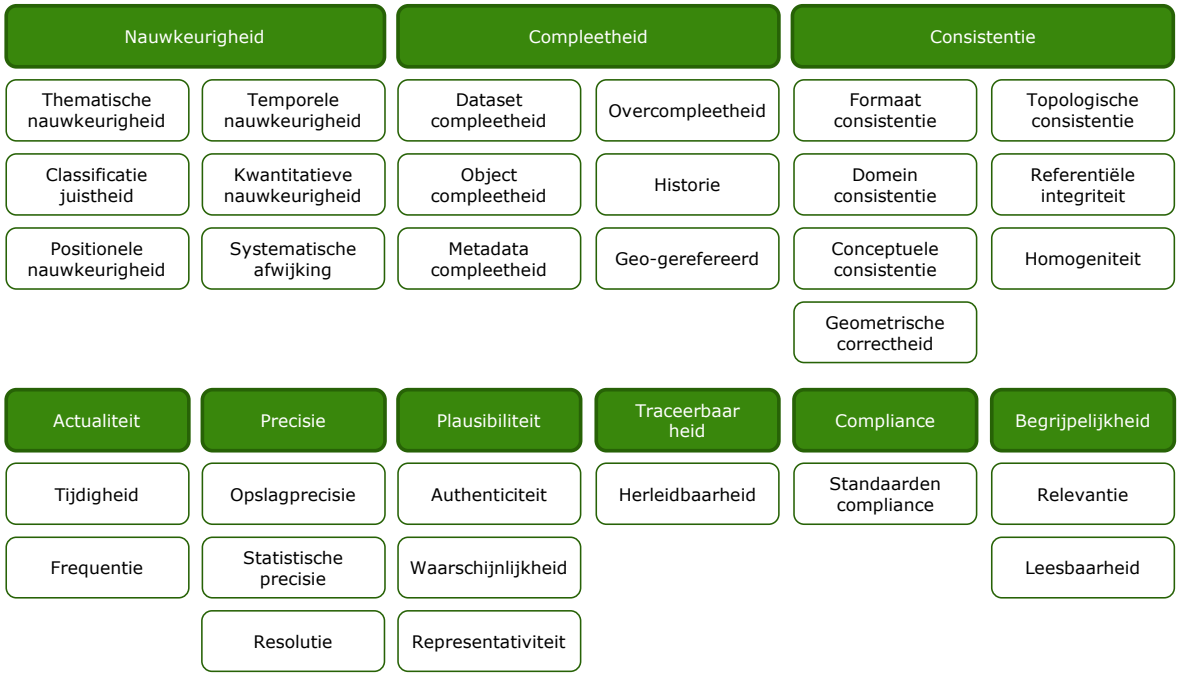
Programma
Aan de slag met de
Omgevingswet

Gegevenskwaliteit in de Omgevingswet

Een raamwerk voor gegevens en informatieproducten

Versie 1.0 Definitief 23-02-2018



Colofon

Titel	:	Gegevenskwaliteit in de Omgevingswet – Een raamwerk voor gegevens en informatieproducten
Versie	:	1.0 Definitief
Datum	:	23-02-2018
Opdrachtgever	:	Project Informatiehuizen
Opdrachtnemer	:	RIVM

Auteurs	:	Danny Greefhorst *+31 6 4523 9926* danny.greefhorst@rivm.nl

Versiehistorie

Versie	Status	Datum	Auteur(s)	Toelichting
0.1	Concept	11-10-2016	D. Greefhorst	Initiële opzet
0.2	Concept	29-11-2016	D. Greefhorst	Tekstuele aanpassingen
0.3	Concept	16-03-2017	D. Greefhorst	Tekstuele aanpassingen
0.41	Concept	19-06-2017	D. Greefhorst	Resultaat van 1e periode
0.5	Concept	01-02-2018	D. Greefhorst	Extern reviewcommentaar verwerkt
1.0	Definitief	23-02-2018	D. Greefhorst	Kleine tekstuele verbeteringen

Goedkeuring

Functie	Naam	Versie	Datum
Regiegroep informatieproducten			

Review

Naam
Kasper Kisjes (Rijkswaterstaat)
Peter Lentjes (Kadaster)
Piet van der Krieke (Kadaster)
Annelies Oversteeg (ILT)
Boris Everwijn (Informatiehuis Water)
John Maaskant (Informatiehuis Water)
Eva Schoenmaker (Informatiehuis Afval)
Remco Schonenberg (Informatiehuis Afval)
Jan Klein Kranenburg (Informatiehuis Bodem)
Andre Batenburg (Business Liason Architect Provincies)
Jan van Langeveld (Business Liason Architect Gemeenten)
Jan-Willem Genuit (Ministerie van IenM)
Martijn Odijk (Ministerie van IenM)
Marjolein Debets (gemeente Amsterdam)
Wim Stolk (voorzitter expertgroep Gegevensmanagement Gemeenten)

Inhoudsopgave

1 INLEIDING 6
1.1 Doel en doelgroep 6
1.2 Documentstructuur 7
2 WAT IS GEGEVENSKWALITEIT? 8
2.1 ISO/IEC 25012:2008 8
2.2 NEN-ISO 19157:2013 9
2.3 ISO 5725-1:1994(E) 10
2.4 DAMA Data management body of knowledge 11
2.5 Code voor informatiekwaliteit 13
2.6 Referentiemodel Datakwaliteit IPO 13
2.7 Kwaliteitsraamwerk Rijkswaterstaat 14
2.8 Relatie tussen raamwerken 14
3 ERVARINGEN BIJ DE OVERHEID 15
3.1 BAG en BGT 15
3.2 AERIUS 16
3.3 Inspectieview Milieu 18
3.4 INSPIRE 19
3.5 Evaluatie van cases 20
4 ONZE VISIE 22
4.1 Vertrouwen als basis voor kwaliteit in de keten 22
4.2 Denken vanuit gebruik 23
4.3 Transparantie naar alle doelgroepen 23
4.4 Kwaliteit organisatorisch en procesmatig borgen 25
4.5 Informatiemodel definieert essentie van gegevenskwaliteit 26
4.6 Gegevenskwaliteit als onderdeel van metadata 26
4.7 Integraal beeld van gegevenskwaliteit 28
5 VAN VISIE NAAR EEN KWALITEITSRAAMWERK 29
5.1 Concepten 29
5.2 Kwaliteitsraamwerk 31
5.3 Indicatoren 32
5.4 Meetfuncties 35
6 VAN VISIE NAAR PROCES 39
6.1 Overzicht 39

6.2 Maken afspraken in de keten 40
6.3 Inrichten organisatie en proces 41
6.4 Definiëren kwaliteitseisen 42
6.5 Beschrijven gegevens 42
6.6 Borgen kwaliteitseisen 42
6.7 Meten en verbeteren 43
7 VAN VISIE NAAR INFORMATIEVOORZIENING 45
7.1 Functionaliteiten 45
7.2 Functionaliteiten gepositioneerd in de Omgevingswet 47
BIJLAGE A: BRONNEN 48
BIJLAGE B: RELATIE TUSSEN ISO 25012 EN DE VOORGESTELDE INDICATOREN EN DE 3 B'S 50
BIJLAGE C: BEGRIPPENLIJST 52
BIJLAGE D: AFKORTINGEN 54
BIJLAGE E: RELATIE TUSSEN INDICATOREN EN STANDAARDEN 55
BIJLAGE F: BETROKKENEN 57
BIJLAGE G: RELATIE TUSSEN KWALITEITSRAAMWERKEN 58
BIJLAGE H: RELATIE VAN CASES MET ISO 25012 60

1 Inleiding

De Omgevingswet is een belangrijke vernieuwing voor zowel burgers als voor de overheid. Het zorgt ervoor dat iedereen meer inzicht krijgt in de leefomgeving en de regels die daarin gelden. Dit stelt hoge eisen aan de kwaliteit van de gegevens over de leefomgeving en de informatieproducten die op basis daarvan gemaakt worden. Gegevenskwaliteit is dan ook vooral gedreven vanuit het perspectief van de afnemer van gegevens. Afnemers moeten n staat zijn om gegevens en informatie te gebruiken in hun eigen context. Dat kan alleen als de gegevens van voldoende kwaliteit zijn; als ze overeenkomen met wensen en verwachtingen.

Dit rapport is het resultaat van een werkgroep met vertegenwoordigers van beoogde informatiehuizen in de context van de Omgevingswet. Het is gestart als een working paper en doorontwikkeld tot een volwaardig kwaliteitsraamwerk. Het geeft een handreiking voor het inzichtelijk maken van gegevenskwaliteit in de context van de Omgevingswet. Het beschrijft een visie en een aanpak voor gegevenskwaliteit op basis van algemene theorie en ervaringen bij de overheid. Het geeft een invulling aan de 3 B's van de Omgevingswet (beschikbaarheid, bruikbaarheid en bestendigheid) op het niveau van gegevenskwaliteit. Het biedt concrete indicatoren op basis waarvan uniform kan worden gerapporteerd over de kwaliteit van brongegevens en daarvan afgeleide informatieproducten.

1.1 Doel en doelgroep

De doelstelling van dit rapport is om inzicht te geven in wat gegevenskwaliteit is en wat dat betekent in de context van de Omgevingswet. Het rapport kan breed worden ingezet in de gehele keten van partijen die betrokken zijn bij de Omgevingswet; van bronhouder tot leveranciers van informatieproducten tot de landelijke voorziening DSO en gebruikers daarvan. Partijen bepalen zelf hoe zij hiermee aan de slag gaan; er worden geen normen gegeven alleen handreikingen. Het onderwerp gegevenskwaliteit wordt vanuit verschillende perspectieven belicht, waardoor het een goede leidraad kan zijn om ermee aan de slag te gaan.

De doelgroep van dit rapport zijn mensen met affiniteit met gegevens- en informatiegerelateerde vraagstukken die betrokken zijn bij het vormgeven van de Omgevingswet en de daarvoor noodzakelijke processen en systemen. Het is specifiek gericht op mensen die beter willen begrijpen wat de kwaliteit van informatieproducten en gegevens betekent. Het is gericht op een brede doelgroep, inclusief gebruikers, projectleiders, informatieprofessionals en data professionals.

Naast dit rapport is er ook een document met een verdere handreiking bij het kwaliteitsraamwerk. Het is gericht op mensen die er concreet mee aan de slag willen. Het biedt meer toelichting bij de voorgestelde indicatoren, maar ook richtlijnen voor het verbeteren van de kwaliteit van informatiemodellen en een generiek stappenplan.

1.2 Documentstructuur

De structuur van dit document is als volgt:

- Hoofdstuk 2 beschrijft algemene theorie op het gebied van gegevenskwaliteit
- Hoofdstuk 3 geeft voorbeelden van hoe met datakwaliteit wordt omgegaan bij de overheid
- Hoofdstuk 4 beschrijft een specifieke visie op gegevenskwaliteit, afgestemd op de context van de Omgevingswet
- Hoofdstuk 5 vertaalt de visie naar kwaliteitsraamwerk, dat gebruikt kan worden om kwaliteitseisen mee te formuleren en uniform over gegevenskwaliteit te rapporteren
- Hoofdstuk 6 vertaalt de visie naar een procesmodel en bijbehorende activiteiten
- Hoofdstuk 7 vertaalt de visie naar een schets van de gewenste informatievoorziening

Lezers die al op de hoogte zijn van standaarden op het gebied van gegevenskwaliteit kunnen hoofdstuk 2 overslaan. Lezers die niet bij de uitvoering van gegevenskwaliteit betrokken zijn kunnen hoofdstuk 5 overslaan. De gebruikte bronnen zijn opgenomen in bijlage A. Bijlage B beschrijft de relatie tussen de ISO/IEC 25012 standaard, de voorgestelde indicatoren en de 3 B's. Bijlage C geeft een beschrijving van de belangrijkste begrippen in dit document. Bijlage D geeft een beschrijving van de gebruikte afkortingen. Bijlage E beschrijft de relatie van de voorgestelde indicatoren met de in dit document gebruikte ISO standaarden. In bijlage F is een lijst van betrokkenen bij dit document opgenomen. Bijlage G beschrijft de relatie tussen een aantal kwaliteitsraamwerken. Bijlage H relateert de ISO/IEC 25012 standaard aan een aantal cases.

2 Wat is gegevenskwaliteit?

Dit hoofdstuk beschrijft een aantal standaarden en raamwerken op het gebied van gegevenskwaliteit. Aan de hand van deze theoretische kaders wordt inzichtelijk gemaakt wat gegevenskwaliteit is en wat relevante perspectieven zijn voor het omgaan met gegevenskwaliteit.

2.1 ISO/IEC 25012:2008

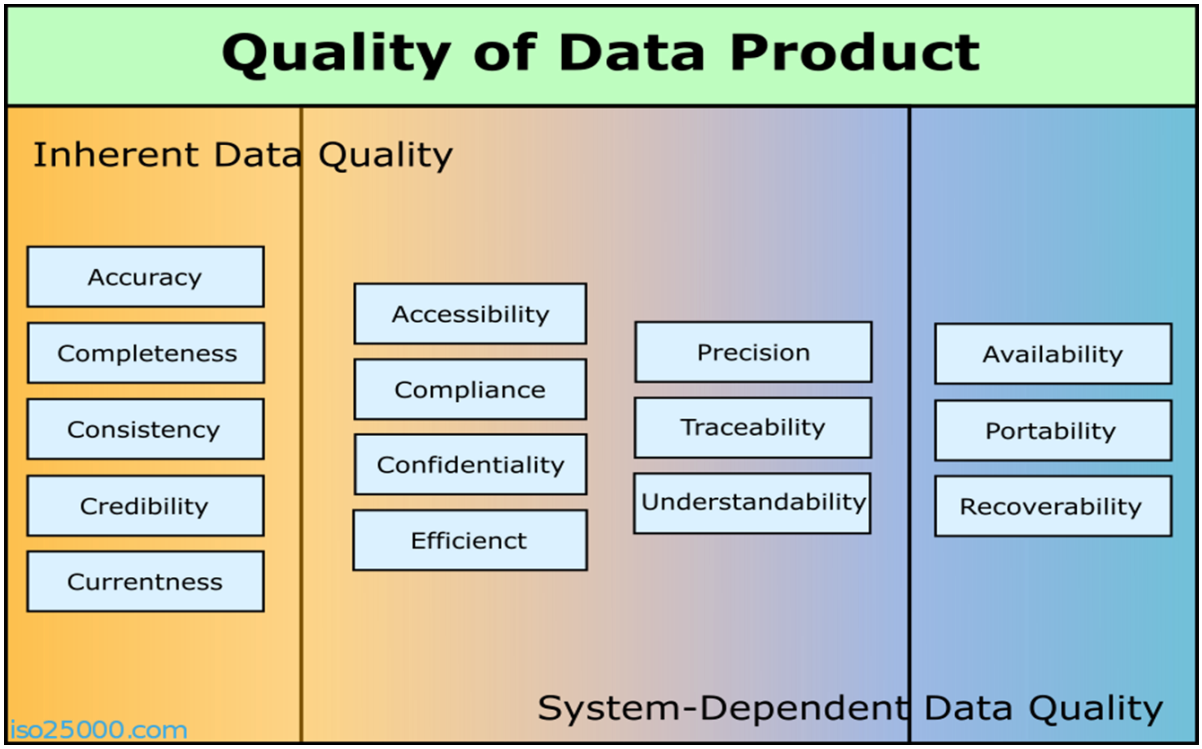
ISO/IEC 25012:2008 [1] is een belangrijke standaard op het gebied van gegevenskwaliteit. De standaard biedt inzicht in een aantal kernbegrippen rondom gegevenskwaliteit. Het begrip "gegevenskwaliteit" wordt gedefinieerd als "de mate waarin de karakteristieken van de gegevens voldoen aan expliciete en impliciete behoeften onder gespecificeerde condities".

Het beheren en verbeteren van de kwaliteit van gegevens is volgens de standaard onder meer belangrijk omdat:

- het veelal onduidelijk is wat de kwaliteit van gegevens is, met name doordat de kwaliteit van het inwinningsproces onduidelijk is;
- foutieve gegevens leiden tot onbevredigende informatie, onbruikbare resultaten en ontevreden klanten;
- gegevens toenemend worden verspreid over verschillende organisaties, terwijl ze veelal niet met dat oogmerk zijn ingewonnen;
- het verwerken van gegevens veelal lastig is vanwege onduidelijke betekenis of inconsistenties;
- er meer behoefte ontstaat aan het frequent uitwisselen van gegevens, maar bestaande systemen veelal onvoldoende in staat zijn gegevens uit te wisselen.

Naast een aantal definities biedt de standaard vooral een raamwerk van kwaliteitsdimensies. Figuur 1 geeft een overzicht van de kwaliteitsdimensies. De standaard maakt onderscheid tussen "inherente gegevenskwaliteit" en "systeemafhankelijke gegevenskwaliteit". De eerste gaat over eigenschappen van de gegevens zelf. De tweede gaat over eigenschappen die niet inherent zijn aan de gegevens zelf, maar afhankelijk zijn van het systeem waarmee ze worden ontsloten.

De ISO/IEC 25012 standaard wordt concreter gemaakt in de ISO/IEC 25024 standaard [2] waarin de bij de kwaliteitsdimensies behorende indicatoren zijn uitgewerkt. Deze indicatoren belichten specifieke deelaspecten van de kwaliteitsdimensies en zorgen ervoor dat de noodzakelijke metingen op een eenduidige wijze kunnen plaats vinden.



Figuur 1 ISO/IEC 25012 kwaliteitsdimensies

2.2 NEN-ISO 19157:2013

NEN-ISO 19157 [3] is ook een standaard op het gebied van gegevenskwaliteit, maar dan specifiek gericht op geo-informatie. Het gaat specifieker in op kwaliteitsdimensies zoals positionele nauwkeurigheid. Het stelt de volgende indeling in kwaliteitsdimensies voor:

- **Volledigheid:** de aanwezigheid of afwezigheid van objecten, hun attributen en relaties. Dit gaat zowel over afwezige gegevens als over teveel gegevens.
- **Logische consistentie:** de mate van overeenstemming met regels voor de gegevensstructuur, voor de toekenning van attributen en voor relaties. Dit gaat over het voldoen aan het conceptuele schema, domeinen, formaten en topologie.
- **Positionele nauwkeurigheid:** de mate waarin features goed zijn gepositioneerd in het ruimtelijk referentiesysteem. Dit gaat over de absolute en relatieve positionering in de ruimtelijke werkelijkheid, en vanuit een grid perspectief.
- **Thematische nauwkeurigheid:** de mate waarin kwantitatieve en niet-kwantitatieve gegevens kloppen, en passen bij de classificatie van objecten en relaties.
- **Temporele nauwkeurigheid:** de mate waarin de tijdsgebaseerde attributen en relaties kloppen. Dit gaat over of tijdmetingen kloppen, de volgordelijkheid van gebeurtenissen kloppen, en de relatie tussen de gegevens en tijd kloppen.
- **Bruikbaarheid:** de mate waarin de registratie kan voldoen aan de gebruikerswensen voor een bepaalde toepassing. Dit gaat om eisen die niet door de overige kwaliteitsaspecten worden afgedekt. Denk hierbij aan uniformiteit van de gegevens.

