

MCTLL

'Het is een ernstige fout
te theoretiseren voor men gegevens heeft.'
(Sir Arthur Conan Doyle)

Hoofdstuk 5.3 Databeheer

V1.19.1 / 01 februari 2019

Auteur: Ton van den Hoogen

Met dank aan alle bedrijven en personen die in de afgelopen jaren bewust en onbewust een bijdrage aan MCTL hebben geleverd.

Tekstredactie: TekstFontein



Geen copyright!

MCTL is in licentie gegeven volgens een Creative Commons Naamsvermelding 3.0 Nederland licentie. Gebaseerd op een werk van www.mctl.nl.

MCTL is geheel Public Domain, er rusten dus *geen* copyrights of auteursrechten op. U mag MCTL (ook commercieel) gebruiken, verwerken, bewerken ... wat u maar wilt. Wanneer iets echter Public Domain is, blijft het Public Domain. Wat u dus niet mag doen is over (delen van) MCTL copyright of auteursrechten claimen, u maakt zich dan schuldig aan copyfraud en bent strafbaar. Indien u zelf overtredingen constateert, vragen wij u dit via www.mctl.nl aan ons te melden.

Wat wij van u vragen is om bij elk gebruik een verwijzing naar de bron: www.mctl.nl op te nemen. De reden hiervan is dat op deze wijze iedereen de oorspronkelijke versie(s) kan vinden.

MCTL – 5.3. Taakgebied Databeheer v1.19.1

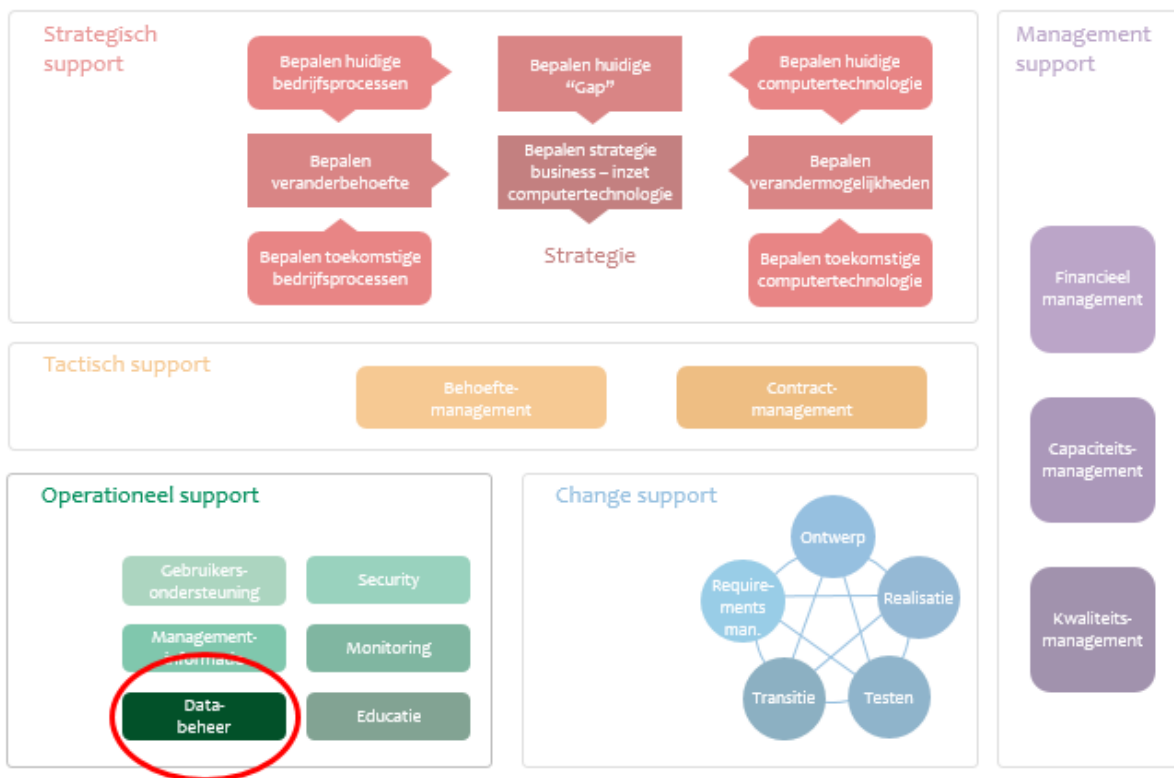
Hoofdstuk 5.3. Taakgebied Databeheer.....	4
Plaats in het MCTL-framework.....	4
Achtergrond	4
Doel van dit taakgebied	6
Hoe weet je dat het doel is bereikt?.....	6
De organisatie van databeheer	7
Verschillende soorten data	9
Taken	9
1. Aanpassen van data.....	10
2. Ondersteuning bij uitwisseling van data	11
3. Uitvoeren van controles op datakwaliteit.....	12
4. Uitvoeren taken big data / data science.....	13
Relaties met andere onderdelen van MCTL	15
Opmerkingen	16
1. Eigenaarschap en beheer van data	16
2. Niet-gedigitaliseerde data	17
3. Basisdata versus parameters	18
4. Big data / data science	18
5. Gestructureerde en ongestructureerde data	20
6. Integratie fysiek verschillende, maar logisch soortgelijke datasets	20
Certificering/proefexamenvragen	21
1. MCTL Foundation - proefexamenvragen	21
2. MCTL Foundation – proefexamenvragen met antwoorden en uitleg.....	22
3. MCTL Advanced-basis - proefexamenvragen	24
Nuttige websites en boeken	24

HOOFDSTUK 5.3. TAAKGEBIED DATABEHEER

Databeheer (en dan vooral het beheer van de in elektronische databases opgeslagen data) is een zaak van de eigenaar van de data: de gebruikersorganisatie. Dit taakgebied omschrijft hoe functioneel support en leveranciers dit beheer kunnen *ondersteunen*. Soms kan de gebruikersorganisatie het gedeeltelijk of geheel *delegeren* aan functioneel support of een leverancier. Laatstgenoemde kan een technologieleverancier zijn; voor het efficiënt beheren van data is computertechnologie noodzakelijk en degene die deze technologie levert, kan dan ook het beheer van de data op zich nemen. Er is nog een andere constructie mogelijk: de ene leverancier levert de computertechnologie, een andere neemt het databeheer op zich. In het laatste geval is sprake van een datamakelaar: een bedrijf dat het verzamelen, verrijken, leveren en verhandelen van data als core business heeft.

PLAATS IN HET MCTL-FRAMEWORK

Het taakgebied Databeheer maakt deel uit van het taakcluster Operationeel support:



ACHTERGROND

In dataopslag en -beheer is de rol van computertechnologie de afgelopen tientallen jaren steeds groter geworden. Al in 1995 verscheen van Nicholas Negroponte het boek 'Being digital'. Daarin huldigt hij het principe: 'Bits moeten geprefereerd worden boven atomen.'

MCTL – 5.3. Taakgebied Databeheer v1.19.1

Razorfish formuleert het net even anders: 'Alles wat digitaal kan worden, zal dat ook worden.' Op het gebied van cijfers en letters kan dat inderdaad voor 100% worden bereikt, evenals voor muziek, beeld en video. Terwijl dat voor bijvoorbeeld een auto of koffiemachine (in ieder geval op dit moment) tot de onmogelijkheden behoort: in bits kun je nu eenmaal niet rijden en bits kun je evenmin opdrinken.

