

Beschrijving datakwaliteiten en kwaliteit verhogende maatregelen

Beschrijving datakwaliteiten en kwaliteit verhogende maatregelen

Bert Dingemans

Beschrijving datakwaliteiten en kwaliteit verhogende maatregelen

Abstract

Datakwaliteiten beschrijven welke eisen aan gegevens gesteld kunnen worden en op welke wijze zij bijdragen aan een adequate bedrijfsarchitectuur. Na een beschrijving van de verschillende data kwaliteiten wordt ingegaan op maatregelen waarmee data kwaliteiten verhoogd worden. Op basis van deze maatregelen gerelateerd aan de data kwaliteiten kan per gegevensverzameling bepaald worden welke niveau van data kwaliteiten relevant is en welke maatregelen daartoe genomen dienen te worden.

Inhoudsopgave

| | |
|--------------------------------------|----|
| Inventarisatie datakwaliteiten | 2 |
| Abstract | 2 |
| Waarom datakwaliteiten?..... | 5 |
| Inleiding | 5 |
| Waarom dit document..... | 5 |
| Overzicht..... | 5 |
| Belang van kwaliteiten..... | 7 |
| Overzicht..... | 8 |
| Dataopslag..... | 10 |
| Precisie..... | 10 |
| Privacy | 11 |
| Referentiele integriteit | 12 |
| Uniekheid..... | 13 |
| Validiteit | 14 |
| Dataintegratie | 16 |
| Consistentie..... | 16 |
| Actualiteit..... | 17 |
| Precisie..... | 18 |
| Privacy | 19 |
| Redelijkheid | 19 |
| Referentiele integriteit | 20 |
| Tijdigheid..... | 21 |
| Validiteit | 22 |
| Datagebruik..... | 23 |
| Accuraatheid | 23 |
| Compleetheid | 24 |
| Consistentie..... | 24 |
| Actualiteit..... | 25 |
| Precisie..... | 25 |
| Privacy | 26 |
| Redelijkheid | 26 |

Beschrijving datakwaliteiten en kwaliteit verhogende maatregelen

| | |
|--------------------------------|----|
| Referentiele integriteit | 27 |
| Tijdigheid..... | 27 |
| Uniekheid..... | 28 |
| Referenties | 29 |

Waarom datakwaliteiten?

Inleiding

Bij de analyse van gegevens, registers en deels bij services en applicaties zijn datakwaliteiten een belangrijk aspect. Wil men bijvoorbeeld een gegevensbron in gaan zetten als authentiek register dan zal een analyse plaats moeten vinden van de datakwaliteiten aan zowel de aanbod- als vraagzijde. Verschillen tussen vraag- en aanbod zullen namelijk de acceptatie van de gegevens aan de vraagzijde bemoeilijken. Deze verschillen kunnen opgelost worden met behulp van technologische implementaties of door het inrichten van beheerprocessen.

Waarom dit document

Dit document biedt een overzicht van de verschillende datakwaliteiten in combinatie met een aantal maatregelen die genomen kunnen worden om de kwaliteit voor een bepaalde situatie op een voldoende hoog niveau te krijgen.

Het document kunnen architecten of ontwerpers gebruiken bij het opstellen van een architectuur of ontwerpdocument. Hierbij wordt in kaart gebracht aan welke kwaliteitseisen voldaan moet worden voor een specifieke situatie of gegevensverzameling. Vervolgens wordt bepaald welke maatregelen genomen kunnen worden om het gewenste kwaliteitsniveau te bereiken als dit nog niet het geval is.

De maatregelen gelden als randvoorwaarde voor het project dat de implementatie gaat uitvoeren. Daarnaast kunnen beheerprocessen noodzakelijk zijn voor het bereiken van voldoende datakwaliteit. Dit zal geborgd worden in de beheerorganisatie zodra de producten opgeleverd worden vanuit het bovengenoemde project.

Overzicht

Binnen het internationale samenwerkingsverband Data Management International (dama.org) zijn een aantal best practices ontwikkeld. Eén van deze best practices is onderstaande lijst van datakwaliteiten.

