



Data Kwaliteiten Raamwerk

Bert Dingemans
Versie 1.0 • Concept

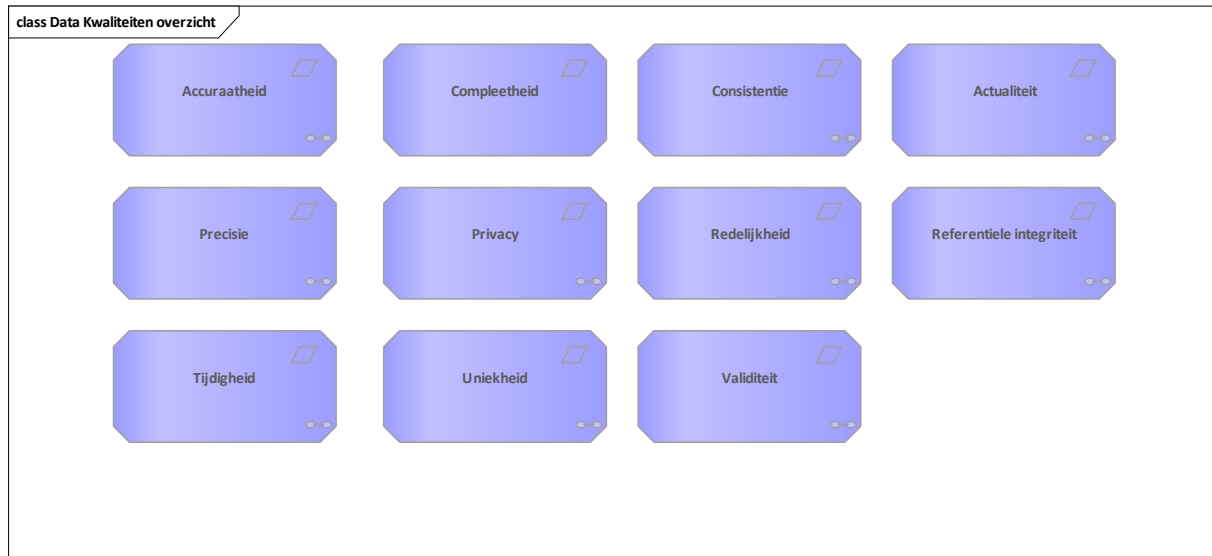
Inhoudsopgave

Data Kwaliteit	3
Data Kwaliteiten overzicht	3
Data Kwaliteit SIPOC (Overzicht)	5
Data Kwaliteit SIPOC (Basis).....	10
Data Kwaliteit Werkproces	12
Data Kwaliteit Score Matrix.....	13
Detail maatregelen	14
Data Kwaliteit Maatregelen Bewustwording.....	14
Data Kwaliteit Maatregelen Werkprocessen	15
Data Kwaliteit Maatregelen RDBMS	16
Data Kwaliteit Maatregelen Software.....	17
Data Kwaliteit Maatregelen Data en Systeemintegratie	19
Data Kwaliteit Maatregelen User interface.....	20
Maatregelen per dimensie.....	21
<i>Data Kwaliteit Maatregelen Accuraatheid</i>	21
<i>Data Kwaliteit Maatregelen Actualiteit</i>	24
<i>Data Kwaliteit Maatregelen Compleetheid</i>	25
<i>Data Kwaliteit Maatregelen Consistentie</i>	27
<i>Data Kwaliteit Maatregelen Precisie</i>	29
<i>Data Kwaliteit Maatregelen Privacy</i>	31
<i>Data Kwaliteit Maatregelen Redelijkheid</i>	33
<i>Data Kwaliteit Maatregelen Ref Integriteit</i>	34
<i>Data Kwaliteit Maatregelen Tijdigheid</i>	36
<i>Data Kwaliteit Maatregelen Uniekheid</i>	37
<i>Data Kwaliteit Maatregelen Validiteit</i>	39

Data Kwaliteit

Het kennisgebied data kwaliteit gaat in op welke wijze een organisatie kan zorgdragen voor data met voldoende kwaliteit afhankelijk van de context. Dit heeft daarmee raakvlakken met onder andere data modelleren, - governance en - architectuur.

Data Kwaliteiten overzicht



Overzicht van soorten data kwaliteiten met een toelichting.

Accuraatheid

Accuraatheid heeft betrekking op de mate waarin een data entiteit de werkelijkheid weergeeft. Accuraatheid kan bepaald worden door een data entiteit te vergelijken met de entiteit in de werkelijkheid.

Actualiteit

Mate waarin een data entiteit de actuele situatie van de werkelijkheid weergeeft.

Compleetheid

Dit heeft betrekking op de mate waarin bepaalde attributen binnen een data entiteit aanwezig zijn. Daarnaast geldt de compleetheid ook voor het altijd voorkomen van een bepaalde set van entiteiten (rijen) in een data set.

Consistentie

Dit heeft betrekking op het feit dat de ene data set van een bepaalde entiteit gelijk is aan een andere dataset. Met andere woorden is een data entiteit onafhankelijk van de bron altijd dezelfde.

Precisie

Mate van detail waarin een data entiteit de werkelijkheid weergeeft. Dit heeft bijvoorbeeld betrekking op de precisie van getallen e.d. Opslag van getallen en datums kunnen onvoldoende nauwkeurig zijn omdat afronding bij opslag of integratie nodig is.

Privacy

Voor sommige data entiteiten is toegangscontrole (autorisatie en authenticatie) of monitoring van gebruik nodig. Denk hierbij aan eisen die gesteld worden aan de ontsluiting van (privacy) vertrouwelijke gegevens.

Redelijkheid

Heeft vooral betrekking op verwachtingen van data entiteiten binnen een bepaalde operationele context.

Referentiele integriteit

Dit is de situatie waarbij verwijzingen vanuit de ene data entiteit altijd correct verwijzen naar de gerelateerde data entiteiten.

Tijdigheid

Is een dataset tijdig beschikbaar binnen de gestelde verwachtingen. Het is het verschil tussen het moment van behoefte en beschikbaarheid.

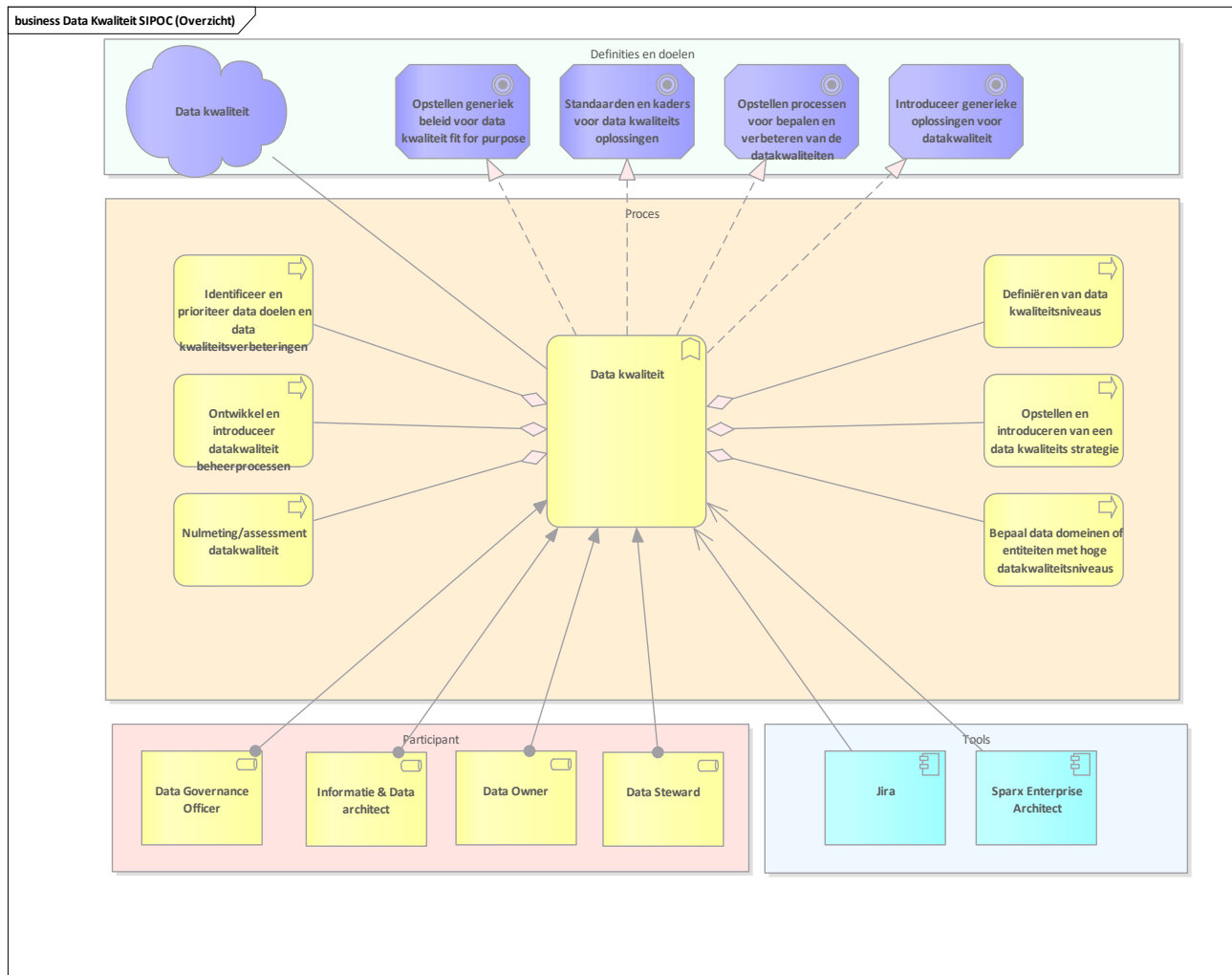
Uniekheid

Uniekheid van een data entiteit is gericht op het feit dat er geen andere entiteiten zijn met dezelfde gegevens. Ook bekend als duplicaten probleem.

Validiteit

Dit is de mate waarin een data entiteit bij opslag en uitwisseling voldoet aan het gewenste formaat. Denk hierbij bijvoorbeeld aan het domein maar ook het datatype van de attributen van een data entiteit. Binnen ketenuitwisseling is dit bijvoorbeeld van het grootste belang.

Data Kwaliteit SIPOC (Overzicht)



Overzicht van de definities en de doelen van data kwaliteit. Daarnaast een uitwerking van het model van de data kwaliteiten op basis van technische implementaties, werkproces en werkinstructies voor de betrokkenen binnen de ondersteunende informatiesystemen en processen.

Bepaal data domeinen of entiteiten met hoge data kwaliteitsniveaus

Voor een data domein of een entiteit kan bepaald worden wat de huidige en gewenste kwaliteitsniveaus zijn op basis van het raamwerk. Is er een verschil tussen de huidige en gewenste kwaliteitsniveaus dan kunnen er maatregelen genomen worden om tot een voldoende hoog kwaliteitsniveau te komen voor dat domein cq die entiteit.

Data Governance Officer

Data governance gaat de hele organisatie aan. Er is altijd een risico dat er versnippering van activiteiten ontstaat. Data governance heeft een nauwe relatie met veel disciplines zoals enterprise architectuur, financiën, control, security en privacy.

Een data governance officers kan de coördinatie en sturing van de inrichting verzorgen. De DGO zal nauw samenwerken met architecten, security en privacy officers en indien aanwezig de CIO office.

Activiteiten

- Werkt nauw samen met de data owners of vervult dit als een combirol
- Signaleert issues binnen de verschillende data management projecten en initiatieven en brengt betrokken bij elkaar en informeert hen
- Communiqueert met de organisatie over data management projecten, roadmap en onderwerpen.
- Betrokken bij de data management projecten (roadmap) of backlog.
- Initieert en neemt deel aan verschillende gremia rond data bijvoorbeeld in de Data governance overleggen en de data governance board.

Vaardigheden

- Kennis van de processen, wensen, behoeften, management en binnen één of meerdere domeinen op management niveau
- Kennis van data management kennisgebieden met name op het gebied van data governance en data kwaliteiten.
- Management vaardigheden en in staat tot delegeren van data activiteiten
- In staat tot samenwerken en communicatieve vaardigheden.

Data kwaliteit

Activiteiten voor kwaliteitsmanagement van data assets zodat het geschikt is voor gebruik en voldoet aan de wensen van de data consumenten (binnen en buiten de eigen organisatie).

Data kwaliteit

Activiteiten voor kwaliteitsmanagement van data assets zodat het geschikt is voor gebruik en voldoet aan de wensen van de data consumenten.

Data Owner

Data Owners of data eigenaren ontwikkelen strategisch beleid rond data sets en nemen hieromtrent de beslissingen, veelal gebaseerd op een opsplitsing in domeinen. Data eigenaarschap is een specifiek onderdeel van een rol als lijn- of divisie manager.

Activiteiten

- Vertaalt data strategie en doelen naar organisatie brede doelen voor motivatie van de data kaders en richtlijnen.
- Verantwoordelijk voor bepalen van de data maturity assessments en aanpak.
- Neemt beslissingen over issues rond data op het vlak van data kennisgebieden (kwaliteit, datagebruik, security).
- Samenwerken met andere data stakeholders (management, architect, X analisten, steward, X officer).
- Communiceert met de organisatie over data management onderwerpen.
- Bepaalt de data management projecten of backlog.
- Initieert en neemt deel aan verschillende gremia rond data bijvoorbeeld in de data governance- of architecture board.

Vaardigheden

- Kennis van de processen, wensen, behoeften, management en binnen één of meerdere domeinen op management niveau.
- Kennis van data management kennisgebieden met name op het gebied van data governance en data kwaliteiten.
- Management vaardigheden en in staat tot delegeren van data activiteiten.
- In staat tot samenwerken en communicatieve vaardigheden.

Data Steward

Stewards richten zich op tactisch en operationeel beleid en zijn gedelegeerd vanuit de Owners. Data steward is een specifiek onderdeel van een rol zoals een X beheerder of een key user.

Activiteiten

- Stelt conventies, kaders en richtlijnen op voor stakeholders rond de data zowel binnen IT als bij de data gebruikers.
- Is betrokken bij het formuleren van de data doelen gebaseerd op de organisatie strategie en vertaalt die naar kaders.
- Signaleert en inventariseert data kwaliteits issues
- Beheert maatregelen voor data kwaliteiten en begeleidt activiteiten om de kwaliteiten te verbeteren
- Beheert de registraties van data governance zoals alle data eigenaarschap (RASCI), data domeinen en - objecten, kaders etc en kwaliteiten en maatregelen.
- Creëren en beheren van de metadata
- Betrokken bij bepalen van de data maturity assessments en aanpak
- Operationele data governance activiteiten
- Ondersteunen business en informatie analisten vanuit data perspectief
- Definieren van metrieken voor data kwaliteit
- Samenwerken met andere data stakeholders (architect, X analisten, owner, officer)
- Communiceert met de organisatie over data management onderwerpen

Vaardigheden

- Kennis van de processen, wensen, behoeften, informatie en data binnen de organisatie binnen één of meerdere domeinen
- Kennis van data management kennisgebieden met name op het gebied van data governance en data kwaliteiten.
- Bij voorkeur kennis van de data en waar deze in de diverse informatiesystemen zijn terug te vinden.
- In staat tot samenwerken en communicatieve vaardigheden.
- Analytische en gericht op vinden van oplossingen rond data (kwaliteit).

Definiëren van data kwaliteitsniveaus

Voor de data kwaliteiten voor data domeinen of - entiteiten kan een huidig en gewenst kwaliteitsniveau bepaald worden. Bijvoorbeeld door het geven van een rapportcijfer van de huidige en gewenste situatie

Definities en doelen

Identificeer en prioriteer data doelen en data kwaliteitsverbeteringen

Op basis van data kwaliteits issues gerelateerd aan het data kwaliteiten raamwerk kan een prioritering opgesteld worden voor het verbeteren van de datakwaliteit op basis van een aantal maatregelen. Veelal wordt dit gecombineerd in een kwaliteitsrelease

Informatie & Data architect

Specifieke architect verantwoordelijk voor data binnen de organisatie. Stelt data architectuur kaders op, maakt data modellen en werkt nauw samen met de andere architecten in de organisatie en de data governance groep.

Introduceer generieke oplossingen voor datakwaliteit

Bijvoorbeeld door het introduceren van generieke oplossingen in werkprocessen, applicaties en het opstellen van procedures en werkinstructies

Jira

Webbased Tool voor projectmanagement binnen een agile perspectief.

Nulmeting/assessment datakwaliteit

Doe een nulmeting voor de datakwaliteiten binnen een domein in de organisatie. Bijvoorbeeld door het bepalen van de huidige en gewenste data kwaliteitsniveaus voor een bepaald domein of entiteit. Desgewenst kan dit gedaan worden op basis van een vragenlijst of op basis van een interactieve workshop waarin meerdere betrokkenen deelnemen

Ontwikkel en introduceer datakwaliteit beheerprocessen

Rond het introduceren van data kwaliteiten van voldoende niveau is een eenvoudig werkproces uitgewerkt.

Opstellen en introduceren van een data kwaliteits strategie

Bepaal wat een strategie is om de data kwaliteiten te verhogen. Dit houdt in dat er een raamwerk wordt opgezet van kwaliteiten en maatregelen. Veelal zullen hierbij ook een aantal kaders gesteld worden bijvoorbeeld als data (kwaliteit) principes.

Opstellen generiek beleid voor data kwaliteit fit for purpose

Opstellen van beleid voor data kwaliteit in de organisatie en dit beleid vervolgens vertalen in kaders en hulpmiddelen om de kwaliteit te bereiken en te handhaven.

Opstellen processen voor bepalen en verbeteren van de datakwaliteiten

Processen, procedures, checklists en werkinstructies voor het zorgdragen voor data van voldoende kwaliteit.

