodat het aansluit bij de vraag vanuit uw organisatie. Bepaal hoe de data verkregen, geïntegreerd en aangeleverd moet worden. Bepaal of u een datawarehouse lokaal (on-premise) of in de cloud wenst. Kijk naar hoe de actuele datastroom verwerkt wordt en op welke manier gegevens gevisualiseerd en geïnterpreteerd kunnen worden.

### **Implementeren**

De implementatie van een datawarehouse bestaat niet alleen uit het beschikbaar maken van de BI tools. Het gaat ook om een verandering in de werkwijze van uw organisatie. Het is van belang dat uw medewerkers stap voor stap bekend raken met het datawarehouse en de mogelijkheden daarvan. Zorg ervoor dat het kennisniveau van uw medewerkers stijgt, naarmate het datawarehouse u meer mogelijkheden biedt. Stel de introductie met de nieuwe datagedreven aanpak niet uit tot het moment dat uw datawarehouse af is. Alleen door uw medewerkers vanaf het begin van het project goed te trainen en bij uw project te betrekken, kunt u ervoor zorgen dat er straks de maximale waarde uit uw data gehaald kan worden.

# In 5 stappen naar een succesvol datawarehouse

Bron: Hot ITem | [hotitem.nl](http://hotitem.nl)

In een datawarehouse wordt data uit verschillende bronnen binnen een organisatie verzameld, gecombineerd en aangeboden. Daarmee kan een datawarehouse waardevolle inzichten opleveren voor de organisatie. In de praktijk maken nog veel organisaties de fout om een datawarehouse te benaderen vanuit de techniek; de organisatie wordt verleid met een mooie tool en vaart blind op de aantrekkelijke visualisaties die in het vooruitzicht gesteld worden. Maar een datawarehouse heeft als belangrijkste doel te kunnen voldoen aan de informatiebehoeften die er zijn binnen de organisatie. Dat dient ook het startpunt te zijn. Waar heeft de eindgebruiker behoefte aan?

Lees hier een praktisch stappenplan voor het opzetten van een succesvol datawarehouse. Ook al is het concept van een datawarehouse intussen alweer zo'n dertig jaar in omloop, ook in de huidige wereld van big data en advanced analytics biedt een datawarehouse nog steeds een duidelijke meerwaarde.

## **Stap 1 - Achterhaal de informatiebehoefte**

Neem de organisatie als uitgangspunt. Werk vanuit de strategie toe naar de belangrijkste informatiebehoeften. Welke inzichten heeft de organisatie nodig om gericht te kunnen sturen? Werk een procesmodel uit: met welke processen en welke gegevens werkt de organisatie? Start ook met het functioneel modelleren van de organisatie: werk een bedrijfsobjectenmodel uit. Bedrijfsobjecten zijn alle gegevensobjecten die van belang zijn binnen een

organisatie. Van deze objecten kunnen eigenschappen worden vastgelegd met behulp van data.

### **Stap 2 - Ga na welke gegevensbronnen nodig zijn**

Welke gegevensbronnen zijn nodig om in de informatiebehoefte van de organisatie te kunnen voorzien? Welke bronnen zijn er al en welke moeten er eventueel nog worden gecreëerd? Bedenk ook hoe de gegevens er uitzien. Zijn ze makkelijk te ontsluiten of niet? Wat is de kwaliteit van de data? Hoe goed zijn alle velden ingevuld? Maar ook: wat voor gegevenselementen bevatten ze? Houd ook rekening met de privacy van de betrokkenen: gaat het om gevoelige gegevens? Bijzondere persoonsgegevens, bijvoorbeeld over iemands etniciteit, gezondheid of levensovertuiging, worden door de wetgever extra beschermd. Je mag alleen bijzondere persoonsgegevens verwerken als je je kunt beroepen op een wettelijke uitzondering.

### **Stap 3 - Stel een roadmap op**

Op basis van de informatiebehoefte binnen de organisatie en de kennis van de gegevensbronnen maak je een afweging: wat pak je als eerste op? Wat is de meest logische keuze? Je kunt bijvoorbeeld starten met een relatief klein deelproject dat naar verwachting snel tot resultaat leidt, om het succes van het datawarehouse te kunnen aantonen. Een valkuil is om eerst alle gegevens te verzamelen en alle data met elkaar te combineren om tot een uniform en veelomvattend model te komen van de gehele organisatie. Deze werkwijze kost veel tijd. Tegen de tijd dat het datawarehouse opgeleverd kan worden, is de kans groot dat de organisatie, de informatiebehoefte en de gegevenshuishouding veranderd zijn. Houd het grotere plaatje voor ogen, maar begin klein en blijf continu herijken.

Bij een urgente informatiebehoefte of weerstand binnen de organisatie kun je kiezen voor een zogenaamd ‘snelspoor’. Je levert dan bijvoorbeeld nog geen historische data en volledige integratie van gegevensbronnen op, maar wél alvast een rapportage op basis van één bronsysteem. Daarmee heb je in een kort tijdsbestek al een informatieproduct beschikbaar, waarmee de eindgebruiker met de urgente behoefte alvast geholpen is. Bovendien kun je de meerwaarde van het datawarehouse hiermee al snel in de praktijk aantonen. Dit vergroot het draagvlak voor vervolgotrajecten.

#### **Stap 4 - Kies een datawarehouse architectuur en technologie**

Bij deze stap komt de techniek om de hoek kijken. Welke modelleringstechnieken kies je voor het datawarehouse? Zet je het datawarehouse op volgens Kimball, Inmon of wordt het een datavault-model volgens Linstedt? Welke lagen creëer je? Denk bijvoorbeeld aan een drielagenmodel, dat aansluit bij de drie functies van een datawarehouse (verzamelen, integreren en aanbieden). En kies je voor een datawarehouse in de cloud of niet? Houd ook rekening met de plaats van ontwikkelingen, zoals realtime data en datavirtualisatie.

#### **Stap 5 - Lever het datawarehouse op**

In de voorgaande vier stappen heb je de blauwdruk van het datawarehouse uitgetekend. Nu is het tijd om het datawarehouse daadwerkelijk te bouwen en op te leveren. Vaak resulteert deze stap in een rapportage of dashboard voor een eindgebruiker. Maar om daar te komen is het belangrijk om de tussenstappen, de verschillende lagen in het datawarehouse, op te leveren als ze klaar zijn. Wacht daar niet mee tot het eindproduct af is. Gebruik een werkwijze als DevOps om regelmatig werkende halffabricaten of producten op te leveren.

Om van het datawarehouse écht een succes te maken binnen de organisatie, is het belangrijk dat de gebruikers ermee leren omgaan. Deze laatste stap draait dus niet alleen om het opleveren van het datawarehouse, maar ook om het implementeren van een datagedreven cultuur. Hoe kunnen de eindgebruikers sturen en verantwoorden op basis van de data? Wat moeten zij weten van het datafundament om er het beste rendement uit te halen?

# Data lake

Bron: ICT informatiecentrum | [ictinformatiecentrum.nl](https://www.ictinformatiecentrum.nl)

Er is steeds meer data uit steeds meer verschillende bronnen beschikbaar. De hoeveelheid beschikbare data is zelfs zo groot dat u over meer data kunt beschikken, dan waar u op dit moment een toepassing voor heeft. Omdat de kosten voor dataopslag relatief laag zijn, kunt u alle data eenvoudig bewaren. Ook de data die u niet direct gebruikt. De verzamelplaats van die data wordt een data lake genoemd.

## **Wat is een data lake?**

De hoeveelheid data neemt toe. Iedere organisatie beschikt over een toenemende hoeveelheid data over klanten, leveranciers, interne bedrijfsprocessen, marktontwikkelingen, financiële transacties of websitebezoek. Vanuit productieprocessen zijn steeds meer gegevens beschikbaar. Social media en andere communicatievormen voegen daar nog veel data aan toe, evenals apparaten die met het internet verbonden zijn (internet of things). Daarnaast is ook vanuit externe bronnen een toenemende hoeveelheid data beschikbaar. Door de combinatie van hoeveelheid, de snelheid waarmee we er iets mee willen doen en grote verscheidenheid waarin deze data voorkomt, spreken we over big data. Zinvol gebruik van deze data is voor analyses die meer inzicht geven in allerlei zaken waar managementbeslissingen over genomen moeten worden.

Voorheen was de beschikbare data voornamelijk gestructureerd. Gestructureerde data kan in logische tabellen worden weergegeven en is afkomstig van operationele systemen, zoals ERP software of CRM software. Met een datawarehouse als centraal verzamelpunt van gestructureerde data uit meerdere systemen, is deze data bruikbaar voor business intelligence en andere data analytics toepassingen. Het grootste deel van alle data die u op dit moment kunt verzamelen is echter ongestructureerd. Hiervoor is een

andere databasetechnologie nodig is. Een data lake biedt de oplossing om alle beschikbare data, gestructureerd en ongestructureerd, op te slaan. Zelfs zonder beoogde toepassing, wordt het data lake gevuld met data. Door de lage kosten van dataopslag hoeft de opslag niet direct verbonden te zijn aan een concreet doel. Uw data in een data lake gebruikt u pas als deze nodig is voor een analyse. De achterliggende onderzoeksvraag daarvan hoeft u nu nog niet te kennen.

### **Data lake versus datawarehouse**

Voor het opslaan van grote hoeveelheden data kunt u kiezen voor een data lake of een datawarehouse. Uw keuze wordt bepaald door het soort data en de manier waarop deze opgeslagen wordt. Bij een data lake gebeurt de opslag zonder speciale structuur. Een enorme hoeveelheid gegevens worden in ruwe en originele vorm centraal opgeslagen. Het idee daarbij is 'baat het niet, dan schaadt het niet'. Zo staan afbeeldingen naast verkoopcijfers en e-mails naast loggegevens.

Een datawarehouse werkt anders en is samengesteld uit data uit specifiek geselecteerde bronnen. Alle gegevens zijn gestructureerd, hebben een vast format en zijn zonder 'vervuiling' opgeslagen. Deze zijn hierdoor gemakkelijker te interpreteren en geven sneller antwoord op business intelligence vraagstukken. Bij een data lake is de data niet gestructureerd. Om er iets mee te doen is dan ook al snel de hulp van een specialist nodig. Het ongestructureerde karakter van een data lake biedt voordelen voor datamining en machine learning, al is het in de praktijk vaak lastig om over de enorme hoeveelheid gegevens overzicht te behouden en daarmee de datakwaliteit te waarborgen. Beide oplossingen vergen onderhoud en brengen kosten met zich mee.

### **Waarom een data lake toepassen?**

Een data lake kan alle informatie bevatten waarover u beschikt en is een combinatie van geordende databases, op zichzelf weinigzeggende losse bestanden en alles daar tussenin. Het gebrek aan structuur biedt voordelen

bij het zoeken naar verbanden tussen verschillende databronnen door middel van machine learning. De oplossing wordt niet beperkt door de vaste structuur, die datawarehouses wel hebben. Hierdoor is een data lake een experimentele plek, waar u bijna grenzeloos uw data kunt onderzoeken, om zo tot nieuwe inzichten te komen. Het is een extra opslaglocatie om naast bestaande databases te gebruiken. Het data lake vervangt deze echter niet.

### **Wat bepaalt de keuze tussen een datawarehouse of data lake?**

Er valt niet te zeggen dat de één een betere oplossing is dan de andere. Wat het beste is, verschilt per organisatie en is afhankelijk van het doel. Er bestaan voldoende analyse oplossingen voor zowel datawarehouses als data lakes. Heeft u vooral behoefte aan overzichtelijke en eenduidige informatie om te gebruiken voor BI tools en datavisualisatie, dan is een datawarehouse een voor de hand liggende keuze. Als u wilt beschikken over veel ongeordende data voor geavanceerde analyses door middel van machine learning, dan is een data lake een betere oplossing. Beide opties hoeven elkaar niet uit te sluiten. U kunt ook kiezen voor oplossingen waarbij een datawarehouse en data lake naast elkaar bestaan en beide optimaal benut worden in hun eigen specialisme.



# Data migratie

Bron: ICT informatiecentrum | [ictinformatiecentrum.nl](https://www.ictinformatiecentrum.nl)

Datamigratie is het proces van dataoverdracht van het ene systeem naar het andere. Voordat deze overdracht mogelijk is, dient de data te worden geselecteerd die voor migratie in aanmerking komt. Daarna moet deze gevalideerd worden om er zeker van te zijn dat deze op de juiste manier wordt ondersteund in de nieuwe applicatie. Veel standaardsoftware beschikt over functies om de migratie van data te ondersteunen en eenvoudig te laten verlopen.

## **Migratie en conversie**

Datamigratie is vaak verbonden aan een systeem migratie. Systeem migraties kunnen nodig zijn als bedrijven fuseren of als bedrijfsactiviteiten worden overgenomen of uitbested. Datamigratie betreft dan het overzetten van de data van die systemen. Doel van de migratie is dat alle data behouden blijft en dat het systeem in de nieuwe configuratie optimaal gebruik kan maken van al eerder beschikbare data. De begrippen datamigratie en dataconversie worden vaak door elkaar gebruikt. Het woord migratie slaat daarbij vaak op het verplaatsen van hele systemen, terwijl conversie in veel gevallen alleen op de verhuizing van de data betrekking heeft.

## **Redenen voor datamigratie**

Er kunnen verschillende redenen zijn voor een migratie van data. Vaak is migratie nodig bij het in gebruik nemen van nieuwe hardware of overgang naar een andere softwareoplossing. Er zijn drie verschillende vormen van datamigratie: migratie tussen dataopslagsystemen, applicaties en databases.

### **Migratie van dataopslagsystemen**

Hierbij gaat het om datamigratie van en naar verschillende typen opslagsystemen, zoals lokale opslag via harde schijven of de cloud.

De reden hiervan is meestal een technologische update, waarbij een nieuwe vorm van opslag in gebruik wordt genomen en waardoor het oude systeem overbodig is geworden. Dit is tevens een goed moment voor het uitvoeren van datavalidatie en -reductie.

Tegenwoordig wordt er veel data naar de cloud gemigreerd. De cloud biedt een combinatie van schaalvoordelen en lagere kosten en krijgt daardoor steeds vaker de voorkeur boven lokale opslag. Deze trend staat niet op zichzelf. Ook bedrijfsapplicaties, zoals ERP, CRM of HRM software, draaien steeds vaker in de cloud. Dat de data vaker moet worden verhuisd, past in deze ontwikkeling.

### **Databasemigratie**

Wanneer er van databaseprovider wordt gewisseld, een update wordt uitgevoerd van de huidige database of een database naar de cloud wordt gemigreerd, is er sprake van een databasemigratie. Het is mogelijk dat hierbij de data zelf wordt gewijzigd, al hoeft dat niet het geval te zijn. Databasemigraties hebben vaak betrekking op het wijzigen van een onderliggend datastructuur, niet op de data in de database zelf.

### **Migratie naar andere applicaties**

Hiervan is sprake als er naar een andere applicatie of platform wordt overgestapt. Dit is een vaak complexe migratie, omdat applicaties binnen een organisatie zonder problemen met elkaar moeten kunnen blijven communiceren. Soms is hiervoor maatwerk noodzakelijk.

### **Oplossingen voor datamigratie**

Het type datamigratie bepaalt het uitvoeringsproces van de migratie zelf. In het geval van een nieuw dataopslagsysteem, zijn er verschillende tools en producten beschikbaar die het proces vergemakkelijken. Bij de keuze voor een oplossing is het praktisch om een checklist te maken voor de functionaliteiten die u nodig heeft in uw bedrijfssituatie.

## **Datamigratie naar de cloud**

Het migreren van data naar de cloud is voor veel organisaties nog relatief onbekend terrein. U heeft keuze uit een private cloud, public cloud of mengvormen. Het beheer van een private cloud vraagt extra middelen van uw organisatie, want het beheer ervan ligt bij uzelf. Dit in tegenstelling tot de public cloud. Het grootste probleem voor organisaties die hun data naar de public cloud willen migreren, ontstaat door de juridische consequenties van de datamigratie. Voordat een organisatie besluit over te gaan tot datamigratie naar de cloud, moet duidelijk zijn wie verantwoordelijk en aansprakelijk is als er iets misgaat met het beheer en beschermen van de data. In ICT zaken gespecialiseerde juridisch adviseurs kunnen u hierbij helpen. Daarnaast zijn er diverse ICT dienstverleners die u kunnen helpen met het migreren van uw data en applicaties. Veel van hen kunnen u bovendien oplossingen bieden voor backups, security- en toegangscontrole.

# Data integratie

Bron: ICT informatiecentrum | [ictinformatiecentrum.nl](http://ictinformatiecentrum.nl)

Data integratie heeft als doel om gegevens die voorkomen in verschillende systemen met elkaar te combineren, zodat deze in uw hele organisatie eenduidig gebruikt kunnen worden. Meestal wordt de gecombineerde data 'schoon' en verrijkt opgeslagen op één centrale plaats. Door data uit verschillende bronnen op een goede manier te integreren, worden de gegevens waarmee u werkt betrouwbaarder en is het mogelijk dat verschillende afdelingen gebruik maken van elkaars data.