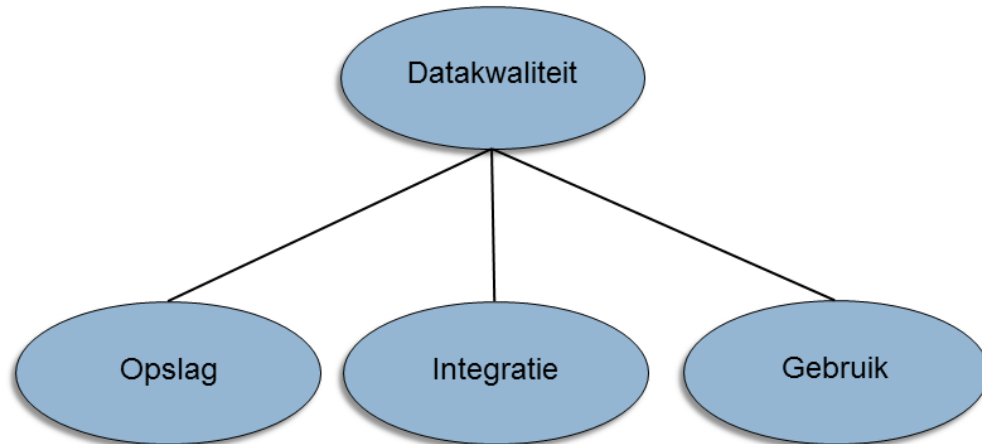
| Omschrijving | Toelichting |
|--------------|---|
| Accuraatheid | Accuraatheid heeft betrekking op de mate waarin een data entiteit de werkelijkheid weergeeft. Accuraatheid kan bepaald worden door een data entiteit te vergelijken met de entiteit in de werkelijkheid. Voorbeeld uit het verleden waarbij een baby een stemkaart ontving op basis van een lage accuraatheid binnen de GBA. |
| 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. Bijvoorbeeld een persoon kan enkel de eigenschap naam hebben, maar ook bestaan uit roepnaam, voornamen, tussenvoegsel, achternaam en meisjesnaam. In het laatste geval is er sprake van een hogere compleetheid |

Beschrijving datakwaliteiten en kwaliteit verhogende maatregelen

| | |
|--------------------------|--|
| 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. Een voorbeeld van een lage consistentie is als er verschillen zitten tussen datasets van dezelfde entiteiten afkomstig uit verschillende bronnen. Replicatie van gegevens is vaak een bron van lage consistentie. |
| Actualiteit | Mate waarin een data entiteit de actuele situatie van de werkelijkheid weergeeft. Veelzeggende voorbeelden zijn overleden mensen die een brief krijgen op basis van een verouderde dataset. Replicatie van gegevens is vaak een bron van lage actualiteit. |
| 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 nodig is. Ook domeinen op eigenschappen kunnen onvoldoende precisie hebben (denk aan een nederlandse postcode in een internationale data-opslag) |
| 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 vertrouwelijke gegevens. In het GBA kent men bijvoorbeeld meerdere niveaus van vertrouwelijkheid. Zo worden bevestigingen van ambtenaren wel gelogd en weergegeven aan de burger maar voor opsporingsambtenaren worden wel gelogd maar niet getoond. |
| Redelijkheid | Heeft vooral betrekking op verwachtingen binnen een bepaalde operationele context. Denk bijvoorbeeld aan het accepteren van een lagere performance bij piekbelasting of aan het langer moeten wachten op een resultset van gearchiveerde data-entiteiten |
| Referentiele integriteit | Dit is de situatie waarbij verwijzingen vanuit de ene data entiteit altijd correct verwijzen naar de gerelateerde data entiteiten. Voorbeelden zijn dubbele sleutels in een dataset waardoor de gekoppelde entiteiten niet kunnen bepalen wat de ouder entiteit is. Ook dangling references waarbij de ouder niet meer bestaat hoort hiertoe |
| Tijdigheid | Is een dataset tijdig beschikbaar binnen de gestelde verwachtingen. Het is het verschil tussen het moment van behoefte en beschikbaarheid. Bijvoorbeeld het opvragen van gegevens in het KCC stelt hoge eisen. Vijf minuten wachten op een response van de gegevensopslag is dan niet acceptabel. |
| Uniekheid | Uniekheid van een data entiteit is gericht op het feit dat er geen andere entiteiten zijn met dezelfde gegevens. Voorbeeld uit de praktijk was een tweeling met dezelfde voorletters, achternaam en geboortedatum. Het onderscheid was daardoor niet te maken door een te lage compleetheid. |
| 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. Men wil niet dat aan het einde van de keten blijkt dat de gegevens niet valide zijn. Dit |

| | |
|--|---|
| | dient in een vroeg stadium geconstateerd te worden. Iedereen kent de voorbeelden uit het verleden waarbij bij webapplicaties na verzenden een melding invalid data terugkwam zonder verdere toelichting |
|--|---|

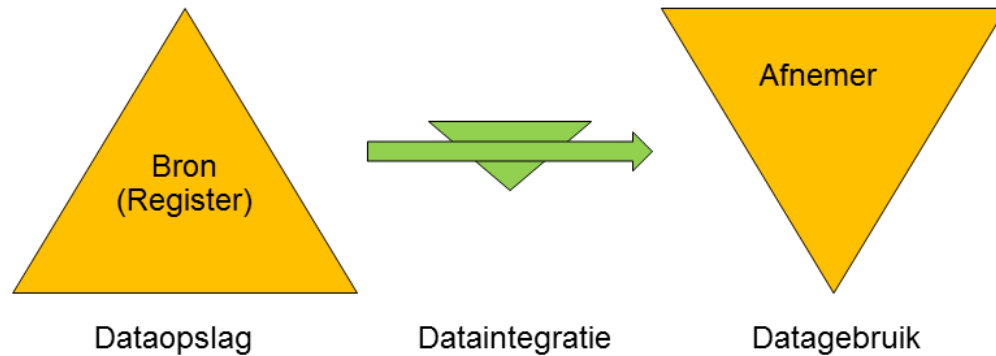
Belang van kwaliteiten



Data kan beschouwd worden vanuit een drietal perspectieven te weten:

- **Dataopslag**, opslag van de gegevens binnen de infrastructuur bijvoorbeeld in databases, filesystemen e.d.
- **Dataintegratie**, gegevens worden getransporteerd van bijvoorbeeld de opslag naar het gebruik of van gebruik naar opslag. Denk aan webservices, filesystemen, API objecten etc
- **Datagebruik**, gegevens worden gebruikt binnen een werkproces, applicatieve toepassing of in een component. Denk bijvoorbeeld aan weergave op een scherm, genereren van een rapport of signalering.

Data architectuur kernindeling



De datakwaliteiten zijn binnen deze indeling niet overal relevant, in onderstaande tabel wordt met een * de relevantie aangegeven.

Overzicht

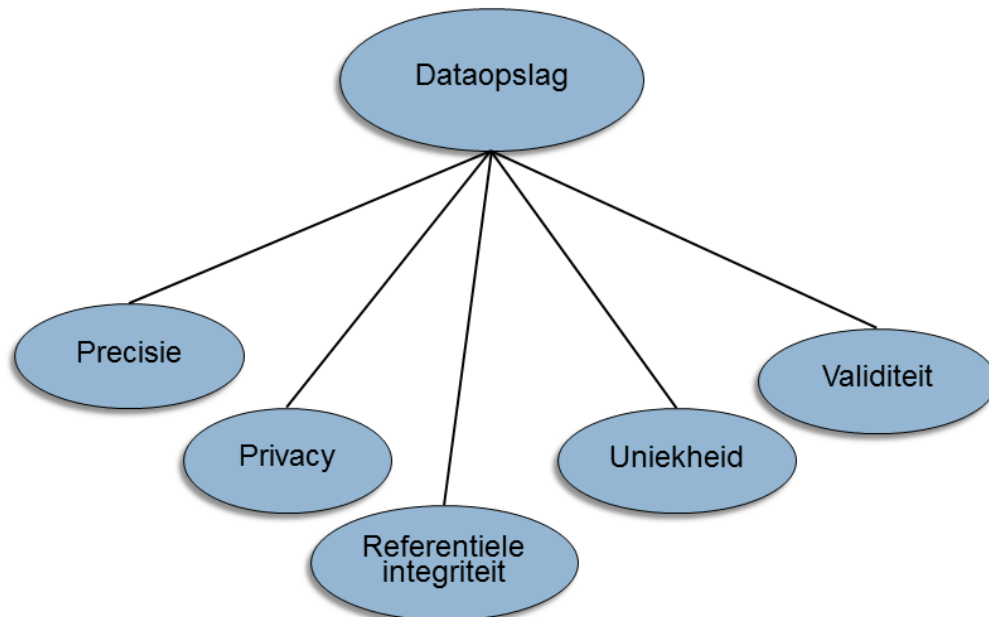
| Omschrijving | Toelichting | Dataopslag | Dataintegratie | Datagebruik |
|---------------------|---|------------|----------------|-------------|
| Accuraatheid | Accuraatheid heeft betrekking op de mate waarin een data entiteit de werkelijkheid weergeeft. Accuraatheid kan bepaald worden door een data entiteit te vergelijken met de entiteit in de werkelijkheid | | | ** |
| 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 | | ** | ** |

Beschrijving datakwaliteiten en kwaliteit verhogende maatregelen

| | | | | |
|---------------------------------|--|----|----|----|
| | bepaalde entiteit gelijk is aan een andere dataset. Met andere woorden is een data entiteit onafhankelijk van de bron altijd dezelfde | | | |
| Actualiteit | Mate waarin een data entiteit de actuele situatie van de werkelijkheid weergeeft | | * | ** |
| Precisie | Mate van detail waarin een data entiteit de werkelijkheid weergeeft. Dit heeft bijvoorbeeld betrekking op de precisie van getallen e.d. | ** | ** | * |
| Privacy | Voor sommige data entiteiten is toegangscontrole (autorisatie en authenticatie) of monitoring van gebruik nodig | ** | ** | * |
| Redelijkheid | Heeft vooral betrekking op consistentie verwachtingen binnen een bepaalde operationele context. Denk bijvoorbeeld aan het accepteren van een lagere performance bij piekbelasting | | ** | * |
| Referentiele integriteit | Dit is de situatie waarbij verwijzigingen 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 | ** | | ** |
| 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. | ** | ** | |

In de volgende hoofdstukken worden eventuele maatregelen voor het realiseren van deze kwaliteiten per data-aspect uitgewerkt.