Participant

Proces

Sparx Enterprise Architect

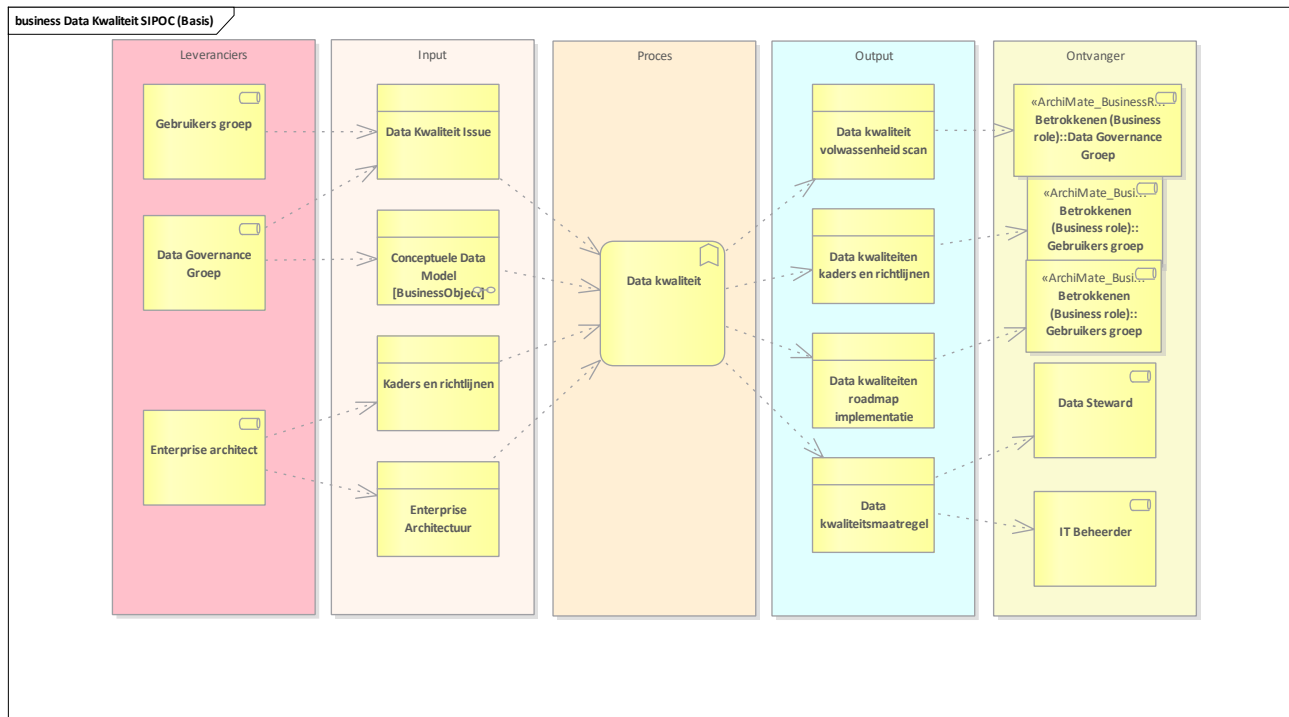
Generieke tool voor het opstellen van modellen, voor data, architectuur, en oa bedrijfsprocessen.

Standaarden en kaders voor data kwaliteits oplossingen

Rond data kwaliteiten kunnen kaders en richtlijnen opgesteld worden op welke wijze data van een bepaald kwaliteitsniveau bereikt en gehandhaafd kan worden. Deze kaders zijn veelal uitgewerkt op basis van de data principes voor de kwaliteit van de data.

Tools

Data Kwaliteit SIPOC (Basis)



In deze SIPOC zie je de verbanden rond het bedrijfsproces voor data kwaliteiten in de organisatie.

Conceptuele Data Model [BusinessObject]

Conceptueel data model omvat de data entiteiten inclusief de bijbehorende definities (inclusief de bovenliggende domeinen) op basis waarvan de data kwaliteiten gerelateerd worden aan deze data elementen.

Data Governance Groep

Groep van rollen in de organisatie die betrokken zijn bij data management in het algemeen en governance in het bijzonder.

Data kwaliteit

Activiteiten voor kwaliteitsmanagement van data assets zodat het geschikt is voor gebruik en voldoet aan de wensen van de data consumenten (binnen en buiten de eigen organisatie).

Data Kwaliteit Issue

Een data kwaliteits issue is een probleem rond een van de data kwaliteiten die een probleem vormen in de bedrijfsvoering en de bedrijfs informatievoorziening. Rond een of meerdere issues rond data kwaliteit wordt een oplossing gezocht. Veelal worden conform de ITIL werkwijze kwaliteits issues gecombineerd tot een probleem waartoe een release gerealiseerd wordt die een oplossing introduceert. Dat kan zowel technisch als procesmatig van aard zijn.

Data kwaliteit volwassenheid scan

Data kwaliteit kent een volwassenheid. Deze volwassenheid is meetbaar met een data kwaliteitsscan of assessment en geeft een resultaat. Dit kan periodiek herhaald worden. Dit geeft focus aan de prioriteit in de data kwaliteits issues die bestaan en welke maatregelen in een prioriteitsvolgorde uitgevoerd (gaan) worden.

Data kwaliteiten kaders en richtlijnen

Kaders en richtlijnen zoals modelleer- en naamgevingsconventies, principes voor de verschillende data kaders en maatregelen welke zorgdragen met data met een kwaliteitsniveau van een voldoende hoog kwaliteitsniveau.

Data kwaliteiten roadmap implementatie

Implementatie van kwaliteitsmaatregelen rond een prioritering van de data kwaliteits issues. Veelal worden issues gecombineerd tot een clustering van een aantal maatregelen om tot een oplossing van deze issues te komen.

Hierbij wordt een groep van maatregelen gecombineerd een release van maatregelen die geïmplementeerd worden om het kwaliteitsniveau naar het wenselijke niveau te brengen.

Data kwaliteitsmaatregel

Een maatregel is een activiteit zowel technisch van aard, bedrijfsproces of werkinstructie die zorg draagt voor het oplossen van een data kwaliteits issue gerelateerd aan het kwaliteitsraamwerk.

Data Steward

Stewards richten zich op tactisch en operationeel beleid en zijn gedelegeerd vanuit de Owners. Data steward is een specifiek onderdeel van een rol zoals een X beheerder of een key user.

Activiteiten

- Stelt conventies, kaders en richtlijnen op voor stakeholders rond de data zowel binnen IT als bij de data gebruikers.
- Is betrokken bij het formuleren van de data doelen gebaseerd op de organisatie strategie en vertaalt die naar kaders.
- Signaleert en inventariseert data kwaliteits issues
- Beheert maatregelen voor data kwaliteiten en begeleidt activiteiten om de kwaliteiten te verbeteren
- Beheert de registraties van data governance zoals alle data eigenaarschap (RASCI), data domeinen en - objecten, kaders etc en kwaliteiten en maatregelen.
- Creëren en beheren van de metadata
- Betrokken bij bepalen van de data maturity assessments en aanpak
- Operationele data governance activiteiten
- Ondersteunen business en informatie analisten vanuit data perspectief
- Definieren van metriecken voor data kwaliteit
- Samenwerken met andere data stakeholders (architect, X analisten, owner, officer)
- Communiceert met de organisatie over data management onderwerpen

Vaardigheden

- Kennis van de processen, wensen, behoeften, informatie en data binnen de organisatie binnen één of meerdere domeinen
- Kennis van data management kennisgebieden met name op het gebied van data governance en data kwaliteiten.
- Bij voorkeur kennis van de data en waar deze in de diverse informatiesystemen zijn terug te vinden.
- In staat tot samenwerken en communicatieve vaardigheden.
- Analytische en gericht op vinden van oplossingen rond data (kwaliteit).

Enterprise architect

Enterprise architect voor het opstellen van organisatie brede architectuur zowel beschrijvend als het opstellen van kaders en richtlijnen gebaseerd op de organisatie strategie.

Enterprise Architectuur

Enterprise bestaat uit zowel een beschrijvende architectuur rond de inrichting van de enterprise op basis van bedrijfs- applicatie- en infra architectuur. Daarnaast omvat de enterprise architectuur ook de kaderstellende architectuur (principes).

Gebruikers groep

Groep van gebruikers van data en IT voorzieningen waarin data verwerkt wordt.

Input

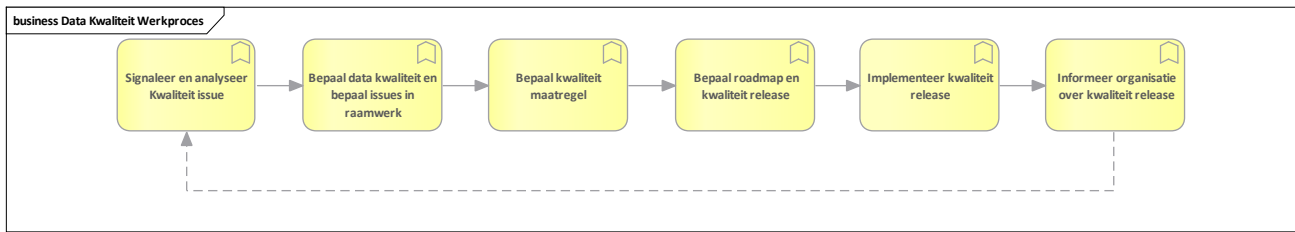
IT Beheerder

Beheert IT artefacten ter ondersteuning van de informatievoorziening voor zowel de primaire- en secundaire werk processen of - domeinen.

Kaders en richtlijnen

Kaders en richtlijnen zoals principes en beperkingen in algemene zin voor de organisatie, wetgevende instanties maar ook specifiek naar data management en governance.

Data Kwaliteit Werkproces



Eenvoudig stappenplan hoe data kwaliteitsissues in een gestructureerde opzet opgelost kunnen worden met behulp van maatregelen en releases.

Bepaal data kwaliteit en bepaal issues in raamwerk

Door de betrokkenen bij data kwaliteiten (data governance groep) worden één of meer issues gekoppeld aan de data kwaliteiten in het raamwerk. In het raamwerk zijn een aantal maatregelen gekoppeld aan deze data kwaliteiten

Bepaal kwaliteit maatregel

Op basis van het raamwerk wordt een aantal maatregelen uit de catalogus geselecteerd en deze worden gecombineerd tot een set om te komen tot een oplossing.

Bepaal roadmap en kwaliteit release

Er wordt een roadmap bepaald welke maatregelen in eenreleasematige uitwerking geïmplementeerd gaan worden. De prioritering in de roadmap is veelal een combinatie van ernst van het issue en het realiseren van quick wins.

Implementeer kwaliteit release

Implementeer de kwaliteitsrelease. Dit is veelal een combinatie van technische IT maatregelen zoals het doen van aanpassingen in informatiesysteem, data modellen en infrastructuur samen met aanpassingen in bedrijfsprocessen of in de werkinstructies voor de betrokkenen.

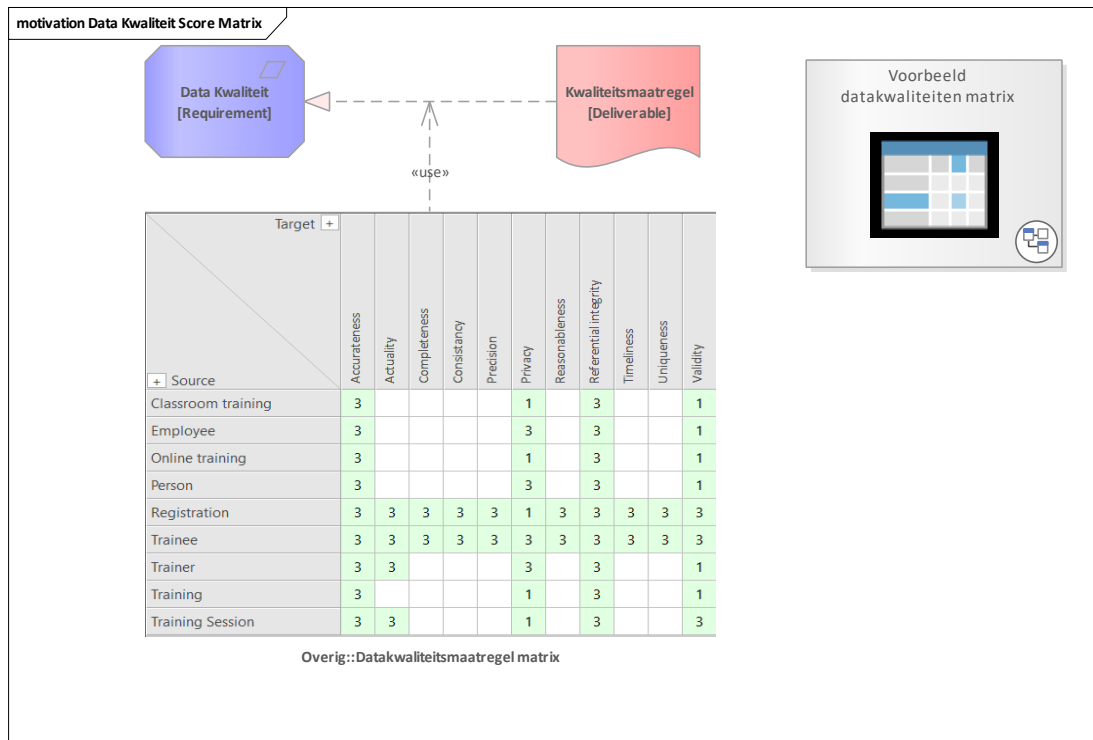
Informeert organisatie over kwaliteit release

Informeert de introductie van de release aan de betrokken binnen en buiten de organisatie.

Signaleer en analyseer Kwaliteit issue

Binnen of buiten de organisatie wordt een kwaliteits issue geconstateerd. Het issue wordt geanalyseerd en gecategoriseerd en vervolgens geregistreerd.

Data Kwaliteit Score Matrix



Voorbeeld van op werkwijze met behulp van een score matrix geregistreerd kan worden wat het kwaliteitsniveau in de huidige en gewenste situatie is. In dit voorbeeld is een score 0 tot 9.

Data Kwaliteit [Requirement]

Datakwaliteit conform het raamwerk van DMBOK

Datakwaliteitsmaatregel matrix

In dit voorbeeld zie je hoe je voor data domeinen of - entiteiten een score kunt opstellen voor de datakwaliteiten. Dit kan zowel voor de huidige als de gewenste situatie gedaan worden.

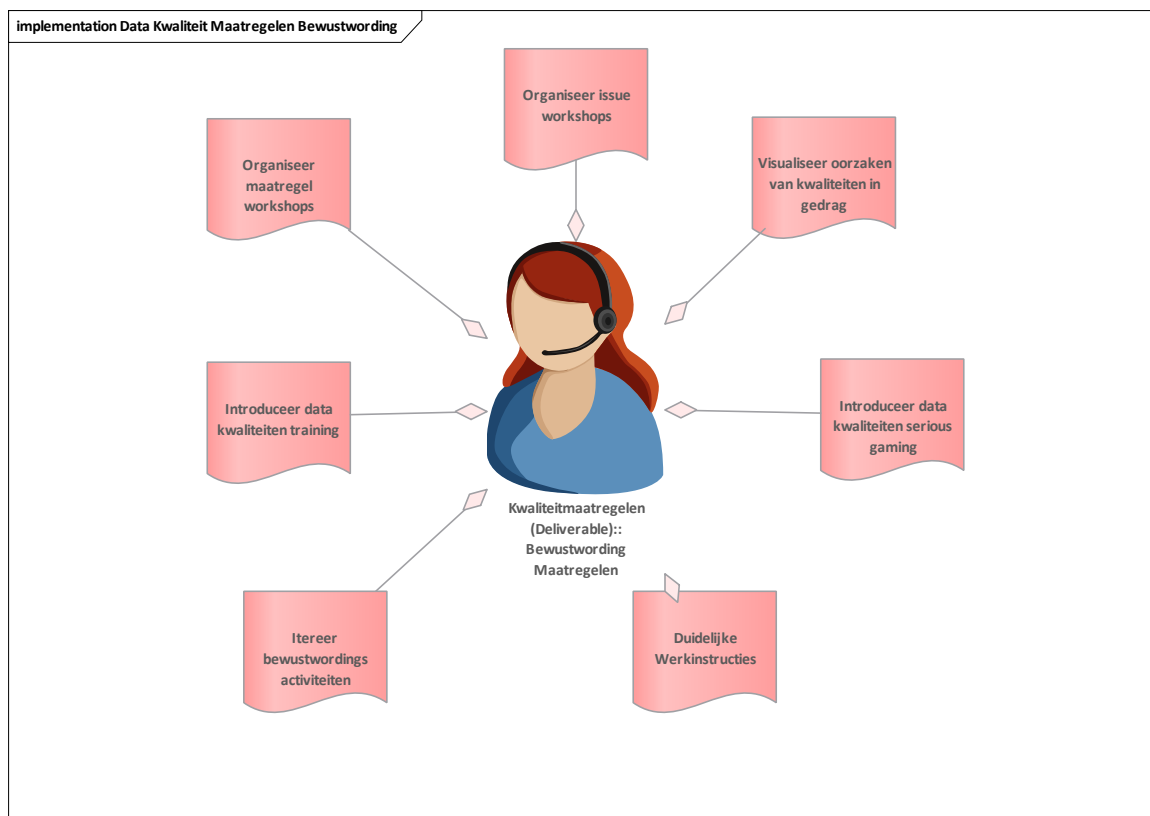
Kwaliteitsmaatregel [Deliverable]

Een maatregel van technische aard, bedrijfsproces of werkinstructie om het bereiken van voldoende kwaliteit gerealiseerd wordt

Detail maatregelen

Extra maatregelen vanuit een bepaalde technische scope zoals databases software user interfaces etc

Data Kwaliteit Maatregelen Bewustwording



Bewustwording Maatregelen

Zorg dat de verschillende betrokkenen in de organisatie zich bewust zijn van maatregelen die de kwaliteit verhogen van de data. Zoerg daarbij dat met zich bewust is van het eigen handelen van de data kwaliteit en het effect wat dat kan hebben op maatregelen

Duidelijke Werkinstructies

Zorg voor duidelijke werkinstructies vanuit het perspectief van data kwaliteit. Veelal wordt in werkinstructies een introductie gedaan van workarounds die een negatief effect kunnen hebben op de kwaliteit van de data

Introduceer data kwaliteiten serious gaming

Data kwaliteiten, bewustwording en verandering van gedrag kan met serious gaming inzichtelijk worden gemaakt. Denk aan interactieve workshops waarin met behulp van een competitief element de kwaliteit van data onder de aandacht wordt gebracht

Introduceer data kwaliteiten training

Bewustwording van kwaliteiten vergt training want het is gericht op bewustwording en gedragsverandering bij de verschillende stakeholders. Dit kan gedaan worden met klassikale, peer to peer en ook e-learning getraind worden

Itereer bewustwordings activiteiten

Bewustwording en gedragsverandering is helaas aan slijtage onderhevig binnen de populatie van een organisatie. Dat houdt in dat er iteratief opnieuw aandacht besteed moet worden aan de bewustwording. Ook hier kan een gaming element van grote hulp zijn.