## **Wat is data integratie?**

Data integratie houdt in dat data uit verschillende systemen wordt samengevoegd. Hierdoor blijft eenduidige data over die gebruikt kan worden door diverse afdelingen in uw organisatie. Data integratie speelt een rol bij data migratie, bijvoorbeeld als twee bedrijven hun databases moeten samenvoegen tot één nieuwe database, maar ook bij het bouwen en beheren van datawarehouses en data lakes of het toepassen van master data management. Ook kan door het integreren van data afdelingen gebruik gemaakt worden van elkaars gegevens. Zo kan het voor uw commerciële afdeling wenselijk zijn inzage te hebben in financiële gegevens over klanten, voordat nieuwe offertes worden uitgebracht of orders worden uitgeleverd. Data uit het financiële pakket moet dan gecombineerd worden met de data uit uw CRM systeem. Bij ERP software speelt dit geen rol, omdat daar de data altijd geïntegreerd en eenduidig is.

## **Waarom is data integratie belangrijk?**

Bedrijfsprocessen zijn gebaat bij juiste en volledige data. Niet alle data die nodig is voor een specifiek bedrijfsproces, wordt gegenereerd in het systeem dat dit proces ondersteunt. Daarvoor zijn koppelingen nodig tussen verschillende databases en moet er bepaald worden welke data de juiste en

dus leidend is. Ook komen dezelfde soort gegevens in verschillende databases voor, die desondanks een verschillende inhoud kunnen hebben. Dit komt vaak voor bij persoons- en contactgegevens. Data integratie richt zich op de koppeling tussen de databronnen en de eenduidigheid van data. Alleen dan kunt u er optimaal gebruik van maken.

### **Voordelen van integreren**

Alle bedrijfsprocessen hebben er belang bij om te kunnen beschikken over alle data die kan helpen bij de optimale uitvoering van die processen. Door klantgegevens te integreren met financiële data, kan bij commerciële activiteiten beter rekening gehouden worden met de kredietwaardigheid of betalingsachterstanden van een klant. U bent eerder op de hoogte van financiële risico's en kan daarop inspelen met leveringsvoorwaarden en aanvullende afspraken. Bij marketingactiviteiten beschikt u over betere en betrouwbaardere contactgegevens, zodat marketingacties succesvoller kunnen zijn. Klanten bereikt u met de juiste gegevens, waardoor de communicatie verbetert. Uw financiële afdeling kan sneller handelen bij facturering en incasso.

### **Actuele klantgegevens**

Een praktisch voorbeeld van data integratie is de mogelijkheid om uw klantdata te laten updaten door daarin gespecialiseerde bedrijven. Een aanzienlijk deel van alle contactgegevens verandert zonder dat u dat in de gaten hoeft te hebben. Bedrijven verhuizen en krijgen nieuwe telefoonnummers. Mailadressen en medewerkers krijgen een andere functie of gaan uit dienst. De in deze data gespecialiseerde bedrijven kunnen u doorlopend voorzien van actuele informatie. Zij hebben daar heel veel databronnen voor en controleren de mutaties. Op vaste momenten wordt uw database automatisch met nieuwe gegevens geactualiseerd. Het is een data integratie met externe bestanden.

### **Data integratie oplossing realiseren**

Bij het realiseren van een oplossing voor data integratie gaat het om twee belangrijke fases: ontwerp en implementatie.

Technisch is veel mogelijk bij het overzetten en combineren van data uit verschillende systemen. Belangrijk zijn de antwoorden op vragen zoals welke data er geïntegreerd moet worden, met welk doel dat gebeurt, vanuit welke systemen de data moet worden geïntegreerd, of alle data wel aanwezig is, wat de kwaliteit daarvan is, om welke volumes het gaat, wie de eigenaar is van de data of welke technologie al in huis is (denk aan datawarehouses). De mogelijkheden en keuzes zijn complex genoeg om u hierbij te laten begeleiden door een specialist op dit gebied. Maak daar gebruik van.

# Data kwaliteit

Bron: ICT informatiecentrum | [ictinformatiecentrum.nl](https://www.ictinformatiecentrum.nl)

Datakwaliteit of data quality heeft betrekking op de mate van geschiktheid van data voor uw uiteindelijke gebruiksdoel. Dit betekent dat data betrouwbaar genoeg moet zijn om er conclusies aan te kunnen verbinden. Is de kwaliteit van uw bedrijfsdata gegarandeerd, dan kunt u daar optimaal gebruik van maken in uw operationele systemen en op basis van de data goed onderbouwde beslissingen nemen. Datakwaliteit stelt niet alleen eisen aan de data zelf, maar ook aan de wijze waarop deze wordt beheerd. Hiermee is dit een belangrijk aspect van data management.

## **Wat is datakwaliteit?**

Het is van groot belang dat uw organisatie beschikt over betrouwbare en kwalitatief hoogwaardige data. Alleen dan bent u in staat om slagvaardig en succesvol te handelen. Organisaties waarbij de data niet op orde is, ondervinden daar beperkingen van. Zo zullen marketingactiviteiten hun effect kunnen missen als deze op verkeerde inzichten gebaseerd zijn. Ook zullen interacties met klanten minder gemakkelijk verlopen als blijkt dat de database fouten bevat. Berichten of producten kunnen naar het verkeerde adres gestuurd worden of onjuist gefactureerd worden. Organisaties die de kwaliteit van hun data niet kunnen garanderen, lopen bovendien extra risico's met het naleven van hun wettelijke verplichtingen. De kwaliteit van data kan uitgedrukt worden in een combinatie van de begrippen volledigheid, accuratesse, uniciteit, validiteit, tijdigheid en consistentie.

## **Onderzoeken van de datakwaliteit**

Om een optimale datakwaliteit te garanderen, zult u eerst moeten onderzoeken wat de huidige status is. Door het nemen van steekproeven krijgt u een beeld van de actuele kwaliteit van uw bedrijfsdata. Meestal is het niet mogelijk om alle data te controleren, omdat dat teveel tijd en geld kost.

Met een steekproef onderzoekt u slechts een deel. Constateert u onjuiste gegevens, dan kunt u verder onderzoeken en beslissen of onjuiste data accepteert, negeert of corrigeert.

### **Bewaking van de kwaliteit**

Als grote hoeveelheden data verspreid zijn over meerdere applicaties, kan het garanderen van datakwaliteit lastig zijn. Kwaliteitscontroles moet u dan automatiseren. Diverse databases moeten op gezette tijden gesynchroniseerd worden. Om te voorkomen dat slechte data in de organisatie gebruikt wordt, is het belangrijk om het probleem bij de bron aan te pakken. Dit gaat het meest gemakkelijk als uw data een 'eigenaar' heeft, die als verantwoordelijke optreedt. Waar voorheen incorrecte data achteraf hersteld werd, richt data kwaliteit zich op preventie. Data moet nu van goede kwaliteit zijn voordat het in systemen en databases terecht komt. Met kwaliteitsbewaking voorkomt u dat onbetrouwbare data al negatieve invloed heeft, voordat u het ontdekt.

### **Verbeteren van de datakwaliteit**

Met het in de praktijk brengen van deze aanwijzingen kunt u de kwaliteit van uw data verbeteren:

- Garandeer dat iedereen in uw organisatie gebruik kan maken van dezelfde gegevens. Een centraal bestand is hiervoor essentieel.
- Leg de datastromen in uw organisatie vast. Ga bij uw verschillende processen en applicaties na hoe data wordt toegevoegd, gebruikt, aangepast of verwijderd.
- Monitor de datakwaliteit doorlopend en maak iemand daar verantwoordelijk voor. Houd vooral de plaatsen in de gaten waar nieuwe data ontstaat of waar de datakwaliteit al eerder niet voldoende is gebleken.
- Gebruik dashboards bij het optimaliseren van de datakwaliteit. Deze geven doorlopend inzicht de huidige status. Betrouwbare data is



niet alleen noodzakelijk voor uw bedrijfsprocessen, ook medewerkers werken prettiger als de data aantoonbaar goed is.

# Datakwaliteit; geen project maar een proces

Bron: In Summa Data Solutions | [datasolutions.nl](https://datasolutions.nl)

Data management staat meer dan ooit hoog op de agenda bij organisaties. De behoefte aan consultancy en oplossingen voor het verwerken van data tot informatie groeit iedere dag. Maar hoe borg je de kwaliteit van die data terwijl de databerg tot ver in de hemel groeit, technologie voortdurend veroudert en informatiebehoeften blijven veranderen? Om deze vraag te kunnen beantwoorden kijken we eerst naar de definitie van datakwaliteit, vervolgens naar de kosten van slechte datakwaliteit, hoe we datakwaliteit kunnen meten en tenslotte naar het proces om de kwaliteit van data te verbeteren.

## Wat is datakwaliteit?

De gebruiker van de data, ook wel knowledge worker of information consumer genoemd, bepaalt of een gegeven product of dienst voldoet aan de verwachtingen. De knowledge worker is dus cruciaal in de processen gerelateerd aan datakwaliteit en zijn/haar verwachtingen moeten de driver zijn. Specificaties (definities en business rules), inhoud en voorstelling bepalen samen de kwaliteit van data.

### Voorbeeld

Gevalideerde bestelling (date/time veld in de database). Wanneer wordt een order gevalideerd? Wat is de definitie? Zwak: een binnenkomende order wordt een gevalideerde bestelling wanneer de order wordt goedgekeurd. Beter: een order wordt gevalideerd wanneer:

1. Het een binnenkomende order is of een afgekeurde bestelling, met alle verplichte informatie.
2. De klant als kredietwaardig wordt beschouwd.
3. De prijs van de order, plus het totaal openstaande krediet van de klant, niet groter is dan de kredietlimiet van die bepaalde klant.

### **Kosten van slechte datakwaliteit**

Wanneer de kwaliteit van data te wensen overlaat, dan brengt dit kosten met zich mee. Denk bijvoorbeeld aan kosten van herwerk, 'work-around' kosten, verminderde productiviteit, boetes voor het niet nakomen van wettelijke verplichtingen, kosten om het vertrouwen van ontevreden klanten terug te winnen, foute beslissingen gebaseerd op incorrecte data en ga zo maar door. De toegenomen kosten kunnen worden verdeeld in de volgende categorieën:

- Impact op tijd.
- Impact op cashflow (boetes, late betalingen, gemiste kansen).
- Impact op materialen/faciliteiten/producten.

Hoe de kosten van slechte datakwaliteit berekend kunnen worden, is niet altijd vanzelfsprekend. Er zijn toch enkele principes die je kan volgen:

- Groepeer gegevens volgens business processen.
- Geef een gewicht aan deze processen (high/medium/low impact).
- Meet de directe kosten van foute gegevens (de TDABC oplossing van In Summa maakt dit inzichtelijk).
- Meet de gemiste kansen door fouten in het proces.
- Identificeer en kwantificeer de risico's.

### **Meten van datakwaliteit?**

Om de kwaliteit van data te meten is het vooral belangrijk eerst te weten

welke onderdelen van de gegevens relevant zijn. Hoe beter de kwaliteit wordt gemeten, hoe meer de onderliggende problemen naar boven zullen komen en hoe groter de impact van herstel zal zijn.

### **Volledigheid**

Een eerste stap is om de volledigheid van de gegevens te onderzoeken. Dit lijkt een gemakkelijke taak, maar toch zijn er een paar valkuilen waarmee rekening gehouden moet worden:

- Is het gemeten data element relevant?
- Zijn er standaardwaarden die in bepaalde systemen gebruikt worden?
- Kan het zijn dat in bepaalde gevallen het normaal is dat een data element niet ingevuld is?
- Indien een volledig record ontbreekt, zijn er eventueel andere databronnen binnen het bedrijf die wel een volledig beeld kunnen geven (data integratie)?

### **Accuraat**

Indien de gegevens compleet zijn, kunnen we gaan meten hoe accuraat de data is. Belangrijk hierbij is dat de data gevalideerd wordt op het moment van ingeven; komen de gegevens overeen met de realiteit? De reden dat deze stap vaak wordt overgeslagen is omdat dit vrij intensief is. 'Waar mensen werken, worden fouten gemaakt'. Vaak kan alleen aan de hand van audits bepaald worden hoe accuraat de gegevens zijn. Er zijn ook statistische methodes die ondersteuning kunnen bieden (bijvoorbeeld bepalen van outliers in een Control Chart).

### **Validiteit**

Nadat de accuraatheid van de meting of ingeving van de gegevens gevalideerd is, kunnen we de validiteit gaan meten. De validiteit

wordt bepaald aan de hand van business rules. Een business rule is een specificatie waaraan de gegevens moeten voldoen om te worden aangezien als correct. Hoe beter de regels gedefinieerd zijn, hoe beter de validatie kan plaatsvinden. Regels vormen de ruggengraat van datakwaliteit en het onderhoud ervan is doorslaggevend om de continuïteit van de datakwaliteit te kunnen garanderen.

### **Eigenaarschap**

Een laatste belangrijke voorwaarde voor datakwaliteit is helder eigenaarschap. Het moet duidelijk zijn wie aangesproken kan worden op de kwaliteit van de data. Zonder aangewezen verantwoordelijkheid heeft het verbeteren van datakwaliteit geen kans van slagen.

### **Verbeteren van datakwaliteit**

Zodra voldaan is aan de voorwaarden van datakwaliteit, is het tijd om datakwaliteit structureel in de organisatie te gaan verankeren en te gaan verbeteren. Een van de fundamentele kenmerken van business intelligence is het beschikbaar stellen van de juiste data, op het juiste moment bij de juiste persoon in het juiste formaat. Business intelligence start dus met het beschikbaar stellen van data. Dit is waar de verbetering van datakwaliteit begint. Door data beschikbaar te stellen, kunnen we op zoek gaan naar foutieve data (root-cause analyse). Veel organisaties hebben de neiging om het dataprobleem later in het proces op te lossen. Dweilen met de kraan open! De gouden regel is: hoe eerder je het probleem oplost in de datastroom, hoe lager de gerelateerde kosten. In bijna alle gevallen doet het probleem van slechte datakwaliteit zich voor in andere delen van de organisatie dan de afdeling waar het initiatief voor het verbeteren van de kwaliteit van data is ontstaan. Daarom is ondersteuning van senior management noodzakelijk om het succes van het verbeteren van datakwaliteit te garanderen.

# Tien basistips voor het verhogen en waarborgen van datakwaliteit

Bron: Cmotions | [cmotions.nl](https://cmotions.nl)

Praktijk en theorie rondom data management hebben zich de afgelopen twintig jaar enorm ontwikkeld. Basisrandvoorwaarden zoals actualiteit, validiteit en compleetheid hebben inmiddels plaats moeten maken voor integraal data management en data governance. Althans, zo lijkt het. Niettemin blijft elementaire data hygiëne van groot belang. Met het gevaar van ‘open deuren’ daarom hieronder toch een aantal basistips voor het verhogen en waarborgen van datakwaliteit.

## **Snelle ontwikkelingen, veel data en meer mogelijkheden**

De wereld van nu ontwikkelt zich steeds sneller. Nieuwe technologieën maken mogelijk dat er veel meer informatie bij de klanten, leveranciers, toezichthouders of andere stakeholders komt en dat beslissingen steeds sneller genomen kunnen worden. Nieuwe technologieën maken het ook mogelijk om informatie hierover vast te leggen, in gestructureerde of ongestructureerde vorm. Bovendien kan data ons helpen in te spelen op de (on)gevraagde en voorspelde behoeften van klanten. Door zo de wereld voorspelbaarder te maken, kan de organisatie marktaandeel behouden en vergroten.

## **Kwaliteit van data is cruciaal**

De kwaliteit van data, in welke vorm dan ook, is hierbij van cruciaal belang. Als je als organisatie data zou willen verzamelen en verwerken, dan is het belangrijk dat data aan kwaliteitsstandaarden voldoet. Denk hierbij aan volledigheid, eenduidigheid, tijdigheid, accuraatheid, consistentie en relevantie. Deze standaarden houden je data immers ‘fit for use’.

## **Tien tips voor continue verbetering van je datakwaliteit**

Waar moet je aan denken als je de datakwaliteit continu wilt verbeteren?

We geven 10 tips:

### **Maak data belangrijk**

Is data een asset in je organisatie of iets wat noodzakelijk is voor de administratieve doeleinden? Het opstellen en monitoren van doelen voor dataverzameling en datagebruik is de eerste stap richting datagedreven werken. En het inbedden ervan in de cultuur waarin beslissingen genomen worden op basis van relevante data, in plaats van op het buikgevoel.

### **Leg alleen relevante data vast**

Ook data heeft zijn levenscyclus. Het definiëren van welke databronnen en welke data vastgelegd dient te worden, maakt je datalandschap overzichtelijk. Bepaal hierbij welke data relevant is en welke data niet. En welke data gewenste toegevoegde waarde heeft voor je primaire en ondersteunende bedrijfsprocessen, zodat het de moeite loont om de data op te slaan en te onderhouden. Wanneer je over een langere periode data verzamelt, verwijder dan data die niet meer relevant is. Wet- en regelgeving kan hierbij ook een essentiële rol spelen, door bewaartermijnen of doelbinding te definiëren.