Dataopslag



Precisie

Precisie bij data opslag heeft betrekking op het feit dat bij het opslaan van gegevens er informatie verloren gaat. Denk bijvoorbeeld aan precisie bij hele grote of hele kleine getallen.

Gegevensopslag kenmerkt zich door structurering en deze structurering kan ervoor zorgen dat de precisie negatief beïnvloed wordt. Enkele voorbeelden van precisie knelpunten bij data opslag:

- Precisie van getallen is onvoldoende bij de fysieke opslag. Denk bijvoorbeeld aan het datatype float dat een afronding van getallen in zich bergt
- Toepassen van compressietechnieken ter beperking van de behoefte aan opslagcapaciteit. Denk aan het comprimeren van fotomateriaal voorafgaand aan de opslag of de conversie naar een PDF formaat van documenten.
- Opslag van Geo data is gebaseerd op vectoren. Bij opslag kan gekozen worden voor een bepaalde mate van precisie. Bij de opslagkeuze dient de precisieafweging gemaakt te worden.

Met behulp van onderstaande maatregelen kunnen risico's rond de datakwaliteit precisie beperkt worden.

Technologie

- Onderzoek of de bestaande datatypes binnen de data opslag omgeving, zoals een database platform, voldoende ondersteuning bieden aan de gewenste 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
- Volstaan bestaande data typen en formaten binnen de data opslag niet dan kan het ontwikkelen van specifieke datatypes en –structuren een oplossing bieden. Houdt er rekening mee dat je hiermee een vorm van platform lock in introduceert

Beheerprocessen

- 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.
- Beschrijf bij een software selectie traject eventuele precisie aspecten op basis van requirements.
- 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.

Privacy

Onder privacy vallen verschillende informatiebeveiligingsaspecten van de gegevensopslag verstaan. Dat is dus niet alleen de bescherming van persoonsgegevens maar ook andere aspecten zoals beschikbaarheid en bronintegriteit. Enkele voorbeelden van informatiebeveiligingsaspecten zijn:

- Gegevensopslag is gesplitst in fysiek gescheiden implementaties, bijvoorbeeld database en bestandssysteem en de beveiligingsinrichting binnen beiden omgevingen is van een verschillend niveau.
- Bij de inrichting van de gegevensopslag worden de autorisaties niet of nauwelijks ingericht waardoor alle gebruikers van de gegevens maximale rechten hebben.

Technologie

- 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)

Beschrijving datakwaliteiten en kwaliteit verhogende maatregelen

- Infrastructuur kent een compartimentering waarbij enerzijds de compartimenten van elkaar gescheiden zijn en anderzijds deze compartimenten verschillende beveiligingsniveaus kennen.
- Dienen zaken voor traceerbaarheid van bewerkingen en bevragingen ingeregeld worden in de vorm van logging en database inrichting.

Beheerprocessen

- 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
- Bepaal in samenspraak met de eigenaar de beveiligingscategorieën en richt op basis daarvan de infrastructuur in en neem de bijbehorende maatregelen in de beheerorganisatie.
- Richt de beheerorganisatie in voor de verschillende IB aspecten bijvoorbeeld door het inrichten van de processen op basis van ISO 27001.

Referentiele integriteit

Referentiele integriteit van gegevens speelt bij opslag een belangrijke rol. Deze integriteit moet gehandhaafd blijven tijdens de opslag voor later gebruik. Daarnaast dienen bij wijzigingen in gegevensverzamelingen binnen de opslag deze integriteitseisen gehandhaafd te blijven. Met name bij opslag in databases zijn een aantal voorzieningen beschikbaar die het afdwingen van deze referentiele integriteit afdwingen. Voorbeelden van mogelijke knelpunten zijn:

- Gegevenssets worden opgeslagen in verschillende opslagsystemen en soms ook in een andere opslagplatform. Veelal wordt referentiele integriteit gebaseerd op basis van sleutels. Als deze sleutels binnen de verschillende platformen op een andere wijze worden opgeslagen is referentiele integriteit afdwingen een grote uitdaging.
- Sleutels inzetten voor verwijzingen (bijvoorbeeld bij inzet van registers) brengt met zich mee dat deze sleutel op dat moment register overstijgend wordt. Dat kan extra eisen stellen aan de opbouw van de sleutel.
- Voorbeelden van dangling references zijn legio. Bijvoorbeeld hyperlinks die niet meer werken in een samenwerkingsomgeving zodra de fysieke naam van een bestand of pagina gewijzigd wordt

Technologie

- 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.
- 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.

- 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

Beheerprocessen

- 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.
- 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.

