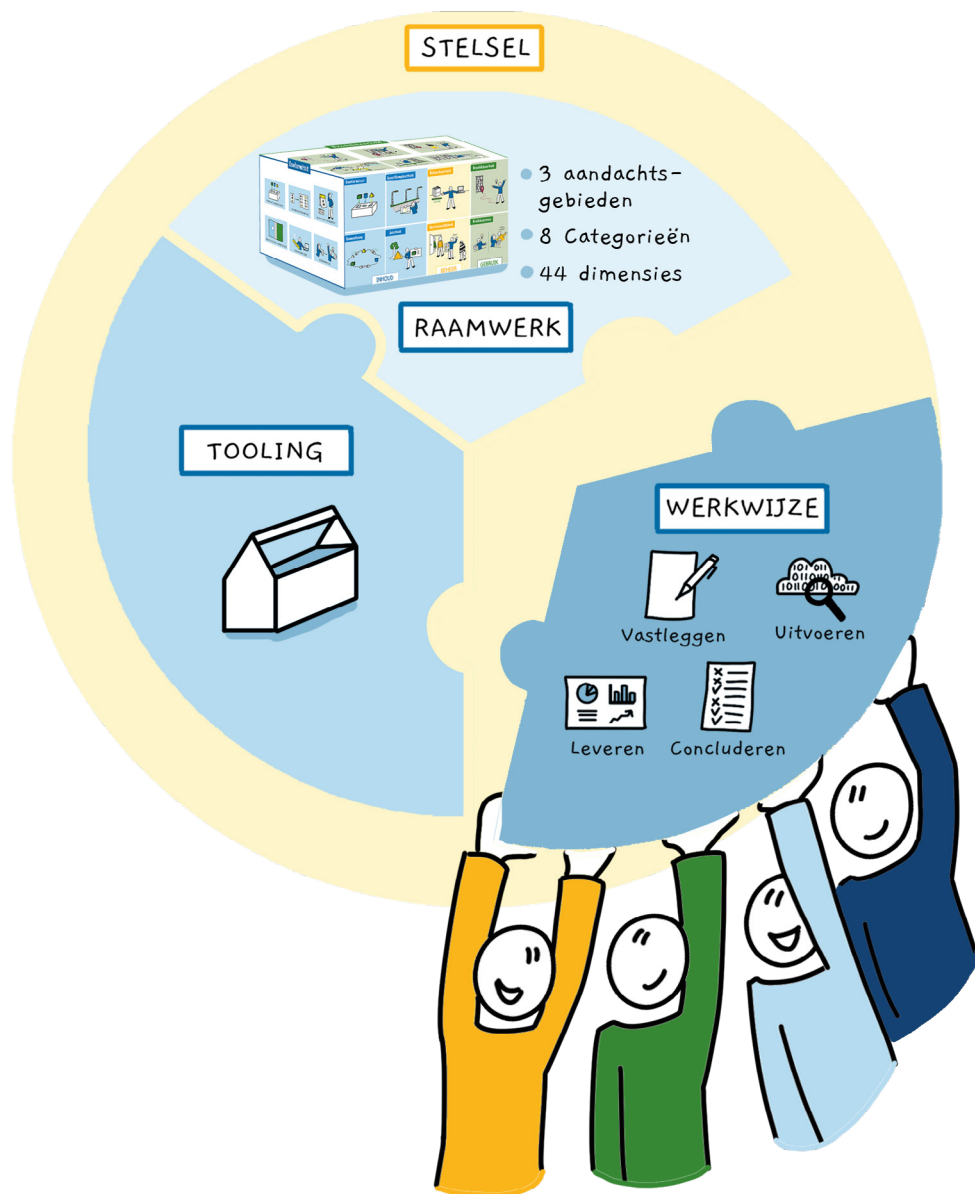






Inleiding

Door RWS is de datakwaliteitstaal ontwikkeld volgens internationale normen (ISO 8000, 25012, 19157)¹, afgestemd met Rijkspartners (o.a. binnen de Omgevingswet beperkingengebied en binnen Talking Traffic). De taal is vormgegeven in het datakwaliteitsraamwerk en toepasbaar gemaakt met twee werkwijzen die beschreven zijn in de richtlijn Datakwaliteitsstelsel. Dit stelsel bestaat uit drie onderdelen:

1. Het datakwaliteitsraamwerk, de datakwaliteitstaal.
2. De werkwijzen voor het toepassen.
3. De tools om datakwaliteit te kunnen toetsen.

Dit document beschrijft het datakwaliteitsraamwerk.

Het datakwaliteitsraamwerk staat centraal in het datakwaliteitsstelsel en bevat drie aandachtsgebieden:

→ Inhoud

1. Compleetheit per object?
2. Uniciteit?
3. Domeinconsistentie?

→ Beheer

1. Processen en systemen, zijn we 'in control'?
2. AVG?
3. Documentatie?

→ Gebruik

1. Hoe geschikt is de data voor gebruik?
2. Tijdigheid?
3. Benaderbaarheid?

Aan de hand van het datakwaliteitsraamwerk zijn vanuit de aandachtgebieden alle data-eisen inzichtelijk te maken. Hiervoor zijn verschillende expertises nodig om dit mogelijk te maken. Leidraad hiervoor zijn de werkwijzen in de richtlijn Datakwaliteitsstelsel.

Inhoud, deze ligt voor de hand, welke waarde hebben de kenmerken (bijv. meetwaarden) in de dataset of -stroom en voldoen die.

Beheer en **Gebruik** gaan over alle andere onderdelen van datakwaliteit die zich buiten de fysieke dataset of -stroom afspelen maar zeker óók van belang zijn.

¹ ISO 8000: Data Quality and Enterprise Master Data
ISO 25012: Software Engineering – Software Product Quality Requirements and Evaluation (SQuaRE) – Data Quality Model
ISO 19157: Geographic information – Data quality

DATAKWALITEITSRAAMWERK

INHOUD		BEHEER	GEBRUIK
CONFORMITEIT	(OVER) COMPLEETHEID	BEHEER-BAARHEID	BESCHIK-BAARHEID
Formaat-consistentie	Compleetheid per object	Wijzigbaarheid	Vindbaarheid
Domein-consistentie	Object-aanwezigheid	Herstelbaarheid	Benaderbaarheid
Temporele validiteit	Historie & frequentie	Controleerbaarheid	Bestendigheid
Geometrische correctheid	Ruimtelijke dekking	Herleidbaarheid tot bronnen	Tijdigheid (van het proces)
Opslagprecisie	Metadata compleetheid	Reproduceerbaarheid	Betaalbaarheid
Leesbaarheid		Documentatie	Retentie
		Rolinvulling	
SAMENHANG	JUISTHEID	VERTROUWELIJKHEID	BRUIKBAARHEID
Referentiele integriteit	Temporele juistheid	Afscherming van onbevoegden	Reputatie van de bron (leverancier)
Conceptuele consistentie	Positionele juistheid	Bescherming persoonsgegevens	Legitimiteit voor gebruiksdoel
Temporele consistentie	Kwantitatieve juistheid	Rubricering/classificatie	Gebruiksgemak
Topologische consistentie	Actuele juistheid		Meet-nauwkeurigheid
Statistische consistentie			Representativiteit
Uniciteit			Koppelbaarheid
			Inzicht in gebruik

Het datakwaliteitsraamwerk bestaat naast de 3 aandachtsgebieden uit 8 categorieën die gezamenlijk 45 dimensies bevatten.

- **Inhoud** bevat 4 categorieën
 1. Conformiteit met 6 dimensies
 2. (over)Compleetheid met 5 dimensies
 3. Samenhang met 6 dimensies
 4. Juistheid met 4 dimensies
- **Beheer** bevat 2 categorieën
 1. Beheerbaarheid met 7 dimensies
 2. Vertrouwelijkheid met 3 dimensies
- **Gebruik** bevat 2 categorieën
 1. Beschikbaarheid met 6 dimensies
 2. Bruikbaarheid met 7 dimensies

In het vervolg van dit document is alles uitgelegd per dimensie.

Tips

In het document 'Datakwaliteitsraamwerk in beeld' is het raamwerk visueel weergegeven. Met de navigatie functionaliteit kan naar de gewenste dimensie 'gesprongen' worden voor een korte uitleg van de betreffende dimensie.

Bij de keuze van de juiste dimensie kan er een grijs gebied zitten tussen twee of meer dimensies. Dat is niet zo erg, kies dan gewoon welke het best passend is (zwaartepuntprincipe).

Per dimensie worden voorbeelden van controleregels gegeven.

Het kan altijd beter, dus als je mooie voorbeelden hebt geef deze dan door aan de adviseur datakwaliteit.

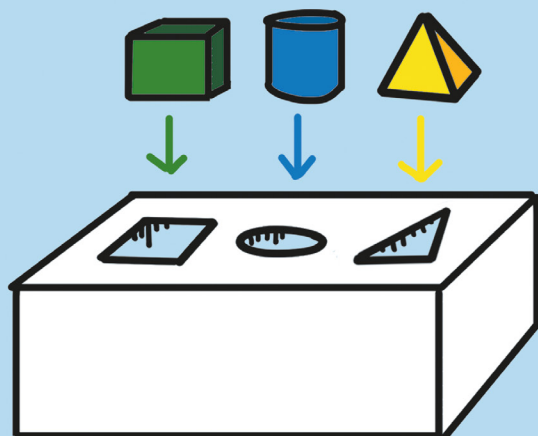
Heb je meer hulp nodig?

Neem dan contact op met de adviseurs datakwaliteit via datakwaliteit@rws.nl
Dit loket is er voor:

- ✓ Het beantwoorden op (hulp)vragen.
- ✓ Het leveren van assistentie.
- ✓ Het melden van verbeteringen. Verbetervoorstellen worden in behandeling genomen door de inhoudelijk beheerder van de richtlijn en verwerkt in de volgende versie.



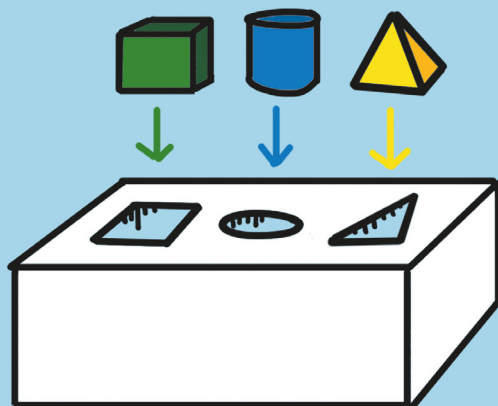
Conformiteit



Conformiteit

De mate waarin de gegevens vanuit verschillende perspectieven een volledig geheel vormen.