### **Dwing invoerstandaarden technisch af**

Veel vrijheid laten aan de kant van invoer leidt tot vervuilde en/of onvolledige data. Dwing daarom technisch af dat data zo veel mogelijk hetzelfde wordt ingevoerd. Wanneer je dit niet doet of niet kunt, zul je additionele technieken (zoals tekstmining) nodig hebben om de data te structureren en de kracht van je data te benutten.

### **Stel duidelijke business rules op voor 'extern' verkregen data**

Het verzamelen van data vindt niet alleen intern plaats. Data kan

ook afkomstig zijn van externe leveranciers. Om hoge datakwaliteit te realiseren, is het belangrijk om standaarden voor uitwisseling af te spreken over formaten, volumes of tijdsperiodes. Bij verwerking van nieuwe data moet de gestandaardiseerde invoer technisch afgedwongen worden.

### **Controleer invoer door het vier-ogen principe**

Je wilt het liefst alleen data bewaren die accuraat is. Om dubbele invoer van klantgegevens te voorkomen, kun je intern een vier-ogen principe toepassen. Laat degenen die verantwoordelijk zijn voor je data akkoord geven op de invoer. Een andere mogelijkheid is om je klant, leverancier of medewerker de acceptatie van invoer en/of eventuele wijzigingen in hun gegevens laten bevestigen.

### **Gebruik contact met je stakeholders om data te actualiseren**

Tijdigheid (of recentheid) is belangrijk als we data willen toepassen om een beeld te krijgen van het moment nu of van het moment in de toekomst. Gebruik bijvoorbeeld een contactmoment met je klant op het callcenter of overleg met product manager om de bestaande data te bevestigen of te wijzigen.

### **Monitor de datakwaliteit**

Stel duidelijke KPI's aan je datakwaliteit en monitor deze door onder andere bounces en retouren te meten. Zet rapportages in via relevante data quality tooling die over deze functionaliteit beschikt.

### **Maak iemand verantwoordelijk**

Een eenmalige data audit kan bestaande issues blootleggen en verbeteracties opleveren. Maar ondertussen stroomt er weer nieuwe data binnen. Om het proces van datakwaliteit continu te bewaken, verdient het de voorkeur hier iemand verantwoordelijk voor te maken. Zo wordt de kwaliteit onderdeel van de kerntaak van die persoon en niet 'een taak erbij'. Zo krijgt het de aandacht die het



verdient. Leg daarbij de verantwoordelijkheid neer op die plek die het meeste directe voordeel heeft van een goede datakwaliteit. Meestal is dit de ‘business kant’ van de organisatie. Ook bij het uitbreiden van (master) data is het de verantwoordelijkheid van de business om de informatie vast te leggen en niet die van de IT afdeling.

### **Doe een periodieke audit met een externe partij**

Gebruik regelmatig frisse blik op zaken. Plan daarom een periodieke audit in. Laat je inspireren door nieuwe inzichten op het gebied van data, wanneer je dit door een externe partij laat doen.

### **Houd focus en wees vasthoudend**

Constant aandacht geven aan je belangrijke asset (data!) werpt zijn vruchten af. Je hoeft nooit meer te staan voor de grote stapels vervuiling. Op de lange termijn is het een kwestie van overzicht hebben van mogelijke issues en wijzigingen bijhouden. Een schone lei maakt ruimte voor nieuwe inspirerende ideeën over wat je met data allemaal kunt.

# Wanneer is data van goede kwaliteit?

Bron: SAP Nederland | [blogs.sap.nl](https://blogs.sap.nl)

De kwaliteit van data laat bij veel bedrijven te wensen over. Dit terwijl deze data juist van groot belang is om de datastrategie van bedrijven te realiseren.

Binnen veel bedrijven wordt geëxperimenteerd met algoritmes, datamodellen en analytische tools. In veel gevallen wordt gewerkt met datasets die speciaal voor dit doeleinden zijn geprepareerd, onder meer door data op te schonen en correct te labelen. Algoritmes, datamodellen en analytische tools kunnen perfect overweg met deze geprepareerde datasets en hiermee de beperkte tactische doelen realiseren die voor experimenten worden gesteld.

## **Meeste data is verre van perfect**

In realiteit is de data waarover het gemiddelde bedrijf beschikt echter verre van perfect. Veel gegevens worden verzameld zonder hierbij een specifiek doel in gedachte te hebben, waardoor de kwaliteit te wensen overlaat. Zo zijn datasets vervuild, gegevens verkeerd gelabeld of in silo's ondergebracht, of zelfs de verkeerde datasets verzameld. Datasets waarover een organisatie beschikt, zijn vaak dan ook niet geschikt om de datastrategie op de lange termijn uit te voeren.

Een datastrategie is de visie of de uitkomstgerichte en domeinspecifieke roadmap van een organisatie om het volledige potentieel van hun data te benutten. Vaak experimenteren bedrijven met nieuwe technologieën, zoals AI, zonder vooraf te weten wat zij met alle data die dit oplevert moeten doen. Hoe moeten deze gegevens bijvoorbeeld georganiseerd, beheerd en gestandaardiseerd worden? Hoe wordt de kwaliteit van data gewaarborgd?

Een goede datastrategie adresseert deze vragen op een holistische manier. De strategie helpt onder meer taken te prioriteren, realistische verwachtingen te scheppen en een meerjarige roadmap op te stellen.

### **Je meest waardevolle strategische bezit**

In het tijdperk van grootschalige dataverzameling is je data het meest kritieke en waardevolle strategische bezit. Zo maken data science en analytics het mogelijk nieuwe inzichten te verkrijgen, innovatie te realiseren en waarde te creëren. Om van de voordelen van data te kunnen profiteren moet de datastrategie gebaseerd zijn op schone data van hoge kwaliteit die correct is gelabeld. Het gaat om deelbare data, die toegankelijk is voor machines en goed gedocumenteerd is met behulp van metadata en een data dictionary. De data is geïnventariseerd, zodat vastgesteld kan worden welke essentiële datasets ontbreken en kan aangevuld worden.

Een goede datastrategie op basis van schone data van hoge kwaliteit is van cruciaal belang voor het realiseren van je bedrijfsmissie, behalen van strategisch succes en creëren van een voorsprong op de concurrentie. Is de kwaliteit van jouw data al op orde?

# Belang van datakwaliteit

Bron: Digital Power | [digital-power.com](https://digital-power.com)

Ga je beslissingen maken op basis van data? Dan moet je ervoor zorgen dat je datakwaliteit op orde is. Goede documentatie volgens een helder proces is hierbij essentieel. Over het waarom en hoe lees je in dit artikel.

## **Verbeteringen in de zorg met behulp van data**

Het UMC Utrecht is sinds enige tijd aan het experimenteren met data science in de psychiatrie. Met data science zoeken ze een manier om de onvoorspelbare psychiatrische patiënt voorspelbaarder te maken.

Het ziekenhuis gebruikt onder andere persoonlijke- en gevoelige data afkomstig uit wearables, sociale data en input van de patiënt en naaste familieleden met onder andere woordgebruik-analyses. Zo kan het algoritme verrijkt worden met contextuele data die veel kan vertellen over mogelijke triggers bij psychiatrische patiënten.

De genoemde voordelen zijn legio: betere én snellere diagnose, betere aansluiting met medicatie en preventieve zorg kunnen verlenen. Maar ook minder administratieve handelingen voor artsen en verplegend personeel, met mogelijk minder benodigd zorgpersoneel, als antwoord op het personeelstekort in de zorg. Enorme verbeteringen, op voorwaarde dat bij het gebruik van dit type data aan de privacywetgeving (AVG) wordt voldaan.

## **Maar wat als er foutjes in de data zitten?**

Het experiment van het UMC Utrecht is een sprekend voorbeeld. En dit is niet de enige zorginstelling die als doel heeft met de best mogelijke zorg de waardigheid van een individu zo lang mogelijk intact te houden. Maar wat als er kleine fouten in de data sluipen? Historische data is mogelijk subjectief. Data input is mogelijk onvolledig en mist relevante data vanuit andere zorginstellingen en overheidsinstanties (overdracht). Data definities zijn

mogelijk verschillend per bron. Data wordt mogelijk verkeerd geïnterpreteerd.

Fouten in de data kunnen, hoe klein ook, potentieel desastreuze gevolgen hebben voor het welzijn van patiënten. Bij een verkeerde diagnose kan het ziektebeeld verder verergeren of kunnen zich andere stoornissen ontwikkelen. Daardoor loopt de kwaliteit van leven voor deze patiënten zover terug dat het doel, goede zorg, een averechts effect sorteert. Het algoritme vergroot niet de menselijke waardigheid, het reduceert het. De potentie alleen is om bij de huiveren. Dit is een scenario dat de EU met de AVG juist zo hard probeert te voorkomen. Daarom wordt gesteld dat transparantie noodzakelijk is voor organisaties die zich bezighouden met het verzamelen en toepassen van data.

### **Focus op datakwaliteit**

Als je als organisatie weet dat slechte, of zelfs maar iets minder goede, datakwaliteit verstrekende negatieve gevolgen kan hebben voor je uiteindelijke afnemer (in dit geval de patiënt), zul je er alles aan willen doen om dit te voorkomen. Er moeten protocollen ingebouwd worden die de kwaliteit waarborgen.

### **Documenteren van data, een cruciale taak**

Om je datakwaliteit te borgen, moet je heldere processen en duidelijke protocollen hebben binnen je organisatie. In de praktijk is het reproduceerbaar documenteren van data vaak de laatste en minst dankbare stap in een proces. Doordat op tempo gewerkt wordt, is dit zelfs de stap die vaak onvolledig wordt uitgevoerd of geheel wordt overgeslagen. Maar wat betekent dat eigenlijk? Ter illustratie zie je hieronder drie verschillende typen resultaten:

Proces stappen	Bouw	Health Analytics	Customer Analytics
Grondstof	Bouwmateriaal	Patiënt data	Klantdata
Ontwerp	Schetsonwerp/ maquette	Gewenste inzichten = requirements	Gewenste inzichten = requirements
Instructie	Bouwbestek	Meetplan = documentatie	Meetplan = documentatie
Bouwproces	Constructie	Verwerking patiënt data	Verwerking klant data
Resultaat	Huis	Diagnose	Gepersonaliseerd aanbod

Als je goed bouwt, heb je een huis. En op basis van betrouwbare patiëntdata kan een arts de juiste diagnose stellen. Met behulp van klantdata kun je een gepersonaliseerd aanbod bieden.

Maar wat als er in de proces stappen een foutje sluipt? Het resultaat ziet er dan een stuk minder rooskleurig uit. Is het eindresultaat dan nog wel betrouwbaar genoeg? Hieronder zie je wat er kan veranderen als de data niet helemaal klopt:

Proces stappen	Bouw	Health Analytics	Customer Analytics
Grondstof fout	Bouwmateriaal	Patiënt data	Klantdata
Ontwerp fout	Schetsonwerp/ maquette	Gewenste inzichten = requirements	Gewenste inzichten = requirements
Instructie fout	Bouwbestek	Meetplan = documentatie	Meetplan = documentatie
Bouwproces fout	Constructie	Verwerking patiënt data	Verwerking klant data
Resultaat	Huis, maar kan je er wel in wonen?	Diagnose, maar voelt de patiënt zich beter?	Gepersonaliseerd aanbod, maar klopt het aanbod wel?

Van de eerste twee voorbeelden kunnen we ons allemaal een levendige voorstelling maken van de te nemen stappen en het resultaat. In het voorbeeld van het UMC zou de impact groot en negatief zijn als de proces

stappen niet secuur en doordacht doorlopen zijn. Het zou nog ziekere patiënten opleveren door toedoen van onjuiste input of verwerking.

In het voorbeeld van de bouw zou een misrekening van een overspanning tussen twee dragende muren een dramatische afloop kunnen hebben. Als het huis niet exact volgens ontwerp en constructie gebouwd wordt, is het geen huis, maar een aaneenschakeling van ramen, stopcontacten en muren zonder enig plan. Is het dan nog wel een huis?

Het laatste resultaat is net zo pijnlijk, maar minder goed zichtbaar. Het is veelal een geaccepteerde foutmarge en daarbij lijken de resultaten redelijk overeen te komen met wat men voor ogen had. Maar als het ontwerp en de instructie een fout bevatten, heeft dat effect op het resultaat. Stel jezelf de volgende vragen:

- Is het minder erg als er budget wordt besteed aan initiatieven en ontwikkelingen die er onderaan de streep niet toe doen?
- Is het minder erg als het niet reproduceerbaar is waarom, hoe en wanneer data tot stand is gekomen?
- Is het minder erg als experimenten, qua data input, zo onbetrouwbaar zijn dat een gok in de wilde weg waarschijnlijk een even groot effect zou sorteren?

De kans is groot dat zal iedere zichzelf respecterende organisatie hier absoluut niet op zit te wachten.

### **Er is teveel focus op zichtbare resultaten**

Toch sluipt het er met lean, agile en scrum werkmethodes snel in. Waarbij waardering vooral tot uiting komt bij het opleveren van zichtbare resultaten. Dat is waar de dankbaarheid vanuit stakeholders in huist. Zij willen resultaten zien, we doen het immers voor hen. Als zij blij zijn, wat is dan het probleem?

Naast het feit dat dit niet slim is en het je potentieel veel geld kost, is het onethisch. Je levert namelijk een ondeugdelijk resultaat. Niet alleen voor jezelf en je collega's (denk aan het extra werk dat het kost om alles te cross checken en work-arounds te bedenken), maar ook voor je organisatie. Deze investeert immers in deze vruchteloze tijdsbesteding.

Ook de eindgebruiker, je klant, wordt benadeeld. Deze wordt namelijk opgezadeld met een resultaat waarbij hij of zij niet centraal staat. Zo wordt de eindgebruiker bijvoorbeeld in een segment ingedeeld waar hij of zij niet thuishoort en matcht de ontvangen service niet (optimaal) met zijn of haar klantgedrag.

### **Hoe documenteren dan wél moet**

Het volgen van een aantal heldere documentatiestappen kan op termijn veel frustratie en vruchteloze tijdsbesteding schelen. Het kost nu moeite om er later de vruchten van te kunnen plukken. Goede documentatie en heldere data processen creëren organisatiebreed heldere definities over metrics, gebruikte regels en de context waarin het gebruikt dient te worden. Het zorgt ervoor dat alle afdelingen dezelfde taal spreken en elkaar begrijpen. Hierdoor zijn KPI's (key performance indicators, targets en operationele actiepunten) logisch en goed te onderbouwen voor alle lagen van de organisatie die met de data werken. Het biedt een fundament waar je samen op kunt voortbouwen.

### **Documenteren: een stappenplan**

Ben je overtuigd van de noodzaak van goede documentatie? Dit stappenplan helpt je op weg om de documentatie als procesonderdeel mee te nemen in je workflow.





### **Business requirements**

Dit beschrijft een gewenst resultaat en komt voort uit wat de business stakeholders moeten weten. De benodigde informatie voor business stakeholders dient in lijn te zijn met de organisatie- en afdelingsstrategie en -doelstellingen. Het is een fase die héél veel afstemming vergt met alle stakeholders, om uiteindelijk tot gezamenlijk gedragen business requirements te komen. Besteed hier het grootste deel van je tijd aan.

### **Dataaag documentatie**

Dit vormt het ontwerp én de instructie van variabelen die vanaf de website dataaag uitgelezen moeten worden om vervolgens gemeten te worden in dimensies en/of metrics van de meetoplossing.

### **Meetoplossing documentatie**

Dit vormt het ontwerp én de instructie van de meetoplossing. Het beschrijft met welke dimensies, variabelen, metrics, business rules én definities de benodigde meting voor de business requirements opgezet moet worden.

### **Beschrijving analytics model**

Dit beschrijft in heldere termen de definitieve meetoplossing. Welke dimensies, variabelen, metrics, business rules en definities zijn

toegepast om de meting tot stand te brengen? Belangrijk is dat hier heldere termen worden gebruikt. Het vormt hét naslagwerk dat door de organisatie geraadpleegd moet worden.

### **Analyse voorbeelden**

Dit geeft ondubbelzinnig weer welke type rapportages en analyses gedaan kunnen worden met dit analytics model.

### **Loop je achter qua documentatie?**

Het kan zijn dat je met achterstallig onderhoud van de documentatie te maken hebt. In dat geval kun je net als bij bug fixes, een groot aantal uren per sprint toekennen aan het documenteren met terugwerkende kracht: start met rondes langs de velden met de checkvraag of de business requirements en huidige metingen nog wel op elkaar aansluiten. Met de uitkomsten daarvan ga je als volgt te werk:

- Start met het goede nieuws en documenteer alle metingen die nog steeds voldoen aan de business requirements, maar nog niet of onvolledig gedocumenteerd zijn.
- Vervolg met metingen waar een mismatch zit tussen business requirements en huidige meetoplossing en zet deze als Request For Change (RFC) op de backlog.
- Eindig met metingen die niet meer voorkomen in de business requirements. Deze zullen uitgefaseerd moeten worden.