Organiseer issue workshops

Binnen de verschillende soorten stakeholders is veelal diepgaande kennis aanwezig van kwaliteitsissues in de data. In workshops is dit een zeer waardevolle bron van problemen in de data. Maar het is ook een mogelijkheid om de bewustwording van datakwaliteit te verhogen

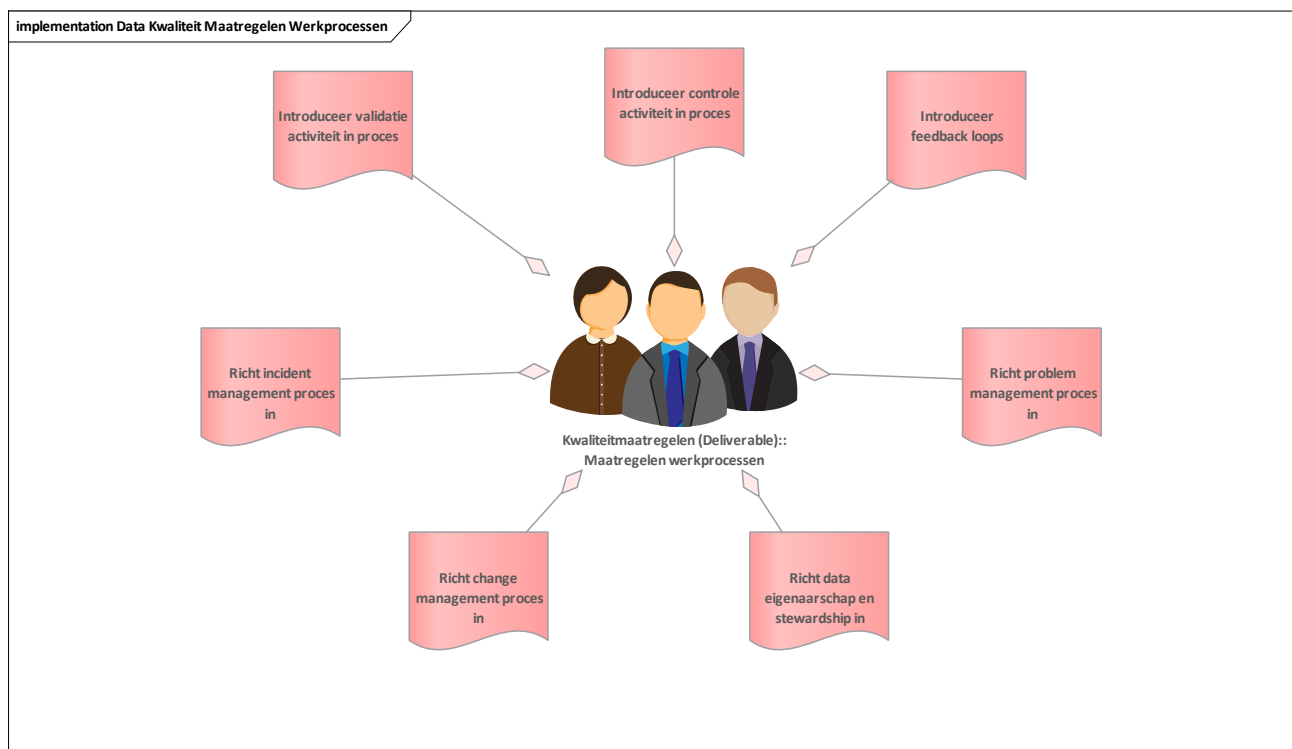
Organiseer maatregel workshops

Naast issue workshops kan het ook interessant zijn om samen met stakeholders op basis van issues te kijken naar relevante en (eenvoudig implementeerbare maatregelen en hiermee de datakwaliteit te verhogen. Deze maatregelen kunnen vervolgens worden opgenomen in het maatregelenregister

Visualiseer oorzaken van kwaliteiten in gedrag

Maak inzichtelijk voor stakeholders die datakwaliteitsproblemen veroorzaken wat het effect is van handelen bij het werken met de data. Denk bijvoorbeeld aan het misbruiken van velden in schermen die daarmee de kwaliteit in de achterliggende databanken kunnen verlagen.

Data Kwaliteit Maatregelen Werkprocessen



Introduceer controle activiteit in proces

Wil je de kwaliteit van data waarborgen dan kun je zorgen dat er een kwaliteitsstap wordt ingebouwd in een werkproces. Hierbij zijn verschillende vormen mogelijk. Zoals dat de activiteit door de data producent zelf gedaan wordt of door een andere anders dan de stakeholder zelf. Bekendste voorbeeld is het twee paar ogen principe

Introduceer feedback loops

Zorg dat bij een validatiestap van de data kwaliteiten in een proces dat er een feedback loop plaatsvindt naar de activiteit waar het issue ontstaan is. Houd er rekening mee dat dit ook mogelijk is naar metadata, datamodel of data architectuur activiteiten in de feedback

Introduceer validatie activiteit in proces

Zorg dat data die geproduceerd of getransformeerd wordt, zeker bij handmatige verwerking van data in een validatie activiteit beoordeeld wordt. Dit biedt een punt in het proces waar issues gesignaleerd kunnen worden, maatregelen genomen kunnen worden of waar signalen zijn dat er in andere data management processen aandachtspunten zijn rond de data kwaliteit.

Maatregelen werkprocessen

Aggreratie van maatregelen die de data kwaliteiten verhogen binnen de categorie werkprocessen

Richt change management proces in

ITIL/BISL processen zijn niet alleen inzetbaar voor ICT voorzieningen maar ook voor data wijzigingen die doorgevoerd moeten worden. Sluit aan bij bestaande inrichting in de organisatie voor wijzigingsbeheer en breid dit uit met specifieke implementaties voor data maatregelen. Echter zorg voor een generieke opzet van dit ITIL/BISL proces

Richt data eigenaarschap en stewardship in

Data governance is essentieel voor data kwaliteiten. Zorg dan ook dat de data governance processen ingericht zijn om sturing en uitvoering te kunnen geven aan processen die de data kwaliteiten kunnen verhogen op basis van maatregelen.

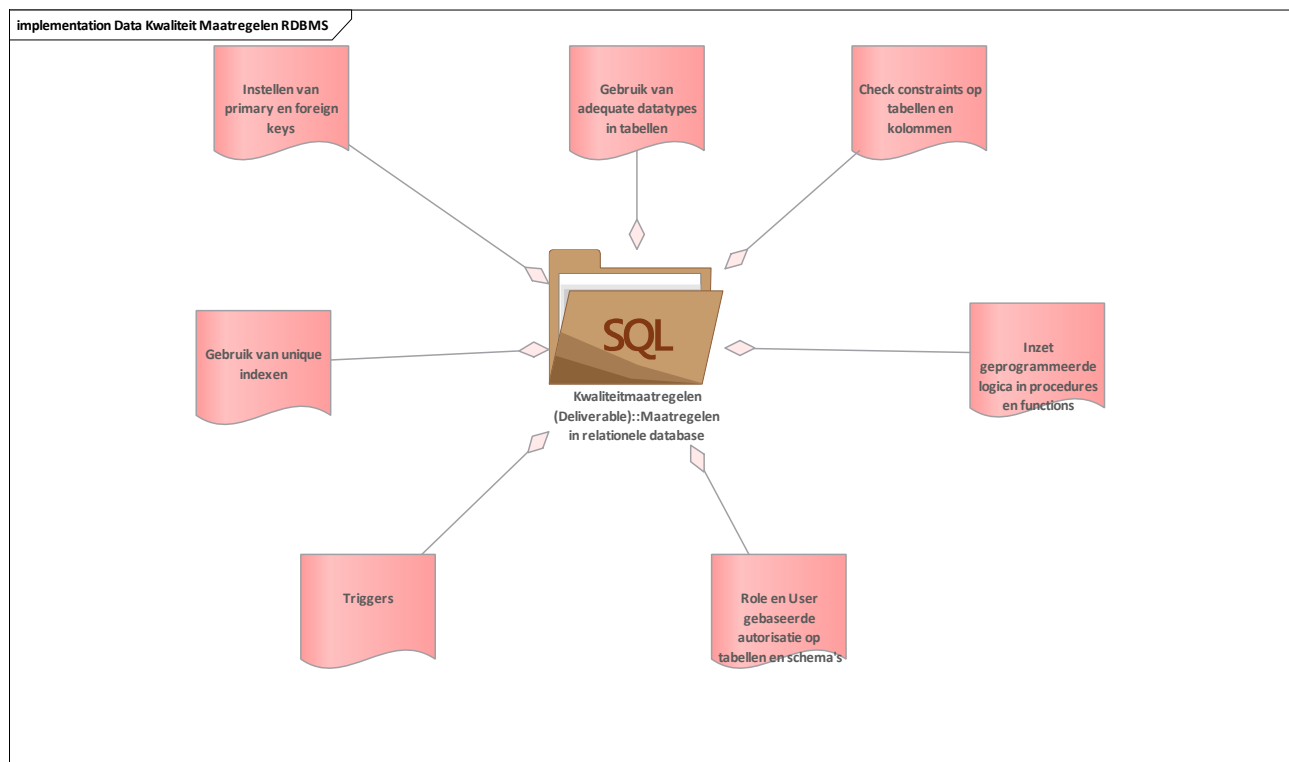
Richt incident management proces in

ITIL/BISL processen zijn niet alleen inzetbaar voor ICT voorzieningen maar ook voor data issues die ontstaan binnen de organisatie. Sluit aan bij bestaande inrichting in de organisatie voor incidentbeheer en breid dit uit met specifieke implementaties voor data issues. Echter zorg voor een generieke opzet van dit ITIL/BISL proces

Richt problem management proces in

ITIL/BISL processen zijn niet alleen inzetbaar voor ICT voorzieningen maar ook voor data wijzigingen die doorgevoerd moeten worden. Sluit aan bij bestaande inrichting in de organisatie voor problem management en breid dit uit met specifieke implementaties voor data maatregelen. Echter zorg voor een generieke opzet van dit ITIL/BISL proces

Data Kwaliteit Maatregelen RDBMS



Check constraints op tabellen en kolommen

Check constraints zijn extra functionaliteit in een database waarin je op basis van een statement extra condities kunt toevoegen aan een of meerdere kolommen. Dit extra controles toe te voegen aan deze kolommen

Gebruik van adequate datatypes in tabellen

Datatypes in de verschillende database platformen maken het mogelijk om op basis van datatypes beperkingen in te bouwen in verschillende datatypes. Zo kun je verschillende numerieke datatypes kiezen zoals integer, decimalen (met het aantal cijfers voor of achter de komma, float, etc. Hiermee zijn extra validaties mogelijk op basis van deze datatypes

Gebruik van unique indexen

Unieke indexen aanleggen kan worden gedaan op basis van de sleutel (de primary keys bijvoorbeeld). Echter als je dat zou willen dan kan dat ook op een combinatie van kolommen voor extra duplicatie validatie checks

Instellen van primary en foreign keys

Primary en foreign keys maken het mogelijk om controles uit te voeren op de sleutels (primaire) in een tabel en vervolgens de verwijzingen vanuit andere tabellen naar deze primaire sleutel af te dwingen. Hiermee voorkom je verwijzingen die niet correct zijn in een goed geconfigureerde database

Inzet geprogrammeerde logica in procedures en functions

Stored procedures en functions maken het mogelijk om transformaties, controles en validaties te automatiseren in de vorm van scripts binnen deze stored procedures en functions. Hiermee heb je vanuit deze programmatuur direct toegang tot de inhoud van de verschillende tabellen.

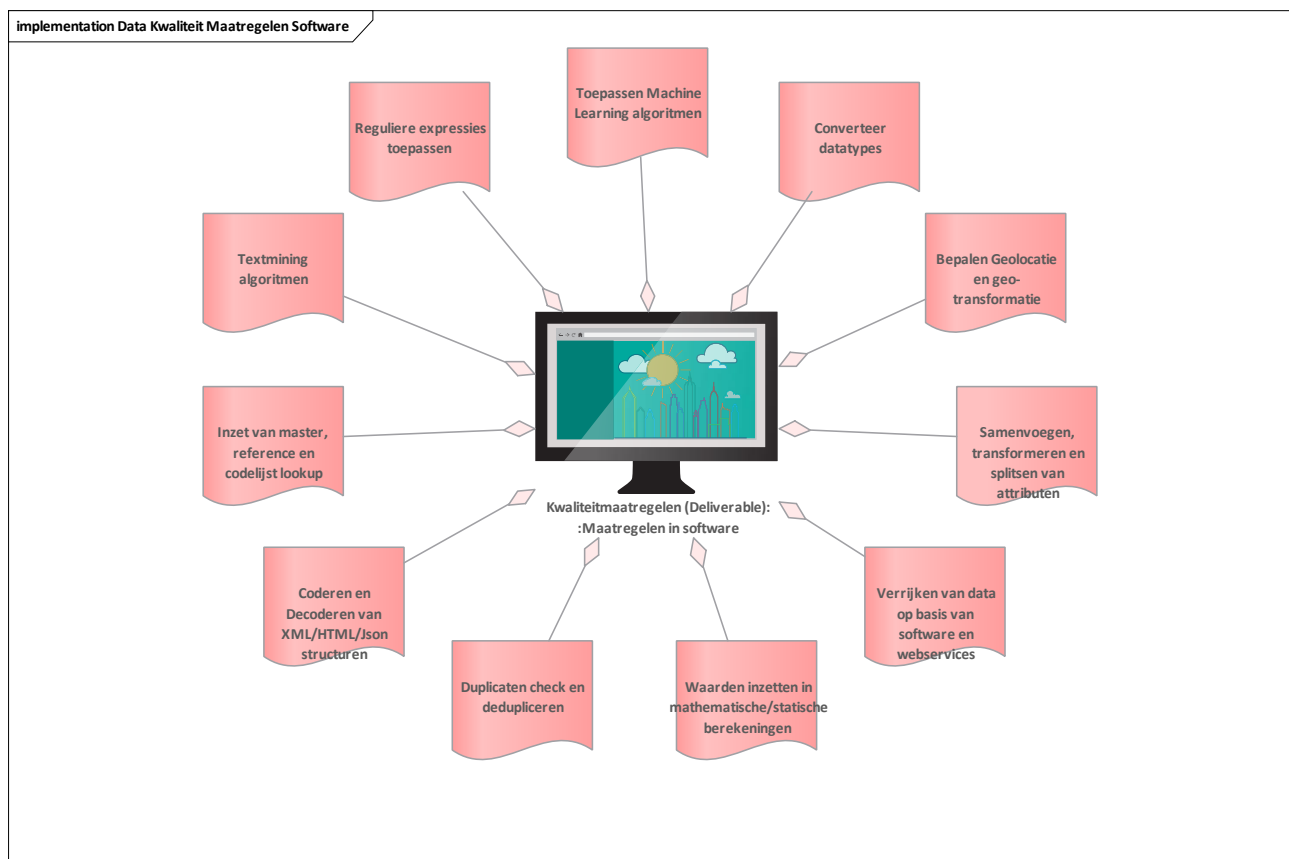
Role en User gebaseerde autorisatie op tabellen en schema's

In databases is op fijnmazig niveau de autorisatie van gebruikers en groepen gebruikers. Dit is gebaseerd op het implementeren van CRUD (Create/Read/Update/Delete) in de database per tabel of schema

Triggers

Triggers maken het mogelijk om op basis van bewerkingen in de database logica aan te roepen om controles te doen, historie op te bouwen of near real time signalen af te geven aan diverse betrokkenen rond de data.

Data Kwaliteit Maatregelen Software



Bepalen Geolocatie en geo-transformatie

Verrijken en transformeren van data van en naar geo locaties bijvoorbeeld door gebruik van database logica en software

Coderen en Decoderen van XML/HTML/Json structuren

Transformeren van datastructuren in XML, HTML en Json naar interpreteerbare en valideerbare data tbv de kwaliteit

Converteer datatypes

Converteren van datatypes naar andere datatypes (tekst <-> Numeriek). Inclusief transformatie naar datatypes op andere platformen zoals database, XML, softwaretalen en localisatie.

Duplicaten check en dedupliceren

Signaleren van duplicaten in sleutels en in combinaties van attributen in dataobjecten.

Inzet van master, reference en codelijst lookup

Verrijken van datasets op basis van kenmerken voor het toevoegen van beschrijvingen, lookups, hierarchien van generiek data attributen

Reguliere expressies toepassen

Validaties in teksten op basis van reguliere expressies. Denk bijvoorbeeld aan het format van een postcode of een emailadres

Samenvoegen, transformeren en splitsen van attributen

Transformeren naar attributen in datasets door omzetten naar samenvoegen tot een geaggregeerd kenmerk. Daarnaast het opsplitsen van delen van een attribuut op basis van een bepaald kenmerk. Splitsen opv komma of spatie.

Textmining algoritmen

Inzetten van algoritmen om te zoeken naar patronen in tekstuele informatie. Denk bijvoorbeeld aan sentimentanalyse. Tellen van voorkomende woorden en nabijheid van woorden

Toepassen Machine Learning algoritmen

Inzetten van machine learning algoritme om de kwaliteit van data te verbeteren. Bijvoorbeeld richting accuraatheid. Denk bijvoorbeeld aan Barneveld en bepalen of de woonplaats in Nederland of de VS is obv van andere attributen in de dataset.

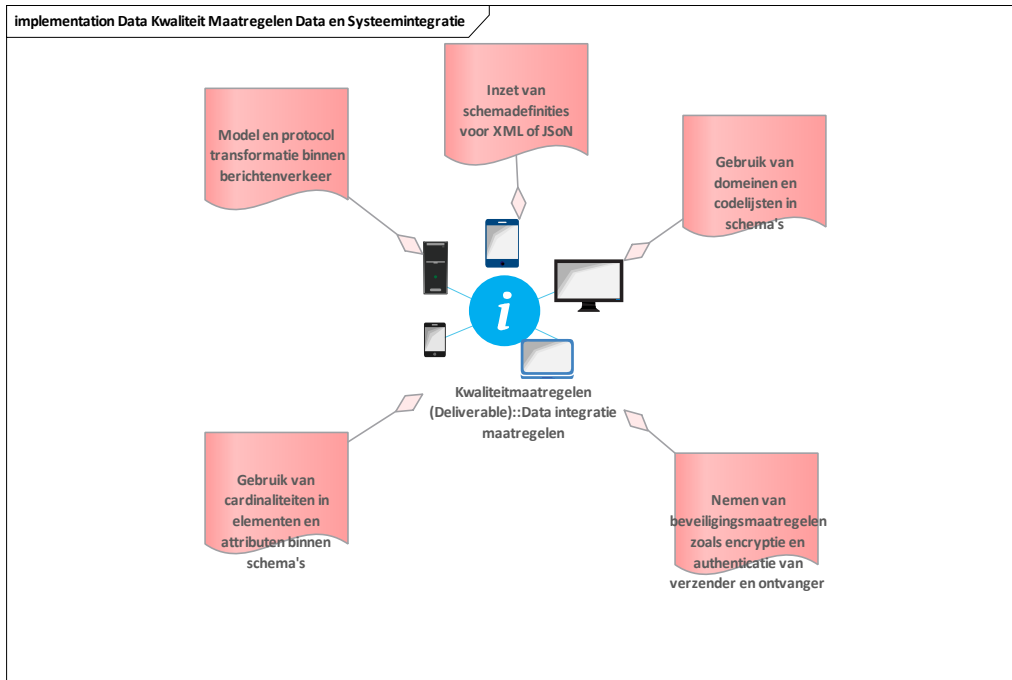
Verrijken van data op basis van software en webservices

Verrijken van data op basis van APIs en webservices voor toevoegen van extra kenmerken. Denk bijvoorbeeld aan postcode.nl voor aanvullen en valideren van adressen op basis van postcode en huisnummer.

Waarden inzetten in mathematische/statische berekeningen

Inzetten van mathematische en statische berekeningen voor bepalen van de kwaliteit maar ook voor het aggregeren of het bepalen van afwijkende waarden in een dataset.

Data Kwaliteit Maatregelen Data en Systeemintegratie



Data integratie maatregelen

Gebruik van cardinaliteiten in elementen en attributen binnen schema's

Gebruik van minimale en maximale cardinaliteiten voor zowel elementen als de attributen. Hiermee kan verplicht worden afgedwongen. Maar ook de mogelijkheid om meerdere elementen vast te leggen voor het realiseren van compleetheid

Gebruik van domeinen en codelijsten in schema's

Gebruik van extra structuren in schema's om extra beperkingen in attributen te implementeren op basis van domeinen (waardelijsten) of codelijsten

Inzet van schemadefinities voor XML of JSON

van schema's voor het valideren van data verzamelingen binnen een XML of JSON gebaseerd bericht of bestand. Hiermee wordt afgedwongen dat de data voldoet aan de regels in het validatieschema.