INHOUD		BEHEER	GEBRUIK
CONFORMITEIT	(OVER) COMPLEETHEID	BEHEER-BAARHEID	BESCHIK-BAARHEID
Formaat-consistentie	Compleetheid per object	Wijzigbaarheid	Vindbaarheid
Domein-consistentie	Object-aanwezigheid	Herstelbaarheid	Benaderbaarheid
Temporele validiteit	Historie & frequentie	Controleerbaarheid	Bestendigheid
Geometrische correctheid	Ruimtelijke dekking	Herleidbaarheid tot bronnen	Tijdigheid (van het proces)
Opslagprecisie	Metadata compleetheid	Reproduceerbaarheid	Betaalbaarheid
Leesbaarheid		Documentatie	Retentie
		Rolinvulling	
SAMENHANG	JUISTHEID	VERTROUWELIJKHEID	BRUIKBAARHEID
Referentiele integriteit	Temporele juistheid	Afscherming van onbevoegden	Reputatie van de bron (leverancier)
Conceptuele consistentie	Positionele juistheid	Bescherming persoonsgegevens	Legitimiteit voor gebruiksdoel
Temporele consistentie	Kwantitatieve juistheid	Rubricering/classificatie	Gebruiksgemak
Topologische consistentie	Actuele juistheid		Meet-nauwkeurigheid
Statistische consistentie			Representativiteit
Uniciteit			Koppelbaarheid
			Inzicht in gebruik



FORMAATCONSISTENTIE



Formaatconsistentie

De mate waarin gegevens syntactisch correct zijn (incl. datatype) en alle verwachte onderdelen bevat zoals mappen, tabellen, kolommen en verplichte velden die nodig zijn voor automatische verwerking.

Waarom belangrijk

Voor correcte (automatische) verwerking verderop in de keten (vaak alleen van toepassing op leveringen: anders vaak afgedwongen in database).

Voorbeeld van controleregels

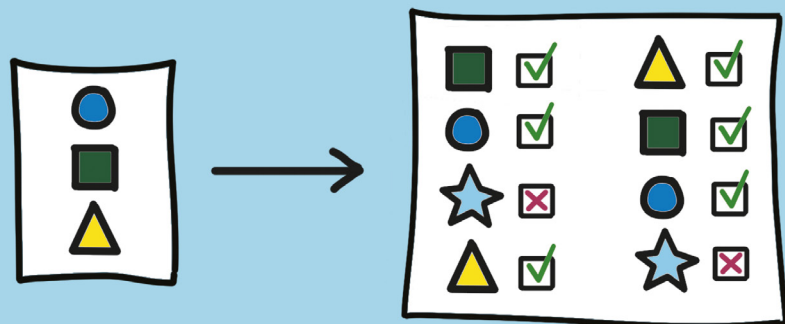
- ✓ Map X aanwezig in dataset.
- ✓ Dataset heeft X kolommen.
- ✓ Bestand is UTF-8 gecodeerd.
- ✓ Regel voor datum/tijdnotatie.
- ✓ Bestandsextensie.

Gerelateerd aan begrip

→ Structuur; Bestandformaat; Opbouw

ISO

- | | |
|---------|--------------------|
| → 8000 | Validity |
| → 25012 | - |
| → 19157 | Format Consistency |



DOMEINCONSISTENTIE

Domeinconsistentie

De mate waarin de inhoud en de lengte van attribuutwaarden consistent zijn met hun domein.

Waarom belangrijk

1. Om te conformeren aan een standaard.
2. Ten behoeve van verwerking verderop in de keten.
3. Om koppeling van gegevens mogelijk te maken.
4. Om onwaarschijnlijke waarden te signaleren.

Voorbeeld van controleregels

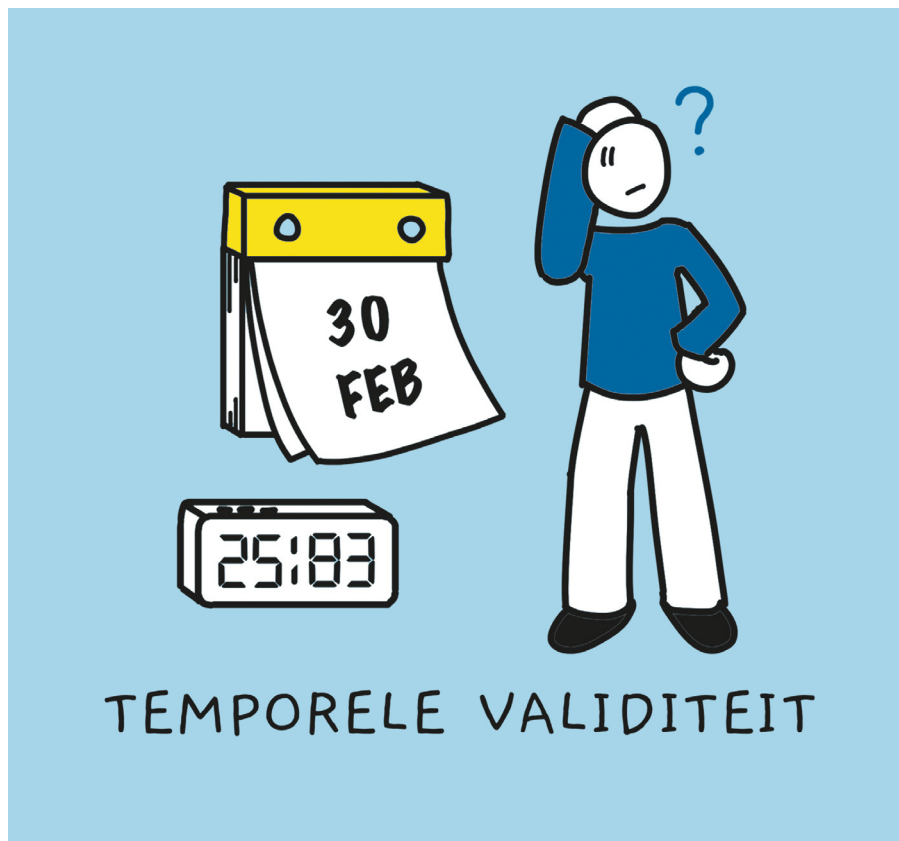
- ✓ Waarde is groter/kleiner dan X, tussen X en Y.
 - ✓ Waarde komt voor in domeinlijst X.
 - ✓ Waarde voldoet aan een patroon.
 - ✓ Waarde bevat geen verboden karakters, bijv. (non-)alfabetisch.
 - ✓ Aantal karakters is niet groter dan X.
- (regels rondom datum/tijd zitten in de dimensie 'Temporele Validiteit')

Gerelateerd aan begrip

- Syntax; Eenduidig (deels); Interpreteerbaar;
- Valide (sommige interpretaties); Consequent;
- Standaardisatie

ISO

- 8000 Validity
- 25012 -
- 19157 Format Consistency



Temporele validiteit

De mate waarin tijdgegevens geldige waarden hebben (bestaande / toegestane datums en tijden).

Waarom belangrijk

1. Niet-bestaande of anderszins niet-toegestane data (datum) kunnen duiden op invoerfouten en kunnen vaak niet goed verwerkt worden
2. Voor sommige tijdreeksen zijn alleen specifieke datums/tijden toegestaan

Voorbeeld van controleregels

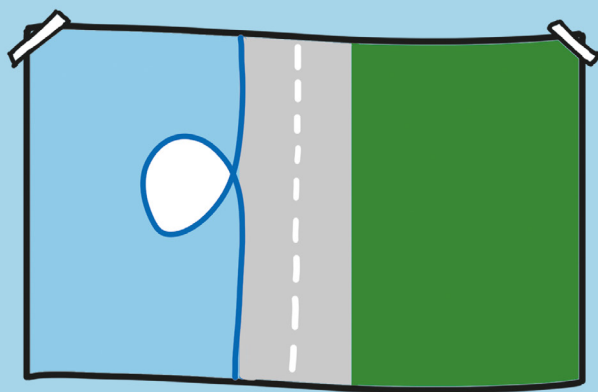
- ✓ Bevat een geldige datum in 2025.
- ✓ Is een geheel uur.
- ✓ Is geen feestdag.

Gerelateerd aan begrip

→ Domeinwaarden (eigenlijk speciaal geval daarvan)

ISO

- | | |
|---------|----------|
| → 8000 | Validity |
| → 25012 | - |
| → 19157 | - |



GEOMETRISCHE CORRECTHEID

Geometrische correctheid

De mate waarin gegevens geometrisch correct zijn (lijnen die zichzelf snijden, polygonen die niet sluitend zijn, etc).

Waarom belangrijk

Geometrische correctheid is van groot belang zodra je de gegevens automatisch wilt gaan verwerken, of plotten op een kaart. Daarnaast kunnen regels binnen deze dimensie ongewenste geometrieën aan het licht brengen, zoals lijnen die een P of T vormen.

Voorbeeld van controleregels

- ✓ Lijnen snijden zichzelf niet; lijnen vormen geen 'P' of 'T'.
- ✓ Features hebben geen dubbele coördinaten; features vormen geen donut.
- ✓ Features hebben ten minste X coördinaten (bijv X=2 bij een lijn).
- ✓ Polygoon binnen polygoon heeft tegengestelde richting.
- ✓ Feature is niet samengesteld (dus geen MultiLine/MultiPolygon etc).
- ✓ Geldige positie binnen het coördinatenreferentiestelsel.

Gerelateerd aan begrip

→ -

ISO

- | | |
|---------|--|
| → 8000 | Validity |
| → 25012 | - |
| → 19157 | Relative internal, Positional Accuracy |



OPSLAGPRECISIE

Opslagprecisie

Het aantal significante decimalen waarmee gegevens zijn geregistreerd.

Waarom belangrijk

Voor sommige gebruiksdoelen is een hoge precisie van belang om de gegevens zo goed mogelijk te kunnen interpreteren. Het kan voorkomen dat ergens in de keten per abuis getallen worden afgerond. Door te kijken naar het aantal decimalen kan dit worden gedetecteerd. Het aantal decimalen zegt echter nog weinig over de Kwantitatieve nauwkeurigheid (onder Juistheid) of de Meetnauwkeurigheid (onder Bruikbaarheid)! Met 10 decimalen ergens ver naast zitten is nog steeds heel erg fout.