# Kosten van slechte data

Bron: Tentive | [tentive.nl](https://tentive.nl)

Het is buitengewoon waardevol als je bij het nemen van strategische beslissingen als organisatie kan vertrouwen op kwalitatief goede data. Al blijft het lastig een financiële waarde toe te kennen aan de kwaliteit van de data. De belangrijkste reden hiervoor is dat organisaties het vaak digitale karakter van hun data lastig vinden te duiden. Dit in tegenstelling tot de traditionele, tastbare 'assets', zoals machines en gebouwen. Hiervan komt de financiële waarde wél tot uitdrukking op de balans. Daar wringt op dit moment nog de schoen. Onze ervaring is dat organisaties wél het belang van datakwaliteit onderschrijven, maar dat de bereidheid om hierin te investeren nog beperkt is. Waarom geldt uitgeven aan iets ongrijpbaars?

## **Data van slechte kwaliteit kost geld**

Organisaties trekken lering uit hun data en gebruiken nieuw vergaarde inzichten om toegevoegde waarde te kunnen leveren. Voorbeelden hiervan zijn het begrijpen van klantgewoonten om een product of dienst te verbeteren of het in kaart brengen van trends in de markt om een betere bedrijfsstrategie te ontwikkelen. Data van slechte kwaliteit heeft een negatieve invloed op deze beslissingen. Bovendien kost data van slechte kwaliteit elke organisatie geld (DAMA International, 2017).

Aan wat voor bedragen moet je dan denken? De schattingen lopen uiteen, maar in de regel kan worden gesteld dat organisaties een substantieel deel omzet mislopen als gevolg van problemen met de datakwaliteit. Ovum Research stelt dat inferieure data voor organisaties resulteert in een inkomstenderving van minstens 30% (Geiger, 2014). Slechte datakwaliteit vernietigt dus bedrijfswaarde. Onderzoek van Gartner (2018) wijst uit dat organisaties van mening zijn dat slechte datakwaliteit verantwoordelijk is voor een verlies van gemiddeld 15 miljoen dollar per jaar. IBM (2016) schat

dat de kosten van slechte datakwaliteit in de Verenigde Staten in 2016 3.1 triljoen dollar bedroegen.

### **Oorzaken**

Op opgelegde boetes na, zijn veel van de kosten die voortkomen uit slechte datakwaliteit verborgen, indirect en daardoor moeilijk meetbaar (Redman, 2016). Kosten komen onder andere voort uit:

#### **Gemiste kansen**

Een organisatie kan een unieke kans laten liggen om een nieuw product te ontwikkelen of in te spelen op hernieuwde klantbehoeften, die door een concurrent met meer volwassen datakennis wél wordt gekapitaliseerd.

#### **Reputatieschade**

Dit kan variëren van kleine, alledaagse schade tot een heuse PR-ramp. Zo kan slechte data bijvoorbeeld bij banken leiden tot het onbedoeld zaken doen met organisaties of groeperingen die op een zwarte lijst staan, alleen omdat de betreffende instelling onvoldoende accurate informatie heeft over de partijen waarmee wordt gehandeld. Naast torenhoge boetes opgelegd door toezichthouders, levert het vaak veel negatieve publiciteit op. Een zorgvuldig opgebouwde reputatie kan daarmee grotendeels om zeep worden geholpen.

#### **Verloren inkomsten**

Slechte data kan op verschillende manier leiden tot verloren inkomsten. Bijvoorbeeld mailings die niet worden omgezet in verkoop, omdat onderliggende klantgegevens niet correct zijn.

### **Voordelen van hoge datakwaliteit**

Hoog tijd daarom om als organisatie de focus te leggen op het verbeteren

van de datakwaliteit. De voordelen van goede datakwaliteit zijn onder meer te behalen op het terrein van:

### **Besluitvorming**

Hoe beter de datakwaliteit, des te meer vertrouwen medewerkers hebben in de resultaten die ze voortbrengen, waardoor het risico in de uitkomsten afneemt en de efficiëntie toeneemt. 'Garbage in, garbage out' geldt ook hier. Het omgekeerde geldt gelukkig ook. Wanneer resultaten betrouwbaar zijn, kunnen giswerk en risico's bij de besluitvorming worden beperkt.

### **Productiviteit**

Data van goede kwaliteit maakt medewerkers productiever. In plaats van tijd te besteden aan het valideren en herstellen van datafouten, kunnen zij zich richten op hun kerntaken.

### **Naleving**

In sectoren waarin de wetgever bepaalt hoe relaties en/of bedrijfsactiviteiten met klanten mag plaatsvinden, kan goede datakwaliteit het verschil zijn tussen naleving (compliance) enerzijds en miljoenen euro's aan boetes anderzijds.

### **Marketing**

Betere data maakt nauwkeurige targetting en effectieve klantcommunicatie mogelijk, vooral in omni-channel omgevingen waar veel organisaties naar streven.

### **Concurrentievoordeel**

Goede datakwaliteit kan leiden tot een concurrentievoordeel doordat een organisatie betere inzicht krijgt over klanten, producten, processen en het sneller marktkansen kan waarnemen.

**Tot slot**

Bovengenoemde kosten en voordelen impliceren dat het beheer van datakwaliteit geen eenmalige activiteit is, maar een voortdurend proces. Data van hoge kwaliteit vereist planning, betrokkenheid en een mindset om die kwaliteit blijvend te borgen in de processen en systemen. Het verbeteren van datakwaliteit zal dus niet gaan zonder de nodige inspanningen, maar uiteindelijk plukt de organisatie er wel de vruchten van.

# Wat is datapreparatie en waarom is het zo belangrijk?

Bron: Totta data lab | [tottadatalab.nl](http://tottadatalab.nl)

Data scientisten zijn gemiddeld zo'n 80 procent van hun tijd kwijt aan datapreparatie. Daarnaast vinden de meeste data scientisten datapreparatie het minst leuke onderdeel van hun baan. Toch zal datapreparatie voorlopig niet verdwijnen. Integendeel. Datapreparatie komt bij elk project wel om de hoek kijken. En laten er nou steeds meer projecten bijkomen, aangezien de hoeveelheid data wereldwijd nog dagelijks groeit. Maar er is goed nieuws voor data scientisten: datapreparatie wordt langzaamaan eenvoudiger. Hierdoor is de kans groot dat zij hier straks minder tijd aan kwijt zijn. Sterker nog: door de komst van selfservice tools hoeft je straks geen data scientist meer te zijn om data te prepareren. Maar waar komt datapreparatie in essentie eigenlijk op neer? En waarom is het zo belangrijk?

## Wat is datapreparatie?

Datapreparatie komt neer op het opschonen en klaarstomen van zogenaamde 'raw data', ook wel ruwe, pure data genoemd. Deze data kan afkomstig zijn van de meest uiteenlopende databronnen. Ruwe data is direct afkomstig van een bron, dus nog volledig onaangetast. Deze data wordt uiteindelijk omgezet naar data die direct en accuraat geanalyseerd kan worden, bijvoorbeeld voor businessdoeleinden. Het 'omzetten' van pure data naar analyseerbare data is echter geen eenvoudig proces. Als je de data niet goed prepareert, dan loop je bij de uiteindelijke data analyse tegen fouten en oneffenheden aan, die vaak lastig terug te draaien zijn.

## Waarom is datapreparatie belangrijk?

Alleen op basis van volledig accurate data kan een business beslissingen nemen. Dit houdt in dat de data van hoge kwaliteit moet zijn en geen

onnodige fouten mag bevatten. Daarnaast kun je betere beslissingen nemen voor je business als je data van betere kwaliteit is. In dit geval kun je namelijk sneller en makkelijker data analyseren en daaropvolgend snellere en efficiëntere beslissingen nemen. Maar voordat je dit kunt doen, moet de data eerst flink onder handen worden genomen. Tijdens dit opschoonproces kun je tegen 2 soorten fouten aanlopen:

### **Systematische fouten**

Deze fouten komen vaak terug bij grote hoeveelheden data. Aangezien grote hoeveelheden data vaak van allerlei bronnen afkomstig zijn, is de kans op reeds aanwezige fouten – zoals missing values en outliers – groter.

### **Individuele fouten**

Deze fouten zijn gemaakt tijdens de invoer van de data(sets) in bijvoorbeeld een datacatalogus. Dit zijn dus geen fouten die oorspronkelijk al aanwezig waren. Aangezien deze fouten alleen bij de invoer zijn gemaakt, is de impact op de rest van de data minimaal.

Als je deze fouten eenmaal hebt opgespoord tijdens de eerste stap(pen) van datapreparatie, dan dien je ze eigenlijk direct op te lossen. Want als je eenmaal een aantal stappen verder bent, dan is het een stuk lastiger om deze eerder gemaakte fouten te herkennen en op te lossen.

### **Datapreparatie volgens de 5 D's**

De specifieke stappen die je moet doorlopen om data te prepareren, verschillen per industrie en organisatie. Toch is er een algemeen toepasbaar stappenplan: het stappenplan van de 5 D's.

#### **Discover**

De meest geschikte data vinden voor een bepaald doel. Een accurate, goed gedocumenteerde datacatalogus creëren is hierbij



essentieel. Hier sla je al je data(bronnen) op, zodat je duidelijk op een rijtje hebt hoe je verzamelde datasets eruitzien en waar deze vandaan komen. Data kan afkomstig zijn uit bestaande datasets of ad hoc worden toegevoegd. Dit laatste komt neer op het eenmalig toevoegen van een bepaalde dataset, aangepast aan de situatie.

### **Detain**

Data selecteren uit je datacatalogus. Je duikt nu echt diep in je data. Je moet erachter zien te komen op welke manier welke data waardevol kan zijn voor een bepaald doeleinde.

### **Distill**

Het opschonen en valideren van data. Het is cruciaal om foutieve data te verwijderen, ontbrekende data aan te vullen en datapoorten te beveiligen. Na deze 'cleanup' moet de data gevalideerd worden door te testen of de data fouten bevat. Hier komt vaak nog meer aan het licht dan men in eerste instantie dacht.

### **Document**

Het verrijken van data. Je dient het format waarin de data zich bevindt, te transformeren naar een up-to-date en begrijpelijk format. Denk hierbij aan het toevoegen van technische definities, metadata, business-termen, etc. Alleen op deze manier kan een goed gedefinieerde uitkomst tot stand komen.

### **Deliver**

De data op de juiste manier opslaan en delen met de eindgebruikers. Dit doe je in bijvoorbeeld een business intelligence tool. Hierdoor kunnen eindgebruikers direct aan de slag met de data.

## **Snel, makkelijk en efficiënt**

Idealiter is datapreparatie een proces dat zichzelf continu herhaalt. Zo maken gedeelde (meta)datasets, gecontroleerde opslag en herbruikbare

transformeer- en opschoontechnieken, datapreparatie efficiënt, consistent en voor herhaling vatbaar. Op deze manier wordt het eenvoudiger voor eindgebruikers om zelf met data aan de slag te gaan. Sterker nog, selfservice datapreparatietools zijn in opkomst, waardoor je straks geen data scientist meer hoeft te zijn om data te prepareren.

# Afwijkende transactiecijfers: wat zijn dat en waar komen ze vandaan?

Bron: Digital Power | [digital-power.com](https://digital-power.com)

Over het aantal transacties dat in een week heeft plaatsgevonden, kan onduidelijkheid bestaan. Zo kan backoffice 12.000 zeggen, het data warehouse 15.000 aangeven en Google Analytics op 11.000 uitkomen. Hoe kan dat? Als verschillende systemen, zoals backoffice systemen, data warehouses en web analytics tools, dezelfde transacties registreren en rapporteren, zou je verwachten dat overal hetzelfde aantal bestellingen wordt geregistreerd. In de praktijk komen de cijfers vaak niet overeen. Waar komen deze verschillen vandaan?

Afwijkingen in het meten van het aantal bestellingen kunnen verschillende oorzaken hebben. De meeste zijn logisch, sommige zijn het gevolg van beperkingen in meetinstrumenten en af en toe kan er een meetfout gemaakt worden. Wanneer cijfers per systeem verschillen, kan het goed dat je het vertrouwen in je systemen verliest. Maar wees gerust: vaak ligt de oplossing in het nalezen van de definities van een bestelling in verschillende systemen. Daarnaast is, afhankelijk van het systeem dat je gebruikt, een klein verschil normaal. Bovendien geven aantallen bestellingen zelfs met een kleine afwijking nog altijd inzicht in trends.

## Wat is afwijking?

Afwijking betekent dat systeem A en systeem B afwijkende getallen rapporteren. Vragen die we regelmatig horen, zijn:

- Web analytics zegt dat er 10.000 bestellingen waren, maar volgens onze backoffice zijn het er 15.000. Hoe kan dat?

- We zien een afwijking tussen verschillende systemen van meer dan 20%. Is dat normaal?

### **Waar komen afwijkende transactiecijfers vandaan?**

Afwijkingen in transactiecijfers kunnen verschillende oorzaken hebben.

Deze verdelen we onder in 4 categorieën:

### **Verschillen in transactiemethodes**

#### **Transacties vanaf andere kanalen**

Web analytics meet alleen transacties op de website, terwijl een ander systeem ook callcenter transacties of transacties in fysieke winkels registreert.

#### **Geannuleerde transacties**

Gebruikers plaatsen een bestelling, maar voldoen niet aan de acceptatiecriteria of annuleren hun bestelling. Deze bestelling wordt dat wel verwijderd uit het ene systeem, maar niet uit het andere.

#### **Testtransacties**

Testtransacties kunnen geregistreerd worden in het ene systeem maar niet in het andere.

#### **Onduidelijke funnels of definities**

Soms is het niet duidelijk wat wel en niet telt als een transactie. Denk hierbij bijvoorbeeld aan het plaatsen van een bestelling namens een klant, een sub-bestelling of een herhaalde bestelling.

### **Instellingen voor opslag (vanuit web analytics)**

#### **Cookie-instellingen (vooral sinds de invoering AVG)**

Wanneer websites gebruikmaken van een 'alles-of-niets cookie-aanbod', kan het voorkomen dat ook het gedrag van gebruikers op

een website helemaal niet geregistreerd wordt. Let wel: dit komt niet vaak voor.

### **Data blokkering/omleiding**

Web analytics data wordt soms geblokkeerd, omgeleid of gelabeld voor bepaalde IP-adressen, bijvoorbeeld die van de eigen organisatie. Dit is meestal zichtbaar in de gebruikersinterface in een analytics tool. Het kan voorkomen dat het onopgemerkt blijft zoals in Adobe's VISTA.

## **Gebruikersvoorkeuren en -instellingen**

### **Adblocker**

Sommige gebruikers installeren adblockers in hun internetbrowsers. Veelgebruikte adblockers blokkeren soms tag managers, waardoor tracking en/of web analytics software niet meer geladen kan worden. Dit moet wel handmatig worden ingesteld door gebruikers. Daarbij kan het zijn dat adblockers veelgebruikte functionaliteiten van web analytics tools uit zichzelf blokkeren.

### **Geblokkeerd door browsers**

Intelligent tracking prevention (ITP) door Safari blokkeert bepaalde cookies en easy referrer acces. Daarnaast bieden andere browsers tegenwoordig ook privacy functionaliteiten aan, zoals anonieme browservensters.

### **Mobiele browsers met data besparingsfunctionaliteiten**

Om dataverbruik te beperken op mobiele telefoons, blokkeren sommige mobiele browsers tracking functionaliteiten.

### **Trage internetverbinding**

Als een gebruiker een trage internetverbinding heeft, worden sommige tags niet geladen of scripts geblokkeerd.

## **Functionele afwijkingen door technische fouten**

### **Implementatiefouten**

Wanneer je nieuwe functionaliteiten bouwt, moeten deze naadloos geïmplementeerd worden binnen het bestaande systeem. Dit kan een uitdaging zijn, omdat een klein foutje al grote gevolgen kan hebben. Zo kan het gebeuren dat bij een fout in de funnel een conversiepagina niet wordt geladen en daardoor een transactie niet wordt opgeslagen.

### **Wanneer je in het ene systeem meer transacties ziet dan in het andere**

Is de transactiedata in de verschillende systemen die je gebruikt niet consistent? Kijk dan nog eens goed naar je definitie van een transactie. Is deze in elk systeem gelijk?

# Management van datakwaliteit: zo pak je dat aan

Bron: Fourpoints | [fourpoints.nl](https://fourpoints.nl)

Problemen in datakwaliteit worden vaak geassocieerd met grote bedrijven. In de praktijk blijkt dit net zo relevant voor het mkb te zijn. Datakwaliteit wordt hier vaak gezien als een nice to have, in plaats van een must have. Hierdoor worden de processen minder beheerst, of er is helemaal niemand vanuit IT of marketing die zicht houdt op het beheer van de data. Hoewel het al een stap is om data te verzamelen, is management van de kwaliteit van de data een minstens zo belangrijke stap.