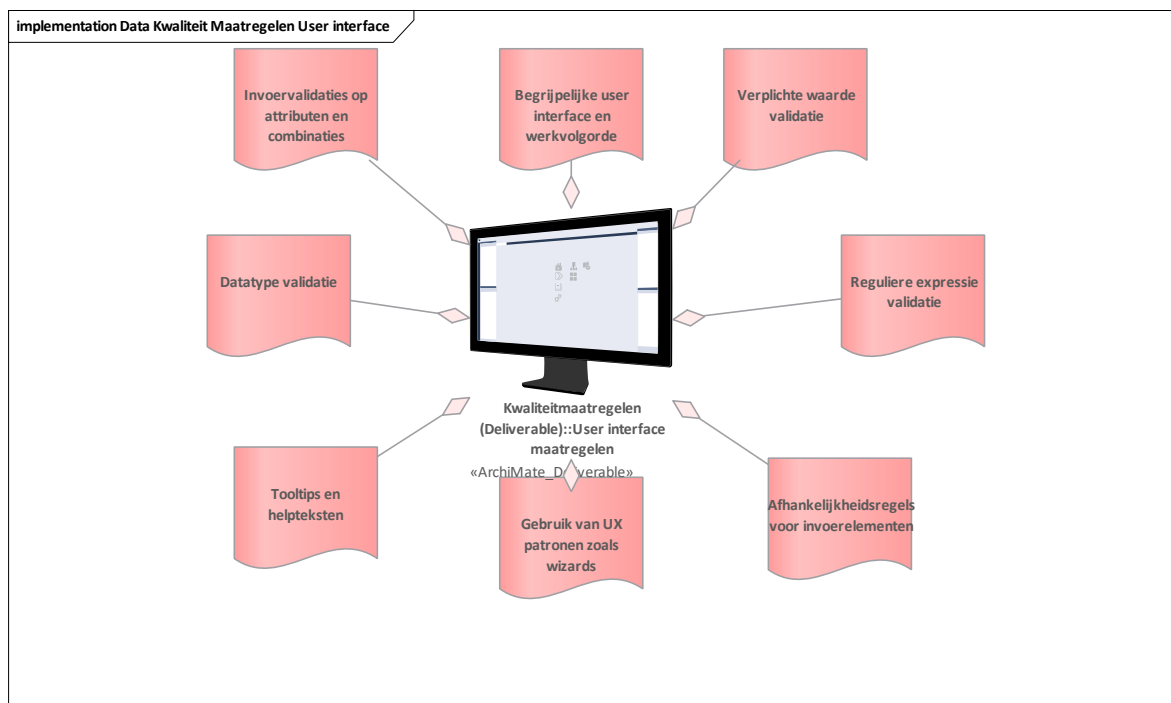
Model en protocol transformatie binnen berichtenverkeer

Transformeren van modellen en protocollen. Bijvoorbeeld van en naar een Canoniek Model transformeren en protocol transformeren XML naar JSON vice versa.

Nemen van beveiligingsmaatregelen zoals encryptie en authenticatie van verzender en ontvanger

Gebruik van beveiligingsmaatregelen binnen de data integratie om zorg te dragen voor de beveiligde uitwisseling van data in berichtenverkeer tussen informatiesystemen en organisaties.

Data Kwaliteit Maatregelen User interface



Afhankelijkheidsregels voor invoerelementen

Op basis van invoervelden kunnen andere invoerelementen beïnvloed worden. Bijvoorbeeld invoerelementen zichtbaar/onzichtbaar of actief/inactief te maken. Bijvoorbeeld na kiezen van anders in een keuzelijst wordt een invoerveld met de titel anders actief gemaakt

Begrijpelijke user interface en werkvolgorde

Een begrijpelijke user interface voor medewerkers en gebruikers zorgt ervoor dat fouten en onduidelijkheden worden voorkomen. Daarnaast is een fijne user interface productiviteit verhogend

Datatype validatie

Bij invoerelementen kunnen near realtime validaties worden toegevoegd. Denk bijvoorbeeld aan, getal en datum tijd validaties of email patroon controls bij invoer

Gebruik van UX patronen zoals wizards

Maak gebruik van interactiepatronen om gebruikers door complexe stappen in het invoeren en muteren van gegevensstructuren te leiden

Invoervalidaties op attributen en combinaties

Invoercontroles voor combinaties van invoervelden, controle validaties en check controles

Reguliere expressie validatie

Gebruik van reguliere expressie validaties in invoervelden om te voorkomen dat er invoerfouten worden gemaakt.

Tooltips en helpteksten

Tooltips en helpteksten geven gebruikers ondersteuning als invoervelden en formulieren onduidelijk zijn of geven een toelichting bij complexe invoercombinaties

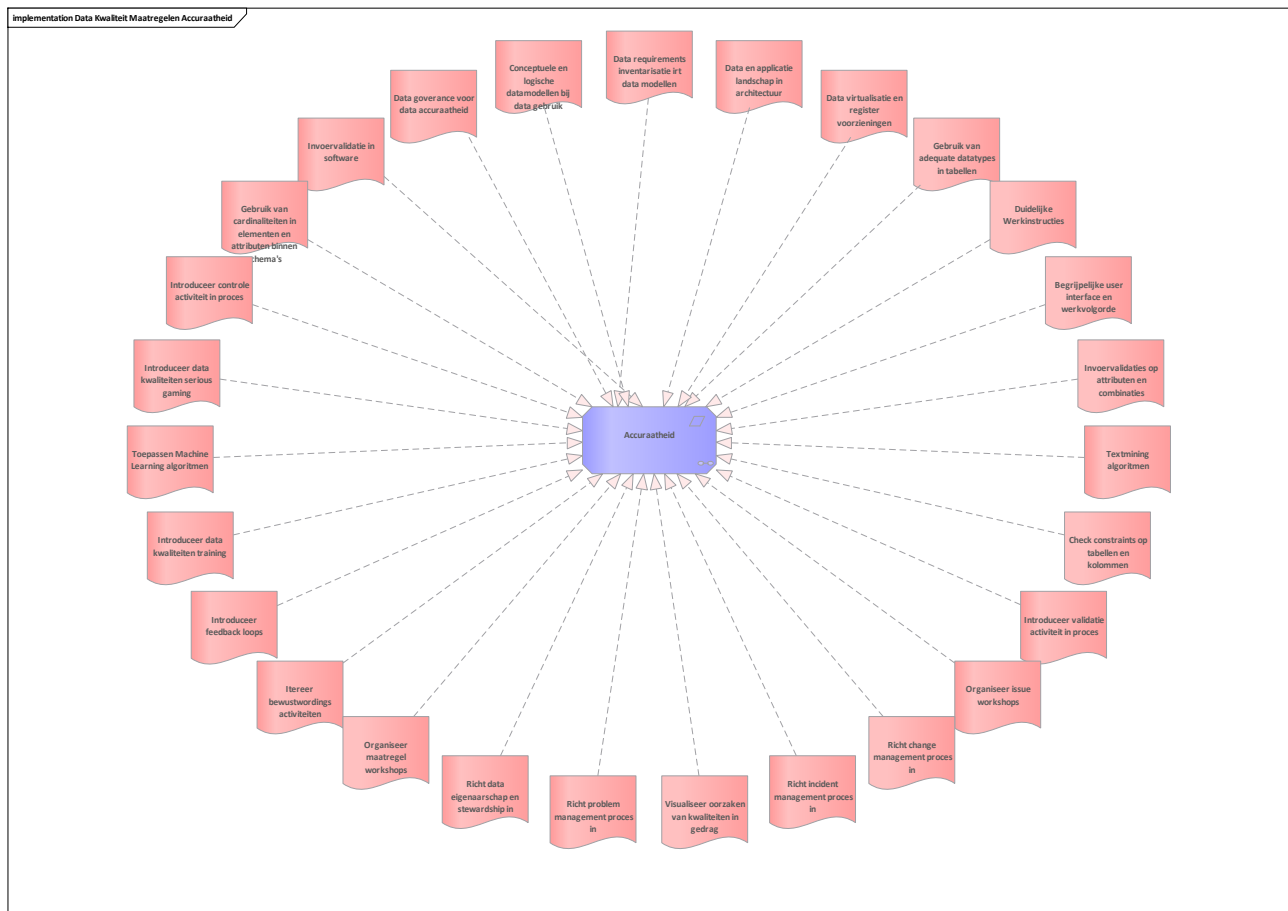
User interface maatregelen

Verplichte waarde validatie

Invoervalidatie op verplichte invoer in het invulformulier in een vroeg stadium

Maatregelen per dimensie

Data Kwaliteit Maatregelen Accuraatheid



Overzicht van de data kwaliteiten en daarbij een aantal maatregelen die genomen kunnen worden om een Data kwaliteit met behulp van deze maatregelen (binnen een release) op een hoger niveau te krijgen.

Accuraatheid

Accuraatheid heeft betrekking op de mate waarin een data entiteit de werkelijkheid weergeeft. Accuraatheid kan bepaald worden door een data entiteit te vergelijken met de entiteit in de werkelijkheid.

Begrijpelijke user interface en werkvolgorde

Een begrijpelijke user interface voor medewerkers en gebruikers zorgt ervoor dat fouten en onduidelijkheden worden voorkomen. Daarnaast is een fijne user interface productiviteit verhogend

Check constraints op tabellen en kolommen

Check constraints zijn extra functionaliteit in een database waarin je op basis van een statement extra condities kunt toevoegen aan een of meerdere kolommen. Dit extra controles toe te voegen aan deze kolommen

Conceptuele en logische datamodelen bij data gebruik

Inzetten van modelleeromgevingen voor het opstellen van modellen van de gegevensbehoefte en het - aanbod.

Data en applicatie landschap in architectuur

Uitwerken en beschrijven van de verschillende architectuur concepten als processen, applicaties, (data) objecten en het in kaart brengen van de onderling relaties.

Data governance voor data accuraatheid

Inrichten eigenaarschap en beheer van gegevensentiteiten en gerelateerde concepten.

Data requirements inventarisatie irt data modellen

Bepalen van de gegevensbehoefte binnen de verschillende activiteiten van werkprocessen en het vergelijken van deze behoefte met de reeds aanwezige gegevensverzamelingen.

Data virtualisatie en register voorzieningen

Inzet van voorzieningen voor data virtualisatie, authenticatie- en kernregisters en ontsluiten van de basisregistraties.

Duidelijke Werkinstructies

Zorg voor duidelijke werkinstructies vanuit het perspectief van data kwaliteit. Veelal wordt in werkinstructies een introductie gedaan van workarounds die een negatief effect kunnen hebben op de kwaliteit van de data

Gebruik van adequate datatypes in tabellen

Datatypes in de verschillende database platformen maken het mogelijk om op basis van datatypes beperkingen in te bouwen in verschillende datatypes. Zo kun je verschillende numerieke datatypes kiezen zoals integer, decimalen (met het aantal cijfers voor of achter de komma, float, etc. Hiermee zijn extra validaties mogelijk op basis van deze datatypes

Gebruik van cardinaliteiten in elementen en attributen binnen schema's

Gebruik van minimale en maximale cardinaliteiten voor zowel elementen als de attributen. Hiermee kan verplicht worden afgedwongen. Maar ook de mogelijkheid om meerdere elementen vast te leggen voor het realiseren van compleetheid

Introduceer controle activiteit in proces

Wil je de kwaliteit van data waarborgen dan kun je zorgen dat er een kwaliteitsstap wordt ingebouwd in een werkproces. Hierbij zijn verschillende vormen mogelijk. Zoals dat de activiteit door de data producent zelf gedaan wordt of door een andere anders dan de stakeholder zelf. Bekendste voorbeeld is het twee paar ogen principe

Introduceer data kwaliteiten serious gaming

Data kwaliteiten, bewustwording en verandering van gedrag kan met serious gaming inzichtelijk worden gemaakt. Denk aan interactieve workshops waarin met behulp van een competitief element de kwaliteit van data onder de aandacht wordt gebracht

Introduceer data kwaliteiten training

Bewustwording van kwaliteiten vergt training want het is gericht op bewustwording en gedragsverandering bij de verschillende stakeholders. Dit kan gedaan worden met klassikale, peer to peer en ook e-learning getraind worden

Introduceer feedback loops

Zorg dat bij een validatiestap van de data kwaliteiten in een proces dat er een feedback loop plaatsvindt naar de activiteit waar het issue ontstaan is. Houd er rekening mee dat dit ook mogelijk is naar metadata, datamodel of data architectuur activiteiten in de feedback

Introduceer validatie activiteit in proces

Zorg dat data die geproduceerd of getransformeerd wordt, zeker bij handmatige verwerking van data in een validatie activiteit beoordeeld wordt. Dit biedt een punt in het proces waar issues gesignaleerd kunnen worden, maatregelen genomen kunnen worden of waar signalen zijn dat er in andere data management processen aandachtspunten zijn rond de data kwaliteit.

Invoervalidatie in software

Inzetten van invoerformulieren die op basis van gestandaardiseerde services zorgdragen voor een correcte, complete en actuele weergave en verwerking van de gegevens ter ondersteuning van een activiteit binnen een werkproces.

Invoervalidaties op attributen en combinaties

Invoercontroles voor combinaties van invoervelden, controle validaties en check controles

Itereer bewustwordings activiteiten

Bewustwording en gedragsverandering is helaas aan slijtage onderhevig binnen de populatie van een organisatie. Dat houdt in dat er iteratief opnieuw aandacht besteed moet worden aan de bewustwording. Ook hier kan een gaming element van grote hulp zijn.

Organiseer issue workshops

Binnen de verschillende soorten stakeholders is veelal diepgaande kennis aanwezig van kwaliteitsissues in de data. In workshops is dit een zeer waardevolle bron van problemen in de data. Maar het is ook een mogelijkheid om de bewustwording van datakwaliteit te verhogen

Organiseer maatregel workshops

Naast issue workshops kan het ook interessant zijn om samen met stakeholders op basis van issues te kijken naar relevante en (eenvoudig implementeerbare maatregelen en hiermee de datakwaliteit te verhogen. Deze maatregelen kunnen vervolgens worden opgenomen in het maatregelenregister

Richt change management proces in

ITIL/BISL processen zijn niet alleen inzetbaar voor ICT voorzieningen maar ook voor data wijzigingen die doorgevoerd moeten worden. Sluit aan bij bestaande inrichting in de organisatie voor wijzigingsbeheer en breid dit uit met specifieke implementaties voor data maatregelen. Echter zorg voor een generieke opzet van dit ITIL/BISL proces

Richt data eigenaarschap en stewardship in

Data governance is essentieel voor data kwaliteiten. Zorg dan ook dat de data governance processen ingericht zijn om sturing en uitvoering te kunnen geven aan processen die de data kwaliteiten kunnen verhogen op basis van maatregelen.

Richt incident management proces in

ITIL/BISL processen zijn niet alleen inzetbaar voor ICT voorzieningen maar ook voor data issues die ontstaan binnen de organisatie. Sluit aan bij bestaande inrichting in de organisatie voor incidentbeheer en breid dit uit met specifieke implementaties voor data issues. Echter zorg voor een generieke opzet van dit ITIL/BISL proces

Richt problem management proces in

ITIL/BISL processen zijn niet alleen inzetbaar voor ICT voorzieningen maar ook voor data wijzigingen die doorgevoerd moeten worden. Sluit aan bij bestaande inrichting in de organisatie voor problem management en breid dit uit met specifieke implementaties voor data maatregelen. Echter zorg voor een generieke opzet van dit ITIL/BISL proces

Textmining algoritmen

Inzetten van algoritmen om te zoeken naar patronen in tekstuele informatie. Denk bijvoorbeeld aan sentimentanalyse. Tellen van voorkomende woorden en nabijheid van woorden

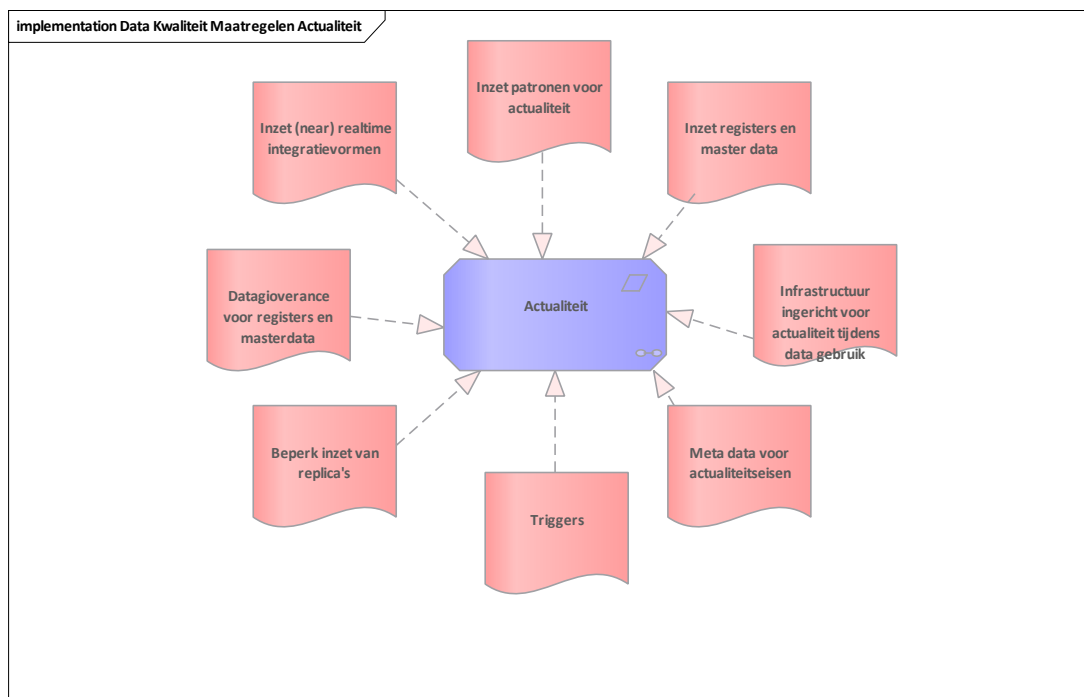
Toepassen Machine Learning algoritmen

Inzetten van machine learning algoritme om de kwaliteit van data te verbeteren. Bijvoorbeeld richting accuraatheid. Denk bijvoorbeeld aan Barneveld en bepalen of de woonplaats in Nederland of de VS is obv van andere attributen in de dataset.

Visualiseer oorzaken van kwaliteiten in gedrag

Maak inzichtelijk voor stakeholders die datakwaliteitsproblemen veroorzaken wat het effect is van handelen bij het werken met de data. Denk bijvoorbeeld aan het misbruiken van velden in schermen die daarmee de kwaliteit in de achterliggende databanken kunnen verlagen.

Data Kwaliteit Maatregelen Actualiteit



Overzicht van de data kwaliteiten en daarbij een aantal maatregelen die genomen kunnen worden om een Data kwaliteit met behulp van deze maatregelen (binnen een release) op een hoger niveau te krijgen.