Voorbeeld van controleregels

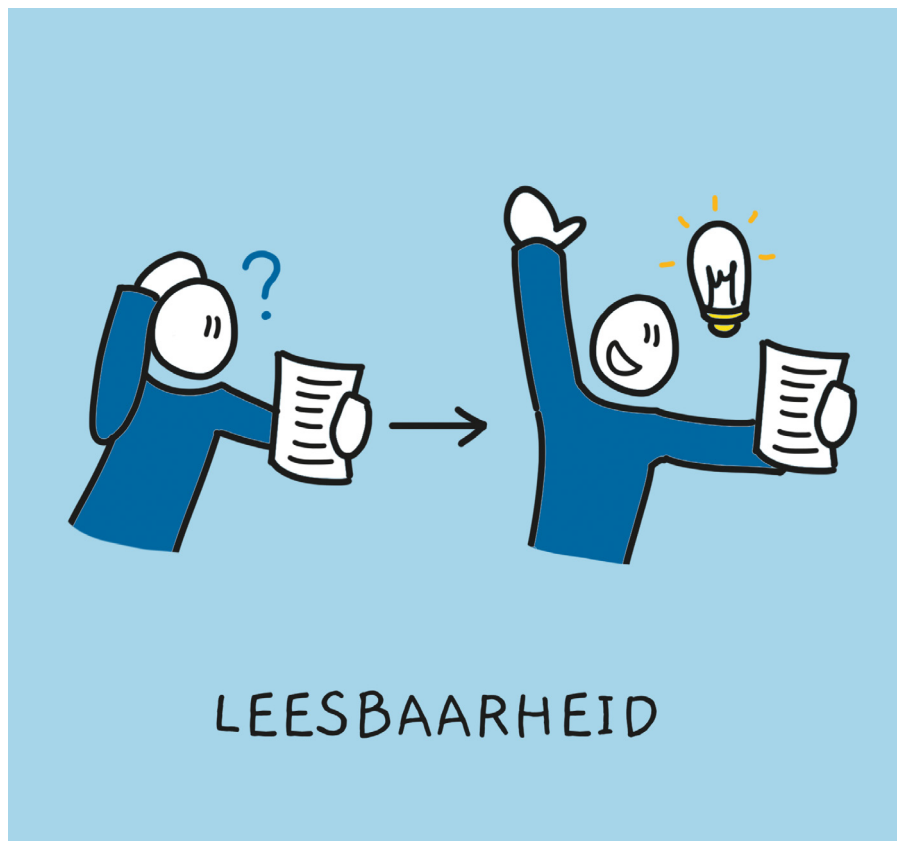
- ✓ Aantal significante decimalen is gelijk aan X.

Gerelateerd aan begrip

→ -

ISO

- | | |
|---------|-----------|
| → 8000 | Precision |
| → 25012 | Precision |
| → 19157 | - |



Leesbaarheid

De mate waarin teksten voor de doelgroep begrijpelijk zijn geformuleerd.

Waarom belangrijk

Om te voorkomen dat gegevens onjuist geïnterpreteerd worden. Het gaat hier om leesbaarheid van de data zelf, bijvoorbeeld van de vrije tekstvelden, en voor zover deze objectief record voor record beoordeeld kan worden. Een (subjectief) oordeel over de leesbaarheid als geheel kun je kwijt onder Gebruiksgemak (onder Bruikbaarheid).

Voorbeeld van controleregels

- ✓ Tekst mag geen waarden bevatten uit bekende afkortingenlijst X.
- ✓ Tekst moet minstens X losse woorden bevatten.

(Deze regels zijn soms ook onder Domeinwaarden te scharen, maar omdat het hier echt gaat om de leesbaarheid kun je ze los zien)

Gerelateerd aan begrip

→ Interpreteerbaar; Begrijpelijk; Eenduidig

ISO

- | | |
|---------|-------------------|
| → 8000 | Clarity |
| → 25012 | Understandability |
| → 19157 | - |

(over)Compleetheid van sets en reeksen

De mate waarin de gegevens vanuit verschillende perspectieven een volledig geheel vormen.

INHOUD		BEHEER	GEBRUIK
CONFORMITEIT	(OVER) COMPLEETHEID	BEHEER-BAARHEID	BESCHIK-BAARHEID
Formaat-consistentie	Compleetheid per object	Wijzigbaarheid	Vindbaarheid
Domein-consistentie	Object-aanwezigheid	Herstelbaarheid	Benaderbaarheid
Temporele validiteit	Historie & frequentie	Controleerbaarheid	Bestendigheid
Geometrische correctheid	Ruimtelijke dekking	Herleidbaarheid tot bronnen	Tijdigheid (van het proces)
Opslagprecisie	Metadata compleetheid	Reproduceerbaarheid	Betaalbaarheid
Leesbaarheid		Documentatie	Retentie
		Rolinvulling	
SAMENHANG	JUISTHEID	VERTROUWELIJKHEID	BRUIKBAARHEID
Referentiele integriteit	Temporele juistheid	Afscherming van onbevoegden	Reputatie van de bron (leverancier)
Conceptuele consistentie	Positionele juistheid	Bescherming persoonsgegevens	Legitimiteit voor gebruiksdooel
Temporele consistentie	Kwantitatieve juistheid	Rubricering/classificatie	Gebruiksgemak
Topologische consistentie	Actuele juistheid		Meet-nauwkeurigheid
Statistische consistentie			Representativiteit
Uniciteit			Koppelbaarheid
			Inzicht in gebruik

Compleetheid per object

De mate waarin attribuutwaarden voor objecten aanwezig zijn (geen lege cellen).

Waarom belangrijk

De attributen die voor objecten/records in een dataset zijn gedefinieerd, zijn dat meestal met een bepaalde reden vanuit (eind)gebruik of beheer. Als bepaalde attributen niet worden voorzien van een waarde, ontstaan er blinde vlekken in onze kennis: we weten bijvoorbeeld niet wat het objecttype is, wie de beheerder of wanneer het contract afloopt. Dat maakt een effectieve operatie, maar ook effectief management, onmogelijk.

Voorbeeld van controleregels

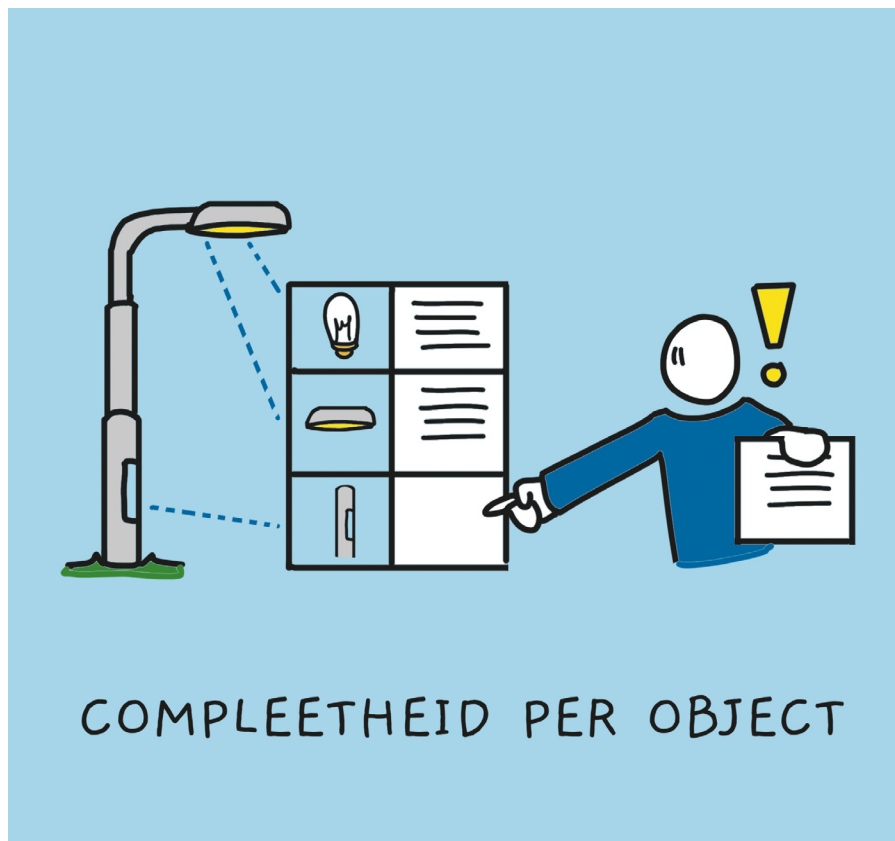
- ✓ Attribuut X is niet leeg, 'null' of anderszins ontbrekend.
- ✓ Van attributen A-E is er ten minste één voorzien van een waarde.

Gerelateerd aan begrip

→ Volledig

ISO

- | | |
|---------|--------------|
| → 8000 | Validity |
| → 25012 | Completeness |
| → 19157 | Completeness |



Objectaanwezigheid

De mate waarin objecten zijn geregistreerd waarvan het bestaan bekend is.

Waarom belangrijk

Gebruikers moeten er op kunnen vertrouwen dat de dataset een volledig beeld van de werkelijke situatie geeft. Zo niet, dan kan een aannemer bijvoorbeeld stuiten op een onbekende ondergrondse kabel of geeft het overzicht van financiën een incompleet beeld.

Voorbeeld van controleregels

- ✓ Alle objectnamen uit referentieset X komen voor in het veld Y.
- ✓ Het totaal aantal records is gelijk aan het aantal X.



Gerelateerd aan begrip

→ Volledig

ISO

- | | |
|---------|--------------|
| → 8000 | Completeness |
| → 25012 | Completeness |
| → 19157 | Completeness |

Historie en frequentie

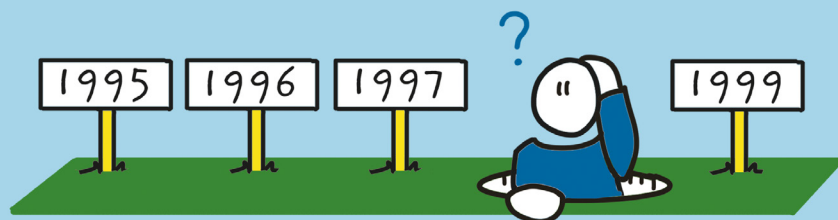
De mate waarin historische gegevens zijn geregistreerd (vroegste registratie, aantal registraties in periode, registraties op specifieke tijdstippen).

Waarom belangrijk

Wanneer gegevens ontbreken voor een bepaald tijdstip of in een bepaalde periode, ontstaan blinde vlekken en kunnen bijvoorbeeld historische analyses minder goed gemaakt worden.

Voorbeeld van controleregels

- ✓ Veld X bevat tenminste Y waarden in iedere maand sinds januari 2012.
- ✓ De dataset bevat een record voor ieder jaar sinds 1970.
- ✓ Bevat een meting voor elk tien-minuten-vak van de afgelopen dag.