De fases in ontwikkeling van opslag en beheer van data

1. Men behandelde digitale data aanvankelijk op dezelfde manier als fysieke data. Net zoals er voorheen voor de papieren versies decentrale dataverzamelingen bestonden, ontstonden die ook voor digitale data;
2. Omdat het maken van een digitale kopie weinig kost, werden in de tweede fase op meerdere plaatsen kopieën van originele datasets gemaakt. Dat had nogal wat vervelende gevolgen. Het bleek vrij lastig om alle kopieën gelijk te houden aan het origineel. Werd bovendien een fout in een kopie ontdekt, dan was het moeilijk te organiseren deze in het origineel en alle andere kopieën te laten bijwerken. Eigenaarschap en beheer waren voorheen zaken die op natuurlijke wijze werden geregeld. Data op papier waren slechts op één plek beschikbaar en het verschil tussen origineel en kopie was doorgaans heel duidelijk (vaak werd de kopie een "doorslag" genoemd, sommige lezers zullen zich het wellicht nog kunnen herinneren). Met de komst van de digitalisering werd dit echter een serieus issue;
3. Hierna is de derde fase ontstaan: die van de interne digitale centrale dataverzamelingen. In plaats van het naar verschillende plaatsen distribueren van data, worden deze op één centrale plaats in de organisatie neergezet. Vandaaruit worden ze via netwerken ter beschikking gesteld. Technisch is daarmee een fraaiere situatie ontstaan, doordat allerlei synchronisatieproblematiek tussen data op verschillende fysieke plaatsen verleden tijd is. Er is duidelijk wie de eigenaar van welke dataset is. Toch blijkt het vaststellen van eigenaarschap nog niet heel eenvoudig bij gecentraliseerde datasets die door veel organisatieonderdelen worden gebruikt. Bij dergelijke data is het lastiger greep te houden op datakwaliteit, consistentie in betekenis (verstaat iedereen hetzelfde onder hetzelfde gegeven) en aanpassingen (meer afstemming, meer verschillende wensen). Daar zit geen eigenaar op te wachten. De betreffende dataeigenaar kan het gevoel krijgen met allerlei problemen rondom het beheer van data opgescheept te zitten, terwijl de rest van de organisatie er alleen maar van profiteert;
4. Langzamerhand ontstaat de vierde fase waarin grote basisdataverzamelingen door een aantal verschillende organisaties/bedrijven gezamenlijk gebruikt gaan worden. Dat is bijvoorbeeld te zien bij persoons- en locatiedata, maar ook bij basistabellen als wettelijke percentages, postcodes en gemeentecodes.

In de publieke sector is het volgende Stelsel van Basisregistraties opgesteld:

- BRP - Basisregistratie personen (bestaat uit ingezetenen en niet-ingezetenen)
- NHR - Handelsregister

MCTL – 5.3. Taakgebied Databeheer v1.19.1

- BAG - Basisregistraties Adressen en Gebouwen (bestaat uit twee basisregistraties)
- BRT - Basisregistratie Topografie
- BRK - Basisregistratie Kadaster
- BRV - Basisregistratie Voertuigen (kentekenregister)
- BLAU - Basisregistratie voor Lonen, Arbeidsverhoudingen en Uitkeringen
- BRI - Basisregistratie Inkomen
- WOZ - Basisregistratie Waarde Onroerende Zaken
- BGT - Basisregistratie Grootchalige Topografie (voorheen GBKN)
- BRO - Basisregistratie Ondergrond (voorheen ook wel DINO)

In andere sectoren zijn identieke ontwikkelingen te zien. Zo bestaan in de zorgsector gezamenlijk gebruikte tabellen met behandelcodes en met data van contra-indicatie, van aandoeningen waarbij bepaalde medicijnen niet mogen worden toegepast en van medicijnen die elkaar negatief beïnvloeden (wisselwerking). Een term die in dit kader steeds meer wordt gebruikt is Master Data Management. Goed beschouwd is gezamenlijk gebruik van data buitengewoon efficiënt maar het leidt tegelijkertijd tot nieuwe, op te lossen problemen. Het wordt steeds lastiger om de betreffende dataeigenaar te bepalen. Bovendien zijn eigenaren steeds meer geneigd het databeheer uit handen te geven. Deze datasets bewandelen daarmee het pad dat we al gezien hebben bij hard- en software: via hosting, cloud-computing etc. wordt steeds minder op locatie zelf geregeld, maar vanuit een leverancier op maat aangeboden. Op deze laatste ontwikkeling wordt binnen dit taakgebied geanticipeerd.

Een buitengewoon aan te bevelen richtlijn is:

“Enkelvoudige registratie, meervoudig gebruik”

Het verzamelen/creëren, beheer, management en uiteindelijk weer verwijderen van data kost veel inspanning. Wat is dan meer voor de hand liggend dan in ieder geval geen data meervoudig vastleggen, maar aan de andere kant wel op zoveel mogelijk zinvolle plaatsen gebruiken?

DOEL VAN DIT TAAKGEBIED

Het ondersteunen van de eigenaar bij het beheer van data dan wel het namens de eigenaar beheren van data.

HOE WEET JE DAT HET DOEL IS BEREIKT?

De volgende indicatoren zijn te benoemen om te weten of bovenstaand doel is bereikt.

- Data blijken bij controle voor 99,9% correct en compleet;

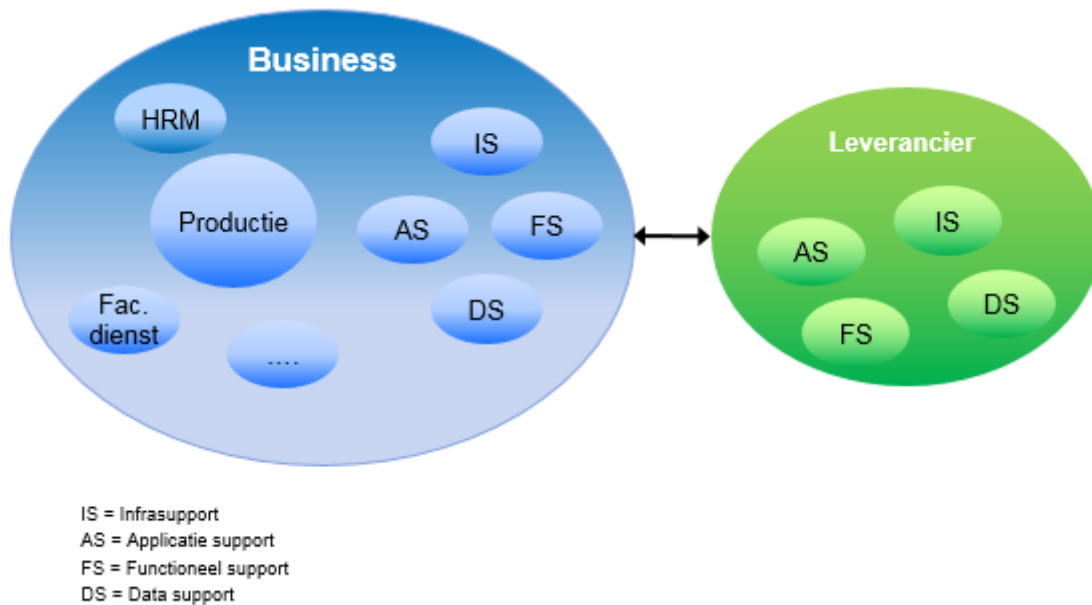
MCTL – 5.3. Taakgebied Databeheer v1.19.1

- Datagebruik blijkt bij controle juist;
- Onnodige dataduplicatie (zelfde data op meerdere fysieke plaatsen) komt in < 5% van de totale dataset voor.

(Genoemde percentages zijn indicatief.)