Uniekheid

Ook uniekheid van gegevens bij opslag is een aspect dat aandacht verdient. Door de evolutionaire ontwikkeling van veel applicatielandschappen is de uniekheid veelal niet optimaal. Dit uit zich dan in verschillende scenario's, zoals:

- Binnen de opslag van entiteiten, bijvoorbeeld in een tabel in een database, wordt niet afgedwongen dat er geen dubbele gegevens worden opgeslagen. Gevolg is dat er meerdere rijen met dezelfde gegevens worden opgeslagen
- Entiteiten worden in verschillende registers opgeslagen, beheerd en ontsloten. Dubbele registratie wordt onderkend door de organisatie maar door de inrichting van de registers wordt dit geaccepteerd. Beheer en eigenaarschap is daardoor niet optimaal maar wordt geaccepteerd.
- Een ernstiger vorm van een beperkte uniekheid is dat entiteiten dezelfde informatie omvatten maar een andere structuur en naamgeving kennen en in één of meerdere registers zijn opgeslagen. Dit staat bekend onder het homo- en synoniemenprobleem.
- Laatste vorm van een lage uniekheid van gegevens is als normalisatie onvoldoende is doorgevoerd. Voor rapportage omgevingen kan hier in het kader van performance overwegingen voor gekozen worden.

Technologie

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

Beschrijving datakwaliteiten en kwaliteit verhogende maatregelen

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

Beheerprocessen

- Pas naamgevingsconventies toe voor entiteiten, attributen en relaties waardoor de kans op het ontstaan van dubbele opslag in registers verkleind wordt.
- 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.
- 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
- Richt processen in voor eigenaarschap en beheer zodat bij nieuwe ontwikkelingen besluitvorming en inrichting eenvoudig gefaciliteerd kunnen worden.

Validiteit

Validiteit van data zit met name op het vlak van opslag van gegevens en de interpretatie van deze gegevens op een later moment. Het is dan ook een knelpunt dat enerzijds bij datagebruik optreedt maar door data opslag veroorzaakt, daarom is het in dit hoofdstuk uitgewerkt

- Denk bijvoorbeeld aan de opslag van een datum in relationele database waarbij geen rekening gehouden wordt met de Amerikaanse of Europese datum notatie. Dat kan bij later gebruik problemen opleveren als de gegevens op een andere wijze geïnterpreteerd worden.
- Domein van een gegevensattribuut wordt niet gevalideerd voor opslag waardoor de gegevens niet overeenkomen met het formaat binnen het datagebruik. Denk bijvoorbeeld aan gegevens die via een data conversie aan een gegevensbestand worden toegevoegd.

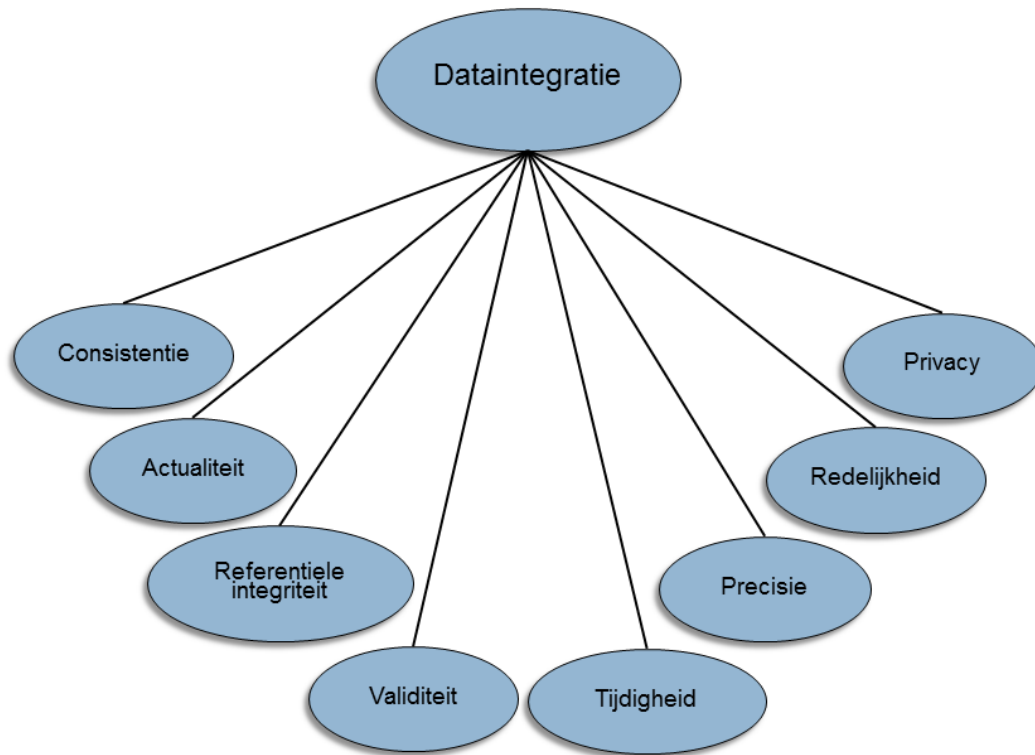
Technologie

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

Beheerprocessen

- 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.
- 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 gegevensintegratieval geïmplementeerd worden

Data-integratie



Consistentie

Bij data integratie wordt de consistentie van gegevenssets voornamelijk bepaald door de bron waar dit vandaan komt. In een aantal gevallen is de bron door een inkapseling in ketens niet altijd bekend, zeker als de bron buiten de eigen organisatiegrenzen ligt.