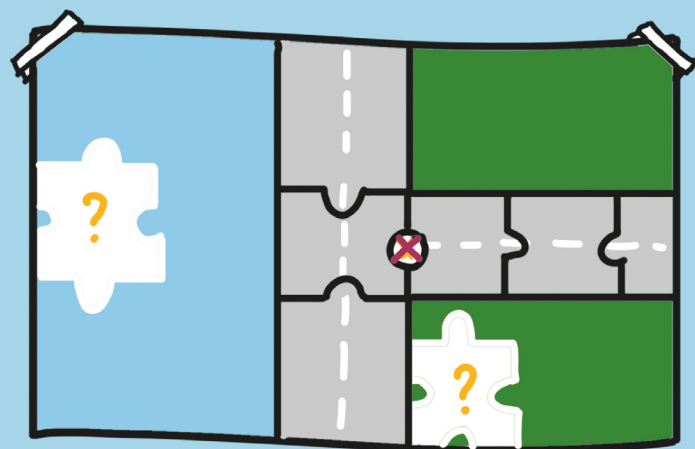
HISTORIE & FREQUENTIE

Gerelateerd aan begrip

→ Volledig

ISO

- 8000 Completeness
- 25012 Completeness
- 19157 Completeness



RUIMTELIJKE DEKKING

Ruimtelijke dekking

De mate waarin de ruimtelijke gegevens ruimtelijk dekkend zijn (geen lege stukken).

Waarom belangrijk

Wanneer een ruimtelijke dataset niet dekkend is, zijn er dus gebieden waar we niets over weten. Dat geeft blinde vlekken, wat leidt tot problemen zodra in die onbekende gebieden bijvoorbeeld werkzaamheden moeten worden uitgevoerd. Of wanneer getracht wordt om objecten ruimtelijk aan elkaar te koppelen.

Voorbeeld van controleregels

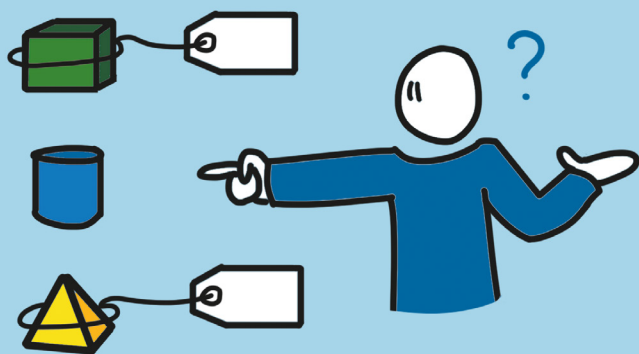
- ✓ De objecten van featureClass X vormen een aaneengesloten geheel zonder gaten.
- ✓ De objecten van featureClasses A-E dekken samen gebied X volledig af..
- ✓ Minstens 1 waarneming in elke gridcel / sector / ... van een geometrisch object.

Gerelateerd aan begrip

→ -

ISO

- 8000 -
- 25012 -
- 19157 -



METADATA COMPLETEHEID

Metadata compleetheid

De mate waarin metadata aanwezig zijn.

Waarom belangrijk

Metadata zijn essentieel om data te kunnen vinden en gebruiken. Zonder metadata kan vaak niet worden herleid wat de oorsprong van gegevens is, wanneer deze door wie zijn aangeleverd, of hoe alle afkortingen en coderingen moeten worden geïnterpreteerd. Let op: dit geldt vooral voor losse datasets / leveringen. Bij databases wordt dit soort zaken afgevangen in de documentatie (zie dimensie Documentatie).

Voorbeeld van controleregels

- ✓ De levering vermeldt een Contactpersoon.
- ✓ De levering bevat een omschrijving van alle velden.
- ✓ De levering bevat gegevens over de gemeten datakwaliteit.
- ✓ De dataset is voorzien van minstens 3 zoekwoorden.

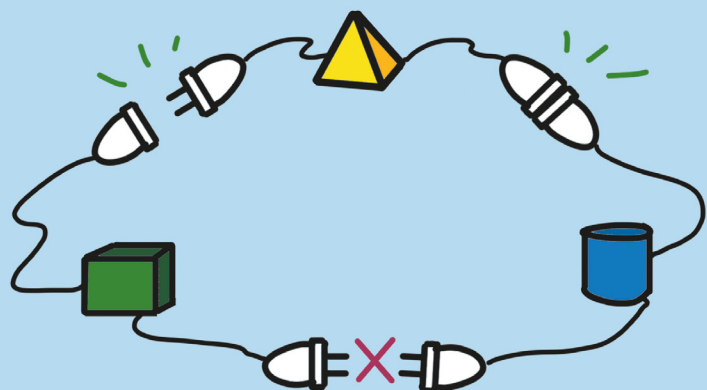
Gerelateerd aan begrip

→ Specificaties; Definities; Interpreteerbaar

ISO

- 8000 -
- 25012 Understandability
- 19157 -

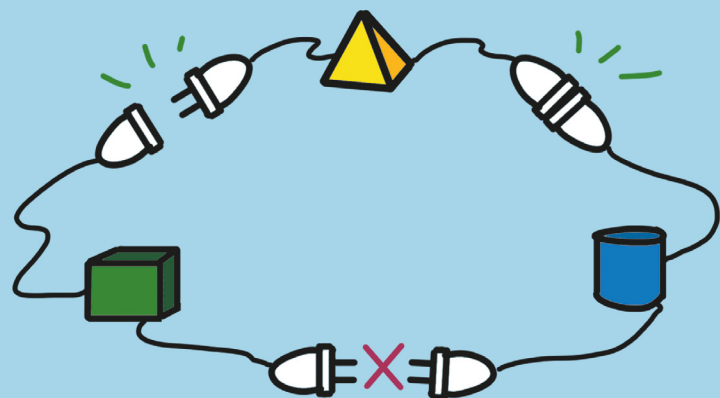
Samenhang



Samenhang

De mate waarin de gegevens elkaar niet tegenspreken in vorm of inhoud.

INHOUD		BEHEER	GEBRUIK
CONFORMITEIT	(OVER) COMPLEETHEID	BEHEER-BAARHEID	BESCHIK-BAARHEID
Formaat-consistentie	Compleetheid per object	Wijzigbaarheid	Vindbaarheid
Domein-consistentie	Object-aanwezigheid	Herstelbaarheid	Benaderbaarheid
Temporele validiteit	Historie & frequentie	Controleerbaarheid	Bestendigheid
Geometrische correctheid	Ruimtelijke dekking	Herleidbaarheid tot bronnen	Tijdsigheid (van het proces)
Opslagprecisie	Metadata compleetheid	Reproduceerbaarheid	Betaalbaarheid
Leesbaarheid		Documentatie	Retentie
		Rolinvulling	
SAMENHANG	JUISTHEID	VERTROUWELIJKHEID	BRUIKBAARHEID
Referentiele integriteit	Temporele juistheid	Afscherming van onbevoegden	Reputatie van de bron (leverancier)
Conceptuele consistentie	Positionele juistheid	Bescherming persoonsgegevens	Legitimiteit voor gebruiksdoel
Temporele consistentie	Kwantitatieve juistheid	Rubricering/classificatie	Gebruiksgemak
Topologische consistentie	Actuele juistheid		Meet-nauwkeurigheid
Statistische consistentie			Representativiteit
Uniciteit			Koppelbaarheid
			Inzicht in gebruik



REFERENTIËLE INTEGRITEIT

Referentiële integriteit

De mate waarin gegevens correct verwijzen naar andere gegevens.

Waarom belangrijk

Daar waar gegevens naar elkaar verwijzen, is het belangrijk dat alle verwijzingen wel intact zijn, dat wil zeggen, verwijzen naar een object dat (nog steeds) bestaat. Als dat niet het geval is vormen de data geen consistent geheel meer en ontstaan er gaten in onze kennis van de objecten en hun samenhang.

Voorbeeld van controleregels

✓ Komt de (sleutel)waarde uit veld X voor in veld Y van tabel Z.

Gerelateerd aan begrip

→ Primary/foreign key; Sleutelwaarden

ISO

→ 8000	Consistency
→ 25012	Consistency
→ 19157	-



CONCEPTUELE CONSISTENTIE

Conceptuele consistentie

De mate waarin combinaties van attribuutwaarden logisch samenhangend zijn.

Waarom belangrijk

Sommige fouten kunnen alleen worden gedetecteerd door gegevens onderling te vergelijken, soms over verschillende datasets / bronnen heen. Je kunt steeds verder gaan in het verzinnen van regels, maar er zijn ook vaak simpele als-dan-constructies die al van alles aan het licht kunnen brengen. De regels binnen deze dimensie vragen vaak om koppeling van gegevens.

Voorbeeld van controleregels

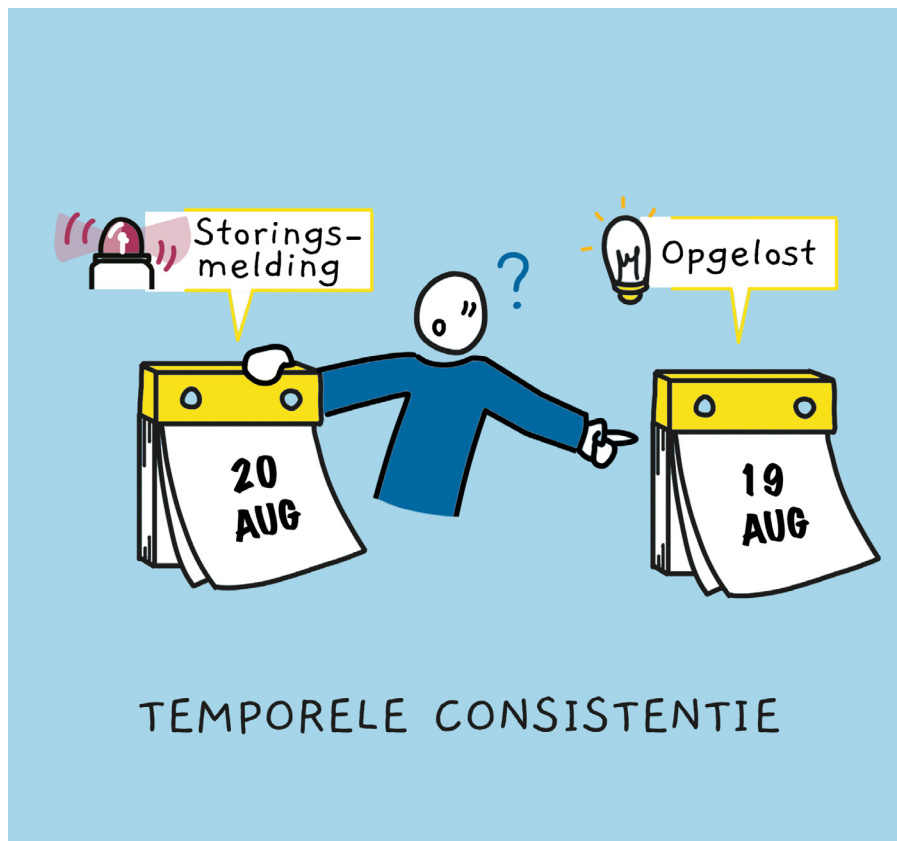
- ✓ Waarde in veld X is groter dan waarde in veld Y.
- ✓ Als Type gelijk is aan X, dan moet Codering beginnen met Y.
- ✓ Het object waar dit objectonderdeel aan is gekoppeld heeft status 'actief'.
- ✓ De waarden in veld X mogen opgeteld niet groter zijn dan Y.
- ✓ De waarden in veld X moeten strikt oplopend zijn over de tijd.