Actualiteit

Mate waarin een data entiteit de actuele situatie van de werkelijkheid weergeeft.

Beperk inzet van replica's

Beperken van replica's. Actualiteit van gegevens wordt door replica's negatief beïnvloed. Zeker bij de aanwezigheid van meerdere replica's van een gegevensset is de laatst bijgewerkte replica de beperkende factor. Zeker in combinatie met vertragende interactievormen zoals batches e.d.

Datagioverance voor registers en masterdata

Richt eigenaarschap en beheerorganisatie in van bronsystemen en koppelvlakken, door voor bronsystemen en koppelvlakken eigenaarschap te benoemen wordt de verantwoordelijkheid belegd. Deze verantwoordelijke zal veelal geselecteerd worden op basis van betrokkenheid bij de gegevensset (bijvoorbeeld door procesrelevantie).

Introduceren van beslismomen voor het bepalen van de behoefte aan actualiteit en beschikbaarheid van gegevenssets

Inrichten van functionele beheerprocessen en het vastleggen van afspraken rond de actualiteitsniveau's van gegevensverzamelingen.

Infrastructuur ingericht voor actualiteit tijdens data gebruik

Inrichten van een infrastructuur die de gewenste actualiteitsbehoefte kan ondersteunen, denk hierbij aan zaken als complexe berichtafhandeling, inrichten van voorzieningen voor tracking en tracing en correcte foutafhandeling.

Inzet (near) realtime integratievormen

Kies integratievormen die een maximale mutatiesnelheid ondersteunen. De ene integratievorm kent een snellere doorlooptijd dan een andere. Zo zal berichtuitwisseling veelal een betere bijdrage leveren aan de actualiteit bij integratie dan bijvoorbeeld bestandsuitwisseling. Voor de inrichting van deze integratievormen kunnen aanvullende infrastructurele voorzieningen noodzakelijk zijn. Denk bijvoorbeeld aan de inzet van servicebus- of orkestratieplatformen

Bepalen van inrichtingswijzen zoals ETL, service technologie en databaseconcepten als views e.d. die de gewenste behoefte aan actualiteit adequaat ondersteunen.

Inzet patronen voor actualiteit

Inrichten van voorzieningspatronen voor het actueel houden van data entiteiten met name binnen de (interne) kernregisters. Bijvoorbeeld door de inzet van voorzieningen als digilevering en componenten ter ondersteuning van het observer patroon.

Inzet registers en master data

Benoem authentieke- en kernregisters, door de inzet van deze registers kunnen maatregelen genomen worden om zorg te dragen voor de actualiteit van de gegevens. Mede door het terugbrengen van het beheer van de inhoud naar één verantwoordelijke heeft een positief effect op de actualiteit.

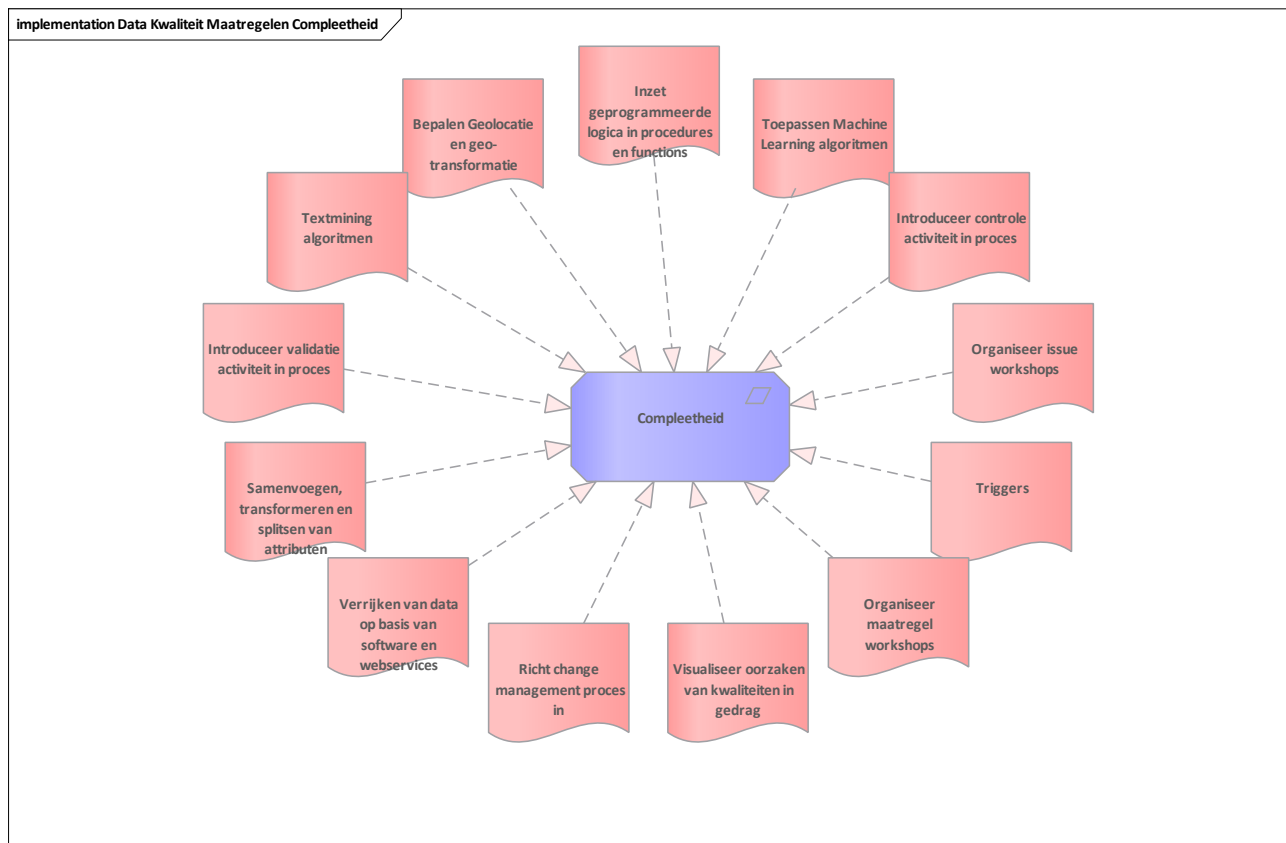
Meta data voor actualiteitseisen

Bepaald actualiteitsbehoefte en -aanbod en registreer dit.

Triggers

Triggers maken het mogelijk om op basis van bewerkingen in de database logica aan te roepen om controles te doen, historie op te bouwen of near real time signalen af te geven aan diverse betrokkenen rond de data.

Data Kwaliteit Maatregelen Compleetheid



Overzicht van de data kwaliteiten en daarbij een aantal maatregelen die genomen kunnen worden om een Data kwaliteit met behulp van deze maatregelen (binnen een release) op een hoger niveau te krijgen.

Bepalen Geolocatie en geo-transformatie

Verrijken en transformeren van data van en naar geo locaties bijvoorbeeld door gebruik van database logica en software

Compleetheid

Dit heeft betrekking op de mate waarin bepaalde attributen binnen een data entiteit aanwezig zijn. Daarnaast geldt de compleetheid ook voor het altijd voorkomen van een bepaalde set van entiteiten (rijen) in een data set.

Introduceer controle activiteit in proces

Wil je de kwaliteit van data waarborgen dan kun je zorgen dat er een kwaliteitsstap wordt ingebouwd in een werkproces. Hierbij zijn verschillende vormen mogelijk. Zoals dat de activiteit door de data producent zelf gedaan wordt of door een andere anders dan de stakeholder zelf. Bekendste voorbeeld is het twee paar ogen principe

Introduceer validatie activiteit in proces

Zorg dat data die geproduceerd of getransformeerd wordt, zeker bij handmatige verwerking van data in een validatie activiteit beoordeeld wordt. Dit biedt een punt in het proces waar issues gesignaleerd kunnen worden, maatregelen genomen kunnen worden of waar signalen zijn dat er in andere data management processen aandachtspunten zijn rond de data kwaliteit.

Inzet geprogrammeerde logica in procedures en functions

Stored procedures en functions maken het mogelijk om transformaties, controles en validaties te automatiseren in de vorm van scripts binnen deze stored procedures en functions. Hiermee heb je vanuit deze programmatuur direct toegang tot de inhoud van de verschillende tabellen.

Organiseer issue workshops

Binnen de verschillende soorten stakeholders is veelal diepgaande kennis aanwezig van kwaliteitsissues in de data. In workshops is dit een zeer waardevolle bron van problemen in de data. Maar het is ook een mogelijkheid om de bewustwording van datakwaliteit te verhogen

Organiseer maatregel workshops

Naast issue workshops kan het ook interessant zijn om samen met stakeholders op basis van issues te kijken naar relevante en (eenvoudig implementeerbare) maatregelen en hiermee de datakwaliteit te verhogen. Deze maatregelen kunnen vervolgens worden opgenomen in het maatregelenregister

Richt change management proces in

ITIL/BISL processen zijn niet alleen inzetbaar voor ICT voorzieningen maar ook voor data wijzigingen die doorgevoerd moeten worden. Sluit aan bij bestaande inrichting in de organisatie voor wijzigingsbeheer en breid dit uit met specifieke implementaties voor data maatregelen. Echter zorg voor een generieke opzet van dit ITIL/BISL proces

Samenvoegen, transformeren en splitsen van attributen

Transformeren naar attributen in datasets door omzetten naar samenvoegen tot een geaggregeerd kenmerk. Daarnaast het opsplitsen van delen van een attribuut op basis van een bepaald kenmerk. Splitsen op komma of spatie.

Textmining algoritmen

Inzetten van algoritmen om te zoeken naar patronen in tekstuele informatie. Denk bijvoorbeeld aan sentimentanalyse. Tellen van voorkomende woorden en nabijheid van woorden

Toepassen Machine Learning algoritmen

Inzetten van machine learning algoritme om de kwaliteit van data te verbeteren. Bijvoorbeeld richting accuraatheid. Denk bijvoorbeeld aan Barneveld en bepalen of de woonplaats in Nederland of de VS is obv van andere attributen in de dataset.

Triggers

Triggers maken het mogelijk om op basis van bewerkingen in de database logica aan te roepen om controles te doen, historie op te bouwen of near real time signalen af te geven aan diverse betrokkenen rond de data.

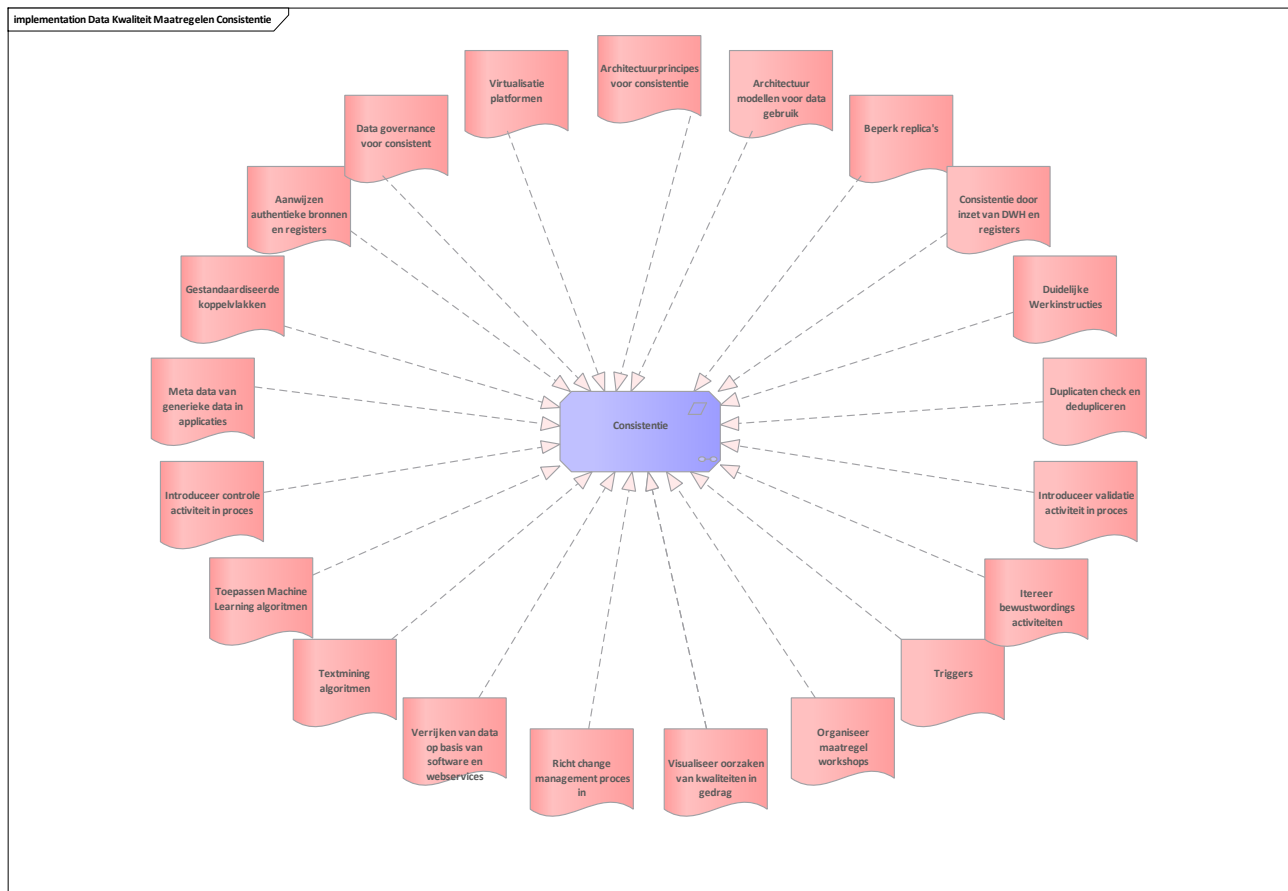
Verrijken van data op basis van software en webservices

Verrijken van data op basis van APIs en webservices voor toevoegen van extra kenmerken. Denk bijvoorbeeld aan postcode.nl voor aanvullen en valideren van adressen op basis van postcode en huisnummer.

Visualiseer oorzaken van kwaliteiten in gedrag

Maak inzichtelijk voor stakeholders die datakwaliteitsproblemen veroorzaken wat het effect is van handelen bij het werken met de data. Denk bijvoorbeeld aan het misbruiken van velden in schermen die daarmee de kwaliteit in de achterliggende databanken kunnen verlagen.

Data Kwaliteit Maatregelen Consistentie



Overzicht van de data kwaliteiten en daarbij een aantal maatregelen die genomen kunnen worden om een Data kwaliteit met behulp van deze maatregelen (binnen een release) op een hoger niveau te krijgen.

Aanwijzen authentieke bronnen en registers

Aanwijzen van unieke bronnen, hiermee wordt bereikt dat de consistentie groter wordt omdat deze unieke bron het aantal replica's en schaduw registraties terugdringt. De data-integratie is vervolgens gebaseerd op het ontsluiten van deze unieke bron.

Architectuur modellen voor data gebruik

Uitwerken van architectuur modellen van de bovengenoemde generieke entiteiten in relatie tot de andere architectuur concepten als functies, processen, actoren en componenten.

Architectuurprincipes voor consistentie

Toepassen van architectuur principes, architectuur principes kunnen bij de uitwerking binnen projecten (bijvoorbeeld in een PSA) zorg dragen dat toegewerkt wordt naar een informatievoorziening die bijdraagt aan een hoge consistentie van data bij opslag, integratie en gebruik.

Beperk replica's

Beperken van replica's brengt met zich mee dat de kans op verschillende inhoud van datasets kleiner wordt. Zeker de situaties waarbij replica's veranderd en verrijkt worden tijdens de data integratie zijn veroorzakers van inconsistentie en moeten daardoor ontmoedigd worden.

Consistentie

Dit heeft betrekking op het feit dat de ene data set van een bepaalde entiteit gelijk is aan een andere dataset. Met andere woorden is een data entiteit onafhankelijk van de bron altijd dezelfde.

Consistentie door inzet van DWH en registers

Introduceren van voorzieningen als het datawarehouse, centrale gegevensvoorziening en gestandaardiseerde webservices. Bij het selecteren van applicaties en (cloud)services rekening houden met het feit dat er gewerkt wordt met authentieke- en kernregisters. Met andere woorden: inzet van applicaties is gebaseerd op een servicelaag waarbinnen generieke data entiteiten worden ingezet.

Data governance voor consistent

Beleggen eigenaarschap en beheerorganisatie. Hiermee wordt gerealiseerd dat de definitie van een gegevensentiteit bewaakt wordt door een eigenaar en dat de beheerorganisatie zorgdraagt voor de bewaking van deze consistentie op basis van de door de eigenaar benoemde definitie. De integratie is vervolgens gebaseerd op deze definitie.

Duidelijke Werkinstructies

Zorg voor duidelijke werkinstructies vanuit het perspectief van data kwaliteit. Veelal wordt in werkinstructies een introductie gedaan van workarounds die een negatief effect kunnen hebben op de kwaliteit van de data

Duplicaten check en dedupliceren

Signaleren van duplicaten in sleutels en in combinaties van attributen in dataobjecten.

Gestandaardiseerde koppelvlakken

Inzetten van gestandaardiseerde koppelvlakken vergroot de kans op hergebruik door een grote groep afnemers. Hierdoor verminder de behoefte om replica's in te zetten. Beheerprocessen kunnen bijdragen aan het introduceren van deze gestandaardiseerde koppelvlakken.

Introduceer controle activiteit in proces

Wil je de kwaliteit van data waarborgen dan kun je zorgen dat er een kwaliteitsstap wordt ingebouwd in een werkproces. Hierbij zijn verschillende vormen mogelijk. Zoals dat de activiteit door de data producent zelf gedaan wordt of door een andere anders dan de stakeholder zelf. Bekendste voorbeeld is het twee paar ogen principe

Introduceer validatie activiteit in proces

Zorg dat data die geproduceerd of getransformeerd wordt, zeker bij handmatige verwerking van data in een validatie activiteit beoordeeld wordt. Dit biedt een punt in het proces waar issues gesignaleerd kunnen worden, maatregelen genomen kunnen worden of waar signalen zijn dat er in andere data management processen aandachtspunten zijn rond de data kwaliteit.

Itereer bewustwordings activiteiten

Bewustwording en gedragsverandering is helaas aan slijtage onderhevig binnen de populatie van een organisatie. Dat houdt in dat er iteratief opnieuw aandacht besteed moet worden aan de bewustwording. Ook hier kan een gaming element van grote hulp zijn.

Meta data van generieke data in applicaties

Bepalen van generieke data entiteiten en bepalen van de mate van consistentie ten opzichte van de aanwezige applicaties.