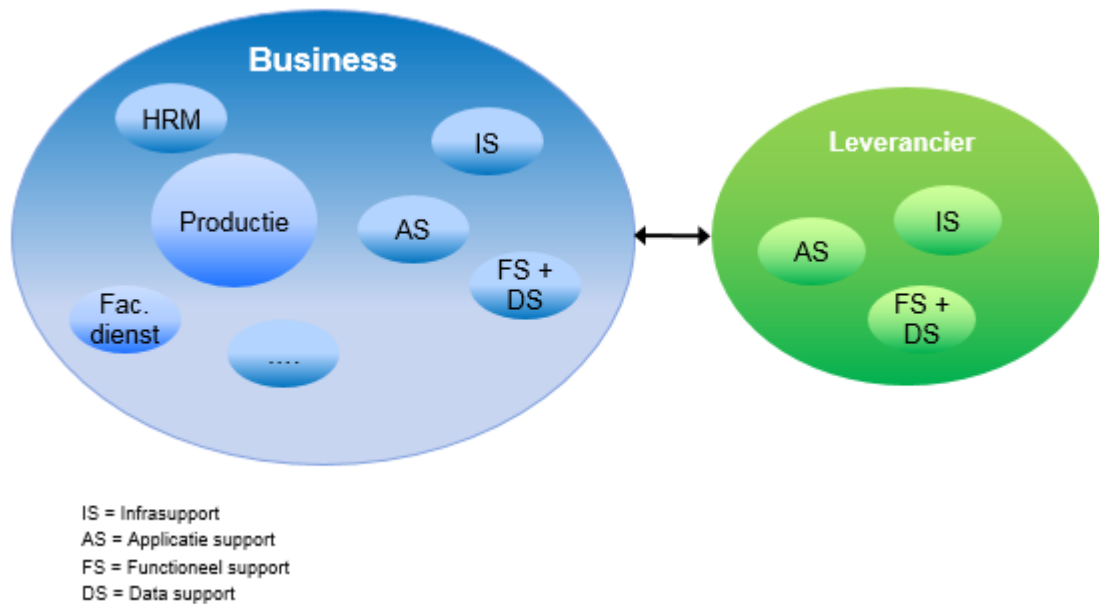
DE ORGANISATIE VAN DATABEHEER

Een illustratie om aan te geven hoe databeheer kan worden georganiseerd:



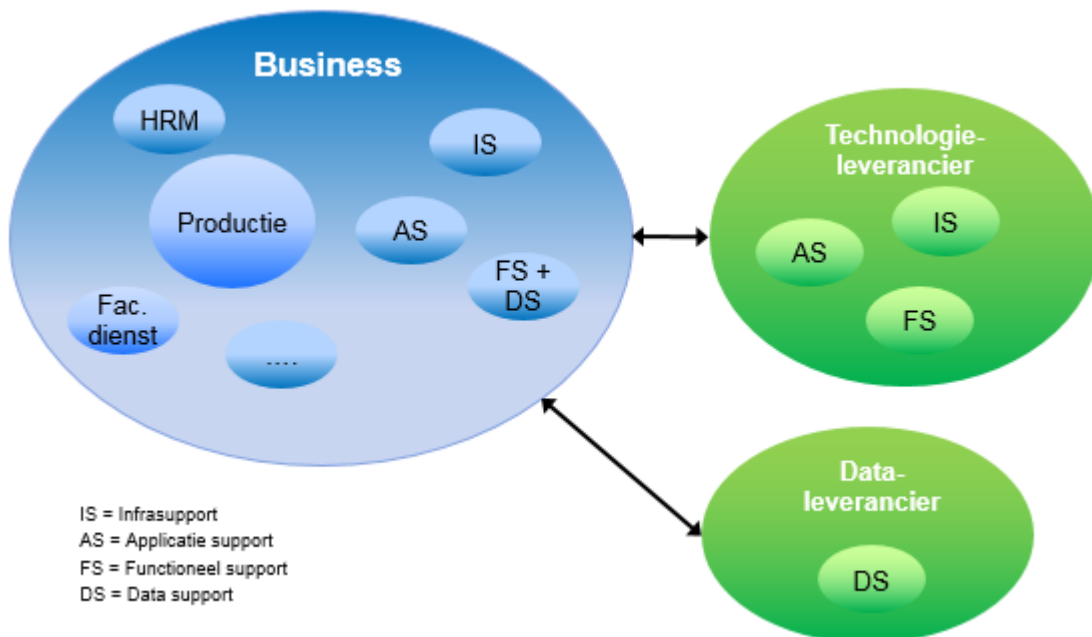
Het is mogelijk functioneel support en datasupport te combineren in één onderdeel (afdeling). Op die wijze ontstaat het volgende organisatieplaatje:

MCTL – 5.3. Taakgebied Databeheer v1.19.1



Hierbij moet worden opgemerkt dat de eigen organisatie en leveranciers natuurlijk los van elkaar staan: het kan zijn dat in de eigen organisatie FS en DS zijn gecombineerd, terwijl bij de leverancier(s) FS en DS zijn gesplitst, of andersom.

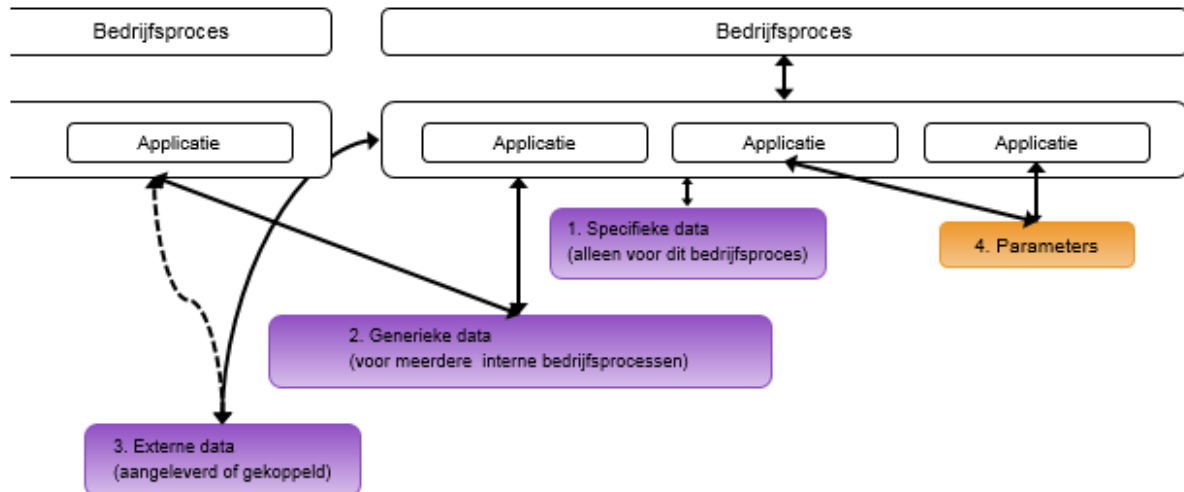
In het navolgende en laatste schema wordt extern onderscheid gemaakt tussen de leverancier(s) die de computertechnologie verzorg(t)(en) en de leveranciers van data.



MCTL – 5.3. Taakgebied Databeheer v1.19.1

VERSCHILLENDE SOORTEN DATA

Binnen een gemiddeld bedrijfsproces worden verschillende soorten data gebruikt. Onderstaande afbeelding illustreert dat.



Toelichting:

1. Allereerst bestaan er data die alleen voor een specifiek bedrijfsproces nodig zijn. Hierbij spelen de kwaliteit van en de relaties met andere data een rol;
2. Vervolgens zijn er generieke data, die in meerdere interne bedrijfsprocessen een rol spelen. Voorbeelden hiervan zijn klant- of orderdata. Hier spelen naast kwaliteit vaak autorisaties en eigenaarschap een rol (wie mag wat);
3. Tot slot bestaan er externe data die worden aangeleverd of via een koppeling ter beschikking staan. Voorbeelden hiervan zijn postcodedata. Hierbij spelen niet alleen kwaliteit en eigenaarschap een rol, maar ook de beveiliging en de tijdstippen waarop data ter beschikking staan en/of worden geactualiseerd;
4. Los hiervan kunnen parameters worden benoemd. Dit zijn ook data – zij het wat bijzondere – die vaak in een tabel binnen een systeem worden opgeslagen. Parameters dienen om applicaties op een bepaalde wijze in te stellen. MCTL bekijkt vooral de functionele instellingen. Voorbeelden hiervan zijn limieten, indeling van data op schermen en taalinstellingen. Aanpassing van parameters leidt tot een functionele wijziging en wordt daarom in het taakcluster Change support als wijziging verwerkt.