Gerelateerd aan begrip

→ -

ISO

- | | |
|---------|------------------------|
| → 8000 | Consistency |
| → 25012 | Consistency |
| → 19157 | Conceptual Consistency |



Temporele consistentie

De mate waarin tijdgegevens logisch samenhangend zijn.

Waarom belangrijk

Als tijdgegevens niet consistent zijn, duidt dat erop dat minstens 1 van de betrokken waarden onjuist is. Dat heeft implicaties voor de herleidbaarheid van gegevens en soms ook voor de verwerkbaarheid: er kunnen bijvoorbeeld negatieve doorlooptijden worden berekend.

Voorbeeld van controleregels

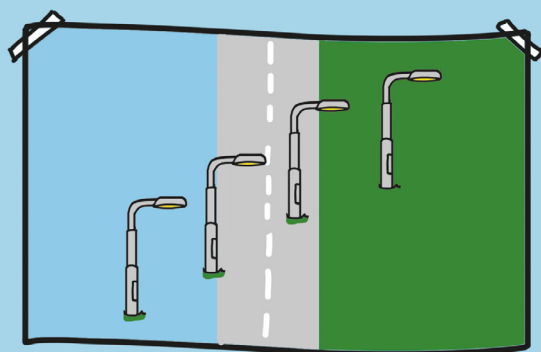
- ✓ DatumTijd X ligt na DatumTijd Y.
- ✓ Er liggen minstens X dagen tussen de startdatum en einddatum.
- ✓ De mutatie van het gegeven heeft plaatsgevonden ná de aanmaak ervan.
- ✓ Het verschil tussen DatumTijd X en DatumTijd Y is maximaal Z seconden.

Gerelateerd aan begrip

→ -

ISO

- | | |
|---------|-------------|
| → 8000 | Consistency |
| → 25012 | Consistency |
| → 19157 | - |



TOPOLOGISCHE CONSISTENTIE

Topologische consistentie

De mate waarin ruimtelijke objecten zich op de juiste wijze tot elkaar verhouden (overlap, aansluiting, etc.).

Waarom belangrijk

Individuele ruimtelijke objecten moeten een logisch geheel vormen, zoals een netwerkstructuur of andersoortige samenhang. Dit is onder andere relevant om afhankelijkheden en consistentie tussen de objecten te kunnen detecteren en beheren. Als dit niet het geval is, is er mogelijk iets mis met de positionering van de objecten.

Voorbeeld van controleregels

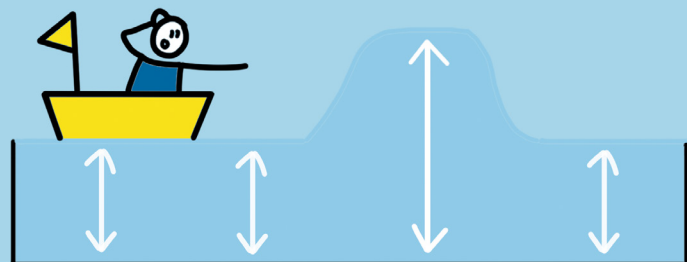
- ✓ Geen overlap tussen objecten van featureClass A en B.
- ✓ Elke lijn sluit aan op een andere lijn.
- ✓ De gekoppelde objecten uit dataset X en Y zijn maximaal x meter verwijderd.
- ✓ Ieder object uit featureClass A grenst aan een object uit featureClass B.

Gerelateerd aan begrip

→ Overlap; Aansluiting

ISO

- 8000 Consistency
- 25012 Consistency
- 19157 Topological Consistency



STATISTISCHE CONSISTENTIE

Statische consistentie

De mate waarin reeksen statistisch gezien een logisch geheel vormen.

Waarom belangrijk

De betrouwbaarheid van een herhaaldelijke meting kan deels worden gecontroleerd door te kijken naar de standaardafwijking. Daarnaast kunnen outliers worden gedetecteerd, die kunnen wijzen op een onjuiste meting / registratie of het begin van een nieuwe trend.

Voorbeeld van controleregels

- ✓ De standaardafwijking van elke set van tien opeenvolgende waarden in veld X mag niet groter zijn dan Y.
- ✓ De waarnemingen in veld X moeten met Y% waarschijnlijkheid een normaalverdeling vormen.
- ✓ Maximaal X% van de waarnemingen is hoger dan de gemiddelde waarde.
- ✓ De gemiddelde absolute afwijking in veld X mag niet groter zijn dan Y.

Gerelateerd aan begrip

→ Aannemelijkheid; Plausibiliteit

ISO

- | | |
|---------|-------------|
| → 8000 | Consistency |
| → 25012 | Consistency |
| → 19157 | - |

Uniciteit

De mate waarin gegevens over hetzelfde object geconcentreerd zijn in één record, en er dus geen (gedeeltelijke) duplicaten in de dataset zitten.

Waarom belangrijk

Als dubbele waarden in een dataset voorkomen waar dat niet wordt verwacht is dit bijna altijd een teken dat er iets fout is gegaan. Dit leidt tot dubbeltellingen, maar kan er ook toe leiden dat de informatie over 1 object op 2 plekken wordt bijgewerkt, waardoor geen centrale waarheid meer bestaat.

Voorbeeld van controleregels

- ✓ Zijn de waarden in veld X uniek?
- ✓ Komen er geen gedeelde identifiers voor tussen tabellen X, Y en Z?

Gerelateerd aan begrip

→ Dubbeling

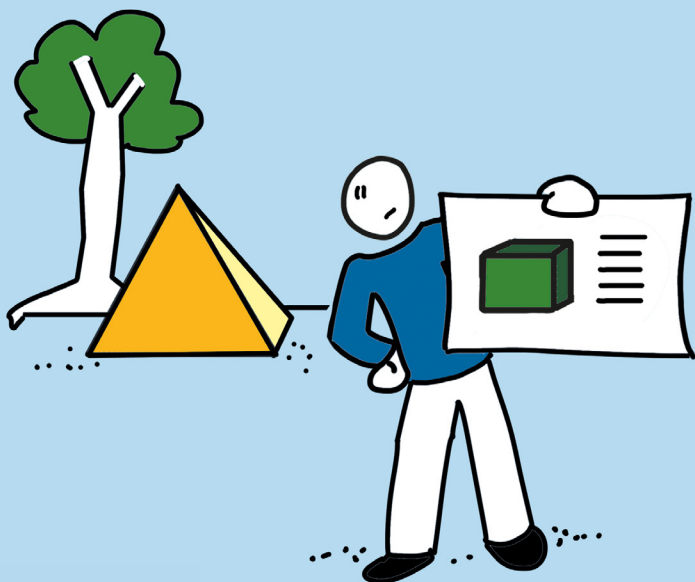
ISO

- 8000 -
- 25012 -
- 19157 Completeness-Omission



UNICITEIT

Juistheid



Juistheid

De mate waarin gegevens overeenkomen met de huidige werkelijkheid.

INHOUD		BEHEER	GEBRUIK
CONFORMITEIT	(OVER) COMPLEETHEID	BEHEER-BAARHEID	BESCHIK-BAARHEID
Formaat-consistentie	Compleetheid per object	Wijzigbaarheid	Vindbaarheid
Domein-consistentie	Object-aanwezigheid	Herstelbaarheid	Benaderbaarheid
Temporele validiteit	Historie & frequentie	Controleerbaarheid	Bestendigheid
Geometrische correctheid	Ruimtelijke dekking	Herleidbaarheid tot bronnen	Tijdigheid (van het proces)
Opslagprecisie	Metadata compleetheid	Reproduceerbaarheid	Betaalbaarheid
Leesbaarheid		Documentatie	Retentie
		Rolinvulling	
SAMENHANG	JUISTHEID	VERTROUWELIJKHEID	BRUIKBAARHEID
Referentiele integriteit	Temporele juistheid	Afscherming van onbevoegden	Reputatie van de bron (leverancier)
Conceptuele consistentie	Positionele juistheid	Bescherming persoonsgegevens	Legitimiteit voor gebruiksdoel
Temporele consistentie	Kwantitatieve juistheid	Rubricering/classificatie	Gebruiksgemak
Topologische consistentie	Actuele juistheid		Meet-nauwkeurigheid
Statistische consistentie			Representativiteit
Uniciteit			Koppelbaarheid
			Inzicht in gebruik



Temporele Juistheid

De mate waarin tijdgegevens overeenkomen met de werkelijkheid.

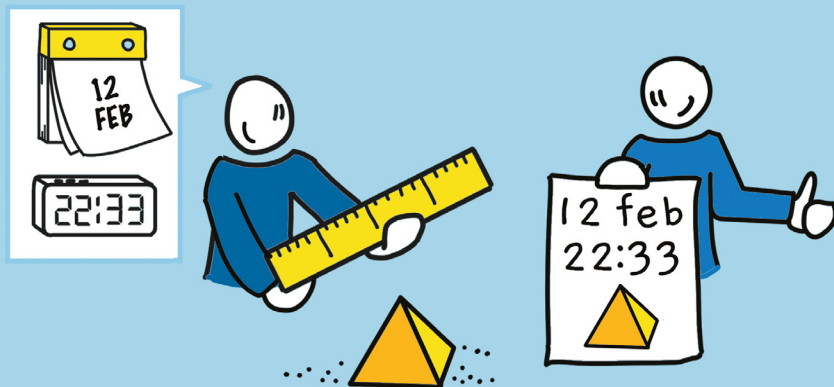
Waarom belangrijk

Onjuiste tijdgegevens leiden tot vertekende trendanalyses, onjuiste conclusies over bijvoorbeeld doorlooptijden en onjuiste besluiten rondom de levenscyclus.