Organiseer maatregel workshops

Naast issue workshops kan het ook interessant zijn om samen met stakeholders op basis van issues te kijken naar relevante en (eenvoudig implementeerbare maatregelen en hiermee de datakwaliteit te verhogen. Deze maatregelen kunnen vervolgens worden opgenomen in het maatregelenregister

Richt change management proces in

ITIL/BISL processen zijn niet alleen inzetbaar voor ICT voorzieningen maar ook voor data wijzigingen die doorgevoerd moeten worden. Sluit aan bij bestaande inrichting in de organisatie voor wijzigingsbeheer en breid dit uit met specifieke implementaties voor data maatregelen. Echter zorg voor een generieke opzet van dit ITIL/BISL proces

Textmining algoritmen

Inzetten van algoritmen om te zoeken naar patronen in tekstuele informatie. Denk bijvoorbeeld aan sentimentanalyse. Tellen van voorkomende woorden en nabijheid van woorden

Toepassen Machine Learning algoritmen

Inzetten van machine learning algoritme om de kwaliteit van data te verbeteren. Bijvoorbeeld richting accuraatheid. Denk bijvoorbeeld aan Barneveld en bepalen of de woonplaats in Nederland of de VS is obv van andere attributen in de dataset.

Triggers

Triggers maken het mogelijk om op basis van bewerkingen in de database logica aan te roepen om controles te doen, historie op te bouwen of near real time signalen af te geven aan diverse betrokkenen rond de data.

Verrijken van data op basis van software en webservices

Verrijken van data op basis van APIs en webservices voor toevoegen van extra kenmerken. Denk bijvoorbeeld aan postcode.nl voor aanvullen en valideren van adressen op basis van postcode en huisnummer.

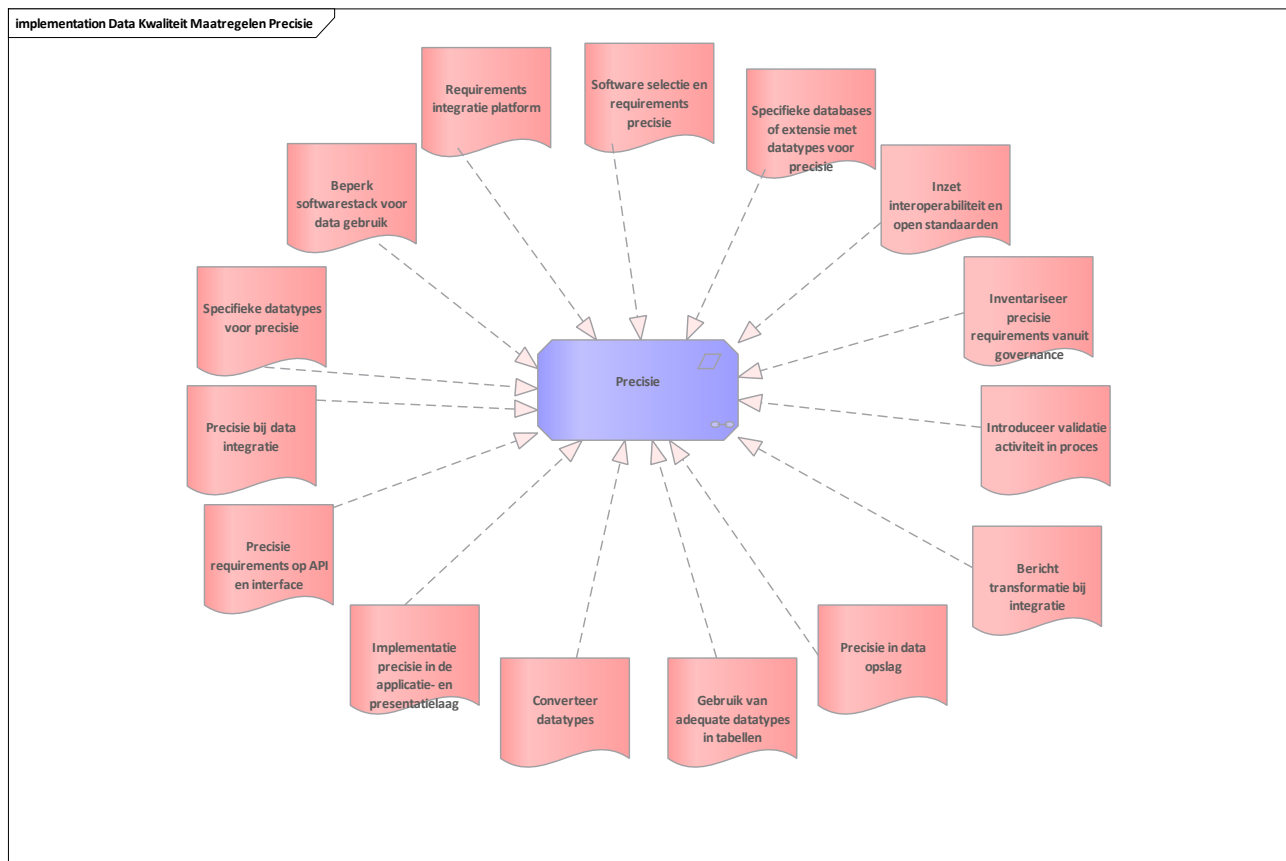
Virtualisatie platformen

Inzet datavirtualisatie, hiermee wordt bewerkstelligd dat er slechts één versie van de gegevens beheerd wordt en beschikbaar wordt gesteld via de generieke voorziening voor datavirtualisatie. Door inzet van datavirtualisatie kan in een later stadium, indien gewenst, een andere dataopslag als unieke bron geselecteerd worden zonder dat dit gevolgen heeft voor de afnemers van deze dataset.

Visualiseer oorzaken van kwaliteiten in gedrag

Maak inzichtelijk voor stakeholders die datakwaliteitsproblemen veroorzaken wat het effect is van handelen bij het werken met de data. Denk bijvoorbeeld aan het misbruiken van velden in schermen die daarmee de kwaliteit in de achterliggende databanken kunnen verlagen.

Data Kwaliteit Maatregelen Precisie



Overzicht van de data kwaliteiten en daarbij een aantal maatregelen die genomen kunnen worden om een Data kwaliteit met behulp van deze maatregelen (binnen een release) op een hoger niveau te krijgen.

Beperk softwarestack voor data gebruik

Beperking van het aantal verschillende softwarestacks en zorgdragen dat deze aanwezige softwarestacks in de infrastructuur gestandaardiseerd worden ingericht. Bijvoorbeeld het dotnet framework als basis binnen de windows server infrastructuur.

Bericht transformatie bij integratie

Transformatie ten behoeve van integratie is een veelvoorkomende bron van precisieproblemen bij integratie. Analyseer risico's en zoek naar componenten en integratievormen die de gewenste precisie tijdens data transport kunnen handhaven.

Converteer datatypes

Converteren van datatypes naar andere datatypes (tekst <-> Numeriek). Inclusief transformatie naar datatypes op andere platformen zoals database, XML, softwaretalen en localisatie.

Gebruik van adequate datatypes in tabellen

Datatypes in de verschillende database platformen maken het mogelijk om op basis van datatypen beperkingen in te bouwen in verschillende datatypen. Zo kun je verschillende numerieke datatypen kiezen zoals integer, decimalen (met het aantal cijfers voor of achter de komma, float, etc. Hiermee zijn extra validaties mogelijk op basis van deze datatypen

Implementatie precisie in de applicatie- en presentatielaag

Gebruik van gestandaardiseerd geconfigureerde componenten in de applicatie- en presentatielaag.

Introduceer validatie activiteit in proces

Zorg dat data die geproduceerd of getransformeerd wordt, zeker bij handmatige verwerking van data in een validatie activiteit beoordeeld wordt. Dit biedt een punt in het proces waar issues gesignaleerd kunnen worden, maatregelen genomen kunnen worden of waar signalen zijn dat er in andere data management processen aandachtspunten zijn rond de data kwaliteit.

Inventariseer precisie requirements vanuit governance

Houdt bij de initiële inrichting van een omgeving rekening met de precisie requirements van (toekomstige) afnemers. Zeker bij opslag van registerdata dient door de eigenaar geïnventariseerd te worden wat de precisiebehoefte van de afnemers is.

Richtlijnen opstellen omtrent inrichting van gegevensbestanden, applicatie componenten en elementen binnen de presentatielaag op het vlak van precisie.

Inzet interoperabiliteit en open standaarden

Selectie van open standaarden, bij het ontwikkelen van open standaarden zal het precisie aspecten aan de orde komen. Hiermee zal de precisie aansluiten op de wensen van een grote groep organisaties. Door het inzetten van deze standaarden bij de integratie zal de precisie voldoende gewaarborgd zijn.

Precisie

Mate van detail waarin een data entiteit de werkelijkheid weergeeft. Dit heeft bijvoorbeeld betrekking op de precisie van getallen e.d. Opslag van getallen en datums kunnen onvoldoende nauwkeurig zijn omdat afronding bij opslag of integratie nodig is.

Precisie bij data integratie

Keuze en inzet integratietechnologie, hierbij moet rekening gehouden worden met de gewenste- en realiseerbare precisie. Dat kan tot gevolg hebben dat een integratievorm ongeschikt is voor een specifieke implementatie. Houdt hierbij ook rekening met het feit dat structuren op basis van de precisie opnieuw opgebouwd moeten kunnen worden, bijvoorbeeld bij het verwerken van object georiënteerde datastructuren.

Precisie in data opslag

Onderzoek of de bestaande datatypes binnen de data opslag omgeving, zoals een database platform, voldoende ondersteuning bieden aan de gewenste precisie.

Precisie requirements op API en interface

Bij registerfunctie van de benoemde gegevensopslag dient bij afwijking van de standaard precisie deze beschreven te zijn in de (interface) documentatie zodat toekomstige afnemers het kwaliteitsniveau kunnen toetsen aan de eigen behoefte.

Requirements integratie platform

Selectieproces integratievorm, bij het uitwerken van de integratievormen kan beschreven worden wat de effecten zijn van een specifieke integratievorm op het vlak van precisie. Vervolgens kan dit vertaald worden in beslisbomen en -documenten rond het selecteren van integratievormen binnen een project.

Software selectie en requirements precisie

Beschrijf bij een software selectie traject eventuele precisie aspecten op basis van requirements. Zoals precisie aspecten meenemen in het ontwerpproces van nieuwe applicaties e.d.

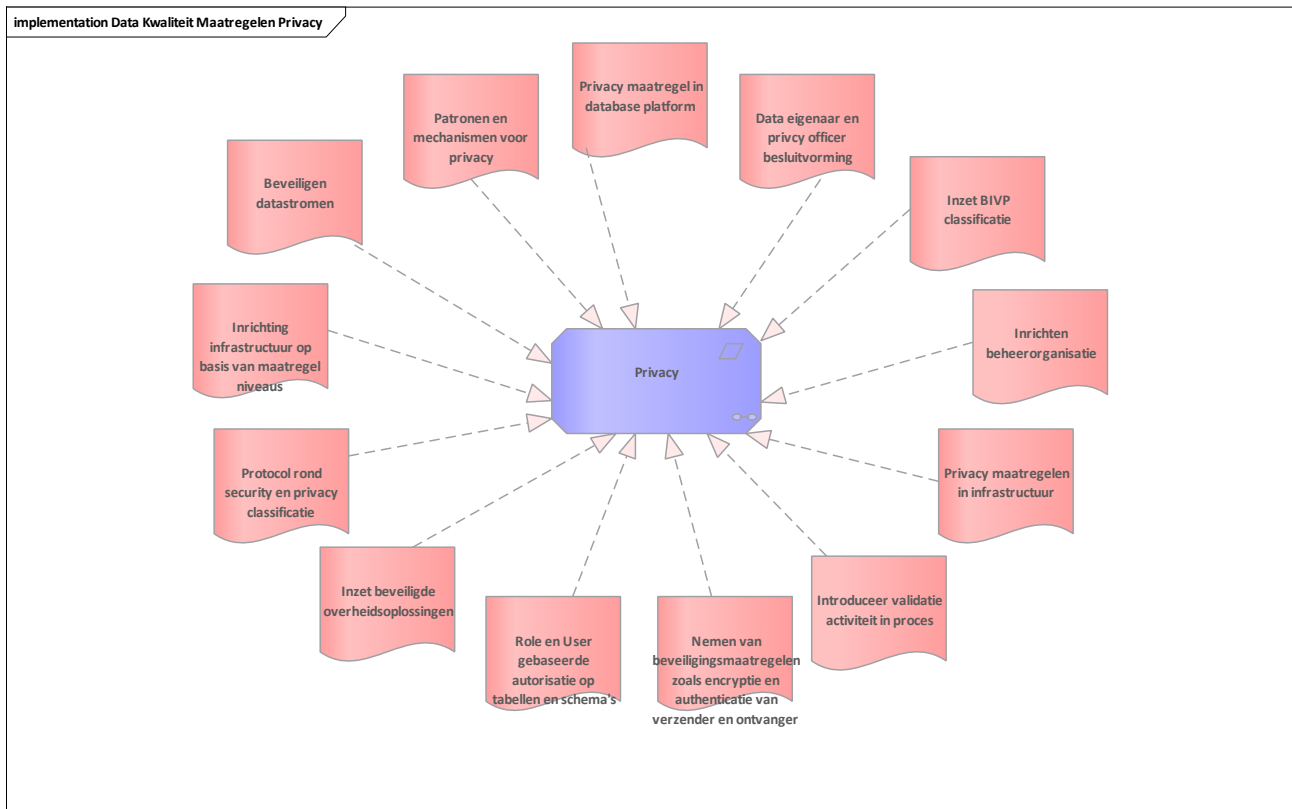
Specifieke databases of extensie met datatypes voor precisie

Pas bij opslag van gegevenscollecties met een hoge mate van precisie of een afwijking van de standaard datatypes gespecialiseerde dataopslagomgevingen toe. Denk aan omgevingen als geo- en objectdatabases. Houdt hierbij wel rekening met de beheeraspecten van een dergelijke gespecialiseerde omgevingen.

Specifieke datatypes voor precisie

Volstaan bestaande data typen en formaten binnen de data opslag niet dan kan het ontwikkelen van specifieke datatypen en -structuren een oplossing bieden. Houdt er rekening mee dat je hiermee een vorm van platform lock in introduceert.

Data Kwaliteit Maatregelen Privacy



Overzicht van de data kwaliteiten en daarbij een aantal maatregelen die genomen kunnen worden om een Data kwaliteit met behulp van deze maatregelen (binnen een release) op een hoger niveau te krijgen.

Beveiligen datastromen

Beveiliging van integratiestromen zal bijdragen aan de beveiliging van de gegevensset tijdens het transport. Dit gebeurt veelal door het versleutelen van de gegevensstroom en het authenticeren en autoriseren van bron- en afnemers. Houdt er rekening mee dat deze activiteiten een negatief effect kunnen hebben op de andere kwaliteiten als actualiteit en tijdigheid. Door het uitwerken van de verschillende beveiligingsvormen kan een adequate keuze van maatregelen gemaakt worden.

Data eigenaar en privy officer besluitvorming

Bepaal in samenspraak met de eigenaar de beveiligingscategorieën en richt op basis daarvan de infrastructuur in en neem de bijbehorende maatregelen in de beheerorganisatie.

Inrichten beheerorganisatie

Richt de beheerorganisatie in voor de verschillende IB aspecten bijvoorbeeld door het inrichten van de processen op basis van ISO 27001.

Inrichten van een beheerrepository voor het ontsluiten van risico's en de bijbehorende te nemen maatregelen, op basis waarvan de werkprocessen op eenvoudige wijze kunnen voldoen aan de beveiligingseisen.

Inrichting infrastructuur op basis van maatregel niveaus

Inrichten van infrastructuur op basis van meerdere niveaus (brons, zilver en goud) voor verschillende informatiebeveiligingsniveaus. De beveiligingsbehoefte door de eigenaar van de gegevens laten bepalen en op basis daarvan de gegevensopslag inrichten.

Introduceer validatie activiteit in proces

Zorg dat data die geproduceerd of getransformeerd wordt, zeker bij handmatige verwerking van data in een validatie activiteit beoordeeld wordt. Dit biedt een punt in het proces waar issues gesignaleerd kunnen worden, maatregelen genomen kunnen worden of waar signalen zijn dat er in andere data management processen aandachtspunten zijn rond de data kwaliteit.

Inzet beveiligde overheidsoplossingen

inzet van (overheids)componenten en standaarden. Overheidscomponenten en -standaarden als SSL, TLS, eherkenning, digid, digikoppeling en digipoort hebben in de basis al een aantal beveiligingsmaatregelen genomen voor het borgen van de privacy tijdens het transport. Hiermee wordt voor de meeste situaties een voldoende hoog beveiligingsniveau gerealiseerd.

Inzet BIVP classificatie

Inrichten dataclassificatie, slechts voor een beperkt aantal typen gegevenssets speelt privacy een rol. Denk bijvoorbeeld aan persoonsgegevens of gegevens over financiën, opsporing en handhaving. In alle andere gevallen zijn bij integratie privacy aspecten van minder belang. In die situatie kan volstaan worden met minder maatregelen voor de beveiliging. Door in kaart te brengen voor welke data objecten privacy relevant is kan eenvoudig bepaald worden in welke projecten met data integratie aanvullende maatregelen nodig zijn.

Nemen van beveiligingsmaatregelen zoals encryptie en authenticatie van verzender en ontvanger

Gebruik van beveiligingsmaatregelen binnen de data integratie om zorg te dragen voor de beveiligde uitwisseling van data in berichtenverkeer tussen informatiesystemen en organisaties.

Patronen en mechanismen voor privacy

Dienen zaken voor traceerbaarheid van bewerkingen en bevestigingen ingeregeld worden in de vorm van logging en database inrichting. Denk hierbij aan patronen bij de inrichting en implementatie.

Privacy

Voor sommige data entiteiten is toegangscontrole (autorisatie en authenticatie) of monitoring van gebruik nodig. Denk hierbij aan eisen die gesteld worden aan de ontsluiting van (privacy) vertrouwelijke gegevens.

Privacy maatregel in database platform

Bij de keuze van het opslagplatform wordt rekening gehouden met de eisen die vanuit informatiebeveiliging gesteld worden denk bijvoorbeeld aan zaken als autorisatie, identificatie en authenticatie, encryptie en beschikbaarheid. De requirements worden getoetst op basis een data classificatie van de gegevens (zie de beheerprocessen).

Privacy maatregelen in infrastructuur

Infrastructuur kent een compartimentering waarbij enerzijds de compartimenten van elkaar gescheiden zijn en anderzijds deze compartimenten verschillende beveiligingsniveaus kennen.

Nemen van infrastructurele maatregelen om tegemoet te komen aan beveiligingsbehoeften tijdens het gebruik. Denk hierbij aan zaken als het beveiligen van verbindingen, het hardenen van infrastructuur of het versleutelen van gegevens binnen een applicatie.