## Verschillende gegevens

Datakwaliteit is vaak een ondergeschoven kindje in het mkb. Jammer, want juist met goede data kun je veel. Ook in het mkb komen heel veel gegevens binnen. Denk hierbij aan je CRM-, contactbeheer- en marketingsystemen die allemaal verschillende bronnen gebruiken en gegevens verzamelen. Als gegevens slecht, inconsistent of onnauwkeurig zijn, kunnen deze processen verstoren of - erger nog - de verkeerde communicatie naar de verkeerde mensen sturen. Dit kan reputatieschade of problemen met regelgevende instanties veroorzaken.

## Wat gebeurt er als je data van goede kwaliteit hebt?

Het ordenen en bijhouden van de data is een uitvoerig proces en kost tijd. Het is dus goed om te weten waarvoor je het doet:

- Je kunt de verkoopprocessen verbeteren, nauwkeurigere en meer gerichte marketing produceren, de klanttevredenheid verhogen en de communicatiekosten naar klanten en kansen laag houden.

- Je kunt marketing automatiseren en workflows maken op basis van nauwkeurige segmenten, zoals klanten die een bepaald product of een bepaalde dienst hebben gekocht. Hierdoor komen er mogelijkheden voor extra verkoop. En dat is zeker belangrijk: waar klanten vroeger bijna altijd trouw bleven aan bedrijven, is dat nu wel anders. Klanten kunnen meer informatie online vinden en besteden hun geld ook gemakkelijk aan de concurrent als ze ontevreden zijn over je product of dienst.
- Je kunt goede beslissingen nemen over de toekomst van je bedrijf, producten en diensten, ondersteund door betrouwbare informatie, doordat je eenvoudige dashboards kunt maken op basis van KPI's.

### **Orde op zaken**

Datakwaliteit is dus van belang om je bedrijf de goede kant op te sturen, dankzij beslissingen op basis van de juiste informatie, maar ook om je interne processen te verbeteren. Doordat steeds meer processen zich online afspelen, is ook het belang van deze gegevens toegenomen. Dit moet gepaard gaan met effectief beheer van de datakwaliteit. De gegevens moeten geschikt zijn voor het doel. Als je dit duidelijk hebt, kun je starten met het orde op zaken stellen.

Het is belangrijk om na te denken over een datastrategie en plek om alle gegevens op te slaan. Houd hierbij in het achterhoofd dat de gegevens relevant moeten zijn en alleen bewaard worden zolang het nodig zijn. De belangrijkste les hierin is dat datakwaliteit begint bij de bron. Dat is namelijk de goedkoopste vorm van datakwaliteitsbeheer die kan worden gecontroleerd en bewaakt. Na het opstellen van een datastrategie is het belangrijk dat er iemand in je organisatie verantwoordelijk wordt voor de naleving ervan. Want alleen op die manier kun je de kwaliteit van de data blijven waarborgen. Samengevat bestaat management van de datakwaliteit dus uit drie stappen:



- Stel doelen op.
- Stel een datastrategie op, inclusief de verschillende bronnen en processen.
- Maak iemand in de organisatie verantwoordelijk voor het waarborgen hiervan.

# Het belang van een data dictionary

Bron: Tentive | [tentive.nl](https://tentive.nl)

De eerste vermelding van data quality software is op zijn minst 20 jaar geleden. In die 20 jaar is er veel veranderd in de markt voor datakwaliteit. Waar datakwaliteit begon als een opzichzelfstaande oplossing, veranderde het langzaam in een commodity. Nu is het veelal een onderdeel binnen andere applicaties. Er zijn nog maar een paar opzichzelfstaande datakwaliteit oplossingen. Hoe komt het dan dat na 20 jaar datakwaliteit de meeste organisaties nog steeds problemen hebben met hun gegevenskwaliteit?

Mogelijk nog verbazingwekkender is dat het min of meer dezelfde soorten problemen zijn als 20 jaar geleden. Het is nogal ongemakkelijk om je te beseffen dat er in die tijd niet veel veranderd lijkt te zijn. We zouden zelfs kunnen concluderen dat wanneer we er in de afgelopen 20 jaar niet in geslaagd zijn de kwaliteit van de gegevens met behulp van tooling te verbeteren, de kans groot is dat wanneer we in de komende 20 jaar hetzelfde blijven doen, de resultaten evenmin overweldigend zullen zijn. Laten we eens een stap terug doen. Waarom hebben we eigenlijk problemen met de kwaliteit van onze gegevens? In de literatuur en verschillende onderzoeken zijn de meest genoemde redenen voor problemen met datakwaliteit:

- Gegevensinvoer door werknemers.
- Datamigratie projecten.
- Batch feeds.
- Gemengde verwachtingen van gebruikers.

## **Gegevensinvoer door werknemers**

Gegevensinvoer wordt steevast genoemd als belangrijkste reden van een slechte gegevenskwaliteit. Bij het invoeren van een nieuwe klant in een CRM-

of ERP systeem komt veel meer kijken dan alleen het invoeren van naam- en adresgegevens. Er is een immense hoeveelheid informatie die benodigd is voor het opvoeren van een nieuwe klant. Vaak moet die informatie aangeleverd worden door verschillende afdelingen en is afhankelijk van bijvoorbeeld het type klant die wordt ingevoerd. Is de nieuwe klant de entiteit waaraan de producten worden verkocht? Of is het slechts een locatie waar de goederen naartoe worden verzonden? Zijn er speciale kortingen van toepassing op dit type klant? Is dit in het verleden al een klant geweest? Zo ja, moet deze dan opnieuw worden aangemaakt of slechts worden geactiveerd?

Het gebruik en de betekenis van de verschillende klantgroepen en bankgegevens zijn vaak complex en de afhankelijkheden met andere informatie niet altijd even duidelijk. Deze onduidelijkheid wordt groter indien nieuwe klanten maar af en toe ingevoerd worden en dit met enige niet regelmaat gebeurt. Met moet dan eigenlijk elke keer opnieuw beginnen met het begrijpen van de klant gegevens. Soms zie je op een afdeling de kennis over hoe (klant)gegevens correct ingevoerd dienen te worden, langzaam afnemen. Een nieuwe medewerker moet misschien nog opgeleid worden en begrijpt de context nog niet die bepaalt hoe en wanneer bepaalde informatie moet worden ingevoerd.

### **Data migratie**

Over het algemeen worden datamigratie trajecten opgeleverd door externe consultants. Deze consultants zijn bekend met datamigratie projecten en de bijbehorende tooling. Ze kennen de uitdagingen die typisch zijn voor datamigratie projecten en hebben hier al vaker mee te maken gehad. In een ideale wereld zijn deze adviseurs ook nog bekend met de bronapplicatie, van waaruit de gegevens worden gemigreerd, én met de doelapplicatie waarnaar de gegevens worden gemigreerd. Ze kunnen zelfs nog bekend zijn met de industrie waarbinnen uw organisatie opereert.

Deze externe consultants zijn echter niet bekend met de manier waarop de gegevens in uw organisatie in het verleden opgevoerd zijn en hoe deze worden gebruikt. Regels en uitzonderingen op die regels zijn bekend bij gebruikers die wellicht niet worden geraadpleegd door het migratieteam. De meeste van deze regels en uitzonderingen worden gevonden door het analyseren van fouten tijdens de migratie. Dankzij meerdere iteraties, langdurige analyses en aannames lukt het uiteindelijk toch nog om de data te laden in de doelapplicatie. Om er op een later tijdstip achter te komen dat de aannames niet klopten en een deel van de data niet bruikbaar is. Hoewel de consultants mogelijk het technische gegevensmodel kennen, is het niet mogelijk alle bedrijfsdefinities en regels te kennen die van toepassing zijn op de data.

### **Batch feeds**

Een batch feed is een exercitie waar grotere hoeveelheden gegevens in één keer worden ingevoerd. Bij batch feeds treedt een combinatie op van de problemen die bij handmatige gegevensinvoer spelen en bij gegevensmigratie. Over het algemeen zijn er twee manieren waarop batchgewijze invoer plaatsvindt:

- Gebruikers die gegevens handmatig in een bepaald formaat voorbereiden en vervolgens een mechanisme gebruiken om al deze gegevens in één keer in de toepassing te laden.
- Een op maat ontwikkelde interface tussen twee applicaties via welke voortdurend gegevens van de ene applicatie naar de andere verstuurd worden, op de achtergrond en zonder dat gebruikers dit merken.

Er zijn een paar zaken die hier mogelijk voor problemen kunnen zorgen:

- Het voorbereiden van de gegevens gebeurt vaak met Microsoft Excel. Dit blijkt steeds vaker een bron van datakwaliteit problemen en niet de oplossing.
- Controles op de invoer, die veelal plaatsvinden wanneer gegevens handmatig worden ingevoerd, worden mogelijk niet toegepast tijdens een batchgewijze invoer. Denk hierbij aan invoerhulp van bijvoorbeeld de straat aan de hand van de postcode en het huisnummer.
- Er is extra inspanning nodig om ervoor te zorgen dat de batchinterface en/of de interface tussen beide toepassingen up-to-date blijft met wijzigingen in de toepassing.

De technische medewerker die verantwoordelijk is voor de batchinterface is zich mogelijk niet bewust dat de zakelijke betekenis van de gegevens in de loop van de tijd is veranderd. Mogelijk is er een nieuw klantsegment toegevoegd dat nog niet wordt ondersteund door de batchinterface. Er is dan een risico dat dat klantsegment verdwijnt tijdens de overdracht van informatie of vervangen wordt door het default klantsegment.

### **Gemengde verwachtingen**

Gemengde verwachtingen komen zowel bij het invoeren van gegevens voor als bij het gebruik van gegevens. Tijdens het invoeren van gegevens kunnen verschillende mensen de gevraagde informatie anders interpreteren. Een voorbeeld hiervan zou een CRM toepassing kunnen zijn, waar bij de invoer van een opportunity een veld 'ETA' is. Sommige medewerkers zijn misschien niet bekend met de afkorting ETA en zullen het veld leeg laten. Anderen herkennen mogelijk dat ETA staat voor 'estimated time of arrival'. In de context van een opportunity; zou dat dan de geschatte tijd betekenen dat de deal wordt gesloten of de geschatte tijd dat de prospect zijn of haar artikelen nodig heeft? De betekenis van beide zijn heel verschillend, maar ze zijn allebei te vertalen naar een datum.

Het mag duidelijk zijn dat wanneer iedereen informatie invoert overeenkomstig de betekenis die zij er zelf aan toekennen, het gebruik van die gegevens in een later stadium erg moeilijk kan zijn. Wat betekent het bijvoorbeeld als er géén datum wordt vermeld? Betekent dit dat:

- Er geen ETA is?
- Het niet duidelijk is wat er ingevoerd dient te worden?
- De medewerker niet over de benodigde informatie beschikt en dus het veld heeft overgeslagen?

### **De enige constante: verandering**

De enige constante in het bedrijfsleven is verandering. Wat het ene jaar nog correct en nodig is om gegevens in te voeren, zal hoogstwaarschijnlijk het volgende jaar anders zijn. De informatie die we nodig hebben voor onze dagelijkse werkzaamheden verandert en daarom verandert de behoefte aan de informatie die we opslaan. Verschillende klanttypen kunnen ontstaan in een poging de markt beter te bedienen. Na verloop van tijd kan dat nieuwe klanttype blijven, (enigszins) veranderen of volledig worden verwijderd. Bij de beslissing over dit type klant is het natuurlijk belangrijk om te weten of het nog steeds moet worden gebruikt of niet. Een discussie rond de type klant en de argumenten die leiden tot een wijziging of verwijdering zijn net zo belangrijk, zo niet nog belangrijker dan de uiteindelijke beslissing.

### **Conclusie**

Datakwaliteit tooling kan goed gebruikt worden voor het in kaart brengen van de problematiek rond datakwaliteit in uw organisatie. Voor een oplossing van deze problematiek moet echter niet naar tooling gekeken worden. Het is van belang dat de betekenis van gegevens, toegestane waardes en uitzonderingen eenduidig en toegankelijk vastgelegd worden. Wat je nu vaak ziet is dat in elk project het wiel opnieuw uitgevonden wordt en het stuk data documenteert wat op dat moment van belang is om het project af te ronden. Deze documentatie verdwijnt na het project of wordt

onder het mom van 'agile' helemaal niet vastgelegd. Een gemiste kans.  
Goede documentatie is de basis voor goede data.

# De waarde van data kwaliteit tooling

Bron: Tentive | [tentive.nl](https://tentive.nl)

Goede data is een *'strategic business asset'* voor de moderne organisatie en datakwaliteit zou een integraal onderdeel moeten zijn van de bedrijfsstrategie. Bij het gros van de organisaties is dit besef nog onvoldoende doorgedrongen.

Data is een waardevol bezit voor organisaties. Met behulp van een doordacht data governance raamwerk kan je dit realiseren. Maar stel niet uit tot morgen wat je vandaag nog kunt doen. De aard en het volume van de data waarover organisaties beschikken, verandert met een ongelooflijke snelheid. Dit betekent dat de potentiële kosten van het niet implementeren van een data governance raamwerk steeds hoger worden.

In kort komt het neer op het volgende. Door de toenemende wet- en regelgeving is het voor organisaties belangrijker dan ooit om data governance toe te passen. Wacht daarom niet op de dingen die gaan komen, maar speel alvast in op toekomstige datavraagstukken die voortkomen uit dit alsmaar veranderende landschap. Daarnaast komen organisaties die werkzaam zijn in minder gereguleerde sectoren steeds meer tot het besef dat concurrentievoordelen te behalen zijn wanneer de eigen data goed wordt gemanaged én begrepen. Organisaties die daarbij een holistische benadering voor data governance hanteren, zullen daarvan het meest profiteren. Maar let wel: initiatieven voor data governance moeten worden geleid door de business en een gestructureerde methodologie gebruiken om succesvol, kosteneffectief en vooral relevant te zijn.



Een goed data governance raamwerk kan eigenlijk niet bestaan zonder technologische ondersteuning. De huidige generatie tooling kunnen dit op de volgende manier doen:

### **Data profiling tooling**

Uit onderzoek blijkt dat slechts 3% van de data voldoet aan de kwaliteitsstandaarden van organisaties (Nagle, 2017). Dit impliceert dat slecht beheerde data organisaties miljoenen euro's aan verspilde tijd, geld en onbenut potentieel kosten. Data profiling tooling bepalen de kwaliteit en de legitimiteit van jouw data. Met behulp van analytische algoritmen worden datasetkarakteristieken tot in het kleinste detail onderzocht.

Veelvoorkomende fouten in databases worden daarbij geëlimineerd. Denk hierbij aan ontbrekende waarden, waarden die niet zouden moeten worden opgenomen, waarden met ongewoon hoge of lage frequentie en waarden die niet de verwachte patronen volgen. Data profiling tooling helpen organisaties met het ordenen en het beter begrijpen van (de onderlinge relaties van) de eigen data. Dit stelt organisaties in staat om essentiële informatie te ontdekken uit de data waarover zij al beschikken. En weten over welke data je beschikt en de betekenis daarvan is natuurlijk elementair voor data governance. Immers, hoe kan je op een goede manier data beheersen als je op voorhand niet eens weet over welke gegevens je beschikt?

### **Data glossaries en metadata tooling**

Een belangrijk onderdeel van data governance is om vast te stellen wat de data is en de betekenis ervan. Het omvat ook het definiëren van business rules en datakwaliteitsregels die uiteindelijk bepalen wat de data goed genoeg maakt voor de eindgebruiker. Het opstellen van een data glossary kan hierbij uitkomst bieden. In een data glossary worden de kenmerken van bedrijfskritische data

elementen in de verschillende systemen beschreven. Deze beschrijving is van belang om ervoor te zorgen dat iedereen weet hoe deze data elementen zijn gedefinieerd, waar ze zijn opgeslagen, en wat de kwaliteit ervan is. (Bij het uitvoeren van projecten zal vaak worden teruggrepen naar de data glossary om te zien hoe data elementen geraakt worden door bepaalde veranderingen (Feltkamp, 2018)). Een data glossary kan worden aangevuld met referentiedata, zoeklijsten en de mapping tussen referentiedata die in verschillende systemen wordt bewaard. Sommige tooling hebben een workflow voor het aansturen van werkstromen. Dit vereenvoudigt het proces voor data eigenaren en data stewards om data definities en referentiedata waarden te beoordelen en goed te keuren.

### **Datakwaliteit monitoring en rapportage tooling**

Dit zijn waarschijnlijk de tooling waarmee eindgebruikers het meest vertrouwd zijn. Zoals de benaming als doet vermoeden, richten deze tooling zich op het beheersen en monitoren van datakwaliteit. Ze houden toezicht op de datakwaliteitsniveaus die zijn gedefinieerd door middel van de business rules. De tooling genereren rapportages waaruit eventuele datakwaliteitsproblemen kunnen blijken. Data eigenaren en data stewards kunnen vervolgens in actie komen om deze gesignaleerde problemen te onderzoeken en op te lossen.