Het is interessant om te zien dat er vanuit ISO twee verschillende kwaliteitsraamwerken bestaan. De door NEN-ISO 19157 voorgestelde

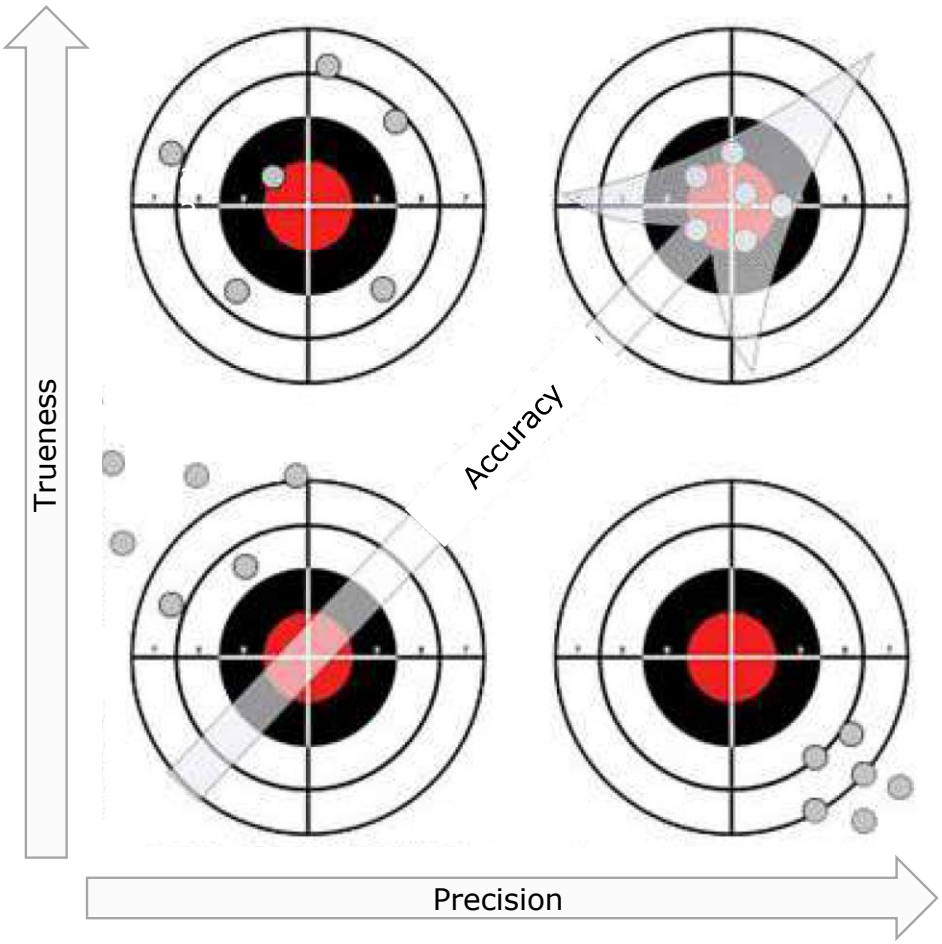
kwaliteitsdimensies zijn allemaal te relateren aan de kwaliteitsdimensies van ISO/IEC 25012. NEN-ISO 19157 gaat in op wat ISO/IEC 25012 inherente gegevenskwaliteit noemt. Daarin is met name veel aandacht voor nauwkeurigheid van gegevens en geometrische aspecten. Het is ook meer dan alleen een kwaliteitsraamwerk: het gaat ook in op het proces van kwaliteitsmeting en biedt onder meer aandacht voor evaluatiemethoden en rapportage.

2.3 ISO 5725-1:1994(E)

De ISO 5725-1:1994 standaard [29] heeft betrekking op de nauwkeurigheid van meetmethoden, resulaten en toepassingen. In tegenstelling tot de ISO/IEC 25012 standaard beschrijft de ISO 5725 standaard de methode om de kwaliteit van het inwinningsproces vast te stellen. Het gaat specifiek in op de relatie tussen nauwkeurigheid, juistheid en precisie. Dit is ook relevant in de context van berekeningen (rekenmodellen). Het vestigt daarmee ook specifiek aandacht op het belang van een goede inwinning van brongegevens. Daarbij wordt onderscheid gemaakt tussen systematische afwijkingen, toevallige afwijkingen en grove fouten. Deze kwaliteitsaspecten zijn nadrukkelijk meer statistisch van aard dan de meer administratieve aspecten zoals ze worden beschreven in de ISO/IEC 25012 standaard. Het gaat vooral over grote hoeveelheden meetgegevens. De standaard biedt dan ook een methode om statistische performance kentallen van de meetgegevens vast te stellen.

Systematische afwijking gaat over de mate waarin het resultaat van een groot aantal metingen of berekeningen overeenkomt met een geaccepteerde referentiewaarde. In ISO 5725 wordt dit “bias” genoemd, maar omdat die term ook negatieve connotaties heeft hanteert de standaard primair de term “trueness”. Dit kun je in het Nederlands “juistheid” noemen. Toevallige afwijking gaat over de mate waarin herhaalde metingen of berekeningen met elkaar oveenkomen. In ISO 5725 wordt dit “precision” genoemd. Het heet ook wel “statistische precisie”. Dat in tegenstelling tot “opslagprecisie”, wat betrekking heeft op het detailniveau waarin gegevens worden opgeslagen. Systematische afwijkingen en toevallige afwijkingen worden ook wel “fouten” genoemd. Een andere categorie van fouten zijn grove fouten, wat zowel onverklaarbaar buitensporige waarden (ook wel: “outliers”) als menselijke vergissingen kunnen zijn.

Nauwkeurigheid is het resultaat van de combinatie van systematische afwijking en toevallige afwijking. Afhankelijk van waar de meting of berekening betrekking op heeft gaat dit bijvoorbeeld over positionele nauwkeurigheid (coördinaten), temporele nauwkeurigheid (tijdmetingen) of kwantitatieve nauwkeurigheid (getallen). De relatie tussen nauwkeurigheid (accuracy), systematische afwijking (trueness) en precisie (precision) is weergegeven in de volgende figuur.

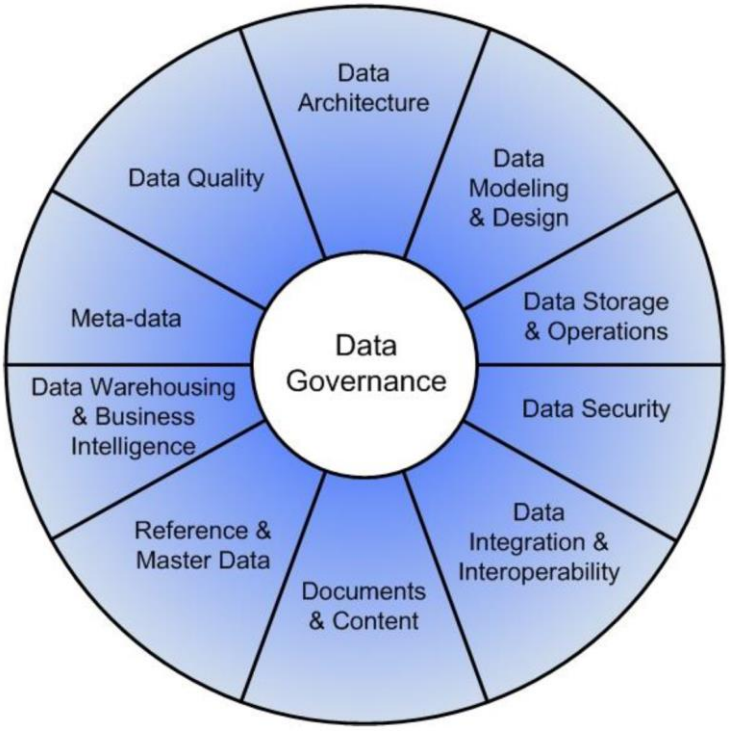


Figuur 2 ISO 5725: de relatie tussen accuracy, trueness en precision

Het voordeel van deze statistische kentallen is dat ze kunnen worden gebruikt bij het vaststellen van de onzekerheid en dus de bruikbaarheid van de informatieproducten die met behulp van de brongegevens worden samengesteld. Bij het aggregeren van gegevens zijn systematisch onzekerheden veel hardnekkiger dan de toevallige onzekerheden (foutenvoortplanting).

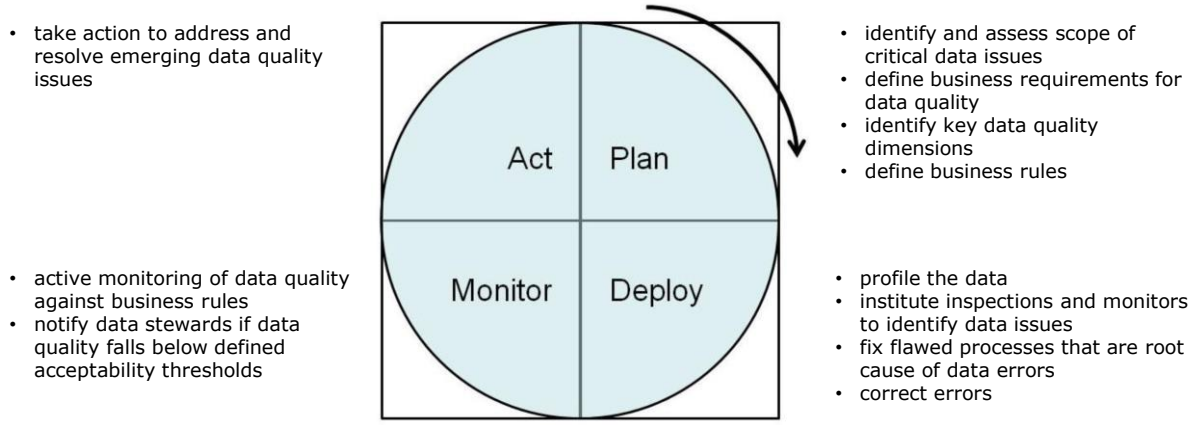
2.4 DAMA Data management body of knowledge

De Data Management Body of Knowledge (DMBoK) [4] is een verzameling van kennis en ervaring op het gebied van "data management". Het is het resultaat van een inspanning van de internationale Data Management Association (DAMA), die ook afdelingen in landen heeft. Het boek waarin dit verzameld is is onderverdeeld in een aantal thema's, waarvan gegevenskwaliteit er één is (zie Figuur 3). Per thema wordt een stuk algemene theorie beschreven, maar wordt ook ingegaan op de processen, rollen, producten en tools. Gegevenskwaliteit heeft in het model een sterke relatie met "data governance", waarin de afspraken worden gemaakt over rollen, taken en verantwoordelijkheden.



Figuur 3 Gegevenskwaliteit in context

De DMBoK definieert ook een verzameling kwaliteitsdimensies, die erg vergelijkbaar zijn met die van ISO/IEC 25012, maar op een paar punten wat specifieker zijn opgesplitst. De basis voor gegevenskwaliteit in DMBoK is een kwaliteitscyclus, vergelijkbaar met de Deming cyclus (zie Figuur 4). Het zorgt ervoor dat kwaliteitsverbetering continu de aandacht heeft. De cyclus start met het definiëren van de belangrijkste uitdagingen op het gebied van gegevenskwaliteit, het uitwerken van de kwaliteitseisen en daarna de kwaliteitsregels. Vervolgens worden dieper liggende problemen opgelost, waarna kwaliteit vooral wordt bewaakt. Dit betekent het meten van de kwaliteit van de gegevens tegen de eerder gespecificeerde eisen en regels. In geval van problemen worden gegevensbeheerders gevraagd deze op te lossen. De centrale rol m.b.t. gegevenskwaliteit in DMBoK is de gegevenskwaliteitsmanager, die voor specifieke zaken wordt ondersteund door de gegevenskwaliteitsanalist.



Figuur 4 kwaliteitscyclus in DMBoK

DMBoK beschrijft dat er allerlei tools ingezet kunnen worden ter ondersteuning van gegevenskwaliteit. Deze tools kunnen ingedeeld worden in de categorieën analyse, schoning, verbetering en monitoring of geleverd worden als geïntegreerde suite.

2.5 Code voor informatiekwaliteit

De Code voor informatiekwaliteit [5] is een denkraam en een verzameling van richtlijnen bedoeld om managers grip te geven op de kwaliteit van te leveren informatie. Het is initieel ontwikkeld door Platform voor Informatiekwaliteit Nederland (NLIQ). Na het opheffen van NLIQ in 2015 is het beheer overgedragen naar de werkgroep Informatiekwaliteit van DAMA NL, de Nederlandse tak van de Data Management Association (DAMA). In de code omvat het begrip "informatie" zowel data als metadata. Het document bestaat vooral uit een set van richtlijnen die aanbieders en afnemers van informatie met elkaar kunnen overeenkomen. Daarnaast beschrijft het de principes en begrippen die aan de richtlijnen ten grondslag liggen. Zo stelt het bijvoorbeeld dat de kwaliteit van de output voor de afnemer centraal zou moeten staan.

De code heeft een selectie gemaakt van kwaliteitsdimensies op basis van een aantal andere raamwerken zoals de DMBoK. De meeste dimensies kunnen worden gepositioneerd onder de kwaliteitsdimensies van ISO/IEC 25012. De code benoemt in de diverse richtlijnen ook kwaliteitseisen aan andere objecten zoals aan de afspraken, rapportages en het verwerkingsproces. Zo legt de code ook een relatie naar archivering, wat in andere raamwerken geen expliciete aandacht krijgt. Het definieert een richtlijn die stelt dat er er processen moeten zijn beschreven en geïmplementeerd voor het bewaren, archiveren en vernietigen van datasets. Daarnaast stelt het dat bewaartermijnen expliciet zouden moeten worden gemaakt. Hierbij wordt gerefereerd naar de Archiefwet en de Baseline Informatiehuishouding Rijksdienst.

2.6 Referentiemodel Datakwaliteit IPO

IPO heeft samen met CGI een referentiemodel voor datakwaliteit ontwikkeld [6]. Het referentiemodel beschrijft de dimensies en eigenschappen waarbinnen gegevenskwaliteit kan worden beschreven. Het model hanteert de volgende kwaliteitsdimensies:

- **Validiteit:** is er actueel en toetsbaar beleid?
- **Verifieerbaarheid:** is er een goed gedocumenteerd informatiemodel?
- **Conformiteit:** voldoen gegevens aan het informatiemodel en zijn ze actueel?
- **Traceerbaarheid:** zijn de transformatieregels van brongegevens tot afgeleide gegevens expliciet vastgelegd?
- **Betrouwbaarheid:** zijn de gegevens overeenkomstig de werkelijkheid?
- **Bruikbaarheid:** kan een gebruiker de gegevens interpreteren en gebruiken?
- **Duurzaamheid:** kan de informatievoorzienig in stand worden gehouden?

Voor elke kwaliteitsdimensie definieert het referentiemodel een aantal kwaliteitsdimensies. Kwaliteitsdimensies worden daarbij in een aantal gevallen ook hergebruikt in meerdere kwaliteitsdimensies. Dit maakt het model relatief uitgebreid. De meeste kwaliteitsdimensies zijn wel te plaatsen in de kwaliteitsdimensies van ISO/IEC 25012, alleen een aantal eigenschappen worden in het referentiemodel van

IPO uitvergroot. Zo gaat het model uitgebreid in op de beschikbaarheid van beleid en over het vastleggen van allerlei soorten regels.

2.7 Kwaliteitsraamwerk Rijkswaterstaat

Het werk van Rijkswaterstaat vraagt om bruikbare en betrouwbare informatie over de toestand en (de effecten van) het gebruik van hun netwerken. Voor de verschillende doelgroepen en klanten wordt gegevenskwaliteit op verschillende manieren aangepakt. Daarbij wordt vooral nog heel veel handmatig gecontroleerd. Op verschillende plekken, onder meer vanuit het NIS (NetwerkManagement Informatie Systeem), een BI omgeving van Rijkswaterstaat, ontstond de behoefte om op een meer generieke en geautomatiseerde manier met gegevenskwaliteit om te gaan. Pas echter in 2016 is RWS-breed gestart met het ontwikkelen van een raamwerk voor gegevenskwaliteit door bestaande standaarden en raamwerken met elkaar te vergelijken [21,22]. Dit raamwerk clustert de kwaliteitsdimensies in de vier hoofddimensies integriteit, juistheid, beheerbaarheid, vertrouwelijkheid, beschikbaarheid en interpretatie. Het raamwerk wordt ondersteund door een set van tools die een deel van de kwaliteitsdimensies volledig geautomatiseerd controleren. De tools leveren ook een kwaliteitsdashboard waarin snel inzichtelijk is wat de kwaliteit van een specifieke gegevensverzameling is. De intentie is om het raamwerk en de bijbehorende tools toe te gaan passen op alle datasets en leveringen van Rijkswaterstaat.

2.8 Relatie tussen raamwerken

In bijlage G is een vergelijking opgenomen van de in dit hoofdstuk beschreven raamwerken. Daar waar dimensies niet expliciet kunnen worden gekoppeld aan de dimensies van ISO/IEC 25012 zijn ze geplaatst bij de meest toepasselijke kwaliteitsdimensie. Door het leggen van deze relaties is ook inzichtelijk welke kwaliteitsattributen het belangrijkst zijn; de attributen die in alle raamwerken terug komen. Merk op dat termen in verschillende raamwerken soms een geheel andere betekenis hebben.

De kwaliteitsdimensies in ISO/IEC 25012 vormen een goede basis voor de Omgevingswet. Het raamwerk is breed genoeg om andere raamwerken in te plaatsen en biedt in combinatie met ISO/IEC 25024 ook een goede mate van concreetheid. Daarnaast is het een breed geaccepteerde, internationale standaard, wat ook het vertrouwen geeft dat het een goede basis vormt en herkenbaar is in een bredere doelgroep. De rest van dit document volgt dan ook de ISO/IEC 25012 standaard. Op het gebied van geo-informatie biedt de NEN-ISO 19157 meer concreetheid dan ISO/IEC 25012. Een combinatie van beide standaarden ligt dan ook voor de hand.

3 Ervaringen bij de overheid

Dit hoofdstuk geeft concrete voorbeelden van hoe met datakwaliteit wordt omgegaan bij de overheid. Er wordt eerst ingegaan op ervaringen met de basisregistraties BAG en BGT en vervolgens op AERIUS, Inspectieview Milieu en INSPIRE.

3.1 BAG en BGT

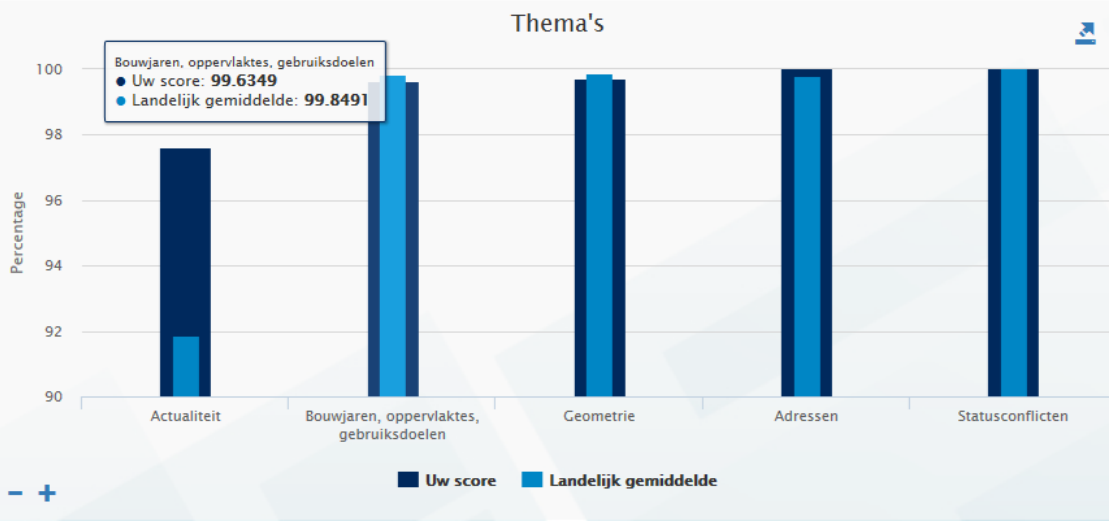
Er zijn veel overeenkomsten tussen de basisregistraties en de informatiehuizen, zeker waar het basisregistraties betreft met decentrale bronhouders (BAG, BGT, WOZ, BRP). Beiden hebben te maken met bronhouders die hun gegevens ontsluiten via een centrale voorziening, die verantwoordelijk is voor validatie. Beiden hebben ook te maken met een terugmeldplicht. Bij de BAG is de bronhouder tevens registratiehouder. De landelijke voorziening is slechts doorgeefluik. Bij de BGT is het Kadaster de registratiehouder, die de 'waarheid' vaststelt met gebruikmaking van gegevens van bronhouders. Daar waar de basisregistraties gericht zijn op een compacte set van authentieke kerngegevens, wordt er in het stelsel voor de Omgevingswet gekeken naar alle relevante informatie. Deze paragraaf beschrijft hoe bij BAG en BGT met gegevenskwaliteit wordt omgegaan. De werkwijze van BGT lijkt erg op die van BAG. Deze laatste is over een wat langere tijd geëvolueerd en recent is besloten tot een belangrijke verschuiving in aanpak.