Protocol rond security en privacy classificatie

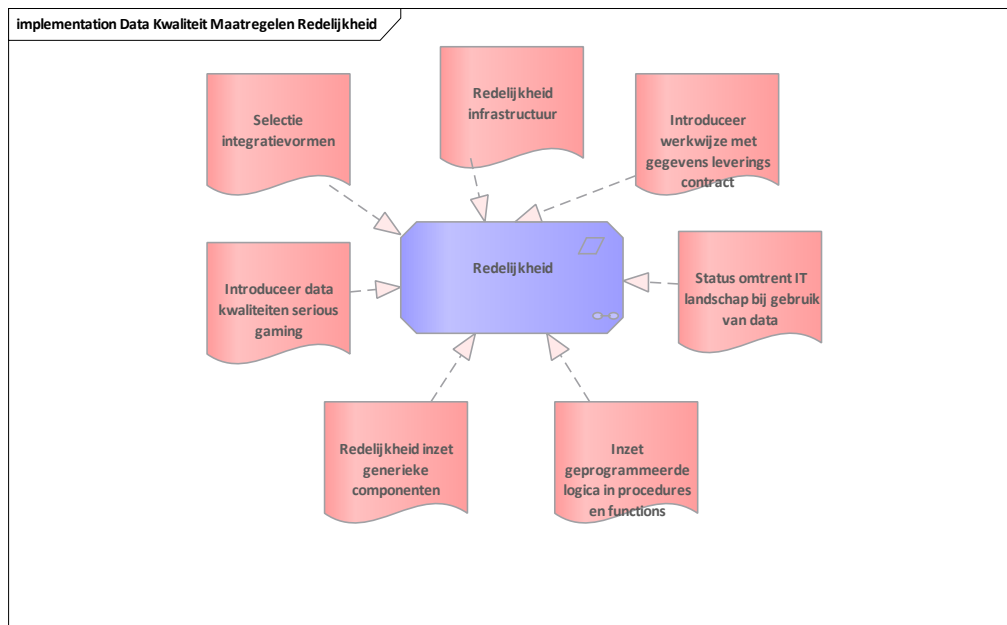
Beslisproces voor beveiliging bij integratie, dit punt sluit aan bij de voorgaande punten. Introduceer een beslisproces om zorg te dragen dat op eenvoudige wijze bepaald kan worden.

Inrichten van beheerprocessen rond informatiebeveiliging. Denk bijvoorbeeld aan het monitoren van beveiligingsaspecten maar ook toetsingsmechanismen zoals checklists inzetten binnen projecten of het doen van beveiligingsaudits.

Role en User gebaseerde autorisatie op tabellen en schema's

In databases is op fijnmazig niveau de autorisatie van gebruikers en groepen gebruikers. Dit is gebaseerd op het implementeren van CRUD (Create/Read/Update/Delete) in de database per tabel of schema

Data Kwaliteit Maatregelen Redelijkheid



Overzicht van de data kwaliteiten en daarbij een aantal maatregelen die genomen kunnen worden om een Data kwaliteit met behulp van deze maatregelen (binnen een release) op een hoger niveau te krijgen.

Introduceer data kwaliteiten serious gaming

Data kwaliteiten, bewustwording en verandering van gedrag kan met serious gaming inzichtelijk worden gemaakt. Denk aan interactieve workshops waarin met behulp van een competitief element de kwaliteit van data onder de aandacht wordt gebracht

Introduceer werkwijze met gegevens leverings contract

Inzet servicecontract of gegevensleveringsovereenkomst. Deze documenten met afspraken kunnen redelijkheid formaliseren en op deze wijze een bijdrage leveren aan een realistische uitwerking van vraag en aanbod op diverse aspecten van redelijkheid bij data integratie. Denk hierbij aan de uitwerking van bijvoorbeeld tijdigheid, actualiteit, consistentie en validiteit
Om redelijkheid te ondersteunen is het van belang dat voor de belangrijkste gebruikgevallen beschreven is wat de redelijkheidsgrenzen dienen te zijn. Dit houdt in dat aangegeven wordt in welke situatie de kwaliteiten dermate laag zijn dat redelijkheid niet meer geldend is.

Inzet geprogrammeerde logica in procedures en functions

Stored procedures en functions maken het mogelijk om transformaties, controles en validaties te automatiseren in de vorm van scripts binnen deze stored procedures en functions. Hiermee heb je vanuit deze programmatuur direct toegang tot de inhoud van de verschillende tabellen.

Redelijkheid

Heeft vooral betrekking op verwachtingen van data entiteiten binnen een bepaalde operationele context.

Redelijkheid infrastructuur

Infrastructurele inrichting, data integratie stelt bij een aantal integratievormen hoge eisen aan de infrastructurele inrichting. Bijvoorbeeld op het vlak van performance, beschikbaarheid en capaciteit. Bij de infrastructurele inrichting dient rekening gehouden te worden met de eisen die vanuit de redelijkheid gesteld worden. Dit brengt een bepaalde configuratie met zich mee per integratievorm. Daarnaast kan de inzet van specifieke componenten noodzakelijk zijn.

Redelijkheid inzet generieke componenten

Inzet van (overheids)componenten. Hierbij dient terdege rekening gehouden te worden met de eisen die vanuit redelijkheid gesteld worden. Veel overheidscomponenten kennen een inrichting die bepaalde integratievormen in combinatie met redelijkheid niet mogelijk maken. Denk bijvoorbeeld aan de inzet van digikoppeling en digipoort in combinatie met de wens voor synchrone (redelijke) interactie.

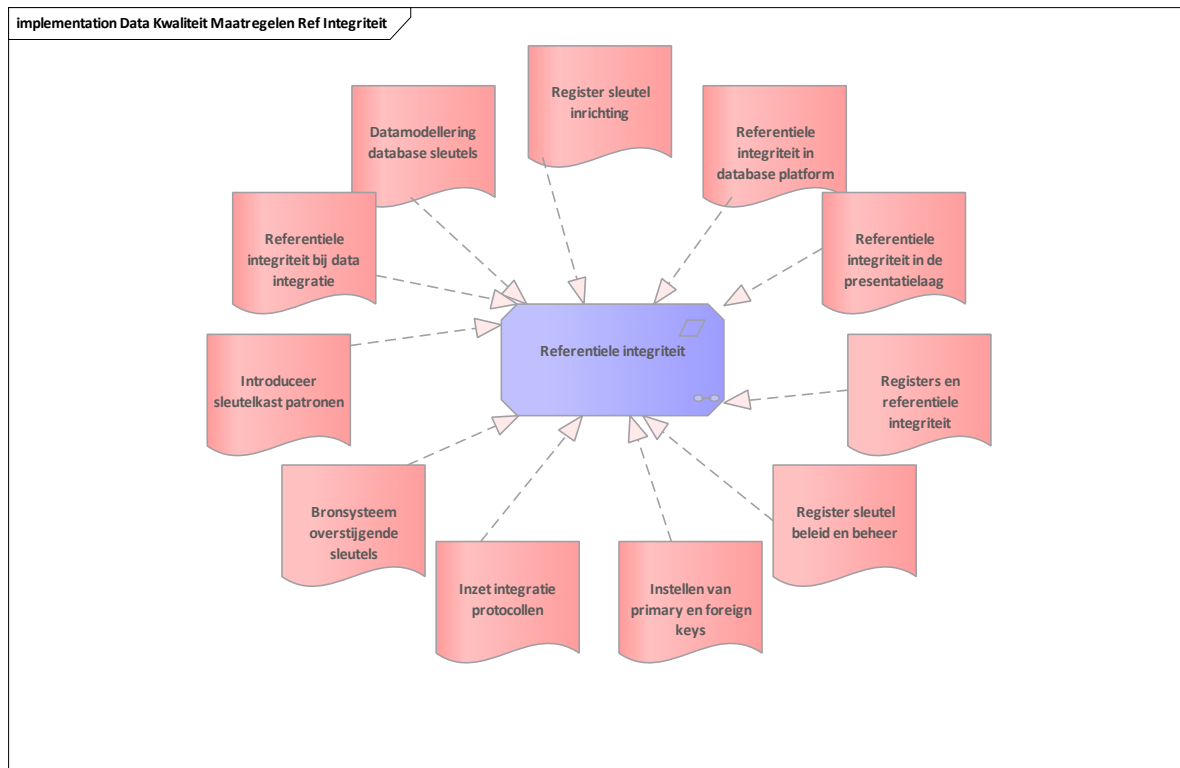
Selectie integratievormen

Integratievorm keuze, vanuit het perspectief van redelijkheid is de keuze van bepaalde integratievormen wel of niet mogelijk. Vanuit architectuur en beheer kunnen beslisbomen en dergelijke ingezet worden om zorg te dragen dat de juiste vorm in relatie tot redelijkheid gekozen worden.

Status omtrent IT landschap bij gebruik van data

Voor de technische voorzieningen is het met name van belang dat men tijdens het data gebruik geïnformeerd wordt over de status waarin de (keten van) voorzieningen zich in bevindt.

Data Kwaliteit Maatregelen Ref Integriteit



Overzicht van de data kwaliteiten en daarbij een aantal maatregelen die genomen kunnen worden om een Data kwaliteit met behulp van deze maatregelen (binnen een release) op een hoger niveau te krijgen.

Bronstelsysteem overstijgende sleutels

Applicatie of organisatie overstijgende sleutels, bij data integratie kan een sleutel vanuit een bronsysteem onvoldoende uniek zijn. Zeker in het geval waar bijvoorbeeld gegevensset gecombineerd worden met andere sets kunnen aanvullende maatregelen noodzakelijk zijn. In die gevallen kan het benoemen van sleutels met voldoende uniekheid vanuit architectuur of de eigenaar van gegevenssets noodzakelijk zijn. Denk bijvoorbeeld aan het BSN en KvK nummer als organisatie overstijgende sleutel.

Datamodellering database sleutels

Beschrijven en toepassen van patronen voor sleutel identificatie en het afdwingen van referentiele integriteit. Zoals het uitbreiden van interne sleutels met een unieke applicatiesleutel of het toepassen van GUID's e.d. Pas deze patronen toe binnen zowel de datawarehouse als de operational datastore omgeving toe.

Instellen van primary en foreign keys

Primary en foreign keys maken het mogelijk om controles uit te voeren op de sleutels (primaire) in een tabel en vervolgens de verwijzingen vanuit andere tabellen naar deze primaire sleutel af te dwingen. Hiermee voorkom je verwijzingen die niet correct zijn in een goed geconfigureerde database

Introduceer sleutelkast patronen

Inzet sleutelkasten, in een aantal gevallen kunnen bij de data integratie de gegevens van een gegevensset verrijkt worden met verwijzingen naar referentiele sleutels zoals die toegepast worden op andere plaatsen binnen de organisatie. Hiermee kan dan op eenvoudige wijze herleidt worden welke identificerende sleutel waar toegepast kan worden. Sleutelkasten worden veelal beschreven binnen de architectuur in samenspraak met de eigenaren van de verschillende registers.

Inzet integratie protocollen

Inzet van protocollen. Denk hierbij bijvoorbeeld aan WS-reliable messaging dat kan zorgdragen voor het compleet afleveren van een gegevensset, het zorgdragen voor de juiste volgorde van de pakketten na ontvangst en dergelijke. Dit soort protocollen kunnen zorgdragen voor een voldoende hoge integriteit

Referentiele integriteit

Dit is de situatie waarbij verwijzingen vanuit de ene data entiteit altijd correct verwijzen naar de gerelateerde data entiteiten.

Referentiele integriteit bij data integratie

Inrichten integratievorm, bij verschillende vormen van integratievormen kan de wijze van inrichting een negatief effect hebben op de referentiele integriteit. Bijvoorbeeld bij integratie op basis van berichten kunnen aanvullende eisen gesteld worden aan de in te zetten componenten, verbindingen en protocollen. Bij de inzet van generieke integratievoorziening dient rekening gehouden te worden met de strengste eisen van integriteit zoals deze binnen deze voorziening ingezet zal worden.

De technische voorzieningen moeten voldoende oplossingen te hebben voor het handhaven van referentiele integriteit. Denk bijvoorbeeld aan voorzieningen als het genereren van unieke sleutels het werken met transacties en rollback mechanismen. Met name in een sterk gedistribueerde omgeving zoals een SOA omgeving is dit een uitdaging.

Referentiele integriteit in database platform

Bepaal de eisen voor referentiele integriteit en kies op basis daarvan het opslagplatform. Richt dit platform op dusdanige wijze in dat de referentiele integriteit maximaal wordt gerealiseerd. Dus als er constraints mogelijk zijn richt deze ook in.

Referentiele integriteit in de presentatielaag

Componenten binnen een applicatie dienen tijdens het gebruik rekening te houden met de referentiele integriteit van de gegevens. Dat betekent dat voldoende uniek identificerende gegevens getoond moeten worden, maar ook dat bijvoorbeeld sleutels tijdens bewerkingssessies onzichtbaar gehandhaafd worden.

Register sleutel beleid en beheer

Met name rond toepassingsgebied overstijgende sleutels dient het beheer en eigenaarschap ingeregeld te worden. Enerzijds bij het selecteren en beschrijven van deze sleutels, anderzijds bij het bewaken van het gebruik van deze sleutels binnen projecten waar deze sleutels ingezet dienen te worden. Het is een logische keuze deze bewaking binnen de taken van de data-architect te beleggen.

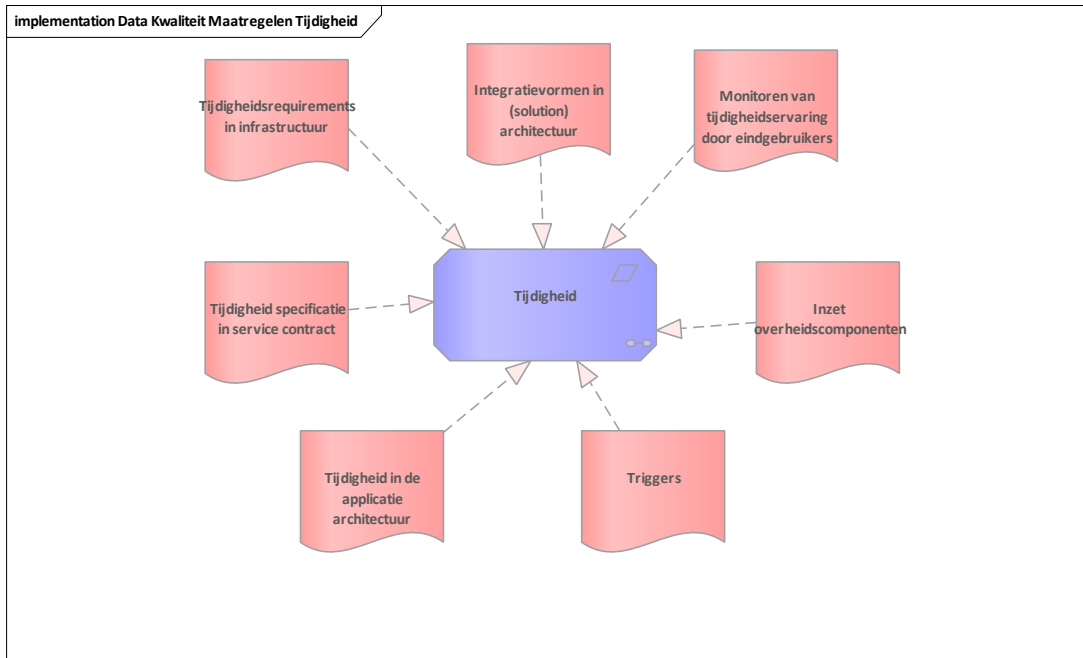
Register sleutel inrichting

Werk eventueel met toepassingsgebied overstijgende sleutels voor het afdwingen van referentiele integriteit. Bijvoorbeeld bij service oriëntatie of keten integratie over de grenzen van een applicatie of organisatie heen. Hierbij kan de inzet van een sleutelkast component of service uitkomst bieden

Registers en referentiele integriteit

Pas repositories en registries toe als er gewerkt wordt met een gegevensopslag die het gebruik van sleutels minder goed ondersteunt. Denk bijvoorbeeld aan het ontsluiten van diverse soorten van bestanden waarbij de bestandsnaam als sleutel wordt gebruikt of waarbij de sleutel door precisieproblemen niet gelijk blijven.

Data Kwaliteit Maatregelen Tijdigheid



Overzicht van de data kwaliteiten en daarbij een aantal maatregelen die genomen kunnen worden om een Data kwaliteit met behulp van deze maatregelen (binnen een release) op een hoger niveau te krijgen.

Integratievormen in (solution) architectuur

Integratievorm keuze bijvoorbeeld door inzet van beslisbomen opgesteld vanuit architectuur kunnen zorgdragen dat op het vlak van tijdigheid voldoende rekening gehouden wordt met de risico's die gelden. Dit zal uitgewerkt worden in de projectdocumenten zoals de PSA e.d.

Inzet overheidscomponenten

Inzet van overheidscomponenten hebben eveneens een nauwe relatie met tijdigheid. Met name bij een keten van overheidscomponenten zal veelal een negatieve uitwerking hebben op de tijdigheid. Bij het uitwerken van een integratieketen zal een risico analyse plaats moeten vinden per component op onder andere het aspect van tijdigheid.

Monitoren van tijdigheidservaring door eindgebruikers

Monitoren van eindgebruikerservaringen bij data gebruik en op basis van deze monitoring bepalen of er optimalisatieactiviteiten ten behoeve van de tijdigheid plaats moeten vinden.

Tijdigheid

Is een dataset tijdig beschikbaar binnen de gestelde verwachtingen. Het is het verschil tussen het moment van behoefte en beschikbaarheid.

Tijdigheid in de applicatie architectuur

Inrichting van de applicatie componenten dient de tijdigheid niet negatief te beïnvloeden. Denk daarbij aan het kiezen van de juiste architectuurkeuzen binnen de componenten, maar ook aan de inrichting van de aan elkaar gerelateerde onderdelen binnen een configuratie.

Paradigma keuze kan invloed hebben op de tijdigheid. Zo zie je meer een meer dat data gebruik webbased is. Echter bij het verwerken van gegevens bijvoorbeeld in (geo) viewers kan een andere paradigma (desktopversie) vanuit tijdigheidsperspectief de voorkeur verdienen.

Tijdigheid specificatie in service contract

Inzet servicecontract of gegevensleveringsovereenkomst. Deze documenten met afspraken kunnen tijdigheid als verbijzondering van redelijkheid formaliseren en op deze wijze een bijdrage leveren aan een realistische uitwerking van vraag en aanbod op aspecten van tijdigheid bij data integratie.

Gebruiksafspraken maken met gebruikers en beheerders om zo voor een betere verdeling van het gebruik te bereiken (zie ook redelijkheid), bijvoorbeeld door contractafspraken te maken omtrent het gegevensgebruik en het uitvoeren van beheeractiviteiten.

Tijdigheidsrequirements in infrastructuur

Infrastructurele inrichting heeft een nauwe relatie met tijdigheid. Beschreven dient te zijn welke eisen op het vlak van tijdigheid en op welke wijze dat met de juiste infrastructurale componenten gerealiseerd kan worden. Hierbij kunnen ook infrastructurale componenten noodzakelijk zijn die in de huidige inrichting nog niet aanwezig zijn. Vanuit architectuur dient dit voldoende bewaakt te worden in samenspraak met de beheerorganisatie.