Bij goede data governance gaat vooral over het toewijzen van de juiste mensen aan de juiste rollen en het opzetten van goede processen. Deze processen kunnen daarbij worden ondersteund door gebruikmaking van de juiste technologie. Tooling alleen is niet het antwoord op data gerelateerde vraagstukken. Dit neemt niet weg dat je met succes tooling kunt gebruiken om sommige van je processen te ondersteunen en te operationaliseren. Als je bijvoorbeeld grotere hoeveelheden data wilt gebruiken, dan heb je alle hulp nodig die je kunt krijgen.

# Wat is data security?

Bron: ICT informatiecentrum | [ictinformatiecentrum.nl](https://www.ictinformatiecentrum.nl)

Data security of dataveiligheid heeft tot doel belangrijke en bedrijfsgevoelige data van uw organisatie te beveiligen en beschermen. Wanneer deze data wordt gestolen of zonder uw medeweten wordt gewijzigd of gewist, heeft dat niet alleen direct grote gevolgen voor uw bedrijfsprocessen, maar ook voor uw reputatie en klantrelaties. Er moet dan ook alles aan gedaan worden om dataverlies of diefstal te voorkomen. Door de steeds grotere rol van internet bij het gebruik van IT systemen, wettelijke verplichtingen en het toenemend gebruik van cloudoplossingen neemt de noodzaak toe om aandacht te besteden aan de beveiliging van uw data.

## Waarom beschermen?

Data security is gericht op het beschermen van uw data tegen het verlies van gegevens door technische storingen, menselijk fouten of criminaliteit. Technische storingen kunnen zich voordoen door problemen met de dataopslag, datadoorgifte of noodzakelijke stroomvoorziening. Menselijke fouten kunnen worden gemaakt als bevoegde of onbevoegde medewerkers in uw organisatie toegang hebben tot data en deze ongewenst aanpassen of verwijderen. Veel aandacht gaat de laatste jaren uit naar criminaliteit, waarbij internet de toegang vormt tot uw data. De toenemende afhankelijkheid van bedrijven van computersystemen en data vergroot de waarde daarvan. Hierdoor wordt diefstal of gijzeling van uw data door criminele organisaties steeds lucratiever.

## Cybercriminaliteit

Alle organisaties die gebruik maken van computersystemen en communicatiemiddelen kunnen te maken krijgen met criminaliteit die tot doel heeft om toegang te krijgen tot bedrijfs- of privacygevoelige informatie. Dit geldt zeker voor de organisaties die hun databeveiliging niet goed op

orde hebben. Door een digitaal lek in de beveiliging van het netwerk, kwetsbare apparatuur en onoplettendheid van medewerkers lopen alle organisaties het risico daarop. Ook u kunt daarom te maken krijgen met cybercriminaliteit. Cyberaanvallen, malware, phishing e-mails of andere technieken waarmee criminelen zich toegang verschaffen tot niet openbare gegevens, zijn een reëel gevaar. Dezelfde netwerken en verbindingen die iedere betrouwbare gebruiker toegang geven tot het internet, worden door criminele organisaties gebruikt voor het stelen, verminken of gijzelen van gegevens. Het is de keerzijde van een toenemende digitalisering, waarbij de afhankelijkheid van IT systemen en waarde van data voor iedere organisatie steeds groter wordt. Data security richt zich op het beveiligen van deze computersystemen en de data.

### **Data security en AVG**

Data security draagt ertoe bij dat uw bedrijfsprocessen ongehinderd kunnen doorgaan en dat u geen schade ondervindt van criminele activiteiten. Naast direct eigen belang om de eigen data te koesteren, stellen ook wettelijke bepalingen eisen aan de veiligheid van data. Zo worden persoonsgegevens beschermd door de AVG, die de verantwoordelijkheid voor het beheer van die persoonsgegevens bij u als eigenaar van de database legt. Als deze gegevens niet goed beschermd worden, bent u dus aansprakelijk voor de schade die daarvan het gevolg is. Ook door deze aansprakelijkheid staat bij veel organisaties het beveiligen van data inmiddels hoog op hun agenda's.

### **Oorzaken en gevolgen van dataverlies en diefstal**

Verlies en diefstal van data kent veel oorzaken. Er kunnen technische problemen zijn met de apparatuur voor dataopslag, er kunnen problemen zijn met de stroomvoorziening, u kunt slachtoffer worden van cyberaanvallen zoals hierboven beschreven en er worden menselijke fouten gemaakt, waardoor data per ongeluk wordt verwijderd. Hoewel bedrijven het belang van een goede netwerkbeveiliging onderkennen, is het nog steeds mogelijk dat er op andere manieren gegevens worden ontvreemd. Zo worden nog altijd veel fysieke databronnen gestolen of verloren, zoals laptops, USB sticks

en externe harde schijven. Inlogcodes en wachtwoorden worden vaak ontvreemd of onderschept, zodat kwaadwillenden eenvoudig toegang hebben tot uw bedrijfsdata. Ook gebeurt het dat personeel verantwoordelijk is voor het stelen of doorverkopen van privacygevoelige informatie. Gegevensverlies als gevolg van fouten van eigen medewerkers komt ook voor en kan worden ondervangen met een goede data security policy. Daarnaast kan data verloren gaan door beschadigde USB sticks, systeemcrashes of stroomuitval.

### **Welke maatregelen kunt u treffen?**

De schade die u ondervindt als gevolg van dataverlies kan erg groot zijn. Uw bedrijfsactiviteiten komen stil te liggen en er kan sprake zijn van reputatieschade en verlies aan inkomsten. Om operationeel te blijven, verlies van data te voorkomen en schade te beperken, zijn goede veiligheidsmaatregelen nodig. In de praktijk bestaan deze uit een combinatie van technische, organisatorische en juridische maatregelen.

#### **Technische maatregelen**

Er is een aantal technische maatregelen die u kunt nemen om uw data beter te beschermen. Het eerste waaraan gedacht kan worden is een goed doordachte back-up strategie, zodat er altijd reserve kopieën van belangrijke data aanwezig zijn in het geval van verlies of diefstal. Back-ups moeten bewaard worden op een externe locatie. Om deze reden wordt er veel gebruik gemaakt van goedbeveiligde cloudoplossingen. Het is logisch dat uw eigen serverruimtes afgesloten en beveiligd zijn, zodat niet iedereen hier zomaar bij kan. Bij gebruik van beveiligingssoftware is het belangrijk altijd tijdig updates te installeren.

#### **Organisatorische maatregelen**

Veel fouten met betrekking tot gegevensverlies zijn menselijke fouten. Wanneer uw medewerkers zich beter bewust zijn van de risico's die bepaalde technologie met zich meebrengt, verkleint dit

de kans op dataverlies en -diefstal. Het is belangrijk om met medewerkers duidelijke afspraken te maken over het wachtwoordgebruik. Lokale opslag van data moet ontmoedigd worden en er moeten maatregelen genomen worden om het gebruik van onbetrouwbare software en apps onder werknemers te voorkomen. Uw medewerkers moet duidelijk zijn hoe zij moeten omgaan met vreemde e-mails of ineens anders functionerende hardware. Door het individueel toekennen van gebruiks- en toegangsrechten wordt de toegang tot bedrijfsgevoelige informatie gereguleerd en zijn onderdelen van de IT infrastructuur selectief toegankelijk. Om beveiligingslekken in software te voorkomen is het tijdig installeren van software-updates noodzakelijk. Daarnaast is het implementeren van beveiligingscertificaten op uw websites nodig om datadiefstal te voorkomen wanneer er online data ingevoerd moet worden.

### **Juridische maatregelen**

De in 2018 in werking getreden AVG verplicht iedere organisatie die persoonsgegevens verwerkt om deze data te beveiligen tegen het verlies en de ongewenste inzage daarvan. Dat doet u door het treffen van de hierboven genoemde technische en organisatorische maatregelen. Iedere organisatie moet hier een oplossing voor bieden. Ook arbeidscontracten kunnen bijdragen aan data security, als hierin wordt vastgelegd wat de rechten en plichten van een werknemer zijn met betrekking tot bijvoorbeeld het delen van bedrijfsgevoelige data.

### **Data security oplossingen**

Uw data kan worden beschermd tegen diefstal of verlies door het beveiligen van uw computers en netwerk. Er bestaan diverse softwareoplossingen die voor dit doel bestemd zijn. Deze software detecteert spyware, faciliteert encryptie en firewalls. In bepaalde branches is een hogere mate van data

security noodzakelijk, dan in andere. Wat in ieder geval van belang is dat de software die u erbij helpt constant up-to-date wordt gehouden.

# Data security tips en adviezen

Bron: ICT informatiecentrum | [ictinformatiecentrum.nl](https://www.ictinformatiecentrum.nl)

Bescherming van data en IT systemen is belangrijk. Niet alleen op uw eigen locatie, maar zeker ook in de cloud, zodat er geen twijfels bestaan over dat uw extern opgeslagen data veilig is. De gevolgen van een datalek kunnen namelijk groot zijn. Dat doet niet alleen afbreuk aan het imago van uw organisatie, maar er kunnen ook flinke boetes aan te pas komen. Als waardevolle data in handen komt van criminelen of concurrentie, kan dat tevens zorgen voor een afgenomen concurrentiepositie en daarmee verlies van inkomsten.

## Tips en adviezen

### Samenwerking

Het beschermen van uw data in de cloud is zowel een taak van de aanbieder van de clouddiensten als van uzelf. U kunt als klant vertrouwen op de best mogelijke beveiliging vanuit uw cloudaanbieder, maar ga hier niet blindelings van uit.

### Veiligheids certificeringen

Let er altijd op dat uw aanbieder over de juiste veiligheids certificeringen beschikt. Daarmee laten ze zien op welke manier zij de data van hun klanten beschermen.

### Back-up

Tref zelf ook veiligheidsmaatregelen. Beschikken over back-ups van uw data is van groot belang. Bij mogelijk dataverlies kunnen er binnen een organisatie veel moeilijkheden optreden. Als dit onverwacht bij uw cloudaanbieder voorkomt, is het noodzakelijk om op een andere plek een back-up van al uw data te hebben staan. De



gevolgen voor zo'n gebeurtenis binnen uw organisatie zijn dan minimaal.

### **Versleutelen**

Bij het verzenden van data van en naar de cloud is het een vereiste om uw data goed te versleutelen. Mocht uw data door hackers toch worden onderschept, dan verkleint u de kans dat de hackers het bestand of toegangscode in kunnen lezen, omdat de 'sleutel' tot de gegevens ontbreekt.

### **Wachtwoordbeleid**

Het hebben van een sterk wachtwoordenbeleid binnen uw organisatie wordt aangeraden. Zorg altijd voor unieke wachtwoorden op verschillende apparaten. Idealiter met behulp van tweefactorauthenticatie. Dat is een extra beveiligingslaag, zodat hackers nog steeds niet kunnen inloggen, ook al beschikken zij over het wachtwoord.

### **Rangschikken van data**

Bepaalde gegevens horen niet extern opgeslagen te worden. Het is daarom van belang om uw data te rangschikken op basis van bijvoorbeeld discretie en integriteit. Daarmee kunt u bepalen welke data wel en niet moeten worden opgeslagen in de cloud en voor welke gegevens uit oogpunt van data security extra aandacht nodig is.

### **Fysieke locatie**

Houd bewust rekening met de fysieke locatie van het datacenter van de cloudaanbieder. In elk land en elke regio gelden andere wetten en regels. Met name als het gaat om gevoelige bedrijfsinformatie, is het van belang om er voor te zorgen dat uw data wordt opgeslagen in een gebied waarin u bekend bent met de wet- en regelgeving die

daar van toepassing is. Enkele grote aanbieders bieden die garantie door uw data op te slaan in een datacenter in een regio naar keuze.

### **Toegangsbeveiliging**

Niet iedereen hoeft alle gegevens in te zien of te bewerken. Zorg daarom voor verschillende rechten, zodat elke medewerker alleen toegang heeft tot de gegevens die bij zijn of haar werkzaamheden passen. Goede cloudoplossingen hebben de mogelijkheid om per gebruiker rechten toe te kennen.

### **Vorbereiding**

Voor het overstappen naar een cloudoplossing kunt u de eerste belangrijke stappen op het gebied van databeveiliging al nemen. Door over een goed beveiligd bedrijfsnetwerk te beschikken, kunnen de eerste gevaren worden vermeden. Het is van belang om data security veel aandacht te geven als u er over denkt om voor een cloudoplossing te kiezen.

### **Updates**

Houd al uw apparaten en software binnen uw organisatie up-to-date. Hoogwaardige producenten en ontwikkelaars voorzien hun producten met regelmaat van updates om beschermd te zijn tegen de meest recente en acute beveiligingsgevaren. Controleer alle instellingen zorgvuldig en schakel de functies uit die binnen uw organisatie niet worden gebruikt.

### **IT-afdeling**

Overleg vanaf het begin van de overstap naar de cloud met de IT afdeling van uw organisatie. Vanuit hun vakgebied zijn ze in staat om eventuele gevaren te beoordelen. Onderzoek gezamenlijk hoe data op eigen locatie beschermd is en kijk of het gewenste beveiligingsniveau ook realiseerbaar is na de overstap naar de cloud.

## **IoT-apparaten**

Als u binnen uw organisatie gebruikt maakt van IoT-apparaten (Internet of Things, ofwel losse 'smart' apparaten en producten die toegang hebben tot internet), is het goed om te weten wat u precies in huis heeft. De beveiliging van IoT-apparaten is vaak minder goed op orde, omdat het ontwikkelen van een goede beveiliging kostbaar is voor de producent van het product. Daarnaast beschikken IoT-apparaten vaak niet over genoeg reken- en opslagkracht om hun data van een complexe versleuteling te voorzien. Controleer of de producent geregeld beveiligingsupdates uitbrengt en welke versleuteling en inlogprotocollen door het IoT-apparaat worden gebruikt. Kijk bij de aanschaf van een product hoe de producent tot nu toe heeft gepresteerd op het gebied van beveiliging.

## **Gebruiker van IoT-apparaten**

In veel gevallen zorgt de gebruiker van IoT-apparaten voor het beveiligingsrisico. Omdat er weinig fysieke interactie met IoT-apparaten is, worden deze vaak vergeten, met als gevolg dat belangrijke beveiligingsupdates niet worden geïnstalleerd en gebruikers het niet doorhebben als er eenmaal ingebroken is. Door nauwkeurig een lijst te maken van alle IoT-apparaten binnen uw organisatie samen met de bijbehorende softwareversie, kunt u dit voorkomen. Een sterk en uniek wachtwoord is juist ook hier van belang, omdat hackers IoT-apparaten vaak zien als de makkelijkste weg naar binnen.

## **Leid de gebruikers op**

In bepaalde gevallen zorgen niet de IT systemen, maar de gebruikers voor het voornaamste beveiligingsgevaar. Train alle gebruikers daarom op het gebied van data security. Op die manier worden uw medewerkers zich bewust van de gevaren van openbare netwerken en de mogelijkheid dat anderen in openbare ruimten kunnen meelezen.

### **Test uw data security**

Laat uw data security testen door beveiligingsexperts in een 'audit'. Specialisten doen zich dan voor als hackers en proberen bij uw IT systemen in te breken om zo zwakheden van uw beveiliging bloot te leggen. Op deze manier kunt u uw beveiliging verbeteren voordat kwaadwillenden echte schade kunnen aanrichten.

### **Ontwerp een goed bedrijfsnetwerk**

Om het hackers zo moeilijk mogelijk te maken data te ontvreemden als ze eenmaal binnen zijn, is het van belang om het bedrijfsnetwerk zorgvuldig te ontwerpen. Zorg ervoor dat de fysieke internet-infrastructuur niet gemakkelijk toegankelijk is, zodat eventuele fysieke inbrekers zichzelf moeilijk toegang kunnen verschaffen tot het bedrijfsnetwerk. Verdeel het netwerk in verschillende zones, zodat hackers - als ze eenmaal binnen zijn - niet direct over het gehele netwerk beschikken, maar data telkens via een nieuwe beveiligde en afgeschermd opslaglocatie moet bemachtigen. Door niet alleen de draadloze, maar ook de bekabelde netwerkverbindingen te versleutelen zorgt u ervoor dat hackers geen data kunnen aftappen van een bekabeld netwerk. Vaak versleutelen bedrijven alleen hun draadloze verbindingen.

### **Wifiverbindingen**

Zorg voor de veiligheid van uw wifiverbindingen. Voorkom dat via het relatief zwak beveiligde wifi gevoelige of waardevolle bedrijfsinformatie wordt verstuurd.

### **Logbestanden**

Door alle netwerkgegevens te verwerken in logbestanden en te analyseren met behulp van speciale software kunnen hackaanvallen en inbraakpogingen tijdig worden gedetecteerd. Het blokkeert de hackers niet rechtstreeks, maar het zorgt er wel voor dat u met uw

organisatie een risicoschatting kunt maken om schade te beperken en volgende pogingen te blokkeren.

# 5 tips voor het opstellen van informatiebeveiligingsbeleid

Bron: ICTrecht | [icrecht.nl](https://www.icrecht.nl)

Het zal niet ieders hobby zijn, maar je kunt er niet omheen: als organisatie die de beveiliging van haar informatie belangrijk vindt, zul je vroeg of laat een beveiligingsbeleid moeten opstellen. Dat kan nog wel eens tot weerstand leiden, omdat degenen die met beveiliging aan de slag gaan, niet altijd veel op hebben met regeltjes en papierwerk. 'Ik kan mijn tijd toch beter besteden aan het veilig houden van informatie, dan aan het schrijven en onderhouden van papieren tijgers?!' is een vaak gehoorde reactie.