Voor beide basisregistraties geldt dat de bronhouders zelf verantwoordelijk zijn voor gegevenskwaliteit. De houder van de landelijke voorziening (Kadaster) heeft geen formele toetsingsrol, maar controleert wel technische aspecten m.b.t. uitwisselbaarheid. Het ministerie is voor beide registraties verantwoordelijk voor toezicht en handhaving. De gegevenskwaliteit van BGT is wettelijk verankerd in een ministeriële regeling onder de wet BGT in de vorm van een gegevenscatalogus. De kwaliteitseisen zijn op het niveau van uitkomsten beschreven, het ‘hoe en wie’ binnen de bronhouder is vrij gelaten. Bronhouders dienen een jaarlijkse zelfcontrole uit te voeren, met een verantwoording aan IenM. Er vindt een driejaarlijkse audit plaats op de landelijke voorziening. De door BAG en BGT gehanteerde kwaliteitsdimensies lijken sterk op elkaar. Voor BAG zijn dat: juistheid, actualiteit, positionele nauwkeurigheid, oppervlakte nauwkeurigheid, volledigheid, compleetheid en tijd. Voor BGT zijn dat: thematische nauwkeurigheid, actualiteit, positionele nauwkeurigheid, volledigheid, logische consistentie en tijd.

Voor BAG geldt dat in de oorspronkelijke wet BAG naast inhoudelijke eisen ook proceseisen bevat die impliciet gaan over actualiteit, compleetheid en juistheid. De kwaliteitseisen zijn feitelijk geoperationaliseerd in een inspectieprotocol. Dat inspectieprotocol was destijds ook verankerd in een wettelijke regeling. Er was een driejaarlijkse audit van de bronhouders op zowel de inhoud als procesinrichting a.d.h.v. inspectieprotocol. De eisen zijn nu nog zowel op niveau uitkomsten beschreven als het ‘hoe en wie’ binnen de bronhouder. Dit stelsel van kwaliteitsborging was weinig doelmatig en effectief en wordt losgelaten n.a.v. wetsevaluatie in 2014. Je zou kunnen zeggen dat het stelsel was gebaseerd op wantrouwen. Bronhouders beleefden het daarmee vooral als een verplichting en werden niet aangesproken op de aanwezige inherente motivatie. De nieuwe werkwijze is gebaseerd op vertrouwen en faciliteren van bronhouders en legt de nadruk op

inhoud in plaats van procesinrichting. Daarmee verdwijnen procedurele eisen t.a.v. hoe de bronhouders intern werken en de formele audits van bronhouder. In plaats daarvan komt een jaarlijkse zelfcontrole. De nieuwe werkwijze is vergelijkbaar met BGT. Het gaat uit van integrale kwaliteitszorg gericht op inhoud en de inrichting van risicogericht toezicht. Deze nieuwe werkwijze wordt ondersteund door een operationeel kwaliteitsmanagementteam vanuit het Kadaster. Dit team ondersteunt bronhouders, met name door inzicht te geven in de huidige kwaliteit. Verder worden nu verspreid staande/impliciete kwaliteitseisen in de regelgeving op één plaats gebundeld op het niveau van ministeriële regeling. Dit is evenals bij de BGT op het niveau van een gegevenscatalogus.

Bij de ondersteuning van bronhouders bij het verbeteren van de kwaliteit van hun gegevens speelt het BAG kwaliteitsdashboard een centrale rol. In dit dashboard wordt inzichtelijk gemaakt hoe de bronhouder scoort op belangrijke kwaliteitsthema's en kwaliteitsregels. Figuur 5 is een uitsnede van dit kwaliteitsdashboard waarin zichtbaar is hoe een individuele bronhouder scoort t.o.v. anderen op de specifieke thema's. Deze thema's zijn geconcretiseerd in kwaliteitsregels die de meer abstracte kwaliteitseisen toetsbaar maken. Voor een deel gaan deze regels over echte fouten in de gegevens (bijv. een verblijfsobject ligt buiten een pand). Voor een ander deel zijn het signaleringen dat er mogelijk een fout zit in de gegevens. In het kwaliteitsdashboard is het mogelijk om dieper door te klikken op de specifieke fouten of signalen tot de individuele gegevensobjecten.

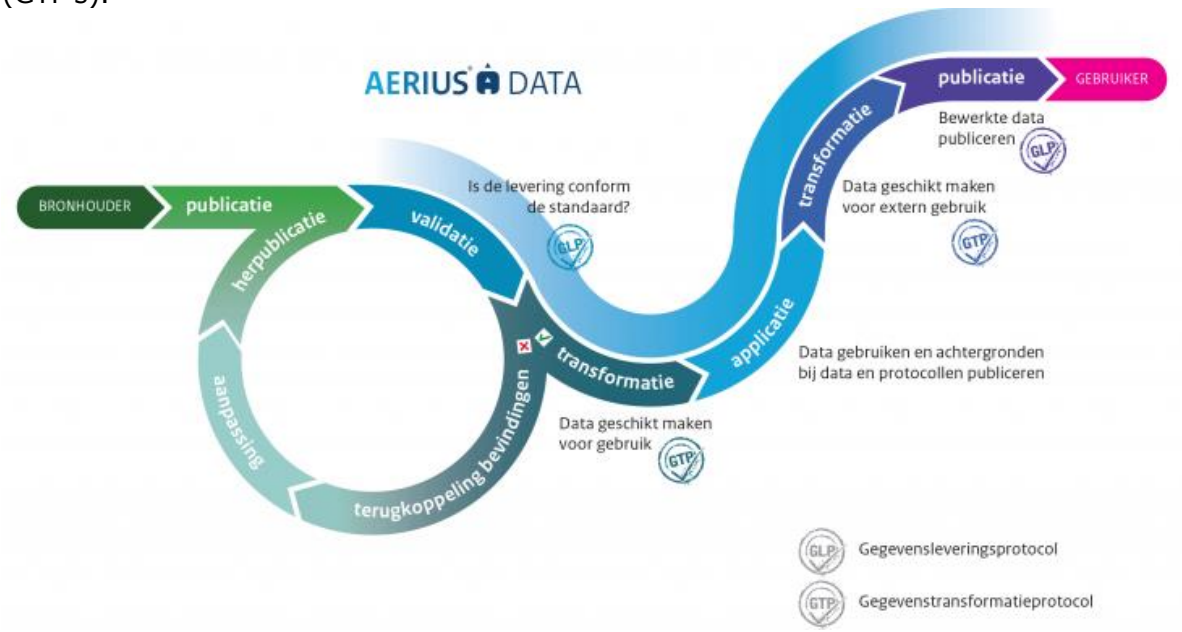


Figuur 5 Thema's in BAG kwaliteitsdashboard

3.2 AERIUS

Het rekeninstrument AERIUS is één van de pijlers van de Programmatische Aanpak Stikstof (PAS). AERIUS is een on-line applicatie voor het berekenen van de stikstofdepositie als gevolg van projecten en plannen op Natura 2000-gebieden. Zodra de PAS in werking treedt, zijn bedrijven wettelijk verplicht om hun vergunningaanvragen voor nieuwe economische activiteiten te onderbouwen met een berekening in AERIUS.

Datamanagement is een belangrijk onderdeel van AERIUS. Aan de basis van de AERIUS-berekeningen staan meer dan vijftig gegevenssets. Data die aantoonbaar van de best mogelijke kwaliteit moet zijn. AERIUS werkt hiervoor met een geformaliseerd model voor het datamanagement. Speerpunt bij datamanagement rondom AERIUS is verifieer- en reproduceerbaarheid. Er is een specifiek model voor datamanagement opgesteld in het kader van AERIUS. Dit beschrijft het volledige proces van inwinnen en bewerken tot publiceren van de data (zie Figuur 6). Steunpilaren van de aanpak zijn de gegevensleveringsprotocollen (GLP's) en de gegevenstransformatieprotocollen (GTP's).



Figuur 6 Model voor datamanagement AERIUS

In samenspraak met leverancier van brongegevens is onder meer een GLP opgesteld voor het samenstellen van een landelijke habitatkaart. Het protocol beschrijft het proces van aanleveren, de vorm en de inhoud van de bestanden en de eisen waaraan de data moeten voldoen. In 2013 hebben de bronhouders zo'n 450 karteringen geleverd, die allemaal aan de hand van het gegevensleveringsprotocol zijn gevalideerd. Bevindingen van de validatie zijn consequent teruggekoppeld naar de leverancier, zodat eventuele rectificaties aan de bron konden worden uitgevoerd. Uitgangspunt is dat de bronhouder verantwoordelijk is voor zijn/haar data en AERIUS geen inhoudelijke aanpassingen aanbrengt in de gegevens. De gegevenstransformatieprotocollen (GTP's) beschrijven de handelingen die het AERIUS-team uitvoert om de data gereed te maken voor gebruik binnen AERIUS. Denk bijvoorbeeld aan het transformeren van gegevens naar verschillende tabellen om de gewenste berekeningen te kunnen doen. Het werken met GTP's zorgt ervoor dat de bewerkingen op de data geborgd en reproduceerbaar zijn. Aanvullend op de GLP's en GTP's worden er ook gegevensleveringsovereenkomsten opgesteld (GLO's). Deze borgen de afpraken m.b.t. gegevensuitwisseling op bestuurlijk niveau zodat ook de noodzakelijke middelen beschikbaar zijn voor het borgen van gegevenskwaliteit.

Voor elke bron wordt bepaald of harmonisatie, validatie of transformatie van de gegevens relevant is. De bronhouder zorgt voor het tijdig, juist en volledig aanleveren van de brongegevens en bijbehorende metadata. Metadata is een essentieel onderdeel van de gegevenslevering en beschrijft de kwaliteit, actualiteit en herkomst van het

databestand. AERIUS controleert de tijdigheid, juistheid en of de inhoud voldoet aan de afspraken in de GLP. Controles vinden gedeeltelijk ook plaats middels inspectie door een inhoudsdeskundige, bijvoorbeeld in de vorm van steekproeven. Gegevenskwaliteit wordt gezien als een gezamenlijke verantwoordelijkheid; afnemers dragen bij aan het verbeteren van gegevenskwaliteit door de bronhouder te informeren als gegevens mogelijk foutief zijn of anderszins niet goed aansluiten bij het gebruik. Dit soort signalen zijn cruciaal om te komen tot continue verbetering. Het AERIUS team beschikt over tools en datasets waar bronhouders niet zelf altijd over kunnen beschikken, maar die wel cruciaal zijn om inzicht te krijgen in gegevenskwaliteit. AERIUS ondersteunt bronhouders daarmee actief in de verbetering van de kwaliteit van hun gegevens.

3.3 Inspectieview Milieu

Inspectieview Milieu (IvM) is een virtueel dossier waarin inspecteurs informatie over inspectieobjecten kunnen raadplegen. Het verhoogt de effectiviteit door risicogericht, informatie-gestuurd programmeren en leidt ook tot meer efficiëntie en een hogere kwaliteit. De informatie van IvM is voor toezichthouders en handhavers toegankelijk via een beveiligde website. De informatie is afkomstig van de inspectiediensten die hun data hiervoor beschikbaar stellen. Het is met de omgevingsdiensten DCMR, OZHZ, het Landelijk Meldpunt Afvalstoffen, de Politie en de Inspectie Leefomgeving en Transport ontwikkeld.

De basis van IvM is vastgelegd in een juridisch kader [11], die de wettelijke (en door de jurisprudentie gepreciseerde) grenzen beschrijft binnen welke IvM kan en mag worden gebruikt. Er wordt onder meer ingegaan op welke gegevens met welke organisatie mogen worden uitgewisseld en met welk doel. Daarnaast beschrijft het hoe wordt omgegaan met de verwerking van persoonsgegevens en welke rechten en plichten daaruit voort vloeien voor actoren. Daarbij wordt onder meer verwezen naar de eisen m.b.t. gegevenskwaliteit die voortvloeien uit de Wet Bescherming Persoonsgegevens. Dit betekent dat persoonsgegevens slechts mogen worden verwerkt voor zover zij – mede gelet op het doel van de verwerking – toereikend, relevant, niet bovenmatig, juist en nauwkeurig zijn. Daarbij is tevens aangegeven dat maatregelen moeten worden genomen die in redelijkheid kunnen worden gevergd. De redelijkheid stelt daarbij, afhankelijk van de soort gegevens die onderwerp van verwerking zijn, de stand van de techniek en de kosten die met de maatregelen gepaard gaan, grenzen aan de te nemen maatregelen.

Naast het juridisch kader is er een informatieprotocol [12] gedefinieerd voor IvM. Dit beschrijft een verzameling van aanvullende afspraken over het gebruik van gegevens. Hierin zijn onder meer vastgelegd welke gegevens worden verwerkt en welk vertrouwelijkheidsniveau geldt. Er staat ook in dat bronhouders streven naar optimale volledigheid en juistheid van de verstrekte gegevens en bij iedere verstrekking de actualiteit van een gegeven melden. Inspectieview Milieu levert de informatie ongewijzigd door. Afspraken over de structuur van uit te wisselen informatie zijn vastgelegd in het Referentie Informatiemodel Handhaving (RIHa). Specifiek voor toepassing in het milieudomein is een toepassingsprofiel RIHa-Milieu opgesteld.
Er zijn bij IvM zes leidende aspecten (kwaliteitsdimensies) gedefinieerd: eenduidigheid, juistheid, actualiteit, volledigheid, rechtmatigheid en beheerd. Deze aspecten zijn doorvertaald naar prestatie-indicatoren en streefwaarden in

dienstverleningsovereenkomsten (DVO's) met bronhouders. Er wordt in de afspraken onderscheid gemaakt tussen het gewenste niveau per ingang van de DVO, over 2 jaar na datum ingang van de DVO en op de lange termijn (3 tot 5 jaar na de ondertekening van de DVO). De afspraken in deze DVO's worden jaarlijks geactualiseerd. Tabel 1 geeft een overzicht van deze aspecten en de prestatie-indicatoren die daarbij zijn gehanteerd.

Aspect	**Prestatie-indicatoren**
Eenduidigheid	• Gebruik van standaarden; • Gebruik van basisregistraties; • Correcte levering van aansluitvoorziening aan Inspectieviews (geautomatiseerd proces). • Correcte verwerking door Inspectieviews (geautomatiseerd proces).
Juistheid	• Gebruik van de basisregistraties; • Juiste invoer van de eigen gegevens: • Correcte levering van aansluitvoorziening aan Inspectieviews (geautomatiseerd proces). • Correcte verwerking door Inspectieviews (geautomatiseerd proces).
Actualiteit	• Periode tussen uitvoeren van inspectie en daadwerkelijke verstrekking aan IV? • Eenmaal in AV-ILT, hoe snel beschikbaar voor afnemers IV?
Volledigheid	• Periode waarover informatie wordt uitgewisseld. • Zo volledig mogelijke invoer van gegevens in het VTH / inspectie systeem. • Zo volledig mogelijke uitwisseling van informatie met IV. • Zo weinig mogelijk uitval in de AV-ILT
Rechtmatigheid	• Uitwisseling door IV vindt plaats conform wettelijke kaders en vastgelegd in afspraken.
Informatiebeheer	• Processen zijn ingericht en geborgd in de eigen organisatie.

Tabel 1 Aspecten van gegevenskwaliteit bij Inspectieview Milieu

3.4 INSPIRE

De Europese richtlijn INSPIRE voorziet in een harmonisatie van geo-informatie in en tussen de lidstaten. De uitvoeringregels van INSPIRE vragen Nederland om een technische en organisatorische infrastructuur waarin geo-informatie van de overheden toegankelijk en interoperabel zijn. INSPIRE stelt de eis dat gegevens worden ontsloten conform het door INSPIRE vastgestelde informatiemodellen (data specifications). Dit vraagt harmonisatie en integratie van brongegevens op zowel semantisch als syntactisch niveau. Hiermee stelt het eisen aan de juistheid en consistentie van gegevens. In de informatiemodellen wordt ook aangegeven hoe wordt omgegaan met temporele aspecten (formele en materiële historie). Daarnaast bevatten de informatiemodellen een hoofdstuk over gegevenskwaliteit. Hierin staat onder meer hoe wordt omgegaan met kwaliteitsdimensies zoals volledigheid, consistentie en nauwkeurigheid van gegevens.

INSPIRE gaat ook specifiek in op metadatering, waarbij wordt verwezen naar de ISO 19115 standaard. In Nederland is afgesproken gebruik te maken van het Nederlands profiel op ISO 19115 voor geografie. Hierin is expliciet aangegeven welke velden moeten zijn ingevuld. INSPIRE stelt daarmee ook eisen aan de volledigheid en consistentie van metadata. Het stelt ook eisen aan de actualiteit van gegevens; ze mogen niet ouder zijn dan een half jaar na de laatste update van brongegevens. Daarnaast stelt INSPIRE eisen aan de technische ontsluiting van gegevens (Quality of Service), waarbij de nadruk ligt op performance, capaciteit beschikbaarheid en betrouwbaarheid van de netwerk services. Een belangrijk aspect van een INSPIRE implementatie is de mogelijkheid deze te valideren en monitoren middels, bij voorkeur, geautomatiseerde processen, dan wel beschreven procedures. Jaarlijks verwacht de Europese Unie een rapport van Nederland waar op basis van monitoring de conformiteit aan de regelgeving beschreven wordt voor de aangeboden datasets en services. Geonovum faciliteert validatie en monitoring door het beschikbaar stellen van een aantal validatietools. Deze tools controleren of de gegevens of services voldoen aan de standaarden en/of profielen en rapporteren de bevindingen. Voor geharmoniseerde datasets is de ambitie om geautomatiseerde validatietesten te realiseren conform de gedefinieerde informatiemodellen.

Publieke Dienstverlening Op de Kaart (PDOK) is het Nederlandse geo-portaal voor de ontsluiting van INSPIRE datasets. Tezamen met het Nationaal Geo Register (NGR) vormt het de toegangspoort naar de nationale INSPIRE-data. Er worden op dit moment circa 100 datasets ontsloten via PDOK. PDOK controleert of ontvangen bestanden voldoen aan de productspecificatie en het datamodel en of de geometrie valide is en topologisch correct. Indien de dataset niet valide is wordt dit teruggerapporteerd aan de data-aanbieder. Deze dient de fouten te herstellen en het bestand opnieuw aan te leveren. Bij het valideren kan een grenswaarde gehanteerd worden voor het aantal fouten. Bij grote datasets kan de impact van een klein aantal fouten beperkt zijn. Als het aantal fouten onder deze grenswaarde blijft, dan wordt het databestand verder verwerkt. De fouten worden wel naar de data-aanbieder teruggekoppeld, zodat deze gecorrigeerd kunnen worden. Na de geautomatiseerde ontvangst wordt visueel gecontroleerd of er geen buitensporigheden in de gegevens zitten. Bij handmatige updates van datasets wordt tevens geautomatiseerd gecontroleerd of de database extreem gegroeid of verkleind is (te weinig of teveel geleverd). De bronhouder is zelf verantwoordelijk voor de aangeleverde gegevens (en metadata) en hun actualiteit. PDOK brengt geen inhoudelijke wijzigingen in de gegevens aan. De transformaties die PDOK uitvoert betreffen alleen de vorm en niet de inhoud.

3.5 Evaluatie van cases

Deze paragraaf evalueert de in dit hoofdstuk beschreven cases, onder meer door te positioneren t.o.v. de ISO/IEC 25012 standaard. Daarbij is alleen gekeken naar documenten m.b.t. de cases die expliciet aangeven dat ze betrekking hebben op gegevenskwaliteit. De relatie is weergegeven in bijlage H, waarbij voor eigen kwaliteitsdimensie in ISO/IEC 25012 de gerelateerde dimensie uit de betreffende casus is weergegeven. Daar waar dimensies niet expliciet kunnen worden gekoppeld aan de dimensies van ISO/IEC 25012 zijn waar mogelijk geplaatst bij de meest toepasselijke kwaliteitsdimensie. Dimensies die niet expliciet zijn voorzien van een naam in de brondocumenten zijn schuingedrukt met de naam van de dimensie zelf

weergegeven. Dimensies die helemaal niet kunnen worden gekoppeld aan de ISO/IEC 25012 dimensies zijn in de onderste rij van de tabel weergegeven.

Uit de vergelijking is af te leiden dat ISO/IEC 25012 een aantal kwaliteitsdimensies die worden gehanteerd in de cases niet expliciet onderkend. Dit zijn allemaal dimensies die niet iets over de gegevens zelf zeggen. Ze hebben vooral betrekking op de inrichting van data governance en de kwaliteit van informatiemodellen. Onafhankelijk van of deze dimensies worden gezien als onderdeel van gegevenskwaliteit, zullen ze wel moeten worden geborgd.

Daarnaast is uit de vergelijking af te leiden dat de precieze invulling van kwaliteitsdimensies niet overeenkomt. Zo wordt bij BAG [32] de kwaliteitsdimensie "compleetheid" relatief breed opgevat en vallen er ook criteria onder die bij ISO/IEC 25012 onder "nauwkeurigheid" vallen.