NB: In tegenstelling tot de andere drie dimensies van Juistheid kan men niet ter plaatse – in de toen heersende realiteit – nagaan of de vastgelegde tijd inderdaad toen juist was. (Tijdmachines werden pas in de 23e eeuw uitgevonden.)

Voorbeeld van controleregels

- ✓ Hebben aantoonbaar gelijktijdig gedane andere metingen in vergelijkbare datasets dezelfde datum en/of tijd?



TEMPORELE JUISTHEID

Gerelateerd aan begrip

→ Correctheid

ISO

- 8000 Accuracy
- 25012 Accuracy
- 19157 Accuracy Of A Time Measurement



Positionele Juistheid

De mate waarin ruimtelijke gegevens positioneel overeenkomen met de werkelijkheid.

Waarom belangrijk

Als objecten niet blijken te liggen op de plek waar ze volgens het systeem liggen, of niet de vorm hebben die in het systeem staat, zijn potentieel rampzalige incidenten mogelijk (stuwen, bruggen, profielen).

NB: de geijkte manier om dit te voorkomen is door steekproefsgewijs her metingen te (laten) doen en de data daarvan te vergelijken met de bestaande data. Dit kan worden aangevuld met een integrale vergelijking van data met een 2e databron.

Voorbeeld van controleregels

Hebben vergelijkbare, relevante en onafhankelijke datasets dezelfde positionele data als de te checken data?

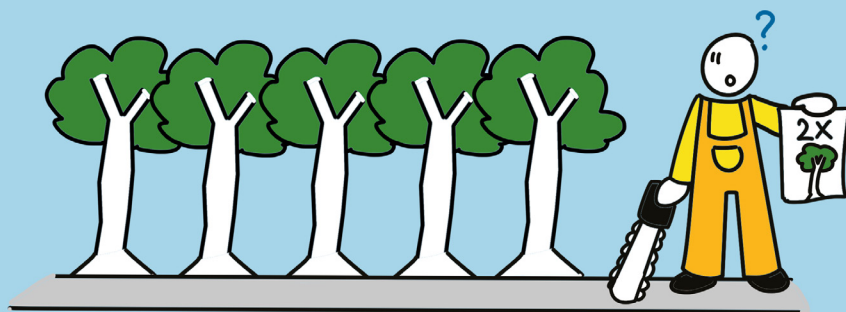


Gerelateerd aan begrip

→ Correctheid

ISO

→ 8000	Accuracy
→ 25012	Accuracy
→ 19157	Absolute External Positional Accuracy



KWANTITATIEVE JUISTHEID

Kwantitatieve Juistheid

De mate waarin kwantitatieve gegevens overeenkomen met de werkelijkheid.

Waarom belangrijk

Als onjuiste getallen zijn vastgelegd, kan dit leiden tot onjuiste conclusies, beslissingen en acties.

NB: de geijkte manier om dit te voorkomen is door steekproefsgewijs hermetingen te (laten) doen en de data daarvan te vergelijken met de bestaande data. Dit kan worden aangevuld met integrale vergelijking van data met een 2e databron.

Voorbeeld van controleregels

- ✓ Hebben vergelijkbare, relevante, onafhankelijke datasets dezelfde kwantitatieve data als de te checken data?

Gerelateerd aan begrip

→ Correctheid

ISO

- 8000 Accuracy
- 25012 Accuracy
- 19157 Absolute External Positional Accuracy



Actuele Juistheid

De mate waarin gegevens overeenkomen met de huidige situatie.

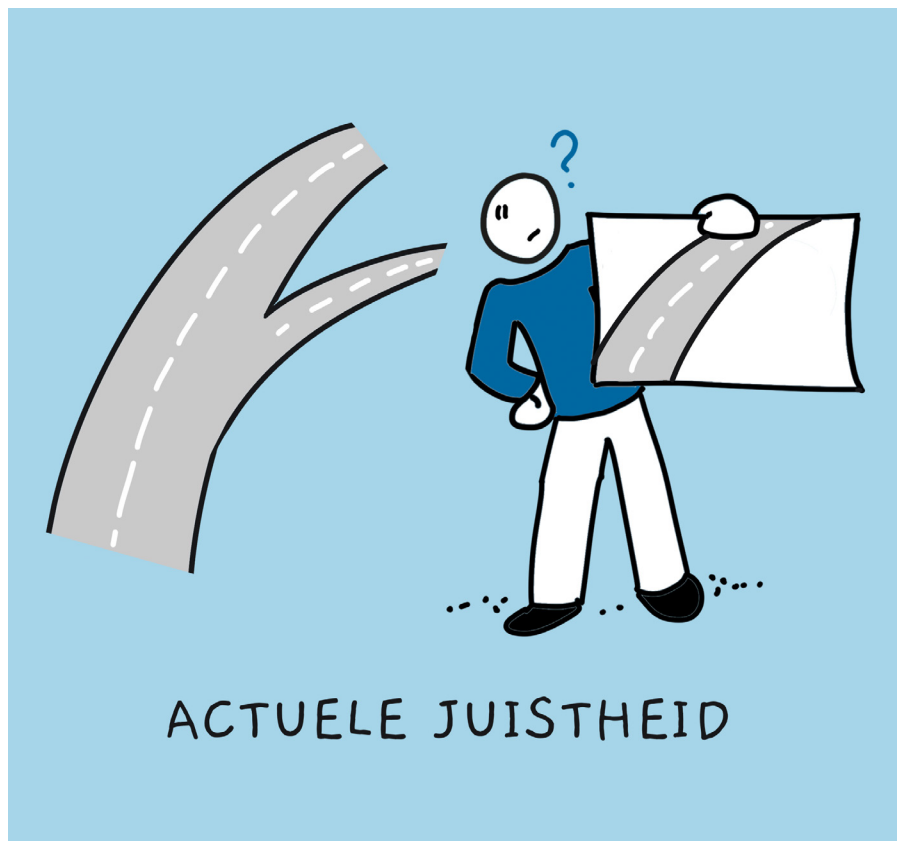
Waarom belangrijk

Hoe langer geleden de gegevens zijn ingewonnen of getoetst, des te groter de kans dat deze niet meer juist zijn. Daarmee wordt bijvoorbeeld een (GWW) project geconfronteerd met behoefte aan hermeting.

Daarnaast is het in sommige gevallen verplicht om een bepaalde actualiteit te handhaven vanwege een wettelijke actualiteitseis.

Voorbeeld van controleregels

- ✓ Het verschil tussen inwindatum en huidige datum is kleiner dan een door de data-eigenaar of gebruiker aangegeven maximum.



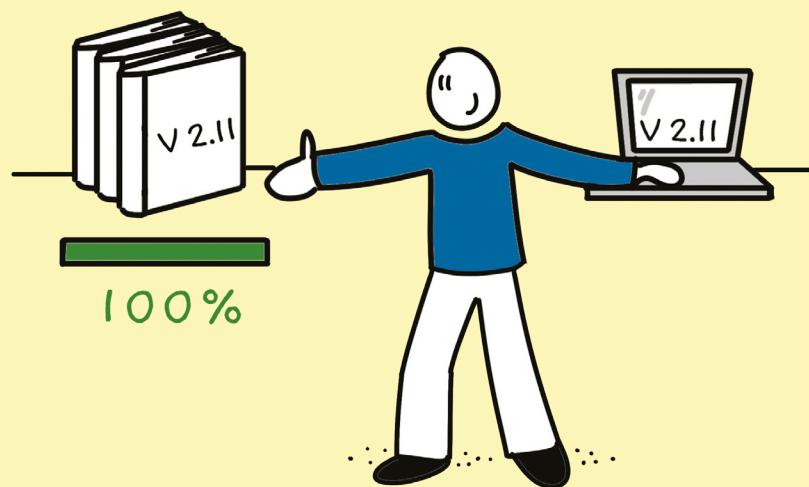
Gerelateerd aan begrip

→ Dataverval; Levensduur; Geldigheid; Recent; Houdbaarheid

ISO

- 8000 -
- 25012 Currentness
- 19157 -

Beheerbaarheid



Beheerbaarheid

De mate waarin de beheerder in staat is zijn taken goed uit te voeren o.a. herkomst, bewerkingen en documentatie.

INHOUD		BEHEER	GEBRUIK
CONFORMITEIT	(OVER) COMPLEETHEID	BEHEER-BAARHEID	BESCHIK-BAARHEID
Formaat-consistentie	Compleetheid per object	Wijzigbaarheid	Vindbaarheid
Domein-consistentie	Object-aanwezigheid	Herstelbaarheid	Benaderbaarheid
Temporele validiteit	Historie & frequentie	Controleerbaarheid	Bestendigheid
Geometrische correctheid	Ruimtelijke dekking	Herleidbaarheid tot bronnen	Tijdigheid (van het proces)
Opslagprecisie	Metadata compleetheid	Reproduceerbaarheid	Betaalbaarheid
Leesbaarheid		Documentatie	Retentie
		Rolinvulling	
SAMENHANG	JUISTHEID	VERTROUWELIJKHEID	BRUIKBAARHEID
Referentiele integriteit	Temporele juistheid	Afscherming van onbevoegden	Reputatie van de bron (leverancier)
Conceptuele consistentie	Positionele juistheid	Bescherming persoonsgegevens	Legitimiteit voor gebruiksdoel
Temporele consistentie	Kwantitatieve juistheid	Rubricering/classificatie	Gebruiksgemak
Topologische consistentie	Actuele juistheid		Meet-nauwkeurigheid
Statistische consistentie			Representativiteit
Uniciteit			Koppelbaarheid
			Inzicht in gebruik

Wijzigbaarheid

De mate waarin gegevens kunnen worden aangepast zonder gevaren voor consistentie (o.a. databasennormalisatie).

Waarom belangrijk

Eenzijds is het belangrijk om gegevens te kunnen wijzigen, bijvoorbeeld wanneer een fout boven water komt of wanneer de werkelijkheid verandert. Anderzijds kunnen wijzigingen zomaar leiden tot allerlei problemen met de interne consistentie of verderop in de keten.

Denk o.a. aan afspraken, procedures en technische maatregelen.

Databasennormalisatie helpt bij het 'eenmalig vastleggen, meervoudig gebruiken'.