Dat mag dan zo zijn, maar een beveiligingsbeleid is méér dan een droge opsomming van getroffen beveiligingsmaatregelen. En het uiteindelijke beleidsdocument is op zichzelf niet het einddoel. Het formuleren van beleid dwingt een organisatie namelijk om goed na te denken over waar zij voor staat, welke doelen daarbij horen en hoe je die kunt bereiken. Daarnaast heeft de documentatie niet alleen waarde als intern naslagwerk, maar kan het ook worden gebruikt om verantwoording af te leggen aan bijvoorbeeld opdrachtgevers, de Raad van Toezicht, de ondernemingsraad, de auditor of de Autoriteit Persoonsgegevens. Maar: hoe kom je tot een beleid waar je ook echt wat mee kunt? Graag geven wij een paar praktische tips, gebaseerd op wereldwijde best practices.

## **Tip 1 - Informatiebeveiliging moet op het hoogste niveau geagendeerd en aangestuurd worden**

Om informatiebeveiliging te laten slagen, moet er betrokkenheid zijn vanuit de directie. Dat betekent ook, dat de directie (minstens op hoofdlijnen) voldoende kennis moet hebben van informatiebeveiliging. Uiteraard kan er van alles gedelegeerd worden, maar

het startpunt van het informatiebeveiligingsbeleid is de directie. De directie bepaalt immers de koers en strategie van de organisatie en draagt eindverantwoordelijkheid.

Het beleidsstuk dat op het hoogste niveau wordt vastgesteld hoeft niet uitputtend en van het grootste detailniveau te zijn. De uitwerking van specifieke beleidsonderwerpen kan op een lager niveau plaatsvinden. Maar het is aan de directie om de aanpak en doelstellingen voor de organisatie te bepalen, en dus ook de doelstellingen met betrekking tot informatiebeveiliging.

### **Tip 2 - Voorkom 'security island' en zorg voor aansluiting op bestaande doelstellingen, strategieën en structuren**

Het is niet alleen inefficiënt om voor je beleid omtrent informatiebeveiliging het wiel opnieuw uit te vinden; het leidt ook tot verwarring en onduidelijkheid. Is er al een kwaliteitssysteem om de bedrijfsprocessen te borgen? Sluit dan daarbij aan voor de inrichting van je managementsysteem voor informatiebeveiliging. Wordt een bepaalde gelaagdheid of structuur gehanteerd in bestaande beleidsstukken, die bekend is in de organisatie? Sluit dan daarbij aan voor de inrichting van je documentatie. Maar bovenal: zorg ervoor dat de doelstellingen en beleidsregels met betrekking tot informatiebeveiliging aansluiten of in het verlengde liggen van de 'algemene' bedrijfsdoelstellingen en missie van de organisatie. Informatiebeveiliging is immers geen doel op zich, maar een middel om de organisatie in staat te stellen haar missie uit te voeren.

### **Tip 3 - Werk van abstract naar concreet**

Veel beleid wordt opgesteld in reactie op een gebeurtenis in de praktijk. Het is daarbij verleidelijk om die praktijkgebeurtenis direct en concreet te adresseren. Het punt is echter dat daardoor een lappendeken kan ontstaan van hele concrete beleidsstukken, die slechts van toepassing zijn in een beperkt aantal gevallen. Met als

gevolg dat voor de organisatie onduidelijk is wat nu precies het beleid is in gevallen die net anders zijn.

Om dit te voorkomen kan ervoor gekozen worden om eerst meer abstracte, richting gevende uitgangspunten te formuleren. Deze principes kunnen altijd worden toegepast, ook in situaties die zich nog niet hebben voorgedaan. Voor situaties die zich vaker voordoen kunnen concrete uitwerkingen van de bovenliggende principes worden opgesteld, bijvoorbeeld in de vorm van een protocol of procedure. Zulke protocollen of procedures moeten altijd in lijn zijn met de bovenliggende uitgangspunten.

#### **Tip 4 - Ken je doelgroep en stem je beleidsstukken daarop af**

Informatiebeveiligingsbeleid wordt zelden in één document gevat. Vaak gaat het om een reeks samenhangende stukken die gezamenlijk het informatiebeveiligingsbeleid vormen. Die stukken zullen elk hun eigen doelgroep hebben. Het is belangrijk om de inhoud daarop af te stemmen. Zo zal een overkoepelend, richting gevend en daardoor tamelijk abstract beleidsdocument nut hebben voor het management, maar weinig waarde hebben als handleiding op de werkvloer. Zo'n handleiding moet specifiek en concreet zijn.

Naast vorm, inhoud en toon is het ook belangrijk om na te denken over het kennisniveau van de doelgroep. Bepaalde doelgroepen zullen over minder of juist meer kennis van informatiebeveiliging beschikken dan de andere. In die situaties zul je het beleid dus moeten ondersteunen met kennisoverdracht, al dan niet in de vorm van trainingen.

#### **Tip 5 - Accepteer dat je beleid niet volledig is**

Het is verleidelijk om eindeloos aan beleidsstukken te blijven schaven, net zolang totdat ze perfect zijn. Maar: de wereld verandert continu, evenals de organisatie. Het is daarom utopisch



om te denken dat het perfecte en volledig complete beleidsstuk een haalbare kaart is. Accepteer daarom dat je nooit volledig kunt zijn. Maar bouw daarvoor ook maatregelen in. Enerzijds kun je beleid maken voor onvoorziene situaties door te werken met principes (tip 3), anderzijds is het belangrijk om het beleid procesmatig te borgen via een plan-do-check-act-cyclus. Kort samengevat: evalueer regelmatig of je beleid nog passend is en het gewenste effect sorteert, en stuur bij wanneer dat niet zo is.

# Data management platform

Bron: ICT informatiecentrum | [ictinformatiecentrum.nl](http://ictinformatiecentrum.nl)

Data management heeft belangrijke toepassingen op het gebied van marketing en sales. Hoe meer data beschikbaar is voor het aansturen van commerciële processen, des te effectiever en efficiënter kunnen deze uitgevoerd worden. De beschikbaarheid van steeds meer klant - en marktdata en de behoefte aan tools voor CRM en marketing automation hebben geleid tot de ontwikkeling van een data management platforms of DMP.

## **Wat is een data management platform?**

U beschikt over steeds meer data over klanten en prospects. Niet alleen in uw eigen systemen zit veel bruikbare data, van buitenaf is er ook veel data beschikbaar. Om van deze commerciële data gebruik te kunnen maken, moet u beschikken over een tool dat voorziet in het beheer van deze data, de analyse van de gegevens en de communicatie met klanten die hiervan een vervolg is. Een data management platform is dat tool.

Met een data management platform verzamelt u, met gebruik van verschillende technologieën, data via verschillende interne en externe kanalen. Hierdoor heeft u beter zicht op wie uw klanten en prospects zijn en waar de grootste kansen voor u liggen. Het DMP kan u inzicht geven in de effectiviteit en resultaten van marketingcampagnes, het koopgedrag van klanten en uw marktkansen. Technisch gezien kunt u een DMP beschouwen als een datawarehouse en business intelligence in één, met een specifieke toepassing voor marketing en sales. Het kan u helpen bij het verkrijgen van een volledig klantbeeld. Daarmee wordt het mogelijk uw commerciële processen beter aan te sturen en zo het rendement van uw marketing en sales te vergroten.

## **Hoe begint u een data management platform?**

De volgende tips en aanwijzingen helpen u met het starten van een data management platform.

- Weet welke doelen u nastreeft. Bepaal wat u met uw data wilt doen en hoe uw salesproces is.
- Ga na of uw medewerkers er ook mee kunnen en willen werken.
- Houd de regelgeving rond persoonsgegevens (AVG) goed in de gaten en zorg dat u eraan voldoet.
- Zorg dat u toegang heeft tot alle voor u belangrijke data.
- Voldoe aan de verwachtingen van uw klanten en verras ze niet met ongewenst gebruik van hun gegevens.
- Beschik over een goed CRM systeem dat goed samenwerkt met het data management platform.
- Bepaal vooraf of door de (soms grote) investering in tijd en geld een DMP lonend voor u zal zijn.

## **Welke data management platform heeft u nodig?**

De keuze van een DMP hangt van veel aspecten af. Uw salesprocessen en de beschikbaarheid van data zijn daarbij belangrijk. Bijna altijd zal u voor de keuze en implementatie van een oplossing afhankelijk zijn van deskundige hulp.

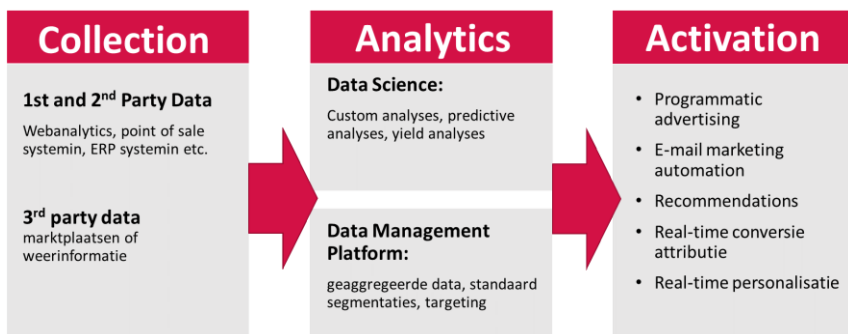
# 5 voordelen van een data management platform

Bron: ISM eCompany | [ism.nl](http://ism.nl)

Door slimme analyses van data maak je betere beslissingen. Maar datadriven marketing is voor veel marketeers een flinke uitdaging. Er is veel data beschikbaar, maar meestal verkeert deze in silo's. De oplossing is een data management platform (DMP).

## Wat is een data managementsysteem?

Een DMP verzamelt, sorteert en bewaart data. Maar nog belangrijker; het helpt je om de complexe en enorme hoeveelheid informatie te vertalen naar marketingacties. De werking van een DMP leggen we uit aan de hand van 3 stappen van onze datadriven strategie.



## Stap 1 - Data verzamelen

De DMP fungeert als een centrale hub voor het verzamelen, aggregeren en beheren van drie soorten gegevens:

- **First Party data** - Dit zijn de gegevens die je zelf hebt verzameld over verschillende gebruikers, klanten, bezoekers. Deze gegevens komen uit verschillende systemen. Bijvoorbeeld uit CRM, Google Analytics of transactiegegevens.
- **Second Party data** - Dit zijn gegevens van partners die hun first-party data met je delen, bijvoorbeeld door te partneren met een andere niet concurrerende retailer (verkoop van huurauto's i.s.m. vakantieaanbieder bijvoorbeeld).
- **Third Party data** - Geanonimiseerde hoeveelheden persoonsdata die je kunt kopen voor bijvoorbeeld segmentatie van je markt.

## Stap 2 - Data analyseren

Een DMP voegt waarde toe aan de verzamelde en opgeslagen gegevens door:

- **Holistische analyses** - In een DMP komt alle data samen, waardoor analyses op basis van deze geaggregeerde inzichten plaatsvinden.
- **Algoritmische modellering** - Hiermee maak je automatisch nieuwe segmenten die je kunt benaderen. Denk aan segmenten op basis van het gedrag van je reeds bekende klanten.
- **Klant- en prospectsegmentatie** - Helpt je om je marketing slimmer in te richten met betrekking tot

targetting en de frequentie waarmee je klanten wilt benaderen.

Geavanceerdere analyses of aangepaste modellering worden doorgaans niet gedaan binnen het DMP, maar een data scientist kan de geaggregeerde data exporteren naar programma's als R en Python. Vervolgens kan hij aangepaste modellering toepassen en deze weer importeren naar je DMP.

### **Stap 3 - Data activeren**

Een DMP ontzorgt je, want hij kan de inzichten uit je analyses vertalen naar acties. Denk aan een toepassing als realtime personalisatie. Met een DMP kun je dus:

- Een centraal klantbeeld creëren welke je kunt verrijken en acties kunt ondernemen over alle digitale kanalen en devices heen.
- Klantsegmenten identificeren en nieuwe communicatiestrategie bepalen voor deze aparte segmenten.
- Nieuwe potentiële waardevolle klanten vinden op basis van je reeds bestaande klanten en hun gedrag.
- Efficiëntie verbeteren over alle kanalen en apparaten.
- Efficiënter targeten, met name in display advertising door het gebruik van audiences.

### **Voordelen van een DMP**

Nu weet je wat een DMP is, maar wat zijn de voordelen? Met de enorme hoeveelheid aan (digitale) touchpoints en interactiemomenten is het voor bedrijven onmogelijk om handmatig alle data optimaal te benutten voor marketingacties. Met een DMP kan dit wel. Je kunt allerlei acties automatiseren en van een persoonlijke aanpak voorzien!

Het gebruik van een DMP geeft je de volgende 5 voordelen:

### **Voordeel 1 - Je bespaart tijd**

Een DMP bespaart je tijd doordat je:

- Gegevens automatisch analyseert en toepast in de processen die volgen.
- Klanten beter kent en begrijpt.
- Campagnes vrijwel realtime organiseert.
- Je concentreert op de meest waardevolle klanten.

### **Voordeel 2 - Hoger rendement**

Een DMP verhoogt je rendement op een aantal verschillende manieren:

- Je maakt aanzienlijke efficiëntieslagen op het gebied van search en display advertising door het gebruik van centraal samengestelde audiences uit het DMP.
- Door het gebruik van segmentatie richt je het marketingbudget geheel anders in. Namelijk op basis van potentiële customer lifetime value van een klant of prospect.
- Je verhoogt de orderwaarde door nieuwe communicatietactieken op basis van de analyses uit het DMP

### **Voordeel 3 - Schaalbaarheid**

Het vermogen van een DMP om de enorme hoeveelheid informatie te verwerken en centraal op te slaan, is een voordeel op zich. Zo kan een DMP verbinding maken tussen gegevens uit je betaalde zoekadvertenties, CRM gegevens, gegevens van derden, webshopgedrag, sociale media, gesprekken in call centers en online

klantenservice en die gegevens tonen. Een systeem dat zoveel verschillende informatiebronnen integreert, kan de verwerking en toepassing van deze data sterk verbeteren.

#### **Voordeel 4 - Betere prospecting**

Wanneer je alle informatie over je klanten en hun gedrag op één plaats opslaat en er vervolgens algoritmische modellen op los laat kun je aan de hand van het gedrag van je bestaande klanten, nieuwe waardevolle klanten herkennen in gebruikers op je search-, display- en socialmediakanalen. Je target dus gemakkelijk relevantere nieuwe segmenten op basis van gedrag in je reeds bestaande segmenten.

#### **Voordeel 5 - De gehele organisatie profiteert!**

Je hele organisatie kan de waardevolle inzichten gebruiken voor de gehele bedrijfsvoering en resultaten. Daarnaast verminder je de isolatie van verschillende afdelingen en maak je gegevens veel makkelijker bereikbaar. Het begrijpen van het gedrag en de voorkeuren van mensen creëert een kans om je inkoop te verbeteren, of om je klantenservice anders in te richten.

### **Een data managementplatform kiezen**

Als jij bent overtuigd en je aan de slag wilt met een DMP kom je voor een dilemma te staan. Want de aanschaf van een DMP kan een flinke investering zijn. Er zijn grofweg twee soorten DMP's: DMP's die zich hebben gespecialiseerd vanuit een bepaalde hoek, bijvoorbeeld display-platformen die ook een DMP-oplossing aanbieden. Vaak zijn de geïntegreerde oplossingen ook de goedkopere oplossingen. Maar vergeet niet dat je je dan ook committeert aan het ecosysteem van deze aanbieder en je vrijheid in onafhankelijke integraties opgeeft. Aan de andere kant is de "stand-alone" DMP's, welke zich op een meer onafhankelijke wijze positioneren. Ze hebben vaak een duurder prijsmodel, maar bieden wel veel vrijheid in integraties en modellering.



Een DMP is niet langer alleen voor de grote ondernemingen bereikbaar. Er komen steeds meer spelers op de markt die zich expliciet richten op het middensegment. Er zijn ook oplossingen in de markt die onderdelen van een DMP bevatten, maar niet de gehele oplossing bieden. Hou dan ook rekening met minder voordelen ten opzichte van een volwaardige DMP oplossing.

**Start met een DMP. Nu is het moment!**

Data speelt een steeds grotere rol in de wereld van de marketeer. Start daarom vandaag met het opstellen van je eerste kleine doelstelling en test deze door het inzetten van datadriven marketing. Is je businesscase positief? Breid dan uit naar een volgend klantsegment.