Er kunnen allerlei zaken geleerd worden uit de beschreven cases waar gebruik van worden gemaakt in de context van de omgevingswet, zoals:

- Het faciliteren van bronhouders in het verbeteren van gegevenskwaliteit blijkt beter te werken dan ze alleen op procesniveau aanspreken (BAG).
- Door gegevenskwaliteit inzichtelijk te maken in een dashboard kunnen bronhouders gericht verbeteracties uitvoeren (BAG).
- Gegevenskwaliteit is sterk gerelateerd aan een informatiemodel en kan integraal worden beschreven (BGT).
- Gegevenskwaliteit kan wettelijk worden verankerd in een ministeriële regeling (BGT).
- Het opstellen van gegevensleveringsprotocollen en gegevenstransformatieprotocollen zorgt voor heldere afspraken tussen leverancier en afnemer van gegevens (AERIUS).
- Gegevenskwaliteit is een gezamenlijke verantwoordelijkheid van de keten (AERIUS).
- Het is belangrijk om het belang en draagvlak voor gegevenskwaliteit ook op bestuurlijk niveau te borgen (AERIUS).
- Het helpt als gegevenskwaliteit wordt doorvertaald naar prestatie-indicatoren en streefwaarden in dienstverleningsovereenkomsten (DVO's) met bronhouders (IvM).
- Gegevenskwaliteit vraagt het inrichten van processen en het borgen ervan in de organisatie (IvM).
- Een deel van gegevenskwaliteit is gerelateerd aan de kwaliteit van meta-data (INSPIRE).
- Er kunnen validatietools worden gebruikt voor het geautomatiseerd controleren van een aantal belangrijke gegevenskwaliteitsaspecten (INSPIRE).

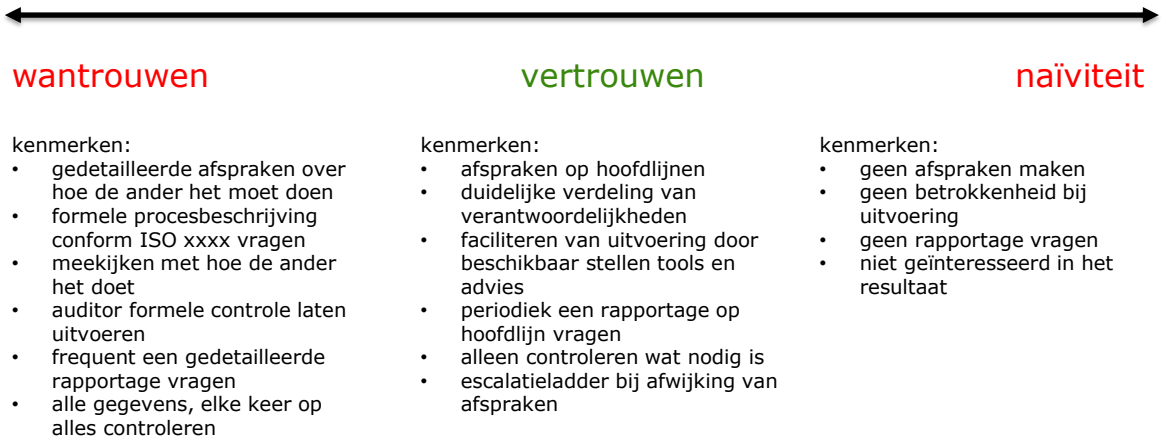
4 Onze visie

Dit hoofdstuk beschrijft een specifieke visie op gegevenskwaliteit, afgestemd op de context van de Omgevingswet. Het hoofdstuk eindigt met een integraal beeld van gegevenskwaliteit waarin alle relevante aspectgebieden zijn benoemd.

4.1 Vertrouwen als basis voor kwaliteit in de keten

Het Digitaal Stelsel Omgevingswet functioneert als een keten. Een grote diversiteit van organisaties werkt samen aan het digitaal ondersteunen van de Omgevingswet. Bronhouders verzamelen gegevens en leveren deze aan partijen die ze verwerken tot informatieproducten. Deze worden doorgeleverd aan centrale voorzieningen in het stelsel die ze via gebruikerstoepassingen ontsluit. Deze gehele keten moet in beschouwing worden genomen bij inrichten van gegevenskwaliteit; het is een verantwoordelijkheid van de keten als geheel. De basis daarvoor zou vertrouwen moeten zijn (zie Figuur 7). Dit laat iedere partij in de keten in haar eigen verantwoordelijkheid en voorkomt overmatige bureaucreatie. Dit betekent dat afspraken vooral op hoofdlijnen worden gemaakt, gericht op uitkomsten en niet gaan over hoe anderen hun proces precies hebben ingericht. Controles zijn beperkt tot wat nodig is en worden niet in elke schakel in de keten opnieuw uitgevoerd. Ze worden bij voorkeur ook zo vroeg mogelijk in de keten uitgevoerd zodat fouten zo snel mogelijk kunnen worden gedetecteerd en gecorrigeerd.

Om vertrouwen te geven is transparantie randvoorwaardelijk. Door anderen inzicht te geven in kwaliteit weten zij waar ze aan toe zijn. Als kwaliteit onvoldoende is dan vraagt dat in eerste instantie een gesprek waarin de oorzaken besproken kunnen worden. Het kan zijn dat partijen onvoldoende middelen hebben om kwaliteit inzichtelijk te maken of te borgen. Pas als dat niet blijkt te werken kan een escalatieladder in werking treden.



Figuur 7 Vertrouwen als basis voor gegevenskwaliteit

4.2 Denken vanuit gebruik

De kwaliteit van gegevens moet aansluiten bij het gebruik; bij wensen en verwachtingen van de gebruikers van de gegevens. Dit wordt ook wel interne kwaliteit genoemd. Daar tegenover staat externe kwaliteit dat vanuit het perspectief van de leverancier kijkt [19]. De externe kwaliteit zou dus leidend moeten zijn. Dit bepaalt ook hoeveel geld zal worden geïnvesteerd in gegevenskwaliteit. Het is voor een belangrijk deel gebaseerd op de processen die worden ondersteund. Een te hoge kwaliteit kan net zo onwenselijk zijn als een te lage kwaliteit, omdat dat een inspanning vraagt die niet in verhouding staat tot wat nodig is voor het gebruik. Het is ook geen toeval dat bij elke kwaliteitsdimensie in de ISO/IEC 25012 standaard expliciet in de definitie van de kwaliteitsdimensies is opgenomen dat deze is afgestemd op de specifieke gebruikscontext. Afnemers kennen hun eigen gebruik het best en zijn dan ook het best in staat de eisen vanuit hun eigen gebruik te definiëren. Daarnaast kunnen zij ook feedback geven op de kwaliteit van gegevens geredeneerd vanuit hun eigen gebruik. Door dit soort mechanismen ontstaat er een feedbackloop van afnemer naar aanbieder.

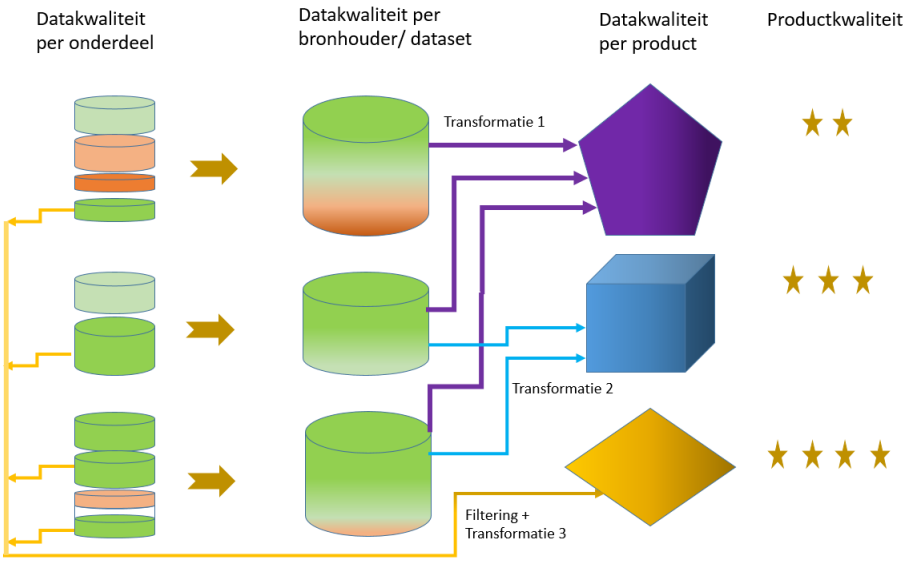
Veel gegevens worden door een grote diversiteit aan afnemers gebruikt wat het erg lastig maakt om in generieke zin te spreken over hun kwaliteit. Ook voor basisregistraties gelden dergelijke uitdagingen en moet er een balans zien te worden gevonden tussen de inspanning die nodig is voor een bepaald kwaliteitsniveau versus de behoeften vanuit gebruik. In de praktijk komen eisen aan dit soort registraties dan ook tot stand middels een dialoog en belangenafweging tussen bronhouders, afnemers en financiers. In de context van de Omgevingswet zal goed moeten worden gekeken naar de eisen die gesteld worden vanuit de verschillende procesfasen (beleidsvorming, oriënteren, checken, indienen, toezicht, handhaving, etc.). Daarnaast zullen eisen per informatieproduct moeten worden bepaald. Een uitdaging daarbij is dat veel informatieproducten nieuw zijn wat het definiëren van kwaliteitseisen lastig maakt. Ervaringen met vergelijkbare informatieproducten en gegevensverzamelingen, alsook inzicht in de huidige kwaliteit kunnen daarbij ondersteunen. Kwaliteitseisen zullen ook gebaseerd moeten zijn op wat bronhouders waar kunnen maken tegen acceptabele kosten.

4.3 Transparantie naar alle doelgroepen

Er wordt in toenemende mate van overheidsorganisaties verwacht dat ze transparant zijn. Overheid, burgers en bedrijven willen inzicht in de wijze waarop besluiten tot stand zijn gekomen. Burgers en bedrijven verwachten een gelijkwaardige informatiepositie zodat ze inzicht hebben in de gegevens die de gemeente van hen heeft en welke gegevens zijn gebruikt om tot een besluit te komen. Inzicht in de kwaliteit van gegevens is een vorm van transparantie die alle gebruikersgroepen ondersteunt.

Afnemers kunnen door inzicht in kwaliteit inschatten of gegevens aansluiten bij hun gebruik. Kwaliteit moet richting afnemers op een eenvoudige manier inzichtelijk zijn. Wij denken dat een sterrensysteem daarbij kan ondersteunen. Elk informatieproduct krijgt dan een score van nul tot vijf sterren, waarbij de precieze opbouw van de sterren gebaseerd is op de kwaliteitsaspecten die voor dat informatieproduct belangrijk zijn. Deze score zou tot stand moeten komen in overleg met bronhouders

en afnemers. Een aanzet tot een dergelijk sterrensysteem is ook gedaan vanuit informatiehuis natuur [28] (zie Figuur 8). Een belangrijke vraag bij het bepalen van een dergelijk sterrensysteem is de grondslag waarop de sterren zijn gebaseerd; een echte meting van de kwaliteit, een globale inschatting (dat is wat er vanuit informatiehuis natuur wordt voorgesteld) of een oordeel van een afnemer.



Figuur 8 Aanzet tot sterrensysteem voor kwaliteit vanuit informatiehuis natuur

Afnemers willen ook graag inzicht in de brongegevens die ten grondslag liggen aan informatieproducten en de transformaties die zijn uitgevoerd om te komen tot een informatieproduct. Ze zouden ook middelen moeten hebben om twijfels die ze hebben over de kwaliteit van gegevens terug te melden. Laagdrempelige terugmeldvoorzieningen zijn belangrijke mechanismen voor de verbetering van gegevenskwaliteit.

Bronhouders en anderen kunnen inzicht in de kwaliteit van gegevens gebruiken om de kwaliteit te verbeteren; ze weten dan beter waar verbetermogelijkheden liggen. Inzicht in kwaliteit van brongegevens geeft informatiehuizen ook het inzicht of dit voldoende is voor de informatieproducten die zij leveren. Het geeft ze ook het vertrouwen dat bronhouders pro-actief bezig zijn met verbetering van de kwaliteit.

Inzicht in gegevenskwaliteit kan ook gebruikt worden voor het optimaliseren van de informatieketen. Als je met statistische kennis kijkt naar de impact van nauwkeurigheid van brongegevens op die van een informatieproduct, dan kun je daarop optimaliseren. Zo kunnen mogelijk bepaalde stappen in de keten worden vereenvoudigd door metingen te vervangen door rekenmodellen. Zo waren er bijvoorbeeld in het verleden zeer omvangrijke meetnetten. Door betere rekenmodellen en andersoortige brongegevens is inmiddels een betere kwaliteit mogelijk met veel minder meetpunten (zie bijvoorbeeld de weersverwachting).

Om inzicht te geven in kwaliteit kan een kwaliteitsdashboard vergelijkbaar met dat voor BAG worden ingericht. Dit geeft inzichten die bronhouders in een aantal gevallen ook zelf niet kunnen genereren omdat hun applicaties daar ook niet altijd mogelijkheden voor bieden. De kwaliteit van gegevens die niet door de validatie van

informatiehuizen heen komt (wordt afgewezen bij ontvangst) is idealiter ook inzichtelijk.

Waarschijnlijk hebben verschillende type gebruikers een variatie in informatiebehoefte over de kwaliteit van gegevens. Een burger of een aanvrager van een eenvoudige vergunning is vermoedelijk het meest geholpen met een eenvoudig sterrensysteem waarbij veel sterren impliceren dat de data zonder al teveel bedenkingen gebruikt kunnen worden. Een kwaliteitsdashboard, indien openbaar, geeft meer ruimte voor specifieke informatie zodat de gespecialiseerde gebruiker meer kan inschatten waarvoor de gegevens gebruikt kunnen worden. Voor situaties die gecompliceerder zijn is mogelijk een meer diepgaand inzicht in scores voor specifieke indicatoren relevant.

4.4 Kwaliteit organisatorisch en procesmatig borgen

Kwaliteit raakt alle organisaties in de keten, die hiervoor kwaliteitsprocessen zullen moeten inrichten. Kwaliteitsverbetering is niet een éénmalige inspanning, maar een continu proces. Het is niet reëel om te verwachten dat de kwaliteit van gegevens direct op het gewenste niveau is. Als het geen aandacht krijgt dan zal het kwaliteitsniveau terugvallen. Daarnaast zal 100% in de meeste gevallen ook niet haalbaar of nodig zijn. Continue verbetering van gegevenskwaliteit zou het uitgangspunt moeten zijn. Dit betekent dat het moet worden geborgd in de governance, processen en rollen. Er zullen allerlei mensen bij betrokken zijn, zodat het wordt gepland, gecoördineerd, uitgevoerd en gecontroleerd. De basis is een Plan-Do-Check-Act cyclus waarin planmatig verbeteringen worden geïdentificeerd, uitgevoerd, gemeten en bijgesteld. Er moeten mechanismen zijn ingericht waarmee in de gehele keten inzicht is in kwaliteit en waarbij partijen elkaar kunnen informeren over verbeteringen.

Een kernactiviteit is kwaliteitsmonitoring: het beoordelen van de kwaliteit van gegevens. Dat kan voor een deel door kwaliteitstoetsen in te bouwen in de gegevensstroom, die bijvoorbeeld controleren of gegevens voldoen aan het informatiemodel (of het daarvan afgeleide schema) en consistent zijn met gegevens in andere datasets. Periodiek kan worden gecontroleerd of een dataset als geheel voldoet aan de voorgedefinieerde kwaliteitsregels en kwaliteitseisen. Voor een ander deel is menselijke analyse van de gegevens zelf nodig. Daarnaast kan gebruik worden gemaakt van data profiling tools om afwijkingen inzichtelijk te maken. Om te toetsen of gegevens overeen komen met de werkelijkheid zijn meestal steekproeven noodzakelijk. Bij al deze analyses is betrokkenheid van een domeinexpert essentieel; deze heeft voldoende kennis om te beoordelen wat de impact van afwijkingen is.

Gegevenskwaliteit vraagt ook data governance; afspraken over taken en verantwoordelijkheden. Dit zijn afspraken zowel op bestuurlijk niveau als op operationeel niveau en zowel in de keten als binnen individuele organisaties. Bestuurders moeten begrijpen dat gegevenskwaliteit belangrijk is en dat daarvoor de middelen beschikbaar moeten zijn. Voor de uitvoering is het belangrijk dat afspraken worden gemaakt over de uit te wisselen gegevens en de daarbij geldende kwaliteitseisen, bijvoorbeeld in de vorm van gegevensleveringsprotocollen. Het is belangrijk goed na te denken waar de eindverantwoordelijkheid voor de inherente gegevenskwaliteit wordt belegd. Het is verstandig deze zoveel mogelijk aan het begin

van de keten te beleggen en wel bij een partij die er zelf een (gebruiks)belang bij heeft. Dit zorgt ervoor dat incentives voor kwaliteitsverbetering op de juiste plaats liggen.

Gegevensbeheer en gegevenskwaliteit verdienen een managementsysteem waarin al dit soort aspecten kunnen worden geborgd. Dat kan op basis van de ISO 9001 standaard [30] of de daaraan ten grondslag liggende ISO High Level Structure [31]. Deze laatste beschrijft een generieke structuur voor managementsystemen. Gegevenskwaliteit sluit bij voorkeur maximaal aan bij de bestaande governance en managementsystemen in de organisatie.

4.5 Informatiemodel definieert essentie van gegevenskwaliteit

Gegevens worden in de keten omgevormd tot betekenisvolle informatie. Informatiemodellen beschrijven de betekenis van gegevens en de relatie met andere gegevens. Betekenis en kwaliteit zijn onlosmakelijk aan elkaar verbonden. Informatiemodellen beschrijven een belangrijk deel van de kwaliteitsregels. Ze geven aan welke gegevens moeten zijn gevuld, welke waarden ze kunnen aannemen en hoe ze aan elkaar kunnen zijn gerelateerd. Dat gaat bij voorkeur verder dan wat er in een standaard gegevensmodel kan worden uitgedrukt. Ook meer complexe verbanden tussen gegevens zouden in de vorm van kwaliteitsregels moeten zijn opgenomen in een informatiemodel. Het is ook de plaats waar traceerbaarheid naar brongegevens, koppelbaarheid met andere datasets, formele en materiële historie en inzicht in de totstandkoming van de gegevens (data provenance) moet worden geborgd. Een informatiemodel is daarmee een belangrijk instrument bij het verbeteren van de kwaliteit van gegevens. Overigens hoeft een informatiemodel er niet per definitie vanaf dag één te zijn; het kan ook het resultaat zijn van consensus die ontstaat nadat eerste ervaringen met gegevensuitwisseling zijn opgedaan.

Informatiemodellen bestaan op meerdere plaatsen in de keten, zowel voor brongegevens als voor informatieproducten. Zij zouden hergebruik moeten maken van standaard informatiemodellen en gebaseerd moeten zijn op standaard modelleerafspraken. Een informatiehuis is verantwoordelijk voor het borgen dat informatiemodellen van brongegevens betekenisvol worden gecombineerd in een informatiemodel voor een informatieproduct. Daarbij is de betrokkenheid van domeinspecialisten cruciaal; zonder voldoende inhoudelijke kennis van de te combineren gegevens worden snel verkeerde aannames gedaan. Beschikbaarheid van een informatiemodel voor de brongegevens is daarbij bijzonder waardevol. Het zorgt dat een belangrijk deel van de kennis over de brongegevens expliciet is gemaakt. Ontvangen brongegevens worden door een informatiehuis gecontroleerd, met name door te controleren of ze voldoen aan het informatiemodel dat behoort bij de brongegevens.

4.6 Gegevenskwaliteit als onderdeel van metadata

Metadata speelt een belangrijke rol bij gegevenskwaliteit. Metadata is een beschrijving van gegevens. Het geeft inzicht in wat gegevens betekenen, hoe ze tot stand zijn gekomen en hoe ze kunnen worden gebruikt. Metadata bestaat op meerdere niveaus (zie Tabel 2). Het kan gericht zijn op de eindgebruikers (business) of meer technisch

van aard zijn. Daarnaast bestaat het zowel op het niveau van de algemene definitie, als op het niveau van een dataset die conform die definitie is ingericht en op het niveau van de individuele gegevensobjecten (feiten, instanties). Gegevenskwaliteit zou direct inzichtelijk moeten zijn in de metadata, zodat je als afnemer direct weet of gegevens bruikbaar voor je zijn. Daarmee ontstaat de transparantie zoals eerder beschreven. Daarnaast is de kwaliteit van de metadata zelf relevant; als deze slechts beperkt is dan is het voor afnemers lastig te bepalen of ze de gegevens kunnen gebruiken.