Digitale data worden nooit direct in een bedrijfsproces gebruikt, maar altijd via de *applicatielaag*. Ook BI-tools die, om de standaardapplicaties heen, data op een andere manier kunnen benaderen en presenteren zijn 'gewone' applicaties, zij het met een enigszins specifieke functionaliteit.

TAKEN

MCTL – 5.3. Taakgebied Databeheer v1.19.1

In zijn algemeenheid is te stellen dat Functioneel support en/of leverancier een rol hebben op het gebied van juistheid, tijdigheid en kwaliteit van data. Op databeheergebied moet daarnaast ook aandacht zijn voor een juiste en veilige fysieke opslag van data, het maken van back-ups etc. Dit is het aandachtsgebied van infra- en applicatiesupport en/of leveranciers (bijvoorbeeld bij cloud-oplossingen).

De taken in dit taakgebied zijn:

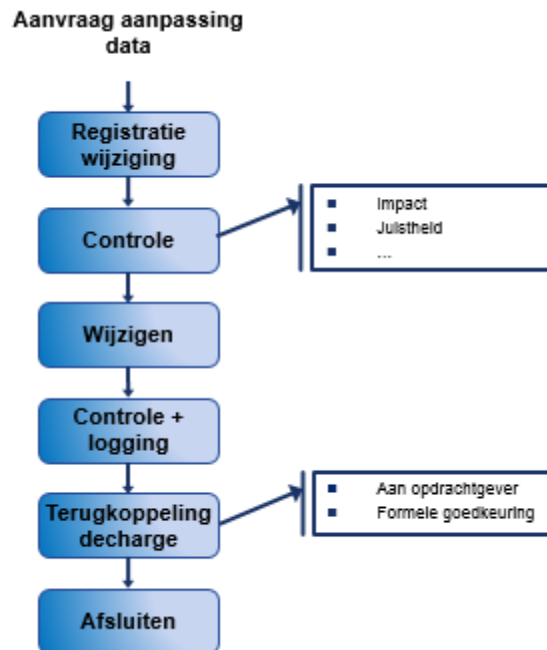
1. Aanpassen van data;
2. Ondersteuning bij uitwisseling van data;
3. Uitvoeren van controles op datakwaliteit;
4. Uitvoeren activiteiten op het gebied van Big Data / data science;
5. Ondersteuning bij archiveren en verwijderen van data.

De taken worden hierna uitgewerkt.

1. AANPASSEN VAN PRODUCTIEDATA

Het aanpassen van productiedata is normaal gesproken een taak van de gebruikers. Dat gebeurt via de normale werkprocedures en systemen. In uitzonderlijke gevallen komt het voor dat data incorrect in de database zijn opgenomen en niet via de normale werkprocedure en het normale systeem te herstellen zijn. In dat soort situaties moet 'om het systeem heen' (vaak rechtstreeks in de database) een aanpassing worden gedaan. Dat moet uiteraard zoveel mogelijk voorkomen worden. Systemen (applicaties) kunnen enerzijds zo worden gemaakt dat zo min mogelijk foute data in databases terechtkomen (invoercontroles). Anderzijds moeten gebruikers veel mogelijkheden krijgen om zelfstandig foutieve data te corrigeren. Toch blijft de kans bestaan dat onjuiste data in de database niet door gebruikers zelf kunnen worden hersteld. Op onderstaande gestructureerde en gecontroleerde wijze kunnen dergelijke foutieve data in de databases worden aangepast.

MCTL – 5.3. Taakgebied Databeheer v1.19.1



De in het schema genoemde acties zijn duidelijk en worden hier niet verder toegelicht. Doorgaans voert de functioneel specialist deze activiteiten uit. Het schema volgt een tamelijk formele werkwijze. Ook indien binnen een bedrijf doorgaans een wat lossere werkwijze wordt gevolgd, is het zaak deze uit te voeren acties relatief volgens een strak regime uit te voeren. Het betreft hier immers productiedata en bijvoorbeeld traceerbaarheid (wie heeft wat, waar en wanneer gedaan) is dan vanzelfsprekend van belang.

Deze acties moeten worden uitgevoerd door personen die:

- Kennis hebben van de benodigde tools (waarmee rechtstreeks in de database wijzigingen kunnen worden aangebracht);
- Kennis hebben van de structuur van de data (om ongewenste consequenties te voorkomen);
- De benodigde autorisaties hebben om dergelijke activiteiten uit te voeren.

Dergelijke acties zijn gelukkig vrij uitzonderlijk. Bij een aanpak zoals in bovengenoemd schema geschetst, zullen ze de goedkeuring van een interne controller of externe auditor kunnen krijgen. Ten overvloede: data- of functioneel support kan hier een ondersteunende rol spelen, maar uiteindelijk blijft de gebruikersorganisatie verantwoordelijk voor de data, ook als deze op genoemde wijze zijn aangepast.

2. ONDERSTEUNING BIJ UITWISSELING VAN DATA

Zoals al eerder betoogd, is het niet ongevoerd dat data worden uitgewisseld (of gezamenlijk gebruikt):

1. Tussen verschillende onderdelen binnen een bedrijfsproces;

2. Tussen verschillende bedrijfsprocessen in de eigen organisatie;
3. Tussen bedrijfsprocessen in verschillende organisaties.

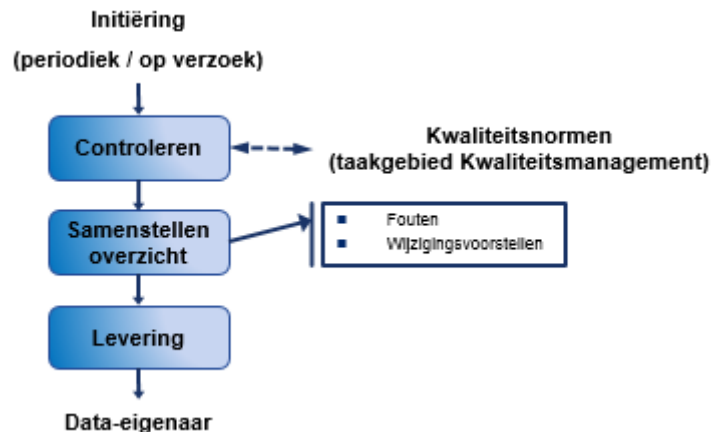
De eigenaren van de systemen en data zijn zelf verantwoordelijk voor deze uitwisseling en/of dit gezamenlijke gebruik. Men noemt dit ook wel *proceseigenaarschap* of *keteneigenaarschap*. Een bedrijfsproces met de bijbehorende data kan echter zo complex zijn dat ondersteuning bij het uitwisselen van data noodzakelijk is. Functioneel support kan dan ondersteunende diensten bieden op het vlak van data-uitwisseling en gezamenlijk datagebruik binnen die bedrijfsprocessen.

Functionele ondersteuning kan bijvoorbeeld nodig zijn indien data uit een externe bron dubbel is aangeleverd, de datastroom plotsklaps “opdroogt” of andere afwijkende karakteristieken vertoont. Het gaat in dit taakgebied dan niet om de monitoring, dat is onderdeel van taakgebied Monitoring, maar wel om de concrete acties die moeten worden ondernomen. Bijvoorbeeld het ontdebelen van data kan een ingewikkelde kwestie zijn die functioneel support eventueel samen met bijvoorbeeld applicatie support uitvoert. Een laatste voorbeeld dat hier ter illustratie kan dienen is het geval dat een datastroom met een doorlopende nummering opeens “gaten” vertoont. Op (nog) onverklaarbare manier ontbreken bepaalde nummers. Ook het analyseren en oplossen van dergelijke situaties kan een functioneel uitdagende taak zijn.

3. UITVOEREN VAN CONTROLES OP DATAKwaliteit

De verantwoordelijkheid voor de kwaliteit van de data ligt bij de eigenaar ervan. Toch kan data-/functioneel support een ondersteunende rol spelen. Het is vaak mogelijk diverse (kruis)controles uit te voeren op data, zodat fouten in die data worden geconstateerd. Daarnaast kunnen fouten worden achterhaald doordat in het ene bedrijfsproces data anders worden gebruikt dan in het andere. In die situatie kan data-/functioneel support een coördinerende rol spelen. In het geval van een ogenschijnlijk incidentele fout kan een databeheerder een nader onderzoek doen naar een eventueel structureel karakter van de betreffende fout.