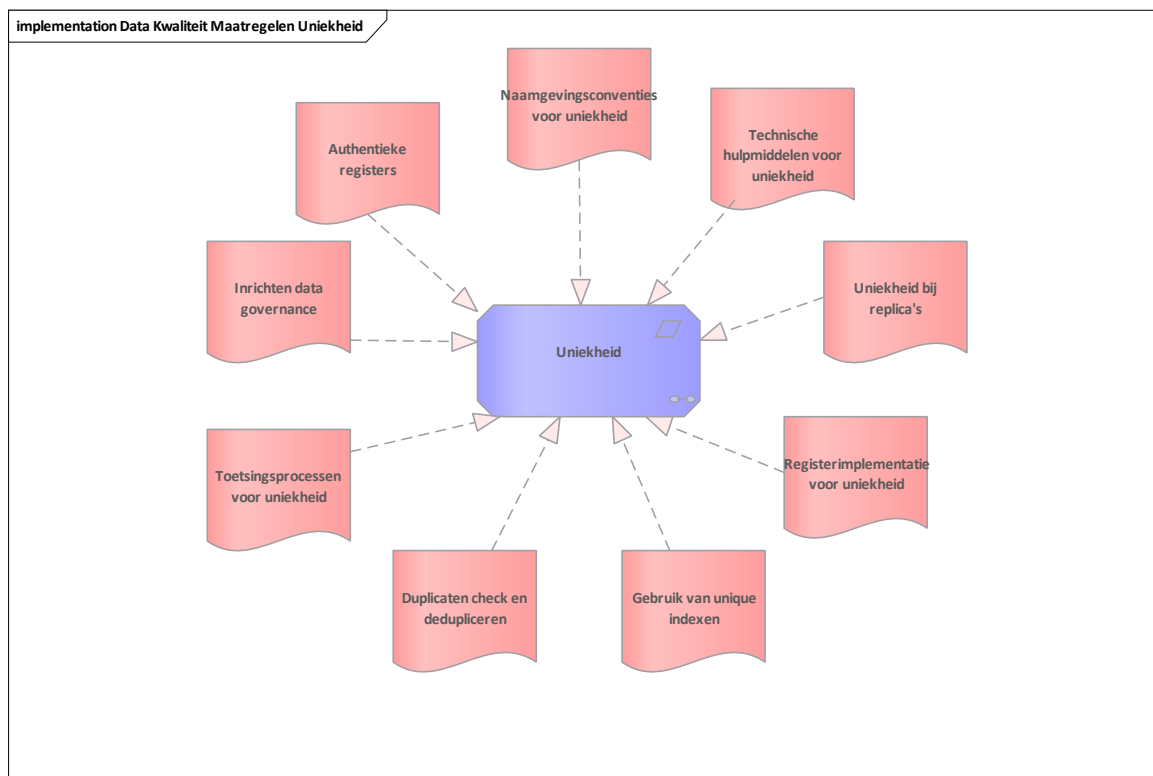
Optimalisatie van de applicatie infrastructuur. De applicatie infrastructuur kan een grote bijdrage leveren aan de tijdigheid. Denk bijvoorbeeld aan het kiezen van virtuele werkomgevingen die ingezet worden voor bewerkingen die hoge eisen stellen aan de onderliggende infrastructuur terwijl de schermwijzigingen beperkt zijn en daarom via (tragere) netwerkvoorzieningen kunnen verwerkt.

Inzetten van elasticiteitshulpmiddelen. Tijdigheid kan negatief beïnvloed worden op momenten dat er veel vraag is naar een bepaalde voorziening. Inzet van elasticiteit bijvoorbeeld cloud technologie, kan bijdragen aan een verbeterde tijdigheid.

Triggers

Triggers maken het mogelijk om op basis van bewerkingen in de database logica aan te roepen om controles te doen, historie op te bouwen of near real time signalen af te geven aan diverse betrokkenen rond de data.

Data Kwaliteit Maatregelen Uniekheid



Overzicht van de data kwaliteiten en daarbij een aantal maatregelen die genomen kunnen worden om een Data kwaliteit met behulp van deze maatregelen (binnen een release) op een hoger niveau te krijgen.

Authentieke registers

Benoem authentieke en kernregisters die aangewezen worden voor de opslag van bedrijfsobjecten en ontsluit deze via een gestandaardiseerde interface op basis van views en/of services.

Duplicaten check en dedupliceren

Signaleren van duplicaten in sleutels en in combinaties van attributen in dataobjecten.

Gebruik van unique indexen

Unieke indexen aanleggen kan worden gedaan op basis van de sleutel (de primary keys bijvoorbeeld). Echter als je dat zou willen dan kan dat ook op een combinatie van kolommen voor extra duplicatie validatie checks

Inrichten data governance

Richt processen in voor eigenaarschap en beheer zodat bij nieuwe ontwikkelingen besluitvorming en inrichting eenvoudig gefaciliteerd kunnen worden.

Beschrijf wat de (unieke) bron is van gegevensentiteiten en de afnemers zijn. Neem daarbij mee wat de actualiteitsbehoefte is van de verschillende afnemers is en richt de beheerprocessen van het bronsysteem hierop in.

Naamgevingsconventies voor uniekheid

Pas naamgevingsconventies toe voor entiteiten, attributen en relaties waardoor de kans op het ontstaan van dubbele opslag in registers verkleind wordt.

Registerimplementatie voor uniekheid

Zet registers en/of repositories in de beschrijving van entiteiten. Hiermee worden projecten en beheerprocessen gefaciliteerd in situaties waar herbruikbaarheid van gegevensdefinitie relevant zijn.

Technische hulpmiddelen voor uniekheid

Zet technische hulpmiddelen in die uniekheid afdwingen, denk hierbij bijvoorbeeld aan relationele databases met voorzieningen als primaire sleutels en clustered indexen.

Toetsingsprocessen voor uniekheid

Draag zorg voor een toetsingsmechanisme voor de uniekheid van entiteiten. Maak iemand verantwoordelijk voor deze uniekheidstoetsing van entiteiten. Kies in eerste instantie voor toetsing van de bedrijfsobjecten en bepaal in welke applicatiefuncties deze entiteiten worden beheerd en gebruikt.

Houdt bij het selecteren van bijvoorbeeld standaardapplicaties rekening met het feit dat delen van de gegevens die gebruikt kunnen worden uit een andere bron kunnen komen. Bijvoorbeeld door het gebruik van een service interface. Veel applicaties zijn hierop niet ingericht waardoor replicatie noodzakelijk wordt.

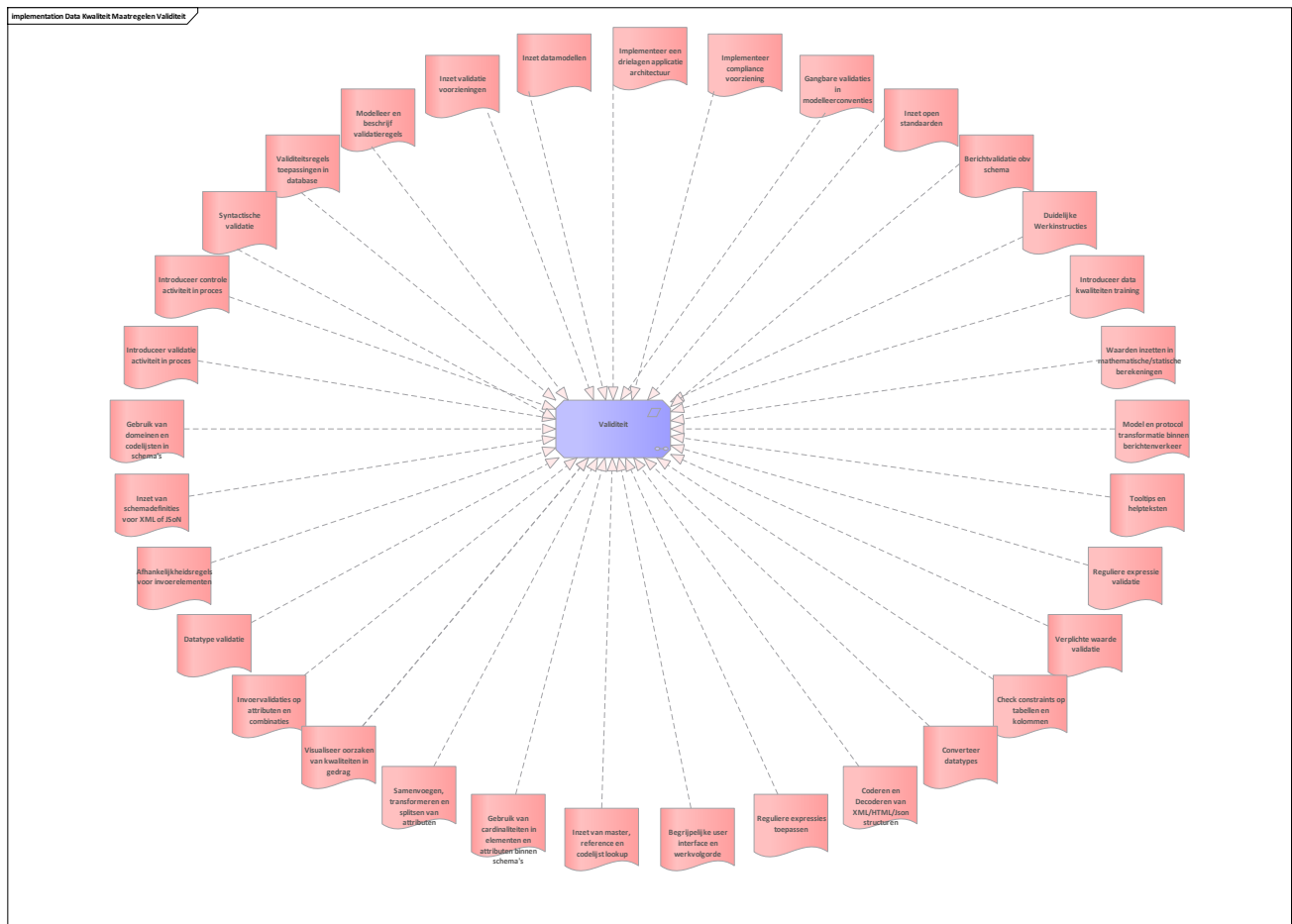
Uniekheid

Uniekheid van een data entiteit is gericht op het feit dat er geen andere entiteiten zijn met dezelfde gegevens. Ook bekend als duplicaten probleem.

Uniekheid bij replica's

Bij het gebruik van gegevens wordt er soms voor gekozen, bijvoorbeeld vanuit performance overwegingen om gegevens tijdelijk op te slaan in een geheugen of op een andere locatie. Dit brengt risico's met zich mee op het vlak van uniekheid op het moment dat de brongegevens wijzigen en dit niet wordt doorgevoerd binnen de replica's. Dit wordt bewaakt binnen governance.

Data Kwaliteit Maatregelen Validiteit



Overzicht van de data kwaliteiten en daarbij een aantal maatregelen die genomen kunnen worden om een Data kwaliteit met behulp van deze maatregelen (binnen een release) op een hoger niveau te krijgen.

Afhankelijkheidsregels voor invoerelementen

Op basis van invoervelden kunnen andere invoerelementen beïnvloed worden. Bijvoorbeeld invoerelementen zichtbaar/onzichtbaar of actief/inactief te maken. Bijvoorbeeld na kiezen van anders in een keuzelijst wordt een invoerveld met de titel anders actief gemaakt

Begrijpelijke user interface en werkvolgorde

Een begrijpelijke user interface voor medewerkers en gebruikers zorgt ervoor dat fouten en onduidelijkheden worden voorkomen. Daarnaast is een fijne user interface productiviteit verhogend

Berichtvalidatie obv schema

Maak gebruik van berichtvalidatie bij het gebruik van services. Denk hierbij aan de inzet van XSD validatie voor het syntactisch testen van berichten voordat de database opslag plaatsvindt. Laat berichtenverkeer waar mogelijk via de bovengenoemde business laag lopen.

Check constraints op tabellen en kolommen

Check constraints zijn extra functionaliteit in een database waarin je op basis van een statement extra condities kunt toevoegen aan een of meerdere kolommen. Dit extra controles toe te voegen aan deze kolommen

Coderen en Decoderen van XML/HTML/Json structuren

Transformeren van datastructuren in XML, HTML en Json naar interpreteerbare en valideerbare data tbv de kwaliteit

Converteer datatypes

Converteren van datatypes naar andere datatypes (tekst <-> Numeriek). Inclusief transformatie naar datatypes op andere platformen zoals database, XML, softwaretalen en localisatie.

Datatype validatie

Bij invoerelementen kunnen near realtime validaties worden toegevoegd. Denk bijvoorbeeld aan, getal en datum tijd validaties of email patroon controls bij invoer

Duidelijke Werkinstructies

Zorg voor duidelijke werkinstructies vanuit het perspectief van data kwaliteit. Veelal wordt in werkinstructies een introductie gedaan van workarounds die een negatief effect kunnen hebben op de kwaliteit van de data

Gangbare validaties in modelleerconventies

Stel generieke eisen aan gangbare validaties zoals datum en numerieke waarden, postcodes etc en beschrijf hoe en waar deze geïmplementeerd en getest moeten worden. Dit kan zowel op gegevensopslag als op gegevensintegratie geïmplementeerd worden.

Gebruik van cardinaliteiten in elementen en attributen binnen schema's

Gebruik van minimale en maximale cardinaliteiten voor zowel elementen als de attributen. Hiermee kan verplicht worden afgedwongen. Maar ook de mogelijkheid om meerdere elementen vast te leggen voor het realiseren van compleetheid

Gebruik van domeinen en codelijsten in schema's

Gebruik van extra structuren in schema's om extra beperkingen in attributen te implementeren op basis van domeinen (waardelijsten) of codelijsten

Implementeer compliance voorziening

Richt indien mogelijk een compliance voorziening in voor gangbare validaties etc.

Implementeer een drielagen applicatie architectuur

Maak gebruik van een drielagenarchitectuur binnen softwarecomponenten waarbij de (middelste) business laag zorg draagt voor validiteitschecks en draag er zorg voor dat alle database mutaties via deze gecentraliseerde business laag verwerkt worden.

Introduceer controle activiteit in proces

Wil je de kwaliteit van data waarborgen dan kun je zorgen dat er een kwaliteitsstap wordt ingebouwd in een werkproces. Hierbij zijn verschillende vormen mogelijk. Zoals dat de activiteit door de data producent zelf gedaan wordt of door een andere anders dan de stakeholder zelf. Bekendste voorbeeld is het twee paar ogen principe

Introduceer data kwaliteiten training

Bewustwording van kwaliteiten vergt training want het is gericht op bewustwording en gedragsverandering bij de verschillende stakeholders. Dit kan gedaan worden met klassikale, peer to peer en ook e-learnings getraind worden

Introduceer validatie activiteit in proces

Zorg dat data die geproduceerd of getransformeerd wordt, zeker bij handmatige verwerking van data in een validatie activiteit beoordeeld wordt. Dit biedt een punt in het proces waar issues gesignaleerd kunnen worden, maatregelen genomen kunnen worden of waar signalen zijn dat er in andere data management processen aandachtspunten zijn rond de data kwaliteit.

Invoervalidaties op attributen en combinaties

Invoercontroles voor combinaties van invoervelden, controle validaties en check controles

Inzet datamodellen

Beschrijving van data objecten en - attributen. Door een gedetailleerde en gestructureerde beschrijving te maken van data objecten die ingezet worden voor data integratie wordt het mogelijk om op basis van deze beschrijving validaties te ontwikkelen en deze te implementeren in bovengenoemde validatie componenten.

Inzet open standaarden

Keuze (open) standaarden, deze zijn veelal gebaseerd op syntactisch in detail uitgewerkte informatiemodellen en kunnen daardoor goed gevalideerd worden door bijvoorbeeld de bovengenoemde validatiecomponenten. Daarnaast bieden deze standaarden zeker bij organisatie overstijgende integratie om deze validatie op een centrale plaats uit te voeren. Bijvoorbeeld in een sectoraal knooppunt of een compliance voorziening.

Inzet validatie voorzieningen

Inzet validatiecomponenten. Voor validatie van bijvoorbeeld berichten zijn componenten en agents beschikbaar die berichten eenvoudig kunnen valideren op syntactisch niveau. Deze kunnen ingezet worden op verschillende plekken in integratieketens indien dit gewenst is. Houdt rekening met de effecten die dit kan hebben op met name de performance binnen een keten.

Inzet van master, reference en codelijst lookup

Verrijken van datasets op basis van kenmerken voor het toevoegen van beschrijvingen, lookups, hiërarchien van generiek data attributen

Inzet van schemadefinities voor XML of JSON

van schema's voor het valideren van data verzamelingen binnen een XML of JSON gebaseerd bericht of bestand. Hiermee wordt afgedwongen dat de data voldoet aan de regels in het validatieschema.

Model en protocol transformatie binnen berichtenverkeer

Transformeren van modellen en protocollen. Bijvoorbeeld van en naar een Canoniek Model transformeren en protocol transformeren XML naar JSON vice versa.

Modelleer en beschrijf validatieregels

Beschrijf de validatieregels van de attributen van data objecten en communiceer deze regels met ontwikkelpartijen. Regel daarnaast een voorziening in voor de ontsluiting van deze regels en draag zorg voor de toetsing.

Reguliere expressie validatie

Gebruik van reguliere expressie validaties in invoervelden om te voorkomen dat er invoerfouten worden gemaakt.

Reguliere expressies toepassen

Validaties in teksten op basis van reguliere expressies. Denk bijvoorbeeld aan het format van een postcode of een emailadres

Samenvoegen, transformeren en splitsen van attributen

Transformeren naar attributen in datasets door omzetten naar samenvoegen tot een geaggregeerd kenmerk. Daarnaast het opsplitsen van delen van een attribuut op basis van een bepaald kenmerk. Splitsen opv komma of spatie.

Syntactische validatie

Inzet syntactische validaties, met name bij berichtenverkeer op basis van XML kunnen berichten binnen de integratieketen op één of meerdere plaatsen gevalideerd worden. Deze validaties zorgen ervoor dat de berichtinhoud gecontroleerd wordt op correctheid op basis van definitiebestanden waarmee voorkomen wordt dat invalide gegevens opgeslagen worden of dat bij verder gebruik problemen in de verwerking van de gegevens ontstaan.

Tooltips en helpteksten

Tooltips en helpteksten geven gebruikers ondersteuning als invoervelden en formulieren onduidelijk zijn of geven een toelichting bij complexe invoercombinaties

Validiteit

Dit is de mate waarin een data entiteit bij opslag en uitwisseling voldoet aan het gewenste formaat. Denk hierbij bijvoorbeeld aan het domein maar ook het datatype van de attributen van een data entiteit. Binnen ketenuitwisseling is dit bijvoorbeeld van het grootste belang.

Validiteitsregels toepassingen in database

Maak gebruik van voorzieningen bij data opslag, bijvoorbeeld in relationele databases die de validiteit van gegevens bij de opslag controleren. Denk bijvoorbeeld aan check constraints, foreign key en domein constraints etc.

Verplichte waarde validatie

Invoervalidatie op verplichte invoer in het invulformulier in een vroeg stadium

Visualiseer oorzaken van kwaliteiten in gedrag

Maak inzichtelijk voor stakeholders die datakwaliteitsproblemen veroorzaken wat het effect is van handelen bij het werken met de data. Denk bijvoorbeeld aan het misbruiken van velden in schermen die daarmee de kwaliteit in de achterliggende databanken kunnen verlagen.

Waarden inzetten in wathematische/statische berekeningen

Inzetten van wathematische en statische berekeningen voor bepalen van de kwaliteit maar ook voor het aggregeren of het bepalen van afwijkende waarden in een dataset.