	Business	**Technisch**
Definitie	• Eigenaar • Begrippen en definities • Informatiemodel • Kwaliteitseisen • Kwaliteitsregels • Bedrijfsregels	• Fysiek gegevensmodel • Mappingregels • Conversieregels
Dataset	• Beheerder • Afbakening dataset • Datum laatste wijzging • Aantal gegevensobjecten • Audit trail • Aanwezige kwaliteit	• Endpoint URL • Job schedule • Backupregels • Technische logfiles • Fysieke omvang
Gegevens-object	• Auteur • Titel • Business identificatie • Locatie van gegevensobject • Wijzigingsgeschiedenis • Aanwezige kwaliteit	• Afzender • Digitale handtekening • Controlegetal (hashcode) • Verzenddatum • Bericht/bestandsidentificatie • Correlatie/transactie-identificatie

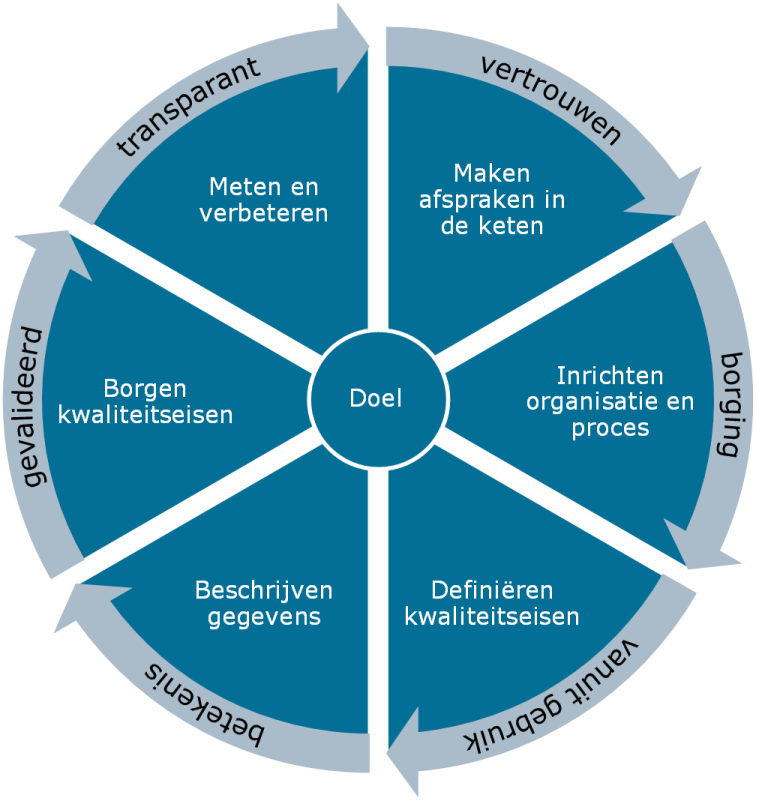
Tabel 2 Categorieën van metadata en voorbeelden van elementen

In de context van de Omgevingswet wordt business meta-data voor een belangrijk deel beschreven in de Stelselcatalogus DSO. Hierin zijn onder meer de informatieproducten, informatiemodellen en begrippen vastgelegd. Gegevenskwaliteit zou ook direct inzichtelijk moeten zijn in deze catalogus, zowel de gewenste kwaliteit als de gemeten kwaliteit. Het W3C heeft recentelijk een standaard gepubliceerd voor het beschrijven van gegevenskwaliteit waarvan gebruik gemaakt kan worden: de Data Quality Vocabulary [20]. Deze standaard biedt ruimte om zowel de indicatoren alsook de gemeten kwaliteit te registreren. Het gebruik van deze standaard is ook een Data on the Web Best Practice van de W3C[1]. Bronhouders en informatiehuizen zullen naast metadata voor informatieproducten ook voor de gegevens die daaraan ten grondslag liggen metadata moeten vastleggen. Metadatabeheer is een functie die expliciet moet worden ingericht op verschillende niveau's in de keten. Metadata zal ook beschikbaar moeten zijn in het Nationaal Geo Register (NGR). In dit register wordt metadata geregistreerd conform het Nederlands profiel voor ISO 19115 [15]. In deze standaard is ook (beperkt) ruimte om een aantal kwaliteitseigenschappen van datasets vast te leggen, zoals compleetheid, geometrische nauwkeurigheid en herzieningsfrequentie.

[1] zie: https://www.w3.org/TR/dwbp/#quality

4.7 Integraal beeld van gegevenskwaliteit

Op basis van de voorgaande beschrijving ontstaat een beeld van de aspecten die relevant zijn als het gaat om gegevenskwaliteit en de sleutelwoorden die daarbij gelden. Figuur 9 geeft hiervan een overzicht. In deze figuur is ook een logische volgordelijkheid aangebracht van aspecten, waardoor het ook gebruikt kan worden als een procesmodel op hoofdlijnen. Het start met het maken van afspraken in de keten over verantwoordelijkheden en het inrichten van organisatie en processen (data governance). Vervolgens zullen eisen moeten worden gesteld aan de gegevens en zullen de gegevens en informatie nader moeten worden beschreven in informatiemodel, kwaliteitsregels en andersoortige metadata. Kwaliteitseisen zullen moeten worden geborgd in processen en systemen, waarna aan een continu proces van meten en verbeteren in gang kan worden gezet. Deze fasering is verder uitgewerkt in hoofdstuk 6.



Figuur 9 Overzicht van aspecten van gegevenskwaliteit

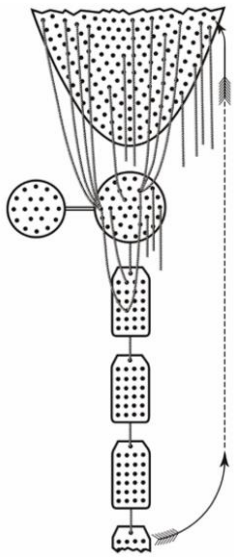
5 Van visie naar een kwaliteitsraamwerk

Dit hoofdstuk concretiseert de visie naar een kwaliteitsraamwerk. Het stelt een standaard verzameling van kwaliteitsdimensies en indicatoren voor. Het hoofdstuk start met een beschrijving van de gehanteerde concepten. In de daarop volgende paragraaf wordt het kwaliteitsraamwerk beschreven, inclusief de uitgangspunten die zijn gehanteerd bij het opstellen ervan. Vervolgens worden de indicatoren verder uitgewerkt en worden voorbeelden gegeven van meetfuncties voor de indicatoren.

5.1 Concepten

De begrippen die zijn gebruikt in dit document zijn beschreven in bijlage C. Deze paragraaf geeft een overzicht en uitleg van een aantal van de kernbegrippen.

Gegevenskwaliteit heeft betrekking op gegevens die zijn vastgelegd over objecten in de werkelijkheid (zie Figuur 10). Daar wordt voor een dataset een zekere afbakening in gekozen; een domein van beschouwing. Naar objecten in de werkelijkheid refereren we door bepaalde termen te gebruiken. Objecten met soortgelijke kenmerken maken deel uit van hetzelfde conceptuele objecttype. De conceptuele objecttypen worden geconcretiseerd naar logische objecttypen die de representatie van gegevens beschrijven. Objecttypen zijn voorzien van attributen; eigenschappen van (gegevens)objecten. Gegevens over een object worden geregistreerd in de vorm van een gegevensobject in een dataset. In deze gegevensobjecten zijn attribuutwaarden vastgelegd voor attributen die zijn gedefinieerd in het objecttype.



Werkelijkheid
- een beeld van de echte of hypothetische wereld
- de echte wereld kunnen wij dus niet direct zien

Domein van beschouwing
- een afbakening van de werkelijkheid waarnaar wordt gekeken

Object
- een fenomeen in de werkelijkheid
- dat deel wat wij ervan waarnemen

Term
- woord dat we gebruiken om te refereren aan objecten in de werkelijkheid
- betekent verschillende dingen voor verschillende mensen en in verschillende contexten

Conceptueel objecttype
- een type van gelijksoortige objecten
- de naam is een term die gebruikelijk is voor dergelijke objecten

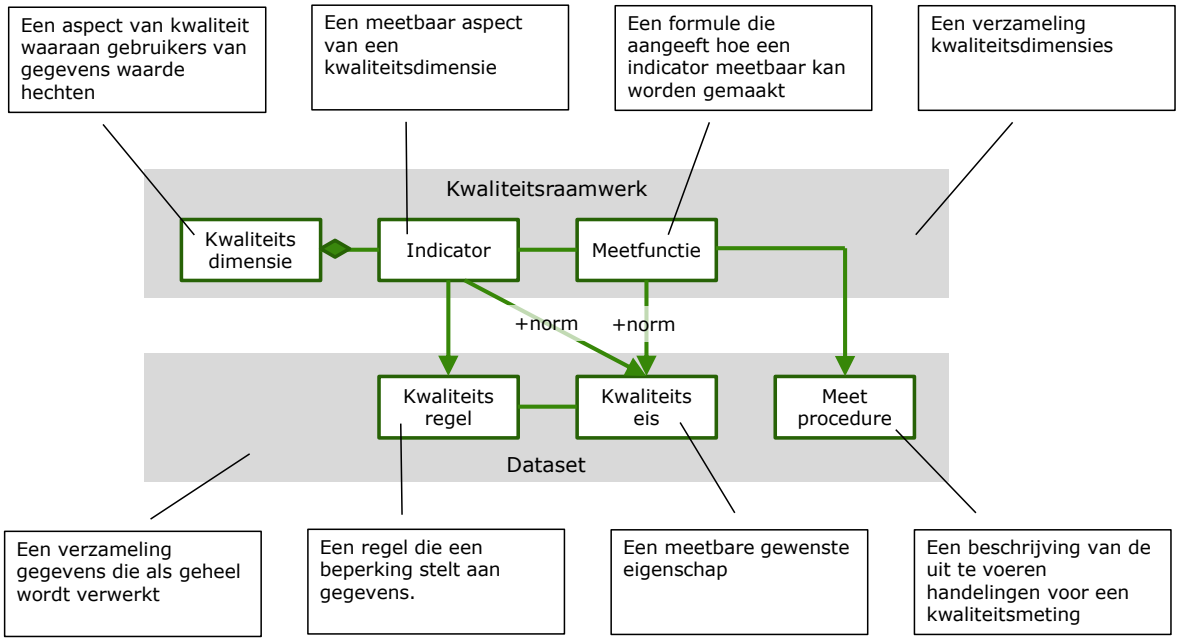
Logisch objecttype
- een type van gelijksoortige gegevensobjecten
- gerelateerd aan één of meer conceptuele objecttypen

Gegevensobject
- een op zichzelf staand geheel van gegevens met een eigen identiteit
- behoort bij een logisch objecttype en daarmee ook bij een conceptueel objecttype
- is gerelateerd aan object in de werkelijkheid

Dataset
- een verzameling van gegevens die als geheel wordt verwerkt

Figuur 10 Overzicht van basisconcepten

Het voorgestelde kwaliteitsraamwerk is gestructureerd in kwaliteitsdimensies, indicatoren en meetfuncties. Deze worden gebruikt om voor een specifieke dataset te komen tot kwaliteitseisen, kwaliteitsregels en meetprocedures (zie Figuur 11).



Figuur 11 Kwaliteitsraamwerk en relatie met specifieke dataset

Een kwaliteitsdimensie is een aspect van kwaliteit waaraan gebruikers van gegevens waarde hechten. Een indicator is een meetbaar aspect van een specifieke kwaliteitsdimensie. Indicatoren zijn de basis voor kwaliteitsrapportages. Bij een indicator horen één of meer meetfuncties die de indicator meetbaar maken door deze te concretiseren vanuit een bepaald perspectief. Een meetfunctie kan zowel een absoluut (bijvoorbeeld aantal objecten dat mag afwijken) of relatief zijn (zoals het percentage van de objecten dat voldoet). Voor een specifieke dataset kan het nodig zijn extra indicatoren of meetfuncties toe te voegen als de standaard indicatoren en meetfuncties in het kwaliteitsraamwerk niet voldoende zijn. Er kunnen dan ook meerdere meetfuncties ontstaan voor een specifieke indicator. Een kwaliteitseis is een meetbare gewenste eigenschap en geeft ook informatie over de norm die worden gehanteerd. Een indicator is daarmee een soort sjabloon die door verdere invulling leidt tot een kwaliteitseis. Een kwaliteitseis kan ook direct worden afgeleid van een indicator. Tabel 3 illustreert de begrippen aan de hand van een voorbeeld.

Kwaliteitsdimensie	Nauwkeurigheid
Indicator	Thematische nauwkeurigheid: de mate waarin gegevens die geen betrekking hebben op locatie, tijd of kwantiteit overeenkomen met de werkelijkheid.
Meetfunctie	voor een specifiek attribuut, het percentage van de gegevensobjecten van het corresponderende objecttype waarvan de waarde van dit attribuut overeenkomt met de eigenschappen van de corresponderende objecten in de werkelijkheid
Kwaliteitseis	95% correctheid

Tabel 3 Voorbeelduitwerking van kwaliteitseis

Naast kwaliteitseisen zullen er kwaliteitsregels moeten worden gedefinieerd. Een kwaliteitsregel is een regel die een beperking stelt aan gegevens. Deze kwaliteitsregels kunnen (deels) zijn opgenomen in een informatiemodel en ook wel "beperkingen" genoemd. Ze zijn ook impliciet aanwezig in een gegevensmodel in de vorm van de relaties en cardinaliteiten. Andere kwaliteitsregels zullen separaat moeten worden gedefinieerd. Dat kan in vrije vorm (natuurlijke taal) of in een meer strikte vorm (formele taal). Een specifieke vorm van kwaliteitsregel is een signaalregel. Als gegevens niet voldoen aan een signaalregel is dat een signaal dat ze mogelijk niet correct zijn omdat ze een onwaarschijnlijke waarde hebben. Er kan in kwaliteitseisen verwezen worden naar kwaliteitsregels.

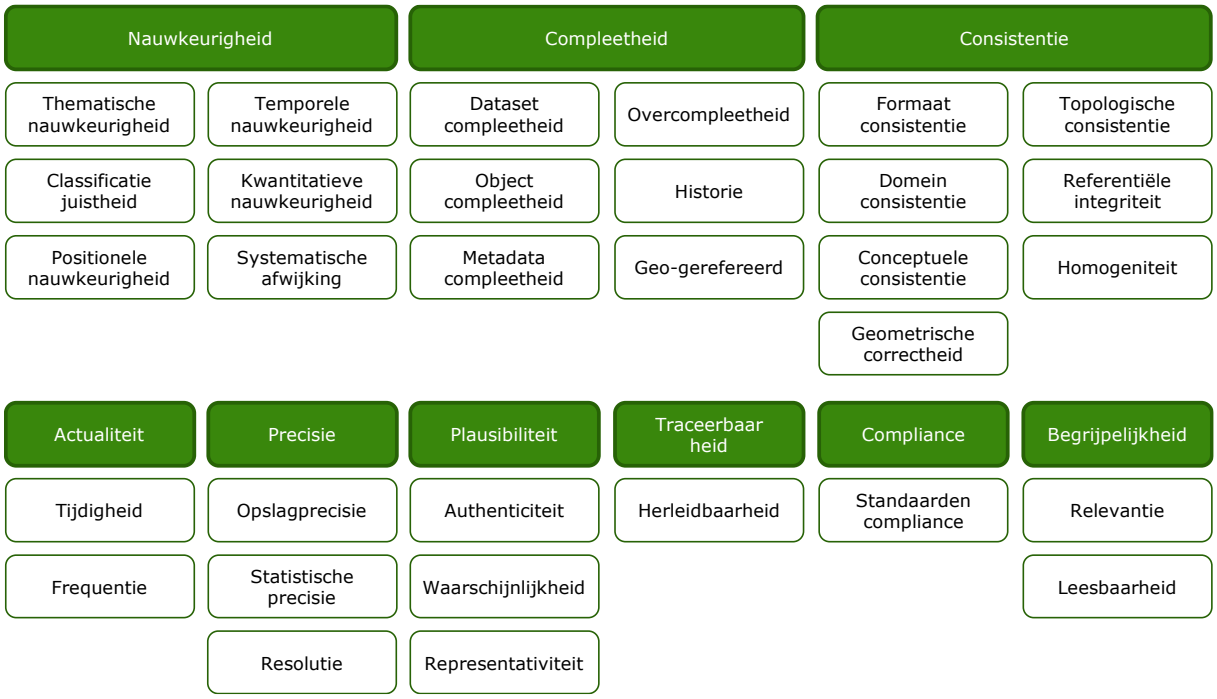
Voor een specifieke dataset zal ook moeten worden bepaald hoe kwaliteit precies zal worden gemeten: de meetprocedures. Deze meetprocedures vertalen de meetfuncties naar specifieke handelingen die uitgevoerd dienen te worden om te bepalen of aan de kwaliteitseisen wordt voldaan. Hierin wordt bijvoorbeeld beschreven of gebruik wordt gemaakt van steekproeven of dat er naar een gehele dataset wordt gekeken en of er een menselijke inspectie wordt uitgevoerd of dat er een geautomatiseerde controle plaatsvindt. Een kwaliteitsmeting kan worden gezien als een vorm van testen. Vanuit dat perspectief kun je onderscheid maken tussen logische en fysieke testgevallen. De eerste beschrijven op functioneel niveau wat er getest moet worden. De tweede zijn de concrete testscripts die gebruikt worden bij het uitvoeren van een test.

5.2 Kwaliteitsraamwerk

Op basis van bovenstaande concepten is een kwaliteitsraamwerk opgebouwd. De volgende uitgangspunten zijn leidend geweest bij het opstellen van dat kwaliteitsraamwerk:

- *ISO/IEC 25012, ISO/IEC 25024, NEN-ISO 19157 en ISO 5725 als basis*: De kwaliteitsdimensies zoals voorgesteld door ISO/IEC 25012 worden als hoofdstructuur gebruikt omdat dit is een internationaal erkende standaard is die breed genoeg is om als basis te dienen. Daarbinnen worden de indicatoren gebaseerd op ISO/IEC 25024, NEN-ISO 19157 en ISO 5725 omdat in deze standaarden de indicatoren verder zijn uitgewerkt. Dit wordt aangevuld met indicatoren uit andere raamwerken.
- *Relevant voor Omgevingswet context*: Dit document is primair gericht op ondersteuning van de Omgevingswet. In de selectie van kwaliteitsdimensies en indicatoren is er daarom vooral gekeken wat er vanuit die context relevant is. Daarom is er in het raamwerk ook expliciete aandacht voor geo-informatie en statistische informatie.
- *Meetbaarheid*: Het is belangrijk om kwaliteit te kunnen meten en bij voorkeur ook op een zo objectief mogelijke manier. Meetbaarheid is dan ook een belangrijk criterium geweest om te bepalen welke indicatoren zijn opgenomen in het raamwerk.
- *Kwaliteit van gegevens zelf*: Alleen de kwaliteitsdimensies en indicatoren die impact hebben op de gegevens zelf zijn uitgewerkt. Systeemafhankelijke kwaliteitseisen zullen in het DSO vooral gesteld worden aan gebruikerstoepassingen. Eisen aan de organisatorische inrichting worden in het DSO vooral juridisch en bestuurlijk geborgd.

De resulterende kwaliteitsdimensies en indicatoren zijn weergegeven in Figuur 12. De relatie tussen deze indicatoren en de genoemde ISO standaarden is beschreven in bijlage E. In het vervolg van dit hoofdstuk worden de indicatoren verder beschreven.



Figuur 12 Kwaliteitsraamwerk

5.3 Indicatoren

In deze paragraaf wordt het raamwerk uitgewerkt door de genoemde indicatoren te beschrijven. Informatiehuizen kunnen de indicatoren aanvullen met extra indicatoren als dat voor hen relevant is. Indicatoren die zijn voorzien van het symbool "✓" kunnen grotendeels geautomatiseerd worden getoetst.