Voorbeeld van controleregels

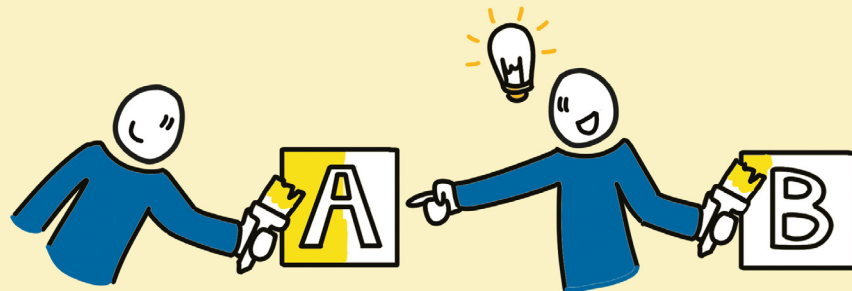
- ✓ Is het datamodel genormaliseerd, d.w.z. wordt elk attribuut van een object maar op een plek vastgelegd (bijv. telnr van een persoon)?
- ✓ Kunnen wijzigingen in deze bron leiden tot inconsistentie met andere bronnen?
- ✓ Is er een procedure om dit te voorkomen?
- ✓ Kunnen ongeautoriseerde personen wijzigingen uitvoeren?

Gerelateerd aan begrip

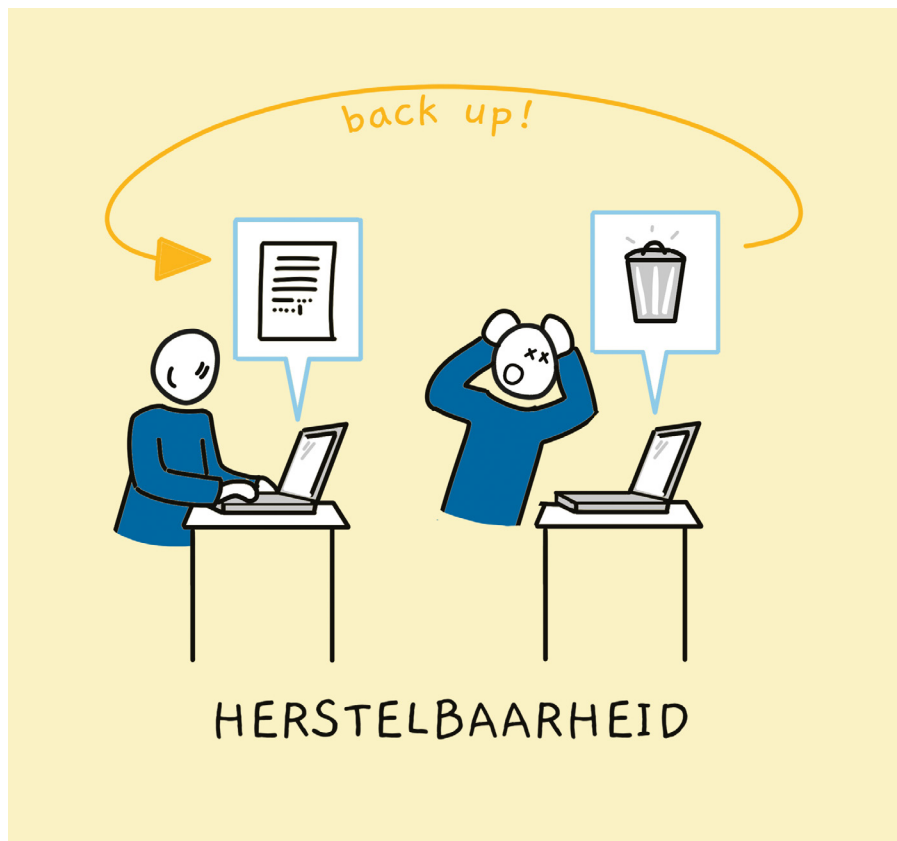
→ Normalisatie; Gegevensordening; Data Lineage

ISO

- 8000 -
- 25012 -
- 19157 -



WIJZIGBAARHEID



Herstelbaarheid

De mate waarin in ieder scenario een mogelijkheid bestaat om terug te gaan naar een vorige versie van de data.

Waarom belangrijk

Als data is overschreven, beschadigd of verdwenen, is het belangrijk om terug te kunnen vallen op een eerdere versie van deze data. Goed proces voor het maken van een back-up en voor versiebeheer speelt hierbij een belangrijke rol.

Voorbeeld van controleregels

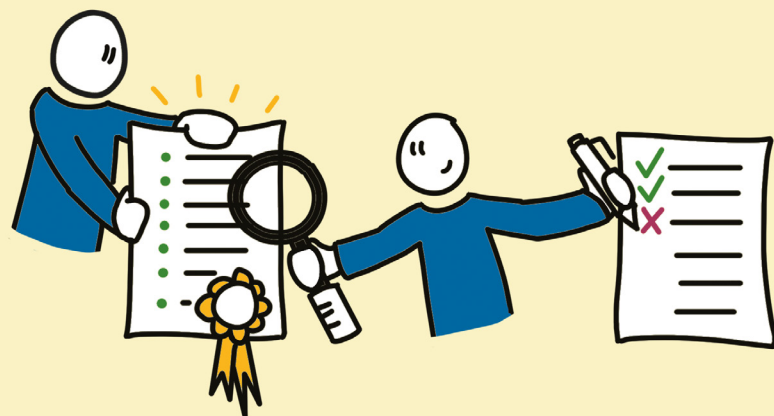
- ✓ Wordt iedere wijziging opgeslagen en gelogd zodat de wijziging kan worden teruggedraaid?
- ✓ Is er een robuuste backup procedure en wordt deze uitgevoerd? Ter verduidelijking de volgende hulpvragen:
 - Wordt de procedure toegepast?
 - Wordt de procedure periodiek herzien?
 - Wordt de backup ergens anders dan de databron, veilig bewaard?
 - Zijn de backup intervallen (dagelijks, wekelijks) toereikend?

Gerelateerd aan begrip

→ -

ISO

- 8000 -
- 25012 Recoverability
- 19157 -



CONTROLEERBAARHEID

Controleerbaarheid

De mate waarin bekend is aan welke kwaliteitsregels de data moeten voldoen, en meting en monitoring mogelijk zijn.

Waarom belangrijk

De organisatie kan alleen 'in control' komen als meetbaar wordt hoe het staat met de (data)kwaliteit, en daaromheen ook de juiste acties worden genomen. Dit vereist allereerst dat alle relevant kwaliteitsregels zijn opgesteld en vastgelegd, maar ook dat de scores op die kwaliteitsregels periodiek beschikbaar zijn, zodat gericht kan worden gestuurd op verbetering.

Voorbeeld van controleregels

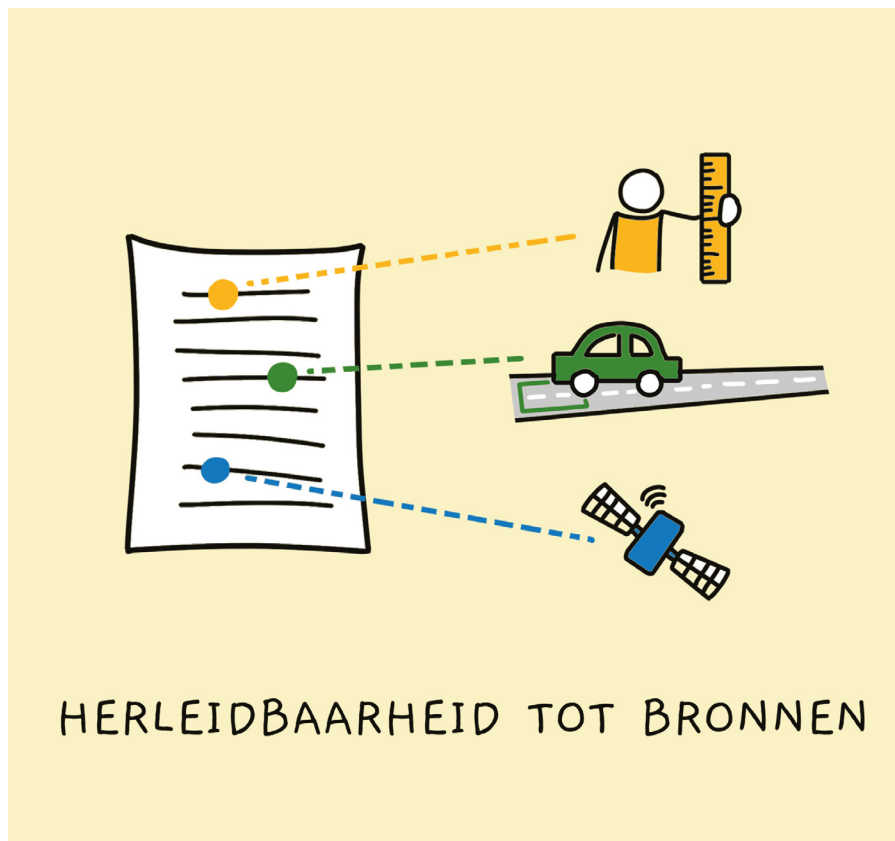
- ✓ Zijn de eisen van alle stakeholders omgezet in datakwaliteitsregels?
- ✓ Zijn alle van toepassing zijnde standaarden, kaders en regelgeving omgezet in datakwaliteitsregels?
- ✓ Wordt de datakwaliteit al (deels) gemonitord?
- ✓ Zijn de productiedata structureel beschikbaar buiten de bron voor data kwaliteitsanalyse en andere analysedoeleinden?

Gerelateerd aan begrip

- Transparant
- Overdraagbaar

ISO

- 8000 -
- 25012 Recoverability
- 19157 -



Herleidbaarheid tot bronnen

De mate waarin valt te achterhalen wie/wat het gegeven waar op welke manier heeft gemeten/vastgelegd.

Waarom belangrijk

Herleidbaarheid tot bronnen gaat over het inzichtelijk hebben van de route van data door de organisatie. Dit speelt bijvoorbeeld een rol bij het correct interpreteren van data, bij het beoordelen van de geloofwaardigheid maar ook bij het herstellen en voorkomen van (herhaaldelijke) fouten.