# Voordelen van een customer data platform

Bron: Roon Hulshoff i.s.m. Graydon Nederland | [graydon.nl](https://graydon.nl)

Geloof u in de voordelen van een datagedreven onderneming? Dan bent u al een heel eind op weg. Legacy systemen zijn de voornaamste drempels in het doorvoeren van de datastrategie op weg naar de datagedreven cultuur. Een praktisch hulpmiddel om te voorkomen dat dit idee strandt op de IT tekentafel, is het customer data platform (CDP). Maar hoe gaat u hiermee aan de slag? Het begint bij het begin: de WHY

Voor veel marketeers en communicatiespecialisten is het denken vanuit de 'why, how en what' van Simon Sinek, ook wel de golden circle, een eye-opener geweest. Deze denkwijze is al jaren leidend voor het opzetten van een marketing communicatie strategie. Het gaat er hierbij niet zozeer om de doelgroep te benaderen van buiten naar binnen, maar van binnen naar buiten. Zo wordt de doelgroep voorzien van context en perspectief, bijvoorbeeld door het maken van een compelling story: een relevant en herkenbaar verhaal. Houd deze gedachte eens vast en projecteer deze dan op het streven naar een datagedreven cultuur. Zou u dan beginnen met het implementeren van een nieuw IT systeem (de 'what') of nadenken over een anders in te richten operationeel proces (de 'how')?

## Strategie is leidend

Het begint bij een doordachte strategie om waarde te creëren uit (operationele) data en deze inzichten te gebruiken voor verbetering van bestaande bedrijfsprocessen en waarde toe te voegen in (nieuwe) businessmodellen. U moet dus een gezamenlijke stip op de horizon creëren om medewerkers te voorzien van context en perspectief. Dit doet u door middel van een compelling data story. Zo verbindt u uw medewerkers en de

afdelingen waartoe ze behoren onderling met elkaar en is er geen sprake van silo-denken. In een later stadium hoeft er pas nagedacht te worden over een plan rondom data governance en master data management, en om deze later in te bedden in de operationele organisatie. En niet verrassend, hier zullen we tegen IT legacy aanlopen die niet praktisch 1-2-3 op te lossen is. Stopt het hier dan?

### **Praktische oplossing: het customer data platform (CDP)**

Een customer data platform (CDP), ook wel klantgegevensplatform, vormt een klantendatabase die toegankelijk is voor andere systemen. Een CDP integreert alle relevante databronnen – intern en extern, offline en online – waarmee profielen en segmenten kunnen worden opgebouwd. Vervolgens kan er met behulp van slimme modellen (artificial intelligence, AI) een advies worden gegenereerd voor een beslissing of vervolgactie via een voorkeurskanaal.

### **Is het customer data platform alleen voor marketeers?**

Wanneer u aan diverse personen vanuit verschillende afdelingen vraagt voor wie het CDP is bedoeld, zult u hoogstwaarschijnlijk ervaren dat de meeste medewerkers het customer data platform zien als een instrument voor de marketeer. Dit is echter een misvatting. Het gaat hier om de disciplines rondom data integratie en management, rondom segmentatie en profilering, inzichten uit analyse en uiteraard rondom activatie om conversie te realiseren. Het grijpt dus in op diverse geledingen en mensen binnen de organisatie. Het is dan ook van wezenlijk belang dat alle betrokkenen de waarde van data bezien vanuit hetzelfde perspectief en voldoende context hebben waarom ze doen wat ze doen.

Vanuit de compelling data story begrijpt eenieder welke kennis zij als individu en als afdeling kunnen toevoegen, en dus waarde bijdragen aan de bedrijfskolom. Data en het platform fungeert dan als smeermiddel van het nieuwe denken. Wat er anders gebeurt, is dat er heel veel data rondgaat

over het platform. Maar betekent gebruik maken van veel data dan ook hetzelfde als datagedreven opereren?

### **Data vanuit een modern perspectief**

Het is een misverstand om vertrouwen te zoeken in het nieuwste systeem of platform op weg naar de datagedreven onderneming: dit zijn slechts hulpmiddelen. Het vertrouwen moet vanuit de interne organisatie komen, die ontdekt waar de potentiële kracht zit van de tot dusverre onontgonnen (interne) data. De kennis van de aanwezige data zit in de hoofden van medewerkers. Veel van deze kennis is echter aan banden gelegd in processen en systemen en is hierdoor minder inzetbaar. Het gevaar ligt dan ook op de loer dat het customer data platform wordt gevoed met data op basis van de oude principes. In welke mate gaat het CDP dan het verschil maken? Gaat het customer data platform dan helpen om ook de 'data mindset' te veranderen en aanzetten tot groei die u voorstaat? Of blijft het bij een nieuwe IT implementatie? Een CDP is een logisch vervolg op een data strategie, niet andersom.

### **Laagdrempelig**

De technologische ontwikkelingen gaan snel. Daardoor is het ook bij de legacy systemen veelal mogelijk om data te ontsluiten en deze te integreren in het CDP. Doordat een CDP vaak modulair is opgebouwd, kunt u met kleine stappen beginnen. Een platform waarop alle (klant)data van zowel binnen als buiten de organisatie centraal binnenkomt en overzichtelijk en toegankelijk is, biedt dus veel steun in uw weg naar datagedreven ondernemen.

# Data hub: informatie, inzicht, actie!

Bron: SAP Nederland | [blogs.sap.nl](https://blogs.sap.nl)

Waar komt in jouw organisatie de data vandaan? Binnen iedere organisatie is informatie verspreid over datawarehouses, big databronnen, datamarts en data lakes. Zowel in de cloud als on-premise. Die data heeft je organisatie nodig om bijvoorbeeld processen op een slimmere manier in te richten of nieuwe businessmodellen te ontwikkelen.

## Hoe komt dat dan?

Veel bedrijven hebben in de afgelopen jaren hun datalandschap en business intelligence (BI) losgekoppeld van de bedrijfsprocessen. Organisaties verzamelen grote hoeveelheden data uit in- en externe bronnen. Die data slaan ze op totdat business users een informatievraag hebben. Met BI tools en data science worden vervolgens analyses, dashboards en rapportages geproduceerd. Op basis daarvan worden conclusies getrokken, maar die worden niet altijd (direct) omgezet in acties.

Vergelijk het met een auto die regen detecteert. Een auto zet zelf de ruitenwissers aan, zodra er druppels op de ruit vallen. Maar stel je voor dat er in plaats daarvan een berichtje op je dashboard verschijnt met de boodschap dat het regent. Totaal overbodig natuurlijk, aangezien je dat met het blote oog ook wel waarneemt. Toch is dat precies wat er gebeurt in heel veel organisaties. In plaats van dat informatie (het regent) direct aanzet tot actie (zet de ruitenwissers aan), denken veel organisaties nog in meldingen en rapporten. Op basis daarvan wordt vaak pas op een later moment actie ondernomen.

## Dat kan anders

Organisaties kunnen projecten en processen versnellen door data, analyse en actie te koppelen en te automatiseren. Daardoor krijgen ze sneller relevante

inzichten en zijn ze in staat datagedreven applicaties te ontwikkelen. Neem een bedrijf dat veevoer produceert en levert aan de agrarische sector. Het bedrijf levert op aanvraag. Boeren houden zelf in de gaten of er voldoende voer is voor de koeien, paarden of varkens. Als de bodem van de voersilo in zicht is, bestellen ze nieuw voer bij de leverancier. Met het risico dat het voer te laat binnen is. Voor een veevoerbedrijf ligt daar een gouden kans. Door IoT-sensoren in de voertank te plaatsen, kan het bedrijf precies meten hoeveel voer er in de tank zit. Op het moment dat er nog maar een bepaalde hoeveelheid voer in de tank zit, wordt er automatisch een signaal verstuurd naar het veevoerbedrijf en een order aangemaakt voor een nieuwe levering. De boer hoeft daardoor nooit meer handmatig te controleren of er genoeg voer is en een bestelling te plaatsen. Hij is altijd verzekerd van genoeg voer.

Voor het veevoerbedrijf zelf levert het ook veel voordelen op. Door de data van honderden boeren te verzamelen en te combineren met enterprise data, kan het volledige productieproces worden geoptimaliseerd. En kun je zelfs voorspellen hoeveel voorraad je volgende maand nodig hebt. Zo kun je klanten op tijd voorzien van hun producten en heb je nooit tekort of overschot aan voer. Kortom, informatie leidt tot inzicht en zet direct aan tot actie.

### **Hoe krijg je dat voor elkaar?**

Je realiseert dit door één centraal punt te creëren om datastromen binnen je bedrijf te coördineren en te monitoren. Een hub, die je in staat stelt inzicht te krijgen in je totale datalandschap en een brug te slaan tussen bedrijfs-systemen en externe databronnen. Hoe dat werkt? Zo'n data hub slaat data zelf niet op, maar verzorgt de flow van (realtime) data tussen bronsystemen en doelsystemen. Met een data hub geef je eigenlijk precies aan wat er met de data moet gebeuren. Zo kun je dus bepaalde informatie van sensoren koppelen aan een geautomatiseerd ordersysteem. Daarbij wordt zo veel mogelijk gebruik gemaakt van de kracht van de bronsystemen, om optimale performance te garanderen. Daarnaast biedt een data hub organisaties inzicht om data goed te kunnen interpreteren. Want als je snapt waar je naar

kijkt, wordt het eenvoudiger om de correctheid van data te garanderen of waar nodig aan te passen. Je kunt letterlijk tot op kolom- en rijniveau zien hoe datasets zijn opgebouwd. Bovendien voldoe je altijd aan wet- en regelgeving, omdat je precies weet wie toegang heeft tot welke data en waar data is opgeslagen.

### **Denk aan de ruitenwissers**

Denk dus bij het volgende dashboard dat je gebruikt nog eens terug aan de automatische ruitenwissers en welke sensorinformatie je auto nog meer direct omzet in actie. Trapt hij voor je op de rem als er een ongeluk dreigt bijvoorbeeld? Welke informatie en inzichten kunnen binnen jouw organisatie direct omgezet worden in actie?

# De juiste dataoplossing: PIM of PDM

Bron: Beeyond | [beeyond.nl](https://beeyond.nl)

Welke dataoplossing je moet kiezen voor welk probleem, is niet per se eenvoudig. Om te weten welke keuze je moet maken, is het belangrijk eerst te begrijpen wat de verschillen zijn tussen deze dataoplossingen en de informatie die deze verzamelen. De verschillende oplossingen op het gebied van data management zijn beschikbaar voor zowel kleine als grote bedrijven en kunnen helpen bij het vereenvoudigen van processen. We leggen de belangrijkste verschillen aan je uit.

## **Wat is het verschil tussen de dataoplossing PIM en PDM?**

Product information management (PIM) en product data management (PDM) worden vaak als dezelfde termen beschouwd, terwijl het twee verschillende dingen zijn. De aanpak is wel hetzelfde: productdata beheren, centraliseren en structuur geven. Maar er zitten significante verschillen in het uiteindelijke doel.

## **Waar gebruiken organisaties PDM voor?**

Product data management is bedoeld voor alle informatie rondom productontwikkeling. Denk bijvoorbeeld aan projectmanagement, productplanning en urenregistratie. PDM software is daarmee een belangrijke data management softwaretool ter ondersteuning van product lifecycle management (PLM). Want PLM staat voor een strategische aanpak, waarbij de procedure centraal staat. Daarvoor put men informatie uit PDM gegevens.

PDM beheert productdata binnen een ontwikkelafdeling, geeft het structuur, zorgt voor centrale opslag van bijvoorbeeld CAD gegevens en 2d tekeningen. Dus wanneer het belangrijk is om de levenscyclus van een product gedetailleerd te volgen, dan komt PDM om de hoek kijken. Informatie uit



PDM geeft antwoord op vragen over *hoe* iets tot stand gekomen is. Is dat de vraag die je graag beantwoord ziet? Dan is een PDM systeem de juiste data-oplossing.

### **Wanneer gebruiken organisaties PIM?**

Product information management (PIM) richt zich met het beheer van productinformatie vooral op de ondersteunende kant van het verkoopproces. Waarbij PDM uitgebreid in gaat op de productcyclus, gaat PIM de diepte in op productniveau. Denk aan kleuren, omschrijvingen, afmetingen, afbeeldingen en handleidingen.

Een PIM systeem vormt de informatiebron van je webshop, je ERP systeem en andere interne of externe systemen die productdata nodig hebben. Het wordt vaak gebruikt door retailers met veel productdata die behoefte hebben aan centralisatie en overzicht in databeheer van verschillende externe partijen. PIM geeft antwoord op alle *wat*-vragen over een product. Als dit helpt bij het organiseren van je informatie, is PIM de juiste data-oplossing voor jou.

### **PIM systeem speciaal voor MKB**

Mkb organisaties hebben bijvoorbeeld voor de informatievoorziening naar hun webshop veel baat bij een geavanceerd PIM systeem. Maar omdat PDM en PIM systemen vaak duur zijn in aanschaf en onderhoud, zijn ze alleen weggelegd voor grote bedrijven.

# Waarmee kunnen wij u helpen?

Laat ons weten wat u zoekt. Wij helpen u graag.

## **Zoekt u meer informatie over dit onderwerp?**

Bezoek dan de sectie over dit thema op de website [ICTinformatiecentrum.nl](http://ICTinformatiecentrum.nl) en vind daar alle informatie die gratis voor u beschikbaar is.

## **Zoekt u meer informatie over bedrijfssoftware?**

Gaat u een selectietraject van bedrijfssoftware starten, vraag dan via [ICTinformatiecentrum.nl](http://ICTinformatiecentrum.nl) de bijbehorende box aan met alle informatie over selectie, implementatie, het aansturen van het project, oplossingen, enz.

## **Zoekt u een oplossing, adviseur of leverancier?**

Bel, mail of chat dan even met ons. Dat is de snelste weg. Omschrijf voor welk IT vraagstuk u een oplossing zoekt en wij gaan voor u aan de slag.

## **Wilt u op de hoogte blijven van nieuws en ontwikkelingen?**

De ICT nieuwsbrief informeert u en 13.000 andere ICT beslissers en belangstellenden iedere twee weken over de meest relevante ontwikkelingen. Informatie, inspiratie, eyeopeners en noodzakelijke kennis.

## **Zoekt u iets anders?**

Onze websites spreken voor zich. Bezoek [ICTinformatiecentrum.nl](http://ICTinformatiecentrum.nl) om uw weg te vinden in alle thema's waarover wij informatie hebben. En als we u kunnen helpen, doen we dat graag. Bel of mail ons gerust!

ICT informatiecentrum, Houten | T 085 40 10 218 | [info@ictinformatiecentrum.nl](mailto:info@ictinformatiecentrum.nl)



# Kennispartners

De inhoud van dit boekje is tot stand gekomen met medewerking van de onderstaande kennispartners. Heeft u een vraag aan hen of over het thema van hun bijdrage, neemt u dan gerust contact met ze op.

## **Tentive**

Bavelselaan 1 | 4835 GP Breda

T +31(0)76 565 80 80 | [info@tentive.nl](mailto:info@tentive.nl) | [tentive.nl](https://tentive.nl)

---

## **Beeyond**

Schoutlaan 21 | 6002 EA Weert, Nederland

T +31(0)49 546 22 22 | [info@beeyond.nl](mailto:info@beeyond.nl) | [beeyond.nl](https://beeyond.nl)

---

## **Hot Item**

Danzigerkade 19 | 1013 AP Amsterdam

T +31(0)20 581 02 00 | [info@hotitem.nl](mailto:info@hotitem.nl) | [hotitem.nl](https://hotitem.nl)

---

## **In Summa Data Solutions**

Oeverkruid 15 | 4941 VV Raamsdonksveer

T +31 (0)16 252 40 40 | [info@datasolutions.nl](mailto:info@datasolutions.nl) | [datasolutions.nl](https://datasolutions.nl)

---

## **Cmotions**

Kosterijland 40 | 3981 AJ Bunnik

T +31(0)33 258 28 30 | [info@cmotions.nl](mailto:info@cmotions.nl) | [cmotions.nl](https://cmotions.nl)

---

## **SAP Nederland**

Amerikastraat 10 | 5232 BE 's-Hertogenbosch

T +31 (0)73 645 75 00 | [info@sap.com](mailto:info@sap.com) | [sap.com](https://sap.com)

**Digital Power**

H.J.E. Wenckebachweg 123-BG03 | 1096 AM Amsterdam

T +31 (0)20 308 43 90 | info@digital-power.com | digital-power.com

---

**Totta data lab**

Joan Muyskenweg 39 | 1114 AN Amsterdam

T +31 (0)20 514 13 28 | info@tottadatalab.nl | tottadatalab.nl

---

**FourPoints**

Anna van Buerenplein 41 | 2595 DA Den Haag

T +31(0)70 219 01 50 | info@fourpoints.nl | fourpoints.nl

---

**ISM eCompany**

Van Nelleweg 1 | 3044 BC Rotterdam

T +31 (0)10 243 60 00 | contact@ism.nl | ism.nl

---

**Graydon Nederland BV**

Hullenbergweg 250 | 1101 BV Amsterdam Zuidoost

T +31(0)20 567 99 99 | service@graydon.nl | graydon.nl

---

**ICTrecht**

Jollemanhof 12 | 1019 GW Amsterdam

T +31(0)20 663 19 41 | info@ictrecht.nl | ictrecht.nl

---