De indicatoren in de kwaliteitsdimensie "nauwkeurigheid" hebben betrekking op de werkelijkheid (wat ook de juridische werkelijkheid). Om te bepalen of een gegeven hieraan voldoet wordt idealiter ook echt gekeken naar de werkelijke objecten waarover gegevens zijn vastgelegd. In de praktijk kan dat vaak alleen door dit middels steekproeven te doen, omdat het te veel inspanning kost naar alle objecten in de werkelijkheid te kijken. Als alternatief kan ook een vergelijking met een andere registratie worden gemaakt om (mogelijke) afwijkingen van de werkelijkheid inzichtelijk te maken. Een dergelijke andere registratie wordt ook wel een referentiedataset genoemd [16]. Voordeel van het gebruik van een referentiedataset is hierdoor een geautomatiseerde kwaliteitsmeting kan plaatsvinden. Het nadeel is dat de betrouwbaarheid van de vergelijking lager is dan een inspectie van de werkelijkheid.

Nauwkeurigheid

De mate waarin gegevens de echte waarde in een specifieke gebruikscontext goed weergeven.

Indicator	**Beschrijving**
Thematische nauwkeurigheid	De mate waarin gegevens die geen betrekking hebben op locatie, tijd of kwantiteit overeenkomen met de werkelijkheid.
Classificatie juistheid	De mate waarin gegevensobjecten zijn geïdentificeerd als het juiste logisch objecttype.
Positionele nauwkeurigheid	De mate waarin locatiegegevens overeenkomen met de werkelijkheid.
Temporele nauwkeurigheid	De mate waarin tijdgegevens overeenkomen met de werkelijkheid.
Kwantitatieve nauwkeurigheid	De mate waarin kwantitatieve gegevens overeenkomen met de werkelijkheid.
Systematische afwijking	De mate waarin de resultaten van metingen of berekeningen overeenkomen met de werkelijkheid.

Compleetheid

De mate waarin gegevens aanwezig zijn in een specifieke gebruikscontext.

Indicator	**Beschrijving**
Dataset compleetheid	De mate waarin objecten waarvan het bestaan bekend is aanwezig zijn.
Object compleetheid	De mate waarin attribuutwaarden bij gegevensobjecten aanwezig zijn. ✓
Metadata compleetheid	De mate waarin metadata aanwezig is. ✓
Overcompleet-heid	De mate waarin gegevensobjecten onterecht aanwezig zijn. ✓
Historie	De mate waarin historische gegevens aanwezig zijn. ✓
Geo-gerefereerd	De mate waarin gegevens zijn voorzien van een geo-referentie. ✓

Consistentie

De mate waarin gegevens niet in tegenspraak zijn met andere gegevens in een specifieke gebruikscontext.

Indicator	**Beschrijving**

Formaat-consistentie	De mate waarin gegevens syntactisch correct zijn en daarmee geautomatiseerd te verwerken.	✓
Domein-consistentie	De mate waarin de inhoud en de lengte van attribuutwaarden consistent zijn met hun waardedomein.	✓
Conceptuele consistentie	De mate waarin de combinaties van attribuutwaarden logisch samenhangend zijn.	✓
Geometrische correctheid	De mate waarin gegevens geometrisch correct zijn.	✓
Topologische consistentie	De mate waarin gegevens topologisch consistent zijn.	✓
Referentiële integriteit	De mate waarin gegevens verwijzen naar bestaande gegevens.	✓
Homogeniteit	De mate van variatie in een dataset.	

Actualiteit

De mate waarin gegevens recent genoeg zijn in een gebruikscontext.

Indicator	Beschrijving	
Tijdigheid	De mate waarin gegevens tijdig zijn geactualiseerd (gemeten of gecontroleerd of ze nog correct zijn).	✓
Frequentie	De mate waarin gegevens met een afgesproken frequentie zijn geregistreerd.	✓

Precisie

De mate waarin gegevens exact of onderscheidend genoeg zijn voor een gebruikscontext.

Indicator	Beschrijving	
Opslagprecisie	De mate van detail waarmee gegevens zijn geregistreerd.	✓
Statistische precisie	De mate waarin metingen of berekeningen bij herhaling dezelfde waarde opleveren.	
Resolutie	De mate van detail waarmee gegevens worden ingewonnen.	

Plausibiliteit

De mate waarin gegevens worden beschouwd als waar en geloofwaardig door gebruikers in een specifieke gebruikscontext.

Indicator	Beschrijving
Authenticiteit	De mate waarin de authenticiteit van gegevens aantoonbaar is.

Waarschijnlijk-heid	De mate waarin gegevens waarschijnlijk zijn voor de situatie. ✓
Representati-viteit	De mate waarin een dataset een goede weergave geeft van het geheel.

Traceerbaarheid

De mate waarin toegang tot gegevens of wijzigingen erin vastgelegd worden in een audit trail in een specifieke gebruikscontext.

Indicator	Beschrijving
Herleidbaarheid	De mate waarin de herkomst, selecties en bewerkingen die hebben plaatsgevonden op de gegevens expliciet zijn vastgelegd. ✓

Compliance

De mate waarin gegevens conformeren aan standaarden, conventies of regelgeving gerelateerd aan gegevenskwaliteit in een specifieke gebruikscontext.

Indicator	Beschrijving
Standaarden compliance	De mate waarin gegevens conformeren aan afgesproken standaarden. ✓

Begrijpelijkheid

De mate waarin gegevens eenvoudig gelezen en geïnterpreteerd kunnen worden door gebruikers, en zijn verwoordt in geschikte talen, symbolen en eenheden in een specifieke gebruikscontext.

Naam	Beschrijving
Relevantie	De mate waarin de gegevens relevant zijn voor de gebruikscontext of daartoe beperkt kunnen worden.
Leesbaarheid	De mate waarin teksten voor de doelgroep begrijpelijk zijn geformuleerd.

5.4 Meetfuncties

In deze paragraaf worden voorbeelden gegeven van meetfuncties voor de in de vorige paragraaf beschreven indicatoren (zie Tabel 4). Deze meetfuncties zijn met name geschikt voor communicatie richting management, met name doordat ze zijn uitgedrukt in percentages. Voor meer specifieke doelgroepen kan het wenselijk zijn om vooral concrete aantallen te meten en inzichtelijk te maken. Daarnaast belichten de voorbeeld meetfuncties slechts een aspect van de indicator.

Indicator	Meetfunctie
Nauwkeurigheid	
Thematische nauwkeurigheid	voor een specifiek attribuut, het percentage van de gegevensobjecten van het corresponderende objecttype waarvan de waarde van dit attribuut overeenkomt met de eigenschappen van de corresponderende objecten in de werkelijkheid
Classificatie juistheid	percentage van de gegevensobjecten dat een objecttype heeft dat overeenkomt met het objecttype van het corresponderende object in de werkelijkheid
Positionele nauwkeurigheid	voor een specifiek objecttype, het percentage van de gegevensobjecten van het corresponderende objecttype waarvoor geldt dat hun geometrie binnen de afgesproken marge overeenkomt met de vorm en ligging van het corresponderende object in de werkelijkheid
Temporele nauwkeurigheid	voor een specifiek attribuut, het percentage van de gegevensobjecten van het corresponderende objecttype waarvoor geldt dat het verschil tussen de geregistreerde waarde en de waarde in de werkelijkheid binnen de afgesproken marge ligt
Kwantitatieve nauwkeurigheid	voor een specifiek attribuut, het percentage van de gegevensobjecten van het corresponderende objecttype waarvoor geldt dat het verschil tussen de geregistreerde waarde en de waarde in de werkelijkheid binnen de afgesproken marge ligt
Systematische afwijking	het verschil in procenten tussen het gemiddelde van resultaten van metingen of berekeningen en de waarde in de werkelijkheid
Compleetheid	
Dataset compleetheid	van de objecten waarvan het bestaan bekend is, het percentage waarover gegevensobjecten zijn opgenomen in de registratie
Object compleetheid	percentage van de gegevensobjecten waarin alle in het informatiemodel als verplicht gemarkeerde attributen zijn voorzien van een attribuutwaarde
Metadata compleetheid	percentage van de metadata attributen die zijn gevuld met een attribuutwaarde
Overcompleet-heid	percentage van de gegevensobjecten die onterecht zijn opgenomen omdat zij: • niet passen bij de overeengekomen scope van de dataset • conform het informatiemodel overlappen met een ander gegevensobject (duplicaten) • onterecht zijn geïdentificeerd als behorend bij een specifiek objecttype (classificatie correctheid)
Historie	percentage van de objecttypes waarvan voor alle momenten van geldigheid van de registratie gegevensobjecten aanwezig zijn

Geo-gerefereerd	van de gegevensobjecten die gegevens over geo-objecten vastleggen, het percentage van de gegevensobjecten dat is voorzien van een georeferentie
Consistentie	
Formaat-consistentie	booleaanse waarde[2] die aangeeft of de dataset voldoet aan het afgesproken fysieke formaat
Domein-consistentie	percentage van de gegevensobjecten waarvan de inhoud en lengte van attribuutwaarden in overeenstemming zijn met de in het informatiemodel gedefinieerde waardedomein
Conceptuele consistentie	percentage van de gegevensobjecten waarvan attribuutwaarden in overeenstemming zijn met de kwaliteitsregels die betrekking hebben op de relatie tussen attribuutwaarden
Geometrische correctheid	percentage van de gegevensobjecten die geometrische attribuutwaarden bevatten en waarvan de attribuutwaarde geometrisch correct is (geen incorrecte ringvolgorde, dubbele punten, juiste aantal punten, onnodig korte lijnen)
Topologische consistentie	percentage van de gegevensobjecten die geometrische attribuutwaarden bevatten en waarvan de attribuutwaarde topologisch consistent is conform NEN-ISO 19157[3]
Referentiële integriteit	percentage van de gegevensobjecten waarvan attribuutwaarden verwijzen naar objecten die zijn geregistreerd in dezelfde of in een andere dataset
Homogeniteit	voor een specifiek attribuut, de afgesproken mate van variatie die de verzameling attribuutwaarden mag vertonen
Actualiteit	
Tijdigheid	percentage van de gegevensobjecten waarvan het verschil tussen nu en het actualiseringsmoment binnen de afgeproken maximale termijn valt
Frequentie	percentage van de gegevensobjecten die conform een afgesproken frequentie zijn geregistreerd
Precisie	
Opslagprecisie	voor een specifiek attribuut, het percentage van de gegevensobjecten van het corresponderende objecttype waarvan de waarde van dit attribuut met het afgesproken detailniveau is geregistreerd
Statistische precisie	voor een specifiek attribuut, de standaardafwijking van de verzameling resultaten van metingen of berekeningen
Resolutie	voor een specifiek attribuut, een getal en eenheid die uitdrukken hoeveel onzekerheid waarden hebben
Plausibiliteit	

[2] Een waarde die aangeeft of iets waar is of niet ("ja" of "nee")
[3] Zie paragraaf D.3.4 in NEN-ISO 19157

Authenticiteit	percentage van de gegevensobjecten waarvan hun bestaan kan worden aangetoond door een formele verklaring van een betrouwbare organisatie
Waarschijnlijk-heid	percentage van de gegevensobjecten waarvan attribuutwaarden in overeenstemming zijn met de signaalregels
Representati-viteit	de statistische variantie van de waarden in een een dataset t.o.v. de waarden de andere datasets behorende bij het geheel
Traceerbaarheid	
Herleidbaarheid	percentage van de gegevensobjecten die zijn voorzien van een audit trail
Compliance	
Standaarden compliance	percentage van de gegevensobjecten die zijn vastgelegd conform afgesproken standaarden
Begrijpelijkheid	
Relevantie	voor een specificiek attribuut met daarin tekst, percentage van de gegevensobjecten waarin de waarde van dit attribuut kan worden ontleed tot kleinere, los identificeerbare elementen
Leesbaarheid	voor een specifiek attribuut met daarin tekst, percentage van de gegevensobjecten waarin de waarde van dit attribuut taalkundig correct zijn geformuleerd en alleen woorden gebruiken die bekend zijn bij de doelgroep of expliciet zijn gedefinieerd

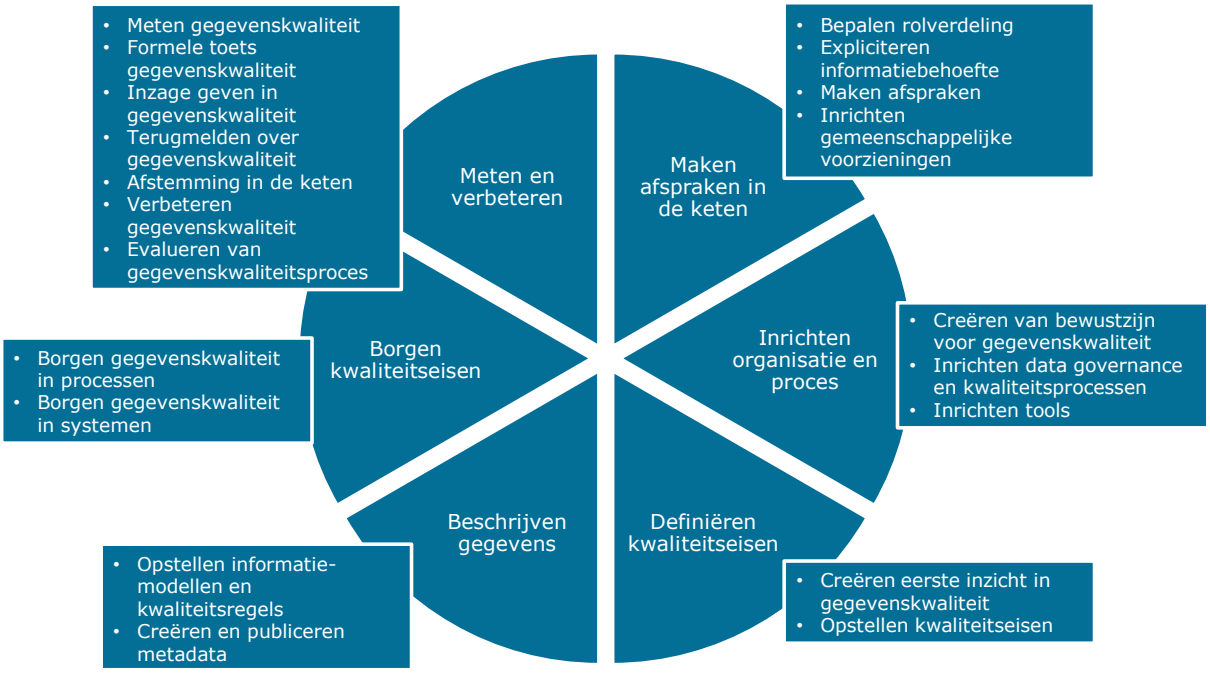
Tabel 4 Voorbeeld meetfuncties voor indicatoren

6 Van visie naar proces

Dit hoofdstuk geeft een handreiking voor het procesmatig borgen van gegevenskwaliteit. Uitgangspunt daarbij is het integrale beeld van gegevenskwaliteit zoals beschreven in de visie. Dit hoofdstuk volgt dan ook de daarin beschreven hoofdstructuur. Binnen deze structuur worden meer concrete activiteiten benoemd, die relevant zijn voor verschillende partijen in de keten. Een ieder zal dan ook zijn eigen vertaalslag moeten maken van deze algemene activiteiten naar de eigen rol in de keten en organisatiecontext.

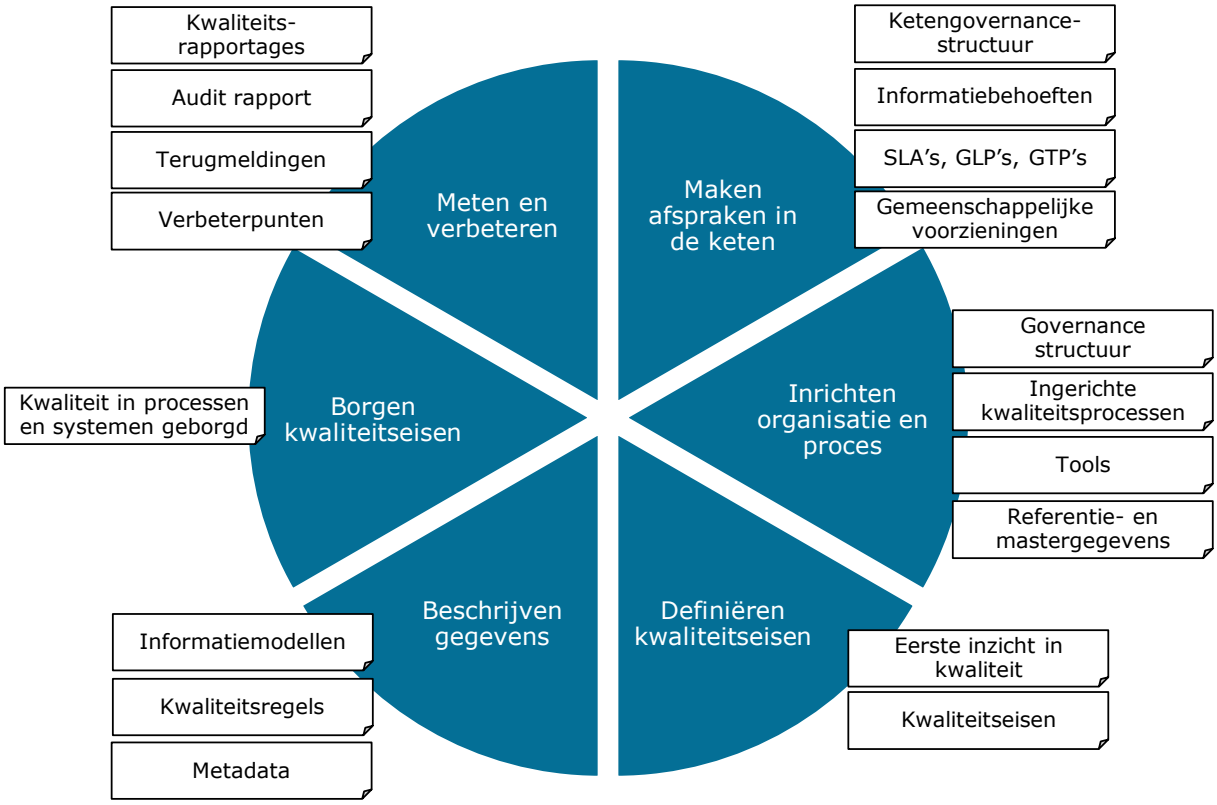
6.1 Overzicht

In hoofdstuk 4 is een visie beschreven met daarin een integraal beeld van de aspecten die belicht dienen te worden m.b.t. gegevenskwaliteit. Deze aspecten zijn fasen die doorlopen kunnen worden. Bij elkaar vormen ze een kwaliteitscyclus waarmee de kwaliteit van gegevens continu kan worden verbeterd. Figuur 13 geeft een overzicht van de fasen en de activiteiten binnen deze fasen. Deze fasen en activiteiten hoeven niet strikt lineair te worden doorlopen, maar geven vooral een logische volgordelijkheid aan. Het is een cyclus waarbinnen iteraties worden doorlopen en waarin het ook mogelijk is om terug te gaan naar voorgaande activiteiten. In paragraaf 6.2 en verder zijn de activiteiten beschreven.



Figuur 13 Overzicht van fasen en activiteiten

Activiteiten in al deze fasen leveren ook producten op. Dat zijn typisch documenten, maar het kunnen ook modellen zijn (zoals een informatiemodel) of andersoortige structureerde gegevens (zoals metadata). Figuur 14 geeft een overzicht van de producten van de verschillende fasen.



Figuur 14 Fasen en producten

6.2 Maken afspraken in de keten

Bepalen rolverdeling – Er is in eerste instantie een goed inzicht nodig in wie welke verantwoordelijkheden heeft in de keten. In de context van de Omgevingswet zal een deel hiervan worden vastgelegd in regelgeving zoals ministeriële regelingen. Daarin is in ieder geval het onderscheid tussen bronhouders, intermediairs (zoals landelijke voorzieningen), leveranciers van informatieproducten en afnemers (incl. DSO) relevant. Sommige organisaties kunnen meerdere rollen hebben, maar het is goed deze expliciet te onderscheiden om zuivere afspraken te kunnen maken. Het kan zijn dat nieuwe datasets moeten worden gecreëerd of dat bronhouders nog onvoldoende bewust zijn van hun rol.