Daarnaast kunnen er problemen optreden op het moment dat gegevenssets één of meerdere keren gerepliceerd worden. Gevolg is dat bij het ontsluiten van deze verschillende (gerepliceerde) bronnen er verschillen kunnen ontstaan.

Technologie

- 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.
- 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.

- 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.

Beheerprocessen

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

Actualiteit

Bij data integratie speelt actualiteit met name een rol bij snel verouderende gegevens. Wordt er bijvoorbeeld gekozen voor een integratievorm met een lage mutatiesnelheid zoals bijvoorbeeld replicatie, bestandsuitwisseling of FTP dan kan de actualiteit een probleem vormen. Een voorbeeld uit de provinciale praktijk is de replicatie van de relatiegegevens vanuit het relatiebeheersysteem naar het document management systeem. Deze kopieerslag vindt tweemaal per jaar plaats. Gevolg is natuurlijk dat de actualiteit door deze lage mutatiesnelheid negatief beïnvloed wordt.

Technologie

- Kies integratievormen die een maximale mutatiesnelheid ondersteunen. De ene integratievorm kent een snellere doorlooptijd dan een andere. Zo zal berichtuitwisseling veelal een goede 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 een servicebus- of orkestratieplatformen
- 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 laatste bijgewerkte replica de beperkende factor. Zeker in combinatie met vertragende interactievormen zoals batches e.d.

Beheerprocessen

- 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
Richt eigenaarschap en beheerorganisatie in van bronsystemen en koppelvlakken, door voorbronsystemen 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).
- Bepaald actualiteitsbehoefte en –aanbod en registreer dit

Precisie

Precisie en data integratie zijn sterk van elkaar afhankelijk, met name vanwege het transformeren van een formaat voor opslag naar een formaat voor transport kan de precisie onvoldoende gehandhaafd blijven.

Veel voorkomende voorbeelden uit de praktijk zijn de uitwisseling van bijvoorbeeld datum- en tijd combinaties waarbij bij de gegevensuitwisseling alleen het datum formaat uitgewisseld kan worden. Gevolg is vanzelfsprekend dat er informatie verloren gaat indien ook de tijddimensie bij het gebruik van de data relevant is.

Technologie

- Keuze 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
- 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.

Beheerprocessen

- 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.
- 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.

Privacy

Bij data integratie spelen beveiligingsaspecten een grote rol. In het ene geval wil men voorkomen dat de gegevens tijdens het transport onderschept worden en gelezen, in het andere geval is het ongewenst dat de gegevens tijdens het transport worden gewijzigd.

Voorbeelden uit de praktijk zijn legio, bijvoorbeeld de beveiliging van berichtenverkeer tussen de provincie en de bank rond betalingsverkeer of het inzetten van de digikoppeling standaard voor de uitwisseling van gegevens met andere overheidspartijen.

Technologie

- 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 veelal een negatief effect hebben op de andere kwaliteiten als actualiteit en tijdigheid. Door het uitwerken van de verschillende beveiligingsvormen kan een adequate keuze van maatregelen gemaakt worden.
- Inzet van overheidscomponenten en standaarden. Overheidscomponenten en –standaarden als 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.

Beheerprocessen

- 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.
- 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.

Redelijkheid

Bij data integratie is redelijkheid een aspect dat met name gericht is op eisen die vanuit de vragende partijen bij het datagebruik gesteld worden. Een voorbeeld uit de praktijk is het inzetten van data integratie vormen via een gebruikers (web) interface. Hierbij is redelijkheid in het geding op het moment dat een synchrone vorm van gegevensuitwisseling noodzakelijk is. Bijvoorbeeld in de situaties waarbij er een bevestigingscode of –nummer geretourneerd moet worden. Dit kan eisen stellen aan de inrichting van de data integratie.

Technologie

- Infrastructurele inrichting, data integratie stelt bij een aantal integratievormen hoge eisen aan de infrastructuur. Bijvoorbeeld op het vlak van performance, beschikbaarheid en capaciteit. Bij de infrastructuur 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.
- Inzet van overheidscomponenten. Hierbij dient ter dege 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.

Beheerprocessen

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

Referentiele integriteit

Bij het uitwisselen van gegevens bij integratie kan de referentiele integriteit van de gegevensset verminderen. Bijvoorbeeld doordat bij transformatie gegevens van sleutels verloren gaan of veranderen. Ander voorbeeld zijn de integratievormen waarbij de gegevensset in stukken wordt verdeeld waarbij in een later stadium de integriteit verminderd als de samenstelling uit de delen niet dezelfde indeling heeft als de initiële set.

Technologie

- 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.
- 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

Beheerprocessen

- 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
- 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.