Het zal duidelijk zijn dat vanuit data-/functioneel support nuttige kwaliteitscontroles te ondernemen zijn om de eigenaar van de data te ondersteunen. Het volgende schema geeft de uit te voeren acties aan.



De eigenaar van de data ontvangt een overzicht met constatering. Indien de betreffende dataset wordt onderhouden door een leverancier, dan krijgt die uiteraard ook dat overzicht.

4. UITVOEREN ACTIVITEITEN OP HET GEBIED VAN BIG DATA / DATA SCIENCE

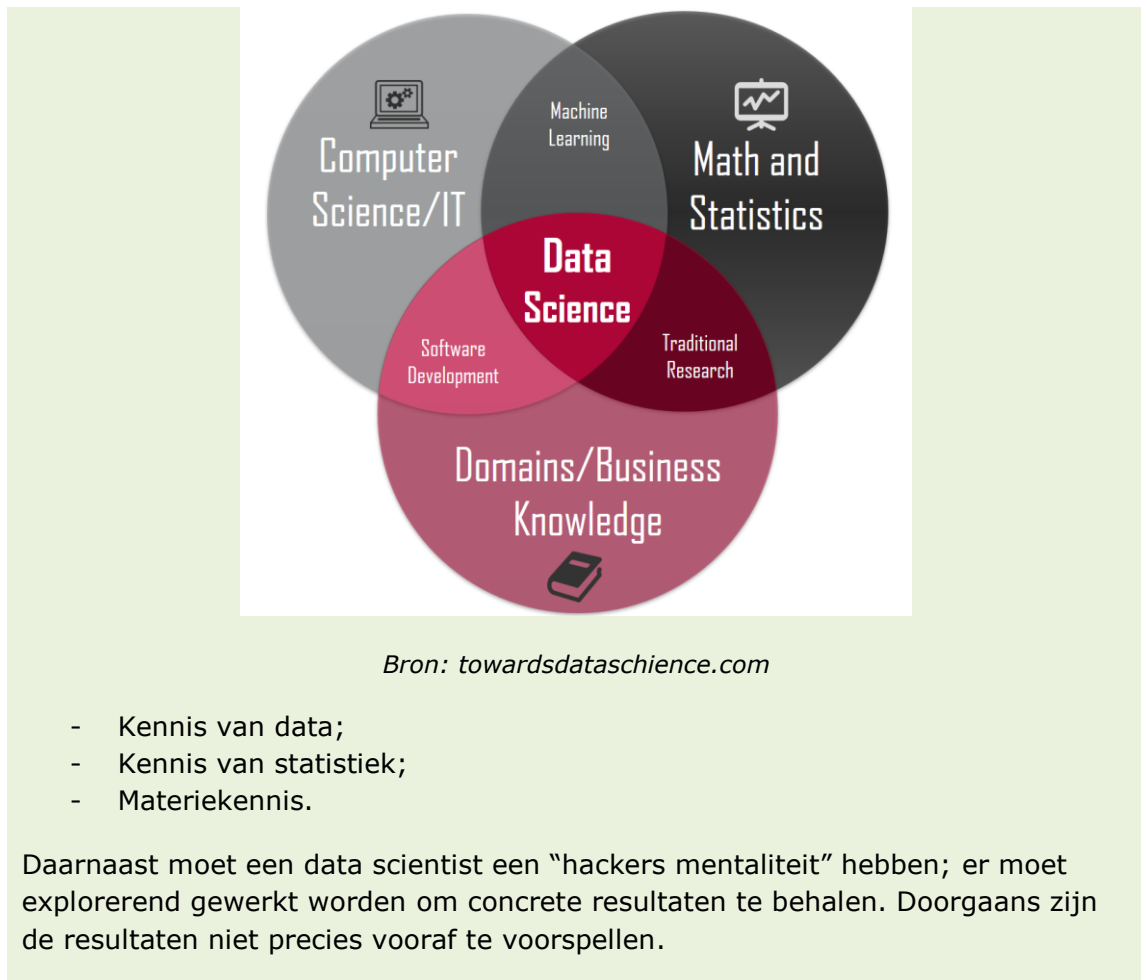
Big Data en Data science is een gebied in ontwikkeling. Activiteiten in dit gebied zijn:

1. Aggregatie van data;
2. Preparatie van data;
3. Analyse;
4. Presentatie van resultaten;
5. Advisering/conclusies op basis van resultaten.

Functioneel specialisten kunnen hierin een goede bijdrage leveren. Sterk punt van functioneel specialisten is immers dat zowel domeinkennis (van de organisatie en de bedrijfsprocessen) als kennis van de data aanwezig is. In dit vakgebied worden deze twee gecombineerd met statistische analyses, wat een vakgebied op zichzelf is. Functioneel specialisten kunnen vanuit hun positie aldus samen met andere specialisten dit gebied invullen.

Data scientist

Een specialist die vaak wordt genoemd in het kader van Big Data / Data science is de data scientist. In onderstaand plaatje is goed te zien welke expertise deze functionaris moet bezitten:



6. ONDERSTEUNING BIJ ARCHIVEREN EN VERWIJDEREN VAN DATA

Data heeft een bepaalde periode van bruikbaarheid. Bovendien moet sommige data na gebruik een bepaalde tijd worden bewaard, aan de andere kant zijn er soms juist ook wettelijke maxima aan die bewaartermijnen gesteld. Ook kan in verband met privacywetgeving soms een verzoek worden gedaan om data te wissen, bijvoorbeeld in het geval van een patiënt over zijn medische gegevens.



Bron: CTC Global Pte

In het algemeen kent data een levenscyclus die loopt van het verzamelen van data, de opslag, beheer en gebruik fase tot de fase van archivering waarbij het actieve gebruik daalt naar 0 en uiteindelijk het verwijderen van data. Het is vaak applicatie en infrasupport die in beeld komt op data te schonen. Maar de zeggenschap over archivering, bewaartermijnen en het bijbehorende schonen of niet schonen van data ligt vanzelfsprekend bij de eigenaar van die data. Vanwege de soms lange bewaartermijnen en wisselingen in de organisatiestructuur en eigenaarschap kan dit echter letterlijk uit beeld verdwijnen; er is dan data waarvan de status onduidelijk is geworden.

Functioneel support kan een rol hebben om zicht te houden op het totaal van alle data in de organisatie en de status ervan. Ze kan de gebruikersorganisatie ondersteunen in het nemen van de juiste beslissingen en namens de gebruikersorganisatie richting applicatie en infrasupport de juiste acties uitzetten.