Expliciteren informatiebehoefte – De basis voor de inrichting van gegevenskwaliteit in de keten zijn de informatiebehoeften die moeten worden ingevuld. Deze komen deels direct uit de wetgeving en deels vanuit afnemers met specifieke behoeften. Een grondige analyse van de wetgeving is noodzakelijk om de informatiebehoeften in de eerste categorie helder te krijgen. Informatiebehoeften in de tweede categorie vraagt met name afstemming met gebruikers van de gegevens en informatie. Dat zijn alle gebruikers van informatie in de keten. Naast het helder krijgen van de informatiebehoefte (vraagarticulatie), zullen soortgelijke informatiebehoeften ook moeten worden gebundeld. De informatiebehoefte zal niet in één keer helder zijn; er zal in iteraties worden gewerkt, waarbij het zicht op de gewenste informatieproducten steeds duidelijker wordt.

Maken afspraken – Afspraken over gegevenskwaliteit zullen moeten worden geborgd in de gehele keten, zowel op bestuurlijk niveau als op uitvoerend niveau. De belangrijkste verantwoordelijkheden van informatiehuizen en bronhouders zullen wettelijk worden geborgd. De aansluitvoorwaarden voor informatieproducten beschrijven de eisen die in het algemeen worden gesteld aan informatieproducten. Dat gaat ondermeer over het beschikbaar stellen van informatiemodellen en andere metadata. Deze metadata zal de kwaliteit van de gegevens zichtbaar maken voor alle partijen in de keten. Er zullen ook dienstenniveauovereenkomsten worden opgesteld voor de levering van informatieproducten. Deze afspraken zullen worden doorvertaald naar afspraken met bronhouders. zoals gegevensleveringsprotocollen en gegevenstransformatie-protocollen.

Inrichten gemeenschappelijke voorzieningen – Er zal bepaald moeten worden welke gemeenschappelijke voorzieningen gegevenskwaliteit in de keten kunnen ondersteunen. Denk bijvoorbeeld aan een terugmeldvoorziening. De voorzieningen zullen moeten worden (door)ontwikkeld en tijdig beschikbaar worden gesteld aan relevante partijen in de keten. Daarnaast zullen reeds bekende voorzieningen expliciet moeten worden ingericht voor gegevenskwaliteit. Denk onder meer aan het borgen van gegevenskwaliteit in de catalogus voor de omgevingswet, zodat kwaliteitsdimensies daar ook in kunnen worden vastgelegd.

6.3 Inrichten organisatie en proces

Creëren van bewustzijn voor gegevenskwaliteit – Gegevenskwaliteit begint met bewustzijn bij betrokkenen van het belang ervan. Daarnaast zal inzichtelijk moeten worden gemaakt wat het allemaal betekent. Dit document draagt bij aan bewustwording in de keten, maar kan ook gebruikt worden voor bewustwording bij individuele organisaties. Mogelijk moeten er in de individuele organisaties nog business cases worden opgesteld om ervoor te zorgen dat er voldoende middelen beschikbaar zijn voor het borgen van gegevenskwaliteit.

Inrichten data governance en kwaliteitsprocessen – Organisaties zullen data governance moeten inrichten waarin gegevenskwaliteit een prominente rol krijgt. Naast gegevenskwaliteit gaat data governance ook over zaken zoals rollen, taken en verantwoordelijkheden m.b.t. (geoorloofde) toepassing, misbruik, beschikbaar stellen en aansprakelijkheid. Idealiter is het ook geborgd in de vorm van een managementsysteem. Er zullen mensen verantwoordelijk moeten worden gemaakt binnen de organisatie voor het borgen van gegevenskwaliteit. Naast een eindverantwoordelijke zullen ook uitvoerende rollen moeten worden toegewezen. Voor alle rollen zal duidelijk moeten zijn wat hun taken en verantwoordelijkheden zijn en hoe datakwaliteit worden geborgd in processen. In de referentie-architectuur voor de informatiehuizen [17] zijn hiervoor de rollen kwaliteitsmanager en functioneel gegevensbeheerder beschreven. In het GEMMA katern tactisch gegevensmanagement [37] zijn ook rollen benoemd die in meer algemene zin bestaan. Mogelijk moeten extra medewerkers worden aangesteld en opgeleid. Om te kunnen sturen op gegevenskwaliteit is de inrichting van monitoring- en rapportageprocessen essentieel.

Inrichten tools – Mensen en organisaties moeten worden gefaciliteerd in het kwaliteitsmanagementproces (zie ook hoofdstuk 7). Voor een deel kan gebruik worden gemaakt van gemeenschappelijke voorzieningen die beschikbaar worden

gesteld vanuit het DSO. Voor een ander deel zullen organisaties aanvullende tools nodig hebben. Met name tools die inzicht geven in gegevenskwaliteit zijn erg waardevol. Dat kunnen data profiling tools zijn, maar bijvoorbeeld ook tools die vergelijkingen kunnen maken tussen datasets en verschillen inzichtelijk kunnen maken. Ook het inrichten van referentie- en master data management vraagt expliciete aandacht. Dit gaat over het expliciet inrichten van het beheer van relatief stabiele gegevens (ook wel: stamgegevens) waarnaar andere gegevens verwijzen. In deze categorie vallen ook de gegevens in de landelijke basisregistraties. Beschikbaarheid van dit soort gegevens is relevant voor het borgen van consistentie, één van de kwaliteitsdimensies. Hiertoe kan het nodig zijn deze gegevens te verspreiden naar applicaties die ervan gebruik maken (ook wel: datadistributie).

6.4 Definiëren kwaliteitseisen

Creëren eerste inzicht in kwaliteit – Voordat kwaliteitseisen kunnen worden gedefinieerd is het noodzakelijk om al een eerste inzicht te hebben in de huidige kwaliteit van brongegevens en de belangrijkste issues daarin. Hierdoor ontstaat een baseline voor verbetering en helder is wat een reëel kwaliteitsniveau is.

Opstellen kwaliteitseisen – Informatiehuizen zullen voor alle informatieproducten die zij gaan ontwikkelen in dialoog met bronhouders en afnemers tot kwaliteitseisen moeten komen. Dat start op het niveau van informatieproduct en vertaalt door naar kwaliteitseisen die worden gesteld aan brongegevens.

6.5 Beschrijven gegevens

Opstellen informatiemodellen en kwaliteitsregels – Informatiemodellen en kwaliteitsregels spelen een belangrijke rol bij het concreet maken van gegevenskwaliteit. Bij elk informatieproduct en bij elke dataset zouden informatiemodellen beschikbaar moeten zijn. Idealiter zijn dat zowel conceptuele informatiemodellen die de werkelijkheid modelleren als logische modellen die de representatie ervan in gegevens modelleren.

Creëren en publiceren metadata – Datasets en informatieproducten moeten worden voorzien van relevante metadata. Deze metadata moet worden gepubliceerd in de relevante catalogi. Naast de catalogus omgevingswet kan dit ook een interne catalogus van een bronhouder of informatiehuis zijn, maar ook andere catalogi zoals data.overheid.nl en het Nationaal Geo Register. Onderdeel van deze metadata is het informatiemodel, maar bijvoorbeeld ook relevante aspecten van gegevenskwaliteit.

6.6 Borgen kwaliteitseisen

Borgen gegevenskwaliteit in processen – Gegevenskwaliteit stelt eisen aan de inrichting van allerlei processen (naast het kwaliteitsmanagementproces), met name aan de kant van bronhouders. Zo zal een specifieke kwaliteitseis m.b.t. nauwkeurigheid impact hebben op de wijze waarop gegevens worden ingewonnen. Daarnaast kunnen er bijvoorbeeld extra controles in een proces moeten worden ingericht om voldoende zekerheid te creëren over de kwaliteit van gegevens.

Borgen gegevenskwaliteit in systemen - Gegevenskwaliteit zal zo vroeg mogelijk in het proces door systemen moeten worden geborgd. Denk bijvoorbeeld aan invoercontroles of validaties bij het uitwisselen van gegevens. Dit stelt ook eisen aan het metadateren van gegevens die worden uitgewisseld. Denk bijvoorbeeld aan metadata die aangeeft wanneer de gegevens zijn ingewonnen of metadata die kan worden gebruikt om te bepalen of de uitgewisselde gegevens volledig zijn (alle geregistreerde gegevensobjecten bevatten). Deze metadata heeft een directe relatie met gegevenskwaliteit. Voor elk koppelvlak zal het genereren van dit soort metadata moeten worden geborgd aan de kant van de zender en het controleren en gebruiken van dit soort metadata door de ontvanger. Kwaliteitseisen en regels zullen ook geconfigureerd moeten worden in tools die gegevenskwaliteit meten.

6.7 Meten en verbeteren

Meten gegevenskwaliteit – Het is wenselijk de kwaliteit van datasets in kaart te hebben, gebaseerd op de opgestelde informatiemodellen en kwaliteitsregels. Dat creëert zicht op de verbeteracties die dienen plaats te vinden. Een groot deel van de kwaliteitsmeting ligt bij bronhouders, maar zij kunnen hierin gefaciliteerd worden door informatiehuizen.

Formele toets gegevenskwaliteit – Er kunnen redenen zijn om een formele toets of audit uit te voeren van de gegevenskwaliteit van een specifieke bronhouder. Hierbij kunnen bijvoorbeeld steekproeven worden genomen om te bepalen in hoeverre de gegevens overeenkomen met de werkelijkheid.

Inzage geven in gegevenskwaliteit – Het resultaat van het meten van gegevenskwaliteit kan gedeeld worden met allerlei partijen in de keten. Dat kan bijvoorbeeld door het beschikbaar te stellen van een kwaliteitsdashboard, kwaliteitsrapportages of audit rapporten. Dit creëert transparantie en daarmee vertrouwen. Afnemers kunnen dan zelf inschatten in welke mate de kwaliteit van de gegevens aansluit bij hun gebruik. Het is ook een trigger om terug te melden om kwaliteit te verbeteren of met elkaar in gesprek te treden in de keten.

Terugmelden over gegevenskwaliteit – Terugmelden is een belangrijke feedback loop van afnemers naar leveranciers van gegevens (bronhouders). Terugmelden op de landelijke basisregistraties is zelfs een verplichting[4]. Het actief terugmelden van alle vorm van twijfel bij de kwaliteit van een gegeven vanuit een specifiek gebruik helpt de leverancier in de verbetering van de kwaliteit. Terugmeldingen kunnen overal vandaan komen; van afnemers, maar ook van informatiehuizen of bronhouders zelf.

Informatiehuizen zullen terugmeldingen moeten beoordelen op de mate waarin zij gerelateerd zijn aan de kwaliteit van de brongegevens of de door het informatiehuis uitgevoerde transformatie. Bronhouders zullen terugmeldingen inhoudelijk moeten beoordelen en vertalen naar gewenste verbeteracties.

[4] zie ook: https://www.vngrealisatie.nl/secties/gegevens-en-berichtenstandaarden/terugmelden-op-basisregistraties

Afstemming in de keten – Niet alle kwaliteitsproblemen kunnen worden opgelost middels terugmeldingen. Sommige problemen vragen een dialoog tussen afnemer en leverancier of mogelijk zelfs breder in de keten. Het kan zijn dat er onvoldoende middelen aanwezig is om kwaliteit goed te borgen. In dat geval dienen extra middelen beschikbaar te worden gesteld.

Verbeteren gegevenskwaliteit – Op basis van inzicht in huidige en gewenste kwaliteit kan gericht gewerkt worden aan het verbeteren van de kwaliteit van gegevens. Daarbij is het belangrijk om de oorzaken die ten grondslag liggen aan de problemen met de huidige gegevenskwaliteit te vinden en te vertalen naar oplossingen. Dit kan bijvoorbeeld het schonen of samenvoegen van gegevens zijn. Daarnaast kan het aanpassingen op procesniveau vragen, bijvoorbeeld in het inwinningsproces.

Evalueren van gegevenskwaliteitsproces – Het kwaliteitsmanagementproces zelf kan mogelijk bijstelling vragen. Periodieke evaluatie ervan, inclusief de daaraan gerelateerde governance en de ondersteuning ervan door systemen is gewenst. Dit kan aanleiding zijn om terug te keren naar één van de eerder genoemde activiteiten en gewenste wijzigingen door te voeren.

7 Van visie naar informatievoorziening

Dit hoofdstuk vertaalt de eerder beschreven visie naar een schets van mogelijke ondersteuning door informatievoorziening. Het beschrijft vooral welke functionaliteiten daarbij relevant zijn.

7.1 Functionaliteiten

Data profiling

Data profiling gaat over het geven inzicht in gegevenskwaliteit, onder meer op basis van statistische analyse. Daarnaast gaat het over het inzichtelijk maken van relaties tussen gegevens. Een data profiling tool kan voor elke kolom in een tabel de frequentie van waarden bepalen waarmee inzicht ontstaat in het type en gebruik van elke kolom. Daarnaast kan het inzicht geven in een aantal kernkarakteristieken zoals minimum, maximum en gemiddelde waarde. Over kolommen heen kunnen afhankelijkheden tussen gegevens inzichtelijk worden gemaakt. Over tabellen heen kunnen overlappende waarden inzichtelijk worden gemaakt en daarmee waarschijnlijke verwijzingen (foreign keys). Over de tabellen heen kunnen ook vergelijkingen met waarnemingen met dezelfde periodes van voorgaande jaren onverklaarbare afwijkingen tonen, wat kan duiden op een gegevenskwaliteitsprobleem. Op deze manier kan data profiling inzicht geven in afwijkingen, waarbij veel tools ook dieper kunnen inzoomen op specifieke aandachtsgebieden in de data zelf. Data profiling tools kunnen ook gebruikt worden om zelfgedefinieerde kwaliteitsregels te controleren en afwijkende gegevensobjecten inzichtelijk te maken.

Standaardisatie

Tools in deze categorie zijn in staat om gegevens op een standaard manier op te maken. Ze werken op basis van patroonherkenning en en regels die detecteren of gegevens goed zijn of niet. Het gaat vooral over gegevens waarbij standaard patronen belangrijk zijn zoals telefoonnummers of bankrekeningnummers. Dit soort gegevens zouden op een standaard manier moeten zijn opgemaakt om verwarring te voorkomen. Mensen blijken erg gevoelig voor standaard patronen. Daarnaast helpt standaardisatie bij het analyseren, schonen en extraheren van waardes uit gegevens.

Transformatie

Gegevens die niet voldoen aan kwaliteitsregels of andersoortige afspraken kunnen in een aantal gevallen worden getransformeerd waardoor ze wel voldoen. Dit gaat verder dan standaardisatie; gegevens kunnen ook in een andere structuur worden gezet, worden gecombineerd, vertaald of op andere wijze worden getransformeerd. Feitelijk gaat het hier over tools voor data-integratie die ervoor zorgen dat gegevens in een andere vorm beschikbaar komen zodat ze voldoen aan wat de ontvanger verwacht.

Matching

Veel problemen met gegevenskwaliteit hebben te maken met redundantie; objecten die meervoudig zijn geregistreerd. Dit creëert allerlei inconsistenties waardoor gebruikers niet meer weten waar ze op kunnen vertrouwen. Een belangrijke functionaliteit is daarom het herkennen van dubbelingen en opschonen van dubbelingen. Dat gaat vaak verder dan een eenvoudige vergelijking en zal op basis van combinaties van attribuutwaarden, patronen en regels plaats vinden. Ook het vergelijken van gegevens over datasets heen leidt tot inzicht in verbetering. Naast redundantie worden ook referentiële integriteitsproblemen inzichtelijk.

Verrijking

Soms is het nodig om gegevens te verrijken met andere gegevens zodat ze beter bruikbaar zijn. Denk bijvoorbeeld aan het verrijken van gegevens met semantische informatie, zoals relaties naar een thesaurus, zodat ze beter kunnen worden gevonden. Ook het verrijken van gegevens met informatie over de bewerkingen die ze hebben ondergaan is wenselijk. Andere potentieel waardevolle verrijkingen zijn het toevoegen van tijdsinformatie, geografische informatie of andersoortige contextuele informatie.

Rapportage

Het is wenselijk om vanuit allerlei perspectieven rapportages of dashboards beschikbaar te hebben over gegevenskwaliteit. Denk daarbij aan rapportages over de mate waarin aan kwaliteitseisen wordt voldaan, maar ook aan rapportages over afwijkingen van specifieke kwaliteitsregels, terugmeldingen, en uitgevoerde bewerkingen op de gegevens. Daarnaast zullen er ook meer ad-hoc rapportagebehoeften zijn.

Gegevensvalidatie

Het controleren of gegevens voldoen aan kwaliteitseisen, standaarden en vormvereisten. Dit betreft zowel een validatie van brongegevens als een validatie van informatieproducten.

Terugmelding

Het terugmelden van gerede twijfel aan een gegeven. Dit kan zowel betrekking hebben op een brongegeven als op een gegeven die is ontstaan als gevolg van een transformatie die is uitgevoerd in een informatiehuis (informatieproduct).

Invoercontrole

Het controleren of gegevens die door een gebruiker worden ingevoerd voldoen aan de vereisten. Dit betreft minimaal een controle op syntax, formaat, verplichte velden en logische verbanden.

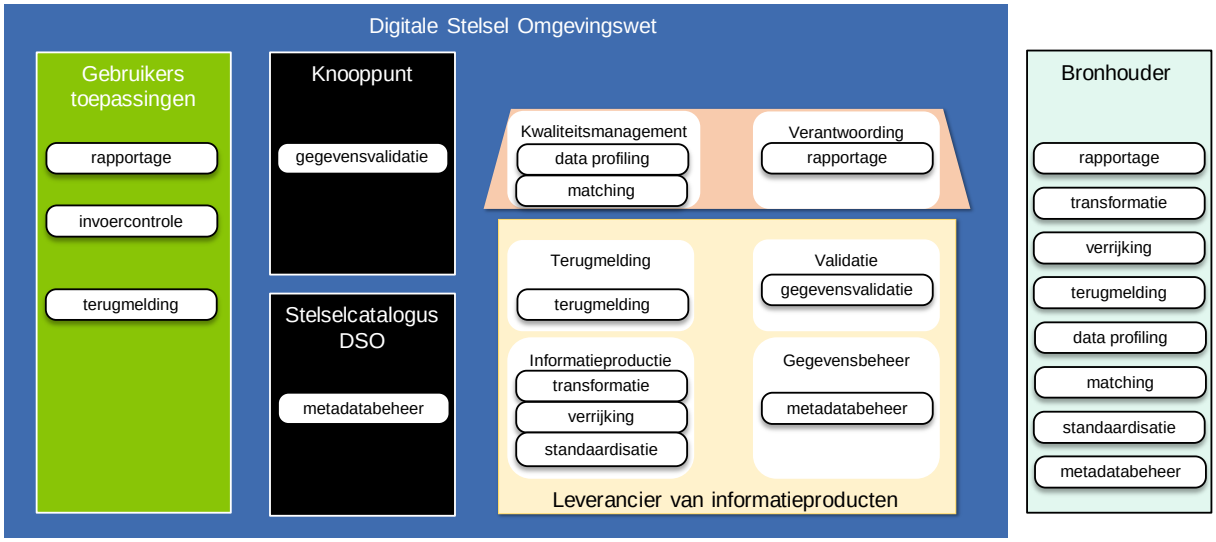
Metadatabeheer

gegevens over gegevenskwaliteit zijn een vorm van business metadata die expliciet moet worden vastgelegd en beheerd. Dit omvat zowel de gewenste kwaliteit als de daadwerkelijk gerealiseerde kwaliteit.

7.2 Functionaliteiten gepositioneerd in de Omgevingswet

Figuur 15 positioneert alle genoemde functionaliteiten in de context van de Omgevingswet. Functionaliteiten relevant voor informatiehuizen zijn geplot op de bedrijfsfuncties zoals benoemd in de referentie-architectuur voor de informatiehuizen. In de figuur is zichtbaar dat veel functionaliteiten zowel voor de bronhouder als voor informatiehuizen relevant zijn. Daarnaast zullen metadatabeheer, terugmelding en rapportage ondersteuning in de gehele keten vragen. Metadata voor informatieproducten zullen vindbaar moeten zijn in de Stelselcatalogus DSO. Terugmeldingen zullen via een gebruikerstoepassing moeten worden ingediend, door een informatiehuis worden beoordeeld en gerouteerd naar de verantwoordelijke bronhouder. Inzicht in gegevenskwaliteit is wenselijk op meerdere plaatsen in de keten, zowel in brongegevens, informatieproducten als overkoepelend op het niveau van het gehele DSO.

Merk op dat alhoewel er geen bedrijfsfuncties zijn beschreven bij de bronhouder, deze ook bedrijfsfuncties zoals kwaliteitsmanagement, verantwoording en gegevensbeheer zal uitvoeren. Hetzelfde geldt voor een deel van de afnemers.