Tijdigheid

Tijdigheid heeft alles te maken met de performance van de ICT inrichting bij integratie en is daarmee een verbijzondering van de kwaliteit redelijkheid. Tijdigheid heeft te maken met met name de inrichting van de aanbodzijde van de data integratie

Technologie

- 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
- 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

Beheerprocessen

- 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 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.

Validiteit

Bij data integratie vindt transformatie plaats van het formaat bij data opslag naar het formaat voor transport en in een aantal gevallen ook weer terug van het formaat tijdens transport naar het formaat bij datagebruik (of –opslag).

Bij deze transformaties kan een vermindering van de validiteit optreden waardoor tijdens het data gebruik ongewenste effecten kunnen optreden. Een voorbeeld waarbij het validiteitsaspect een grote rol speelt is de inzet van bijvoorbeeld een compliance voorziening (bij digikoppeling). Dat is een voorziening waarop gegevensset op validiteit gecontroleerd worden voordat de data integratie daadwerkelijk ingezet wordt.

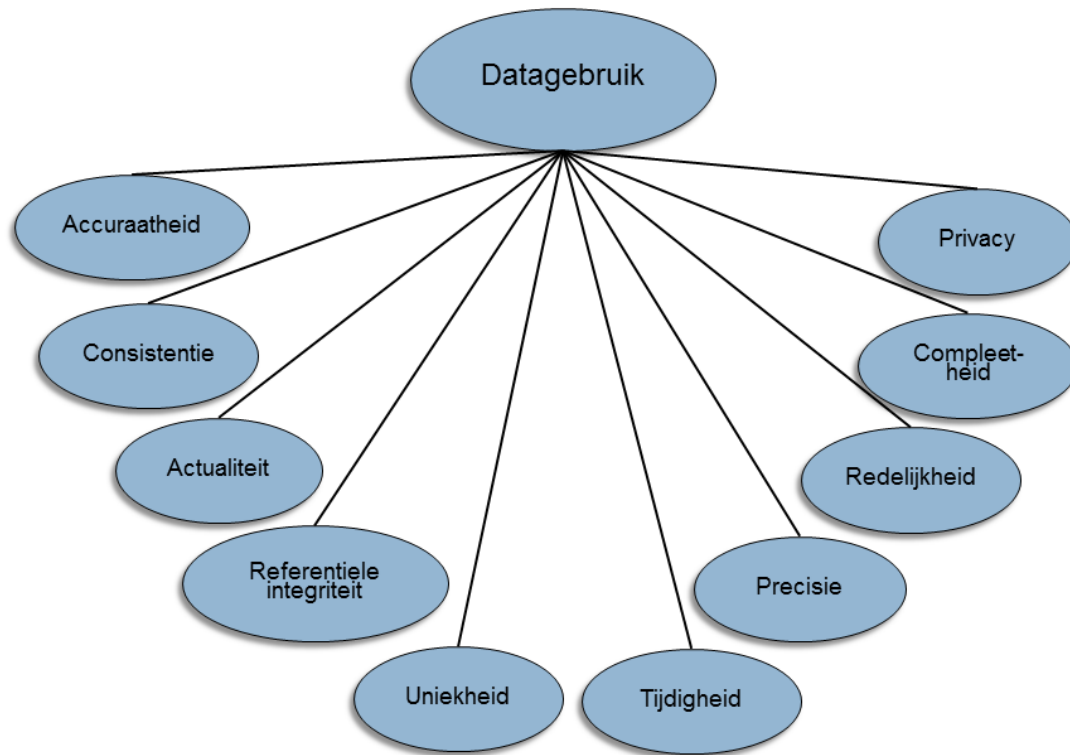
Technologie

- 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.
- 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.

Beheerprocessen

- 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
- 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.

Datagebruik



Datagebruik gaat over het verwerken van gegevens binnen werkprocessen en applicaties. Het is de plaats waar de feitelijke bewerkingen op de gegevens plaats vinden. Niet verwonderlijk dat voor bijna alle datakwaliteiten er een requirements zijn binnen het datagebruik.

Accuraatheid

Datagebruik en accuraatheid binnen een versnipperd applicatielandschap is een grote uitdaging. Zo is het lastig om de verschillende weergaven van de werkelijkheid binnen de data objecten in kaart te brengen en te bepalen in welke mate zij een accurate weergave van de werkelijkheid zijn.

Technologie

- Inzet van voorzieningen voor data virtualisatie, authentieke en kernregisters en ontsluiten van de basisregistraties.
- Inzetten van modelleromgevingen voor het opstellen van modellen van de gegevensbehoefte en het – aanbod

Beschrijving datakwaliteiten en kwaliteit verhogende maatregelen

- Inzetten van e-formulieren 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.