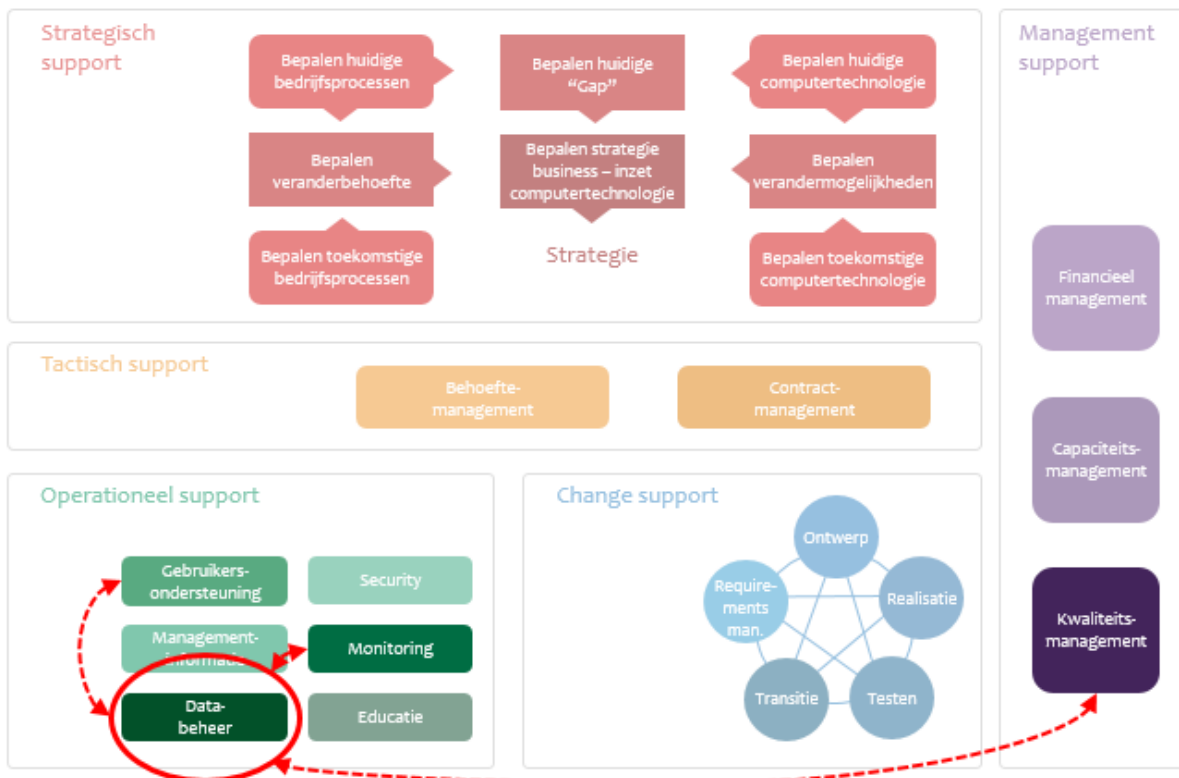
Data voorzien van "te gebruiken tot"

De combinatie van een toenemende hoeveelheid data, de soms lage gestructureerdheid ervan en de lange bewaartermijnen geeft de nodige problemen. Een niet onaardige oplossing voor de lange termijn zou kunnen zijn om elk logisch samenhangend stukje data te voorzien van een digitale stempel "te gebruiken tot". Na het verstrijken van deze datum wordt dit stukje data automatisch verwijderd. Overigens zou, maar dat is een onderwerp dat meer past bij Security, op dezelfde wijze een digitaal stempel kunnen worden geplaatst "te gebruiken door", waardoor de beveiliging op dataniveau wordt geregeld.

RELATIES MET ANDERE ONDERDELEN VAN MCTL

Dit taakgebied kent de volgende belangrijke relaties.

MCTL – 5.3. Taakgebied Databeheer v1.19.1



Ten eerste is er een relatie met **Gebruikersondersteuning**. Daar zullen aanvragen betreffende het aanpassen van data binnenkomen. Ten tweede is er een relatie met het proces **Monitoring** omdat uitwisseling van data tussen verschillende systemen (en vooral met systemen buiten de eigen organisatie) moet worden gemonitord. Eventueel benodigde ingrepen vinden vervolgens in dit taakgebied (Databeheer) plaats.

Tot slot is er een belangrijke relatie met **Kwaliteitsmanagement**. In dat taakgebied zijn kwaliteitsnormen opgesteld, vanzelfsprekend ook op het gebied van datakwaliteit. Aan de hand van deze kwaliteitsnormen kunnen controles op de data worden uitgevoerd waarmee de gebruikersorganisatie, als eindverantwoordelijke voor de kwaliteit van de data, haar voordeel kan doen.

OPMERKINGEN

De volgende opmerkingen zijn over het taakgebied Databeheer te maken.

1. EIGENAARSCHAP EN BEHEER VAN DATA

Data zijn in principe eigendom van de gebruikersorganisatie. Bij gemeenschappelijk gebruikte data ontstaat de vraag: waar precies moet het eigenaarschap terechtkomen? Veel gekozen oplossingen zijn:

MCTL – 5.3. Taakgebied Databeheer v1.19.1

1. De afdeling die het meeste belang heeft bij de (juistheid) van de data, wordt eigenaar;
2. De afdeling die de data het meest gebruikt wordt eigenaar;
3. De meest voor de hand liggende afdeling wordt eigenaar. De afdeling financiële administratie wordt bijvoorbeeld altijd eigenaar van financiële data;
4. Er wordt een aparte afdeling opgericht (data support) die het beheer van gemeenschappelijke data voor haar rekening neemt.

Vervolgens moet het beheer worden belegd. Vaak kiest men ervoor het eigenaarschap en het uitvoerend beheer te combineren, maar dat hoeft niet. Het beheer van data kan worden gedelegeerd aan anderen, onder andere aan een leverancier.

Een verdere kwestie is of datasupport een onderdeel van functioneel support zou moeten zijn. Data vormen geen functionaliteit, bieden die ook niet, en dus is hierop geen Functioneel support van toepassing. Aldus zou naast infra-, applicatie en functioneel support een aparte eenheid datasupport kunnen worden ingericht. Is dit organisatorisch niet wenselijk, dan is de samenvoeging van data- en functioneel support het meest logisch.

Beheer betekent niet dat de eindverantwoordelijkheid kan worden overgenomen door de beheerder. Dataconversies kunnen bijvoorbeeld worden uitgevoerd onder auspiciën van de beheerder; de eigenaar blijft inhoudelijke verantwoordelijk.

De verdeling taken tussen de gebruikersorganisatie en functioneel support is visueel als volgt weer te geven:



2. NIET-GEDIGITALISEERDE DATA

Naast de in dit taakgebied geschetste data die in digitale vorm zijn opgeslagen in databases op computers, bestaan er niet-gedigitaliseerde data (ook wel met de niet zo fraaie term *niet-geautomatiseerde informatievoorziening* aangeduid. Deze term wordt in MCTL niet gebruikt). Zolang deze niet-gedigitaliseerde data nog bestaan, moet de relatie met de gedigitaliseerde data worden onderhouden. Een voorbeeld is een ziekenhuis waar alle

patiëntdata in een elektronisch dossier zijn opgenomen, maar toch nog een papieren dossier bestaat.

Omdat men ervan uitgaat dat 'alles wat digitaal kan worden, ook digitaal zal worden,' zijn niet-gedigitaliseerde data een aflopende zaak. Het is dus een goede zaak daar zo min mogelijk tijd, geld en energie in te steken.

3. BASISDATA VERSUS PARAMETERS

MCTL ziet een parameter als een instelling waarmee de functionele werking van een systeem binnen het bedrijfsproces kan worden beïnvloed. Voorbeelden zijn een instelling waarmee het maximumaantal rollen per werknemer kan worden ingesteld, de presentatie van data op een scherm en de volgorde waarin op het scherm door deze data kan worden gelopen.

Basisdata als tarieven, afdelingscoderingen etc. beïnvloeden weliswaar de uitkomsten van bedrijfsprocessen (zo geeft een ander btw-percentages andere eindbedragen), maar veranderen de functionele werking van het systeem niet. Daarom beschouwt MCTL dergelijke data als 'gewone' data, die wel specifiek (generiek of extern) voor een bedrijfsproces aangeleverd kunnen zijn (zie ook het schema aan het begin van dit taakgebied). Basisdata worden ook wel *stamgegevens* genoemd (in MCTL wordt de term stamgegevens niet gebruikt).

De gebruikersorganisatie kan basisdata zelf wijzigen. Indien hulp op dit vlak gewenst is, kan data-/functioneel support haar via dit taakgebied ondersteunen. Parameterwijzigingen worden níét in dit taakgebied uitgevoerd. Dergelijke wijzigingen betreffen een aanpassing van de functionaliteit. Zij worden daarom afgehandeld via het taakcluster Change support.