Figuur 15 Positionering van functionaliteiten in Omgevingswet context

Bijlage A: Bronnen

[1] ISO/IEC: "ISO/IEC 25012:2008 - Software engineering — Software product Quality Requirements and Evaluation (SQuaRE) — Data quality model", First edition, 15 december 2008.
[2] ISO/IEC: "ISO/IEC 25024:2015 - Software engineering — Software product Quality Requirements and Evaluation (SQuaRE) — Measurement of data quality", First edition, 15 oktober 2015.
[3] NEN: "Nederlandse norm NEN-EN-ISO 19157 - Geographic information - Data quality", januari 2014.
[4] M. Mosley et al.: "Guide to the Data Management Body of Knowlegde", first edition, The Data Management Association (DAMA International), ISBN 978-0-9771400-8-4, United States of Amersica, 2009.
[5] P. Vieveen et al.: "Code voor informatiekwaliteit 2016", Versie 2.1, DAMA NL, 10 april 2016.
[6] IPO/CGI: "Referentiemodel Datakwaliteit - Handreiking voor het meten, verbeteren en borgen van datakwaliteit", versie 1.0, 28 september 2014.
[7] H. Berkhout en M. van Langen: " Omgevingswet: Aanbevelingen voor verdere concretisering kwaliteitscriteria (3 B's) - kwaliteitseisen in het digitaal stelsel omgevingswet", versie 0.99, RIVM Rapport 2016-0091,
[8] Y. Ellenkamp en M. Rietdijk: "Kwaliteit van de basisregistraties adressen en gebouwen - Verdiepingsinformatie kwaliteit en de meting daarvan tijdens de toelatingsaudit", versie 2010, Ministerie van VROM, 2010.
[9] BAG BAO: "Toelichting BAG Kwaliteitsdashboard", versie juli 2016, juli 2016.
[10] L. van den Brink et al.: "Basisregistratie Grootschalige Topografie - Gegevenscatalogus BGT", versie 1.1.1., Ministerie van Infrastructuur en Milieu, juli 2013.
[11] W. Schoots en C. Coolen: “Juridisch kader voorziening Inspectieview Milieu”, versie 1.2, Programma Informatie-uitwisseling Milieuhandhaving, augustus 2013.
[12] “Informatieprotocol Inspectieview Milieu (IvM)”, Programma Informatie-uitwisseling Milieuhandhaving, 20 november 2013.
[13] ISO: “ISO 19107:2003 - Geographic information -- Spatial schema“, 1 mei 2003.
[14] ISO: “ISO 19115-1:2014 - Geographic information -- Metadata -- Part 1: Fundamentals“, 1 april 2014.
[15] Geonovum: “Nederlands profiel op ISO 19115 voor geografie”, versie 2.0.0, 9 november 2017.
[16] P. van Oort: “Spatial data quality: from description to application”, proefschrift, Wageningen UR, ISBN 90-8504-339-5, januari 2006.
[17] D. Greefhorst: "Referentie-architectuur voor de informatiehuizen”, versie 0.4, 13 december 2017.
[18] D. Greefhorst en T. Sloos: “Globale Architectuur Schets Aansluiten Informatiehuizen”, versie 1.51, Programma Implementatie Omgevingswet, Ministerie van IenM, 2 maart 2017.
[19] W. Vullings, F. Rip: “Praten over kwaliteit van ruimtelijke data”, Geo-Info, juni 2015.
[20] W3C: “Data on the Web Best Practices: Data Quality Vocabulary”, W3C Working Group Note, 30 augustus 2016.
[21] K. Kisjes: “Dimensies voor Datakwaliteit”, Rijkswaterstaat, 31 maart 2016.
[22] J. Takkenberg: “Onderbouwing dimensies HAT - Het stuur op datakwaliteit: hoe goed zijn de data?”, Rijkswaterstaat, 18 juli 2016.

[23] D. Loshin: "Enterprise Knowledge Management - The Data Quality Approach", ISBN 0-12-455840-2, Academic Press, 2010.
[24] M. Meijer en L.A.E. Vullings: "Kwaliteit van ruimtelijke data in relatie tot het LPIS – kwaliteitsaspecten rondom het beheer van ruimtelijke data", Alterra Wageningen UR, Alterra-rapport 2285, ISSN 1566-7197, juni 2012.
[25] Strategische en Tactische Baseline Informatiebeveiliging Nederlandse Gemeenten (BIG), KING, 30 april 2014.
[26] "Visie Digitaal Stelsel Omgevingswet 2024", versie 1.0, 10 november 2016.
[27] "Doelarchitectuur Digitaal Stelsel Omgevingswet 2024", versie 2.0, 10 november 2016.
[27] "Globaal Programma van Eisen Digitaal Stelsel Omgevingswet", versie 1.0, 10 november 2016.
[28] R. Bekker, P. Marschalk en C. Nikkels: "Informatiehuis Natuur - Verdiepende analyse datakwaliteit", december 2016.
[29] ISO: "ISO 5725-1:1994(E) – Accuracy (trueness and precision) of measurement methods and results", first edition, 15 december 1994.
[30] ISO: "ISO 9001:2015(E) – Quality management systems – requirements", 15 september 2015.
[31] ISO: "ISO/IEC Directives, Part 1, Consolidated ISO Supplement", 2014.
[32] A. de Boer, F. Kooij en L. Jessen: "Basisregistratie Adressen en Gebouwen - Gegevenscatalogus BAG", versie 0.4, Ministerie van Infrastructuur en Milieu, 10 oktober 2016.
[33] C. Batini, C. Cappiello, C. Francalanci, & A. Maurino: "Methodologies for data quality assessment and improvement", ACM Comput. Surv., 41(3), 1–52, 2009.
[34] Y. Lee, L. Pipino, J. Funk, & R. Wang: "Journey to data quality", Computer (Vol. 1), 2006.
[35] R. Wang, M. Reddy, & H. Kon: "Toward quality data: An attribute-based approach.", Decision Support Systems, 13(3–4), 349–372, 1995.
[36] "Katern GEMMA Strategisch Gegevensmanagement", VNG Realisatie, https://www.gemmaonline.nl/index.php/Katern_GEMMA_Strategisch_Gegevensmanagement.
[37] "Katern GEMMA Tactisch Gegevensmanagement", VNG Realisatie, https://www.gemmaonline.nl/index.php/Katern_GEMMA_Tactisch_Gegevensmanagement.

Bijlage B: Relatie tussen ISO 25012 en de voorgestelde indicatoren en de 3 B's

In Tabel 5 is de relatie weergegeven tussen de 3 B's en ISO/IEC 25012. Schuingedrukt weergegeven zijn de kwaliteitsdimensies uit ISO/IEC 25012. In Tabel 6 is deze verdiept tot op het niveau van de voorgestelde indicatoren. Merk op dat deze indicatoren slechts een deel van de kwaliteitsdimensies uit ISO/IEC 25012 dekken. Hierdoor zijn een deel van de aspecten van de 3 B's wel gerelateerd aan een kwaliteitsdimensie, maar niet aan een indicator.

Beschikbaarheid	**Bruikbaarheid**	**Bestendigheid**
Openheid *Toegankelijkheid*	Actualiteit *Actualiteit*	Werkterrein
Duidelijkheid *Compleetheid*	Consistentie *Consistentie*	Organisatorische basis
Bereikbaarheid *Compliance*	Juistheid *Nauwkeurigheid*	Aanspreekpunt
Vindbaarheid *Toegankelijkheid*	Nauwkeurigheid *Nauwkeurigheid* *Precisie*	Besturing/regie
	Volledigheid *Compleetheid*	Aanbod
	Ondubbelzinnigheid *Compleetheid* *Begrijpelijkheid*	Kwaliteitsborging
	Geo-gerefereerd *Compleetheid*	Data licentie
		Financiering producten
		Vraagsturing
		Aansprakelijkheid
		Herleidbaarheid *Traceerbaarheid*

Tabel 5. Relatie tussen de 3 B's en ISO/IEC 25012 en

Beschikbaarheid	**Bruikbaarheid**	**Bestendigheid**
Openheid	Actualiteit *Tijdigheid* *Frequentie*	Werkterrein
Duidelijkheid *Metadata compleetheid*	Consistentie *Domeinconsistentie* *Standaarden compliance* *Homogeniteit*	Organisatorische basis
Bereikbaarheid *Standaarden compliance*	Juistheid	Aanspreekpunt
Vindbaarheid	Nauwkeurigheid *Positionele nauwkeurigheid* *Temporele nauwkeurigheid* *Kwantitatieve nauwkeurigheid* *Systematische afwijking* *Opslagprecisie* *Statistische precisie* *Resolutie*	Besturing/regie
	Volledigheid *Dataset compleetheid* *Object compleetheid* *Overcompleetheid* *Historie*	Aanbod
	Ondubbelzinnigheid *Metadata compleetheid* *Relevantie*	Kwaliteitsborging
	Geo-gerefereerd *Geo-gerefereerd*	Data licentie
		Financiering producten
		Vraagsturing
		Aansprakelijkheid
		Herleidbaarheid *Herleidbaarheid*

Tabel 6. Relatie tussen de 3 B's en de voorgestelde indicatoren

Bijlage C: Begrippenlijst

Term	Definitie
Attribuutwaarde	De waarde die een logisch attribuut aanneemt.
Conceptueel attribuut	Een kenmerk van een object.
Conceptueel objecttype	Een type van gelijksoortige objecten.
Data	*Zie gegevens.*
Dataset	Een verzameling van gegevens die als geheel wordt verwerkt.
Gegevens	Weergave van een feit, begrip of aanwijzing, geschikt voor overdracht, interpretatie of verwerking door een persoon of apparaat.
Gegevenskwaliteit	De kwaliteit van gegevens.
Gegevensobject	Een op zichzelf staand geheel van gegevens met een eigen identiteit.
Geo-object	Een object dat direct of indirect geassocieerd is met een locatie relatief ten opzichte van het aardoppervlak.
Georeferentie	De locatie van een geo-object vastgelegd in een ruimtelijk referentiesysteem.
Indicator	Een meetbaar aspect van een kwaliteitsdimensie.
Informatie	Gegevens waaraan vanuit een bepaalde context betekenis wordt toegekend.
Informatiemodel	Een formele definitie van objecttypen, attributen, relaties en regels.
Informatieproduct	Bij ministeriële regeling aangewezen digitaal product met informatieve waarde dat voldoet aan de daarbij gestelde kwaliteitseisen en standaarden.
Kwaliteit	De mate waarin een geheel van eigenschappen en kenmerken van een gegevensobject voldoet aan impliciete of expliciete eisen.
Kwaliteitsdimensie	Een aspect van kwaliteit waaraan gebruikers van gegevens waarde hechten.
Kwaliteitseis	Een meetbare gewenste eigenschap.
Kwaliteitsmeting	Een meting die aangeeft in welke mate gegevens voldoen aan kwaliteitseisen en/of daaraan gerelateerde kwaliteitsregels.
Kwaliteitsraamwerk	Een verzameling kwaliteitsdimensies.
Kwaliteitsregel	Een regel die een beperking stelt aan gegevens.

Logisch attribuut	Een kenmerk van een gegevensobject.
Logisch objecttype	Een type van gelijksoortige gegevensobjecten.
Meetfunctie	Een formule die aangeeft hoe een indicator meetbaar kan worden gemaakt.
Meetprocedure	Een beschrijving van de uit te voeren handelingen voor een kwaliteitsmeting.
Meetresultaat	Het resultaat van een kwaliteitsmeting.
Metadata	Gegevens over gegevens.
Model	Een abstractie van de werkelijkheid.
Object	Een fenomeen in de werkelijkheid.
Signaalregel	Een kwaliteitsregel die aangeeft of een gegeven of de relatie tussen gegevens waarschijnlijk correct is.
Waardedomein	Een verzameling van waarden die een attribuut kan aannemen.
Werkelijkheid	Een beeld van de echte of hypothetische wereld.

Bijlage D: Afkortingen

Afkorting	Betekenis
BAG	Basisregistratie Adressen en Gebouwen
DAMA	Data Management Association
DMBoK	Data Management Body of Knowlegde
BGT	Basisregistratie Grootschalige Topografie
BRP	Basisregistratie Personen
DSO	Digitaal Stelsel Omgevingswet
DVO	Dienstverleningsovereenkomst (zie ook: SLA)
IPO	Interprovinciaal Overleg
IvM	Inspectieview Milieu
GLO	Gegevensleveringsovereenkomst
GLP	Gegevensleveringsprotocol
GTP	Gegevenstransformatieprotocol
NGR	Nationaal Geo Register
NLIQ	Platform voor Informatiekwaliteit Nederland
PDOK	Publieke Dienstverlening Op de Kaart
SLA	Service Level Agreement (zie ook: DVO)
WBP	Wet Bescherming Persoonsgegevens
WOZ	Waardering Onroerende Zaken

Bijlage E: Relatie tussen indicatoren en standaarden

In de volgende tabel is weergegeven wat de relatie is tussen de voorgestelde indicatoren en (groepen van) indicatoren in de in dit document gebruikte ISO standaarden. Deze relatie is in de meeste gevallen niet exact.

Indicator	**ISO/IEC 25024**	**NEN-ISO 19157**	**ISO 5725**
Thematische nauwkeurigheid	Syntactic data accuracy Semantic data accuracy	Non-quantitative attribute correctness	
Classificatie juistheid		Classification correctness	
Positionele nauwkeurigheid	Data accuracy range	Positional accuracy	Accuracy
Temporele nauwkeurigheid	Data accuracy range	Accuracy of a time measurement	Accuracy
Kwantitatieve nauwkeurigheid	Data accuracy range	Quantitative attribute correctness	Accuracy
Systematische afwijking	Data accuracy range		Trueness
Dataset compleetheid	Data file completeness	Omission	
Object compleetheid	Attribute completeness	Omission	
Metadata compleetheid	Metadata completeness	Omission	
Overcompleetheid		Commission	
Historie		Omission	
Geo-gerefereerd	Attribute completeness	Omission	
Formaatconsistentie		Physical structure conflict rate	
Domeinconsistentie		Value domain conformance rate	
Conceptuele consistentie	Semantic consistency	Conceptual consistency	
Referentiële integriteit	Referential integrity	Conceptual consistency	

Geometrische correctheid		Physical structure conflict rate	
Topologische consistentie		Topological consistency	
Homogeniteit			
Versheid	Timeliness of update		
Frequentie	Update frequency		
Opslagprecisie	Precision of data values		
Statistische precisie			Precision
Resolutie			
Authenticiteit	Source credibility		
Waarschijnlijkheid	Values credibility		
Representativiteit			
Herleidbaarheid	Traceability of data values		
Standaarden compliance	Regulatory compliance of value and/or format		
Relevantie	Data values understandability		
Leesbaarheid	Semantic understandability		

Bijlage F: Betrokkenen

De werkgroep waarin dit document tot stand is gekomen bestond uit:
- Danny Greefhorst (RIVM/Kadaster) - redacteur
- Martijn van Langen (RIVM)
- Wiard Ruisdonk (RCE)
- Jan Skornsek (RIVM)
- Jeroen van der Veen (Kadaster)
- Edwin Biegel (Rijkswaterstaat)
- Renée Bekker (IPO/BIJ12)
- Hans Berkhout (RIVM)
- Carla Nikkels (IPO)

De werkgroep die is verder gegaan in het kader van het werkplan Informatiehuizen 2017 bestaat uit:
- Ronald Hoogerbrugge (RIVM)
- Edwin Biegel (Rijkswaterstaat)
- Danny Greefhorst (RIVM/Kadaster)
- Hans van Diën (Rijkswaterstaat)
- Renée Bekker (IPO/BIJ12)
- Hans Berkhout (RIVM)
- Kasper Kisjes (Rijkswaterstaat)
- Marcel Boons (VNG)
- Dick Bergmans (RIVM)

Bij de totstandkoming van dit document zijn onder meer gesprekken gevoerd met:
- Kasper Kisjes (Rijkswaterstaat)
- Bert Mast (Kadaster)
- Peter Lentjes (Kadaster)
- Michiel Schram (RIVM)
- Hennie Duijst (RCE)
- Elsbeth Pluim (Dataland)
- Richard Slagter (Dataland)
- Marjan Bevelander (Ministerie van IenM)
- Jan-Willem Genuit (Ministerie van IenM)
- Stan van Bosbeek (Ministerie van IenM)
- Huibert-Jan Lekkerkerk (Waterschapshuis)
- Bart Maessen (Kadaster)

Bijlage G: Relatie tussen kwaliteitsraamwerken

ISO/IEC 25012	NEN-ISO 19157	ISO 5725	DMBoK	Code voor informatiekwaliteit	IPO	Rijkswaterstaat
Nauwkeurigheid	Thematische nauwkeurigheid Positionele nauwkeurigheid Temporele nauwkeurigheid Logische consistentie	Juistheid Nauwkeurigheid	Accuraatheid	Juistheid van data	Juistheid	
Compleetheid	Volledigheid		Compleetheid	Bruikbaarheid van de inhoud van een dataset Bruikbaarheid van records in een dataset Volledigheid van de data in een dataset Volledigheid van de records van een dataset	Volledigheid	Compleetheid Aanwezigheid specificatie
Consistentie	Logische consistentie Temporele nauwkeurigheid		Consistentie Referentiële integriteit Uniekheid Validiteit	Consistentie van data Integriteit van data Koppelbaarheid van records in een dataset Stabiliteit van een dataset met data over een periode of tijdstip	Juistheid Referentiële integriteit	Domeinwaarden Validiteit Referentiële integriteit Uniciteit Normalisatie Logische verbanden
Plausibiliteit			Representativiteit	Plausibiliteit van data		Reputatie Waarschijnlijkheid
Actualiteit	Temporele nauwkeurigheid		Actualiteit Tijdigheid	Punctualiteit van het tijdstip van levering van een dataset Tijdigheid van het tijdstip van levering van een dataset	Tijdigheid	Historie Frequentie Actualiteit / dataverval

						Tijdigheid
Toegankelijkheid	Bruikbaarheid			Toegankelijkheid van een dataset	Beschikbaarheid	Toegankelijkheid
Compliance						Standaardisatie
Vertrouwelijkheid			Privacy			Classificatie Versleuteling Afgrendeling Privacynormen
Efficiëntie			Validiteit		Beschikbaarheid	
Precisie		Precisie	Precisie			Nauwkeurigheid
Traceerbaarheid					Transformatieregels	Traceerbaarheid
Begrijpelijkheid						Eenduidigheid Bondigheid
Beschikbaarheid					Beschikbaarheid	
Overdraagbaarheid					Migreerbaarheid	Overdraagbaarheid
Herstelbaarheid					Leveringszekerheid Betrouwbaarheid Beheerbaarheid	
					Actualiteit Eigenaarschap Toetsbaarheid Consistentie Controleerbaarheid Presentatieregels Regels voor uitval Uitwisselbaarheid Herstelbaarheid Begrijpelijkheid Bruikbaarheid Herstelbaarheid	Wijzigbaarheid Rollen bekend en ingericht Vindbaarheid Gebruiksverplichtingen Relevantie / publieke taak

Bijlage H: Relatie van cases met ISO 25012

ISO/IEC 25012	BAG	BGT	AERIUS	IvM	INSPIRE/PDOK
Nauwkeurigheid	Juistheid Positionele nauwkeurigheid Oppervlakte nauwkeurigheid Tijdnauwkeurigheid	Thematische nauwkeurigheid Postionele nauwkeurigheid Tijdnauwkeurigheid	Juistheid	Juistheid	*Juistheid*
Compleetheid	Volledigheid Compleetheid Tijdconsistentie	Volledigheid Tijdconsistentie	Volledigheid	Volledigheid	*Compleetheid*
Consistentie	Compleetheid Tijdconsistentie	Logische en temporele consistentie Tijdconsistentie	Integriteit		*Consistentie*
Plausibiliteit	Juistheid				
Actualiteit	Actualiteit	Actualiteit	Actualiteit	Actualiteit	*Actualiteit*
Toegankelijkheid					
Compliance				Eenduidigheid Rechtmatigheid	*Compliance*
Vertrouwelijkheid					
Efficiëntie					
Precisie		*Precisie*			*Precisie*
Traceerbaarheid					
Begrijpelijkheid					
Beschikbaarheid					*Beschikbaarheid*
Overdraagbaarheid					
Herstelbaarheid					
				Beheerd	