Beheerprocessen

- Bepalen van de gegevensbehoefte binnen de verschillende activiteiten van werkprocessen en het vergelijken van deze behoefte met de reeds aanwezige gegevensverzamelingen.
- Uitwerken en beschrijven van de verschillende architectuur concepten als processen, applicaties, (data) objecten en het in kaart brengen van de onderling relaties.
- Inrichten eigenaarschap en beheer van gegevensentiteiten en gerelateerde concepten

Compleetheid

Compleetheid binnen data gebruik is met name gericht op de compleetheid van deze benodigde gegevens in een activiteit binnen een (werk)proces. Met andere woorden in hoeverre de informatievoorziening de activiteit ondersteunt met complete gegevens. Hierbij dient opgemerkt te worden dat een overvloed aan informatie ook ongewenst is, het gaat om een voldoende model ter ondersteuning van de activiteiten.

Technologie en beheerprocessen zijn gelijk aan de uitwerking binnen accuraatheid.

Consistentie

Problemen met consistentie treden vooral op in situaties waar meerdere bronnen beschikbaar zijn die dezelfde gegevenssets bevatten. Denk hierbij aan meerdere registers waarin dezelfde entiteiten beheerd worden, maar ook aan situaties waar gegevens gerepliceerd (en getransformeerd) worden.

Technologie

- 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

Beheerprocessen

- Bepalen van generieke data entiteiten en bepalen van de mate van consistentie ten opzichte van de aanwezige applicaties
- Uitwerken van architectuurmodellen van de bovengenoemde generieke entiteiten in relatie tot de andere architectuur concepten als functies, processen, actoren en componenten.
- Inregelen van eigenaarschap en beheerprocessen

Actualiteit

Actualiteit van gegevens is wederom relevant bij de ondersteuning van activiteiten binnen werkprocessen. Zijn binnen die activiteiten de gegevens niet actueel dan is de kans dat de activiteit correct wordt afgehandeld klein. Actualiteit van gegevens speelt een belangrijk rol in processen als vergunningen, handhaving maar ook op het vlak van subsidieverstrekking

Technologie

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

Beheerprocessen

- Introduceren van beslisbomen 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

Precisie

Precisie binnen het gebruik van data heeft met name betrekking op hoe applicaties omgaan met de gegevens vanuit de verschillende (bron)registers. Bij het gebruik, bijvoorbeeld in de presentatielaag van applicaties, van data zijn met name de transformaties die nodig zijn beruchte bronnen van precisieproblemen. Denk hierbij bijvoorbeeld het verkeerd representeren van getallen en datums door transformatie bij de weergave (het Amerikaans datumformaat in een Nederlandse context).

Technologie

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

Beheerprocessen

- Richtlijnen opstellen omtrent inrichting van gegevensbestanden, applicatie componenten en elementen binnen de presentatielaag op het vlak van precisie
- Precisie aspecten meenemen in het ontwerpproces van nieuwe applicaties e.d.

Privacy

Privacy moet bij het datagebruik in de breedste zin van het woord beschouwd worden, wat inhoudt dat het op een aantal beveiligingsaspecten betrekking heeft. Denk hierbij aan concepten als authenticatie en autorisatie, logging en monitoring etc.

Technologie

- Inzet van generieke beveiligingscomponenten. Voor authenticatie en autorisatie zoals active directory maar ook Digid en dergelijke. Bij monitoring en logging inzetten van generieke voorzieningen zoals event- en error logs en dergelijke.
- 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.

Beheerprocessen

- 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
- 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

Redelijkheid

Is een belangrijk aspect binnen het datagebruik omdat het een kwaliteit is die gericht is op de vraagzijde en niet op de aanbodzijde. Hierbij geldt dat redelijkheid betrekking heeft op de andere kwaliteiten waarbij de kernkwaliteiten zoals tijdigheid, actualiteit en accuraatheid.

Technologie

- 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.

Beheerprocessen

- Om redelijkheid te ondersteunen is het van belang dat voor de belangrijkste gebruiksgevallen 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.

Referentiele integriteit

Referentiele integriteit speelt zich niet alleen bij opslag en integratie af maar ook bij gebruik.

Technologie

- 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.
- 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.

Tijdigheid

Tijdigheid is met name een aspect van data gebruik, In combinatie met de integratie wordt bepaald of een gegevensset tijdig beschikbaar is binnen de gebruikssituatie. Naast de aspecten die reeds genoemd zijn bij de tijdigheid zijn de volgende aspecten specifiek voor data gebruik.

Technologie

- 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.
- 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.
- 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
- 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.

Beheerprocessen

Beschrijving datakwaliteiten en kwaliteit verhogende maatregelen

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

Uniekheid

Technologie

- 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
- Introduceer componenten en voorzieningen die zorgdragen voor de uniekheid van gegevens. Denk hierbij bijvoorbeeld aan de centrale gegevensvoorziening maar ook aan zaken als een servicebus e.d.

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.

Beheerprocessen

- Inzet van replica's van gegevens is niet altijd te voorkomen echter draag daarbij zorg dat in kaart gebracht is welke replica's er zijn en beschrijf daarbij de werkwijze waarop het actualiseren van de gegevens plaatsvindt
- 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.

Referenties

Mosley, Mark et al (2010). *DAMA guide to the data management body of knowledge*. Bradley beach: Technics Publications LLC.