4. BIG DATA / DATA SCIENCE

In het kader van databeheer mag big data niet ontbreken. Big data kenmerkt zich door:

- Volume (hoeveelheid) – omvang van de gegevensverzameling;
- Variety (diversiteit) – in soorten gegevens;
- Velocity (snelheid) – in (tijdelijke) waarde van gegevens.

Daarnaast worden de volgende kenmerken vaak toegevoegd:

- Value (waarde) – bijv. aan de hand van persona's waarde aan gegevens geven;
- Veracity (kwaliteit) – kwaliteit en oorsprong gegevens;
- Variability (variatie) – hoever en hoe snel is de structuur van de data te veranderen?

Big data/data science kent vaak een focus op specifieke te bereiken doelen. Echter, ook het zoeken naar onverwachte relaties in data is een activiteit die veel kan opleveren. Helaas zijn daarvan de resultaten vooraf natuurlijk niet voorspelbaar. Juist uitgevoerd databeheer zodat een juiste dataset als basis ontstaat is vanzelfsprekend essentieel.

Toepassing van Big Data

Toepassing van Big Data staat pas in de kinderschoenen. Een paar voorbeelden van toepassing zijn:

Het voorspellen van een griepepidemie aan de hand van concrete zoekcriteria per regio

Het voorspellen van drukte in een restaurant aan de hand van telefoongegevens. Aan de hand van deze gegevens is zelfs goed te bepalen wat het aantal zitplaatsen in een restaurant is; als er op bepaalde momenten mobieltjes weer "weglopen" is kennelijk het maximum bereikt.

Wordt Big Data/data science ook opdrachtgever?

Aan de hand van Big Data / data science zijn verbetermogelijkheden te identificeren. Op dit moment is er nog een menselijke beoordeling; willen we deze verbeteringen wel of niet? Een niet geheel onrealistische gedachte is dat in de toekomst Big Data / data science dusdanige uitkomsten geeft dat zonder tussenkomst van een menselijke opdrachtgever (systeem-eigenaar) deze verbeteringen worden uitgevoerd.

Gebruik in- en externe data

Bij toepassing van Big Data / data science wordt gebruik gemaakt van zowel in- als externe data. Problemen met externe data kunnen zijn dat deze met een ander doel zijn verzameld en verwerkt. Daarmee zijn ze niet meer neutraal; er zijn aannames en overtuigingen in die data verwerkt. Een ander veel voorkomend probleem is dat externe data op een ander abstractie niveau beschikbaar zijn, bijvoorbeeld in plaats van leeftijd is de data dan beschikbaar in bepaalde leeftijdscategorieën (0-12 jaar, 13-20 jaar, etc.). Dit maakt de combinatie van verschillende datasets lastig of onmogelijk.

Welke data is er, en welke juist niet?

Soms is ontbrekende data minstens zo interessant als de data die wel ter beschikking staat. Bijvoorbeeld van een enquête ontbreekt veel data uit een

bepaalde postcode. Ook dat kan na analyse mogelijk een interessante conclusie opleveren!

5. GESTRUCTUREERDE EN ONGESTRUCTUREERDE DATA

Gestructureerde data (opgeslagen in databases en daarbinnen in tabellen en records) krijgt doorgaans voldoende aandacht in beheer en op het gebied van security. Daarnaast bestaat ook ongestructureerde data (in de vorm van bijv. brieven, tekeningen, memo's, e-mailtjes, foto's, video, audio). Laatstgenoemde is vaak lastiger te beheren. Bijvoorbeeld alleen al het tijdig schonen van dergelijke data is lastig omdat de oorsprong soms onduidelijk is, de waarde moeilijk te bepalen en ook de bewaartermijn kan een ingewikkelde zaak zijn. Bijvoorbeeld van facturen is het in verband met wettelijke bewaartermijnen doorgaans een heldere zaak. Echter, stel dat een video wordt gemaakt waarin een bepaalde casus wordt besproken en intern wordt gebruikt voor trainingsdoeleinden. Hoelang mag/moet deze worden bewaard? Of stel dat in een interne opleiding digitaal audiomateriaal wordt verwerkt van een derde partij met een licentie. Op welke manier mogen dan kopieën en back-ups worden gemaakt, onderdelen worden hergebruikt, verknipt, in hoeverre mag het materiaal van de derde partij intern worden bewerkt, wat gebeurt er als de licentie verloopt? Ongestructureerde data vergt, zo zal duidelijk zijn, de nodige aandacht vanuit beheer. Van alle ongestructureerde data zou in ieder geval bekend moeten zijn:

- Oorsprong;
- Type data (tekst, afbeelding, video, audio, virtual reality (vr), animatie);
- Aard data (privacygevoelig, bron- of afgeleide data etc.);
- Eigenaar;
- Bewaartermijn;
- Samenstelling.

Op basis van deze gegevens kan samen met de eigenaar het beheer worden geregeld.

Toepassing van ongestructureerde data

Om ongestructureerde data hangt altijd het zweem van geringe(re) betrouwbaarheid dan gestructureerde data. Maar stel dat u wilt weten waar iemand woont. Wat is dan het meest actueel en betrouwbaar: de gegevens in de Basisregistratie personen van de gemeente, of de locatiegegevens van een mobiele telefoon?

6. INTEGRATIE FYSIEK VERSCHILLENDE, MAAR LOGISCH SOORTGELIJKE DATASETS

Bij een fusie of overname of door omstandigheden die vaak een oorzaak in het verleden hebben, moeten diverse datasets soms worden geïntegreerd tot een set. Dit vergt altijd bijzondere inspanning van applicatie- en functioneel support, eventueel van leveranciers en van de eigenaar van deze data. Conversie en automatische ontdebblingen kunnen heel wat handwerk voorkomen. Het blijft de eigenaar die er uiteindelijk voor moet (laten) zorgen dat een dataset van een toereikend kwaliteitsniveau ontstaat.

CERTIFICERING/PROEFEXAMENVRAGEN

Voor MCTL kunt u zich certificeren op *foundation*, *advanced* en *expert* niveau. Het foundationniveau toetst uw kennis van MCTL. Het advanced en expert level toetsen uw vaardigheid in het toepassen van MCTL. In een apart onderdeel, 'MCTL Certificering', vindt u alle informatie over de drie niveaus. Hierna vindt u proefexamenvragen op foundationniveau. Aansluitend treft u een aantal vragen aan op advanced-basisniveau.

1. MCTL FOUNDATION - PROEFEXAMENVRAGEN

Voor dit hoofdstuk zijn de volgende proefexamenvragen beschikbaar. Maak deze zonder terug te bladeren. De uitleg vindt u direct hierna.

8-1. Het doel van Databeheer is:

- het voor de leverancier intern beheren van data.
- het namens de eigenaar beheren van data.
- het beheren van data, ongeacht voor of van wie die data zijn.
- het ondersteunen van de eigenaar bij het beheer van data (of: het namens de eigenaar beheren van data).

8-2. Binnen Databeheer wordt onderscheid gemaakt tussen externe en interne data omdat:

- het eigenaarschap van externe data buiten de organisatie ligt en daar bij het beheer rekening mee gehouden moet worden.
- de data een ander formaat kunnen hebben.
- externe data een andere betekenis kunnen hebben dan interne data.
- interne data eenvoudiger kunnen worden opgesplitst of samengevoegd.

8-3. Indien een functioneel specialist door de gebruikersorganisatie wordt gevraagd enkele data aan te passen, dan:

- doet die dat nooit, maar helpt hij/zij de vraagsteller het zelf te doen.
- helpt hij/zij de vraagsteller het zelf te doen. Indien nodig past de functioneel specialist ook zelf data aan.
- verwijst de functioneel specialist de vraagsteller naar een applicatiespecialist, die verder zal helpen.
- vraagt de functioneel specialist de key-user waarom hij/zij dit niet heeft opgepakt.

8-4. Verantwoordelijk voor de kwaliteit van data afkomstig van een externe bron:

- is de functioneel specialist die bij die externe bron verantwoordelijk is voor het systeem aldaar.

- b. is de eigenaar van die data.
- c. is degene die bij het inlezen van die data een kwaliteitscontrole uitvoert.
- d. zijn het management van de interne en van de externe organisatie samen.

8-5. Worden functionele parameters binnen Databeheer aangepast, wanneer dat nodig is?

- a. Ja, functionele parameters zijn ook data en vallen daarom onder Databeheer.
- b. Ja, functionele parameters zijn wel bijzondere data, maar worden indien nodig aangepast binnen Databeheer.
- c. Nee, het aanpassen van functionele parameters is een simpel service request en valt daarom onder Gebruikersondersteuning.
- d. Nee, het aanpassen van functionele parameters is een wijziging en valt daarom onder Change support.

2. MCTL FOUNDATION – PROEFEXAMENVRAGEN MET ANTWOORDEN EN UITLEG

Hierna vindt u de proefexamenvragen direct daarachter de antwoorden en uitleg.

8-1. Het doel van Databeheer is:

- a. het voor de leverancier intern beheren van data.
 - b. het namens de eigenaar beheren van data.
 - c. het beheren van data, ongeacht voor of van wie die data zijn.
 - d. het ondersteunen van de eigenaar bij het beheer van data (of: het namens de eigenaar beheren van data).
- a. Onjuist. Leveranciers zijn in het algemeen geen dataeigenaars en daarom worden voor hen geen data beheerd.
- b. Onjuist. Dit is slechts een deel van het antwoord.
- c. Onjuist. Het maakt wel degelijk uit van wie data zijn.
- d. Juist. Dit is precies de definitie die MCTL hanteert. Zie hoofdstuk 8.

8-2. Binnen Databeheer wordt onderscheid gemaakt tussen externe en interne data omdat:

- a. het eigenaarschap van externe data buiten de organisatie ligt en daar bij het beheer rekening mee gehouden moet worden.
 - b. de data een ander formaat kunnen hebben.
 - c. externe data een andere betekenis kunnen hebben dan interne data.
 - d. interne data eenvoudiger kunnen worden opgesplitst of samengevoegd.
- a. Juist. Het eigenaarschap heeft een rol in het beheer en zodra data op een andere plaats terechtkomen, moet daar ook beheer worden geregeld. Zie hoofdstuk 8.
- b. Onjuist. Data kunnen een ander formaat hebben, maar dat kan intern evengoed het geval zijn. Bovendien is het een technische aangelegenheid, geen functionele.
- c. Onjuist. Data kunnen een verschillende betekenis hebben, maar de bron van data maakt daarbij vanuit beheer gezien geen principiële verschil.
- d. Onjuist. Voor splitsing of samenvoeging maakt de bron van de data niet uit.

8-3. Indien een functioneel specialist door de gebruikersorganisatie wordt gevraagd enkele data aan te passen, dan:

- a. doet die dat nooit, maar helpt hij/zij de vraagsteller het zelf te doen.
- b. helpt hij/zij de vraagsteller het zelf te doen. Indien nodig past de functioneel specialist ook zelf data aan.
- c. verwijst de functioneel specialist de vraagsteller naar een applicatiespecialist, die verder zal helpen.
- d. vraagt de functioneel specialist de key-user waarom hij/zij dit niet heeft opgepakt.

- a. Onjuist. Functioneel specialisten kunnen en mogen wel degelijk data aanpassen.
- b. Juist. In principe moet de gebruikersorganisatie dit zelf doen, eventueel met enige hulp. Indien nodig kan een functioneel specialist data aanpassen, maar wel conform een bepaalde procedure zodat bijvoorbeeld later is terug te zien welke data op welk moment zijn aangepast. Zie hoofdstuk 8.
- c. Onjuist. Een applicatiespecialist heeft wel bemoeienis met databases (zoals met de structuur ervan), maar niet met de inhoud.
- d. Onjuist. Een key-user kan eventueel ook een gedelegeerde rol hebben in het aanpassen van data, maar er is geen sequentie (eerst key-user, daarna functioneel specialist) noodzakelijk.

8-4. Verantwoordelijk voor de kwaliteit van data afkomstig van een externe bron:

- a. is de functioneel specialist die bij die externe bron verantwoordelijk is voor het systeem aldaar.
- b. is de eigenaar van die data.
- c. is degene die bij het inlezen van die data een kwaliteitscontrole uitvoert.
- d. zijn het management van de interne en van de externe organisatie samen.

- a. Onjuist. Die functioneel specialist kan de eigenaar van die data ondersteunen, maar is nooit verantwoordelijk.
- b. Juist. De eigenaar van die data is verantwoordelijk. Zie hoofdstuk 8.
- c. Onjuist. De eigenaar van de brondata is verantwoordelijk voor de kwaliteit. Een ingangscntrole kan heel nuttig zijn om eventuele kwaliteitsproblemen tijdig te signaleren.
- d. Onjuist. Het zou kunnen dat het management eigenaar is van de data en daarmee verantwoordelijk voor de kwaliteit, maar dit betreft dan wel het management bij die externe bron.

8-5. Worden functionele parameters binnen Databeheer aangepast, wanneer dat nodig is?

- a. Ja, functionele parameters zijn ook data en vallen daarom onder Databeheer.
- b. Ja, functionele parameters zijn wel bijzondere data, maar worden indien nodig aangepast binnen Databeheer.
- c. Nee, het aanpassen van functionele parameters is een simpel service request en valt daarom onder Gebruikersondersteuning.
- d. Nee, het aanpassen van functionele parameters is een wijziging en valt daarom onder Change support.

- a. Onjuist. Functionele parameters zijn weliswaar data, maar aanpassing ervan leidt tot een functionele wijziging en wordt daarom in Change support afgehandeld.
- b. Onjuist. Functionele parameters zijn weliswaar bijzondere data, maar aanpassing ervan leidt tot een functionele wijziging en wordt daarom in Change support afgehandeld.

MCTL – 5.3. Taakgebied Databeheer v1.19.1

- c. Onjuist. Functionele parameters aanpassen valt buiten de scope van een Service request, en kan daarom niet in Gebruikersondersteuning worden afgehandeld.
- d. Juist. Aanpassing van functionele parameters leidt tot een functionele wijziging en wordt daarom in Change support afgehandeld. Zie hoofdstuk 8.

3. MCTL ADVANCED-BASIS - PROEFEXAMENVRAGEN

Voor dit hoofdstuk zijn de volgende proefexamenvragen op advanced-basisniveau beschikbaar. Het zijn open vragen waarop u de antwoorden in de tekst van dit hoofdstuk kunt terugvinden. Om veel herhaling te voorkomen is daarom hier geen aparte uitleg per vraag opgenomen.

Vraag 1 (5 punten): Geef een verschil aan tussen specifieke en generieke data.

Vraag 2 (5 punten): Hoe verhoudt taakgebied Databeheer zich tot taakgebied Monitoring als het gaat om externe datauitwisseling?

Vraag 3 (5 punten): Wanneer wordt data door functioneel support aangepast?

Vraag 4 (5 punten): Is functioneel support verantwoordelijk voor de kwaliteit van de data? Zo nee, wie is dan verantwoordelijk?

Vraag 5 (5 punten): Worden parameters in het taakgebied Databeheer aangepast? Zo nee, geef dan aan waar dit wel plaatsvindt.

NUTTIGE WEBSITES EN BOEKEN

Vanuit functioneel perspectief zijn de volgende websites interessant voor het taakgebied Databeheer.

- www.mctl.nl
MCTL.nl – Website met alle informatie over MCTL; de achtergrond, een beschrijving van het model, video's, artikelen, etc. etc. Alle documenten, waaronder dit document, zijn vanaf deze website te downloaden.
- www.bisl.nl
BiSL.nl – Website met alle informatie over BiSL. BiSL is, als voorganger van MCTL, interessant vanwege de verzameling Best Practices, whitepapers en artikelen die op deze website zijn te vinden.

Vanuit functioneel perspectief is het volgende boek interessant voor het taakgebied Databeheer.

- Roodink, C. (2013). *Nederland opent*.