Voorbeeld van controleregels

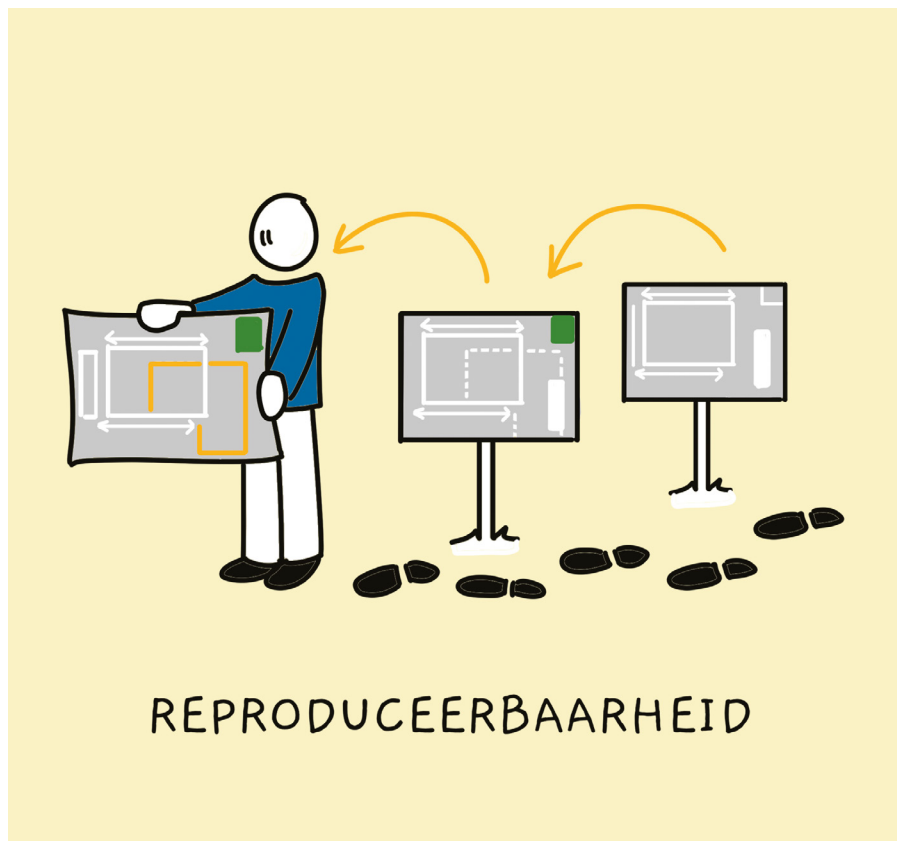
- ✓ Is in de praktijk met beperkte inspanning te achterhalen wat de oorspronkelijke bronnen zijn van de data?
- ✓ Is herleidbaar wie/wat sinds de oorspronkelijke vastlegging mutaties heeft uitgevoerd?

Gerelateerd aan begrip

→ Transparant; Traceerbaar; Tijdrefereerbaar; Data lineage

ISO

- | | |
|---------|--------------|
| → 8000 | Provenance |
| → 25012 | Traceability |
| → 19157 | - |



Reproduceerbaarheid

De mate waarin eenduidig wordt vastgelegd welke bewerkingen op de data zijn uitgevoerd.

Waarom belangrijk

Het vastleggen van de uitgevoerde bewerkingen is onder andere een vorm van kennis- en continuïteitsborging. Daarnaast maakt het mogelijk om data na te rekenen (bijvoorbeeld in een steekproef) en kan het helpen bij het interpreteren van resultaten.

Voorbeeld van controleregels

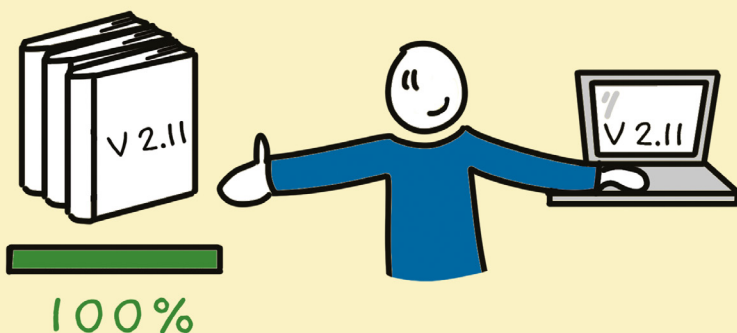
- ✓ Wordt vastgelegd welke bewerkingen hebben plaatsgevonden (bijv. koppelingen, filters, schoningsacties, (de)coderingen?)
- ✓ Is het voor de gebruiker mogelijk om te herleiden hoe de data tot stand zijn gekomen?
- ✓ Kan de data opnieuw berekend worden na constatering van fouten?

Gerelateerd aan begrip

→ Transparant; Overdraagbaar; Data lineage

ISO

- 8000 -
- 25012 Portability
- 19157 -



DOCUMENTATIE

Documentatie

De mate waarin alle vereiste documentatie aanwezig, up-to-date en beschikbaar is.

Waarom belangrijk

Goede documentatie is van vitaal belang om de data en het datamodel op langere termijn goed te kunnen beheren, bijvoorbeeld in het geval van structurele wijzigingen. Daarnaast helpt documentatie bij het invoeren van de juiste data. Verder helpt het vanuit organisatieperspectief als de bron is aangemeld bij de dataadministratie en voorzien van de gevraagde data. Alleen dan kunnen de data goed in samenhang met andere bronnen worden beheerd.

Voorbeeld van controleregels

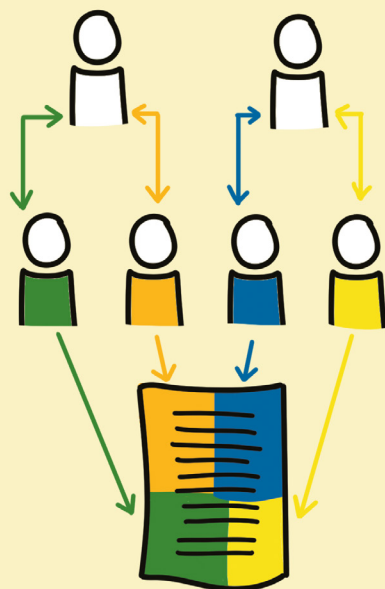
- ✓ Is de documentatie eenvoudig toegankelijk voor een geïnteresseerde?
- ✓ Welke van de onderstaande soorten documentatie zijn aanwezig en up to date? (beantwoorden in subvragen per verwacht document)

Gerelateerd aan begrip

→ Specificaties

ISO

- 8000 -
- 25012 -
- 19157 -



ROLINVULLING

Rolinvulling

De mate waarin de benodigde activiteiten zijn vastgelegd in rolverdelingen en ook worden uitgevoerd.

Waarom belangrijk

De benodigde taken ten behoeve van de beheerbaarheid en de daarbij behorende verantwoordelijkheden zijn essentieel voor de beheerbaarheid van de databron. Als taken niet zijn belegd dan kunnen ze wel of niet worden uitgevoerd. Dat is een risico voor kwaliteit van de data.

Voorbeeld van controleregels

- ✓ Huidige situatie: Zijn bij een eigen data governance rollen bekend en toegewezen?
- ✓ Toekomstige situatie: Zijn alle data governance Rijkswaterstaat rollen bekend en toegewezen (DRoM-R)?
- ✓ Is het FB georganiseerd volgens het vastgestelde uniforme FB proces van RWS?
- ✓ Is bekend wie de servicemanager is?
- ✓ Is bekend wie de informatiemanager is?
- ✓ Is Troux volledig en goed gevuld met rollen?
- ✓ Monitort en stuurt het management daadwerkelijk op verbetering?

Gerelateerd aan begrip

→ Governance

ISO

- 8000 -
- 25012 -
- 19157 -

Vertrouwelijkheid



Vertrouwelijkheid

De mate waarin wordt voldaan aan alle eisen rondom afscherming, gevoeligheid en classificatie.

INHOUD		BEHEER	GEBRUIK
CONFORMITEIT	(OVER) COMPLEETHEID	BEHEER-BAARHEID	BESCHIK-BAARHEID
Formaat-consistentie	Compleetheid per object	Wijzigbaarheid	Vindbaarheid
Domein-consistentie	Object-aanwezigheid	Herstelbaarheid	Benaderbaarheid
Temporele validiteit	Historie & frequentie	Controleerbaarheid	Bestendigheid
Geometrische correctheid	Ruimtelijke dekking	Herleidbaarheid tot bronnen	Tijdigheid (van het proces)
Opslagprecisie	Metadata compleetheid	Reproduceerbaarheid	Betaalbaarheid
Leesbaarheid		Documentatie	Retentie
		Rolinvulling	
SAMENHANG	JUISTHEID	VERTROUWELIJKHEID	BRUIKBAARHEID
Referentiele integriteit	Temporele juistheid	Afscherming van onbevoegden	Reputatie van de bron (leverancier)
Conceptuele consistentie	Positionele juistheid	Bescherming persoonsgegevens	Legitimiteit voor gebruiksdoel
Temporele consistentie	Kwantitatieve juistheid	Rubricering/classificatie	Gebruiksgemak
Topologische consistentie	Actuele juistheid		Meet-nauwkeurigheid
Statistische consistentie			Representativiteit
Uniciteit			Koppelbaarheid
			Inzicht in gebruik

Afscherming van onbevoegden

De mate waarin de data elektronisch of fysiek alleen toegankelijk zijn voor bevoegden.

Waarom belangrijk

Afscherming van onbevoegden is belangrijk voor het voorkomen van lekken en diefstal van gegevens en voor het voldoen aan privacywetgeving. Onder deze dimensie valt ook het autorisatieproces, dat cruciaal is om gebruikers te voorzien van de juiste informatie (en niet meer dan dat) en het overzicht vormt van wie bij welke data mag.

Voorbeeld van controleregels

- ✓ Bestaan er voldoende mogelijkheden om selectief toegang te verlenen aan gebruikers/afnemers?
- ✓ Bestaat een overzicht van autorisaties voor de applicaties en data-sets?
- ✓ Worden de toegekende rechten regelmatig herzien, bijv. o.b.v. gebruik?
- ✓ Vind er periodiek een audit plaats op het autorisatieproces?
- ✓ Worden wachtwoorden regelmatig vernieuwd?
- ✓ Worden er periodiek security (penetratie) testen uitgevoerd?

Gerelateerd aan begrip

→ Veiligheid; Beveiligingsniveau; Afgrendeling

ISO

- | | |
|---------|-------------------------|
| → 8000 | Accessibility; Security |
| → 25012 | - |
| → 19157 | - |



AFSCHERMING VAN ONBEVOEGDEN



BESCHERMING PERSOONSgegevens



Bescherming persoonsgegevens

De mate waarin de juiste maatregelen getroffen zijn om privacy te waarborgen.

Waarom belangrijk

Afscherming van persoonsgegevens is een wettelijke verplichting (AVG: algemene verordening gegevensbescherming). Daarnaast zorgt het voor grote imagoschade wanneer persoonsgegevens openbaar worden. Let op: het is vaak de combinatie van data die informatie over individuen bloot kan leggen.

Voorbeeld van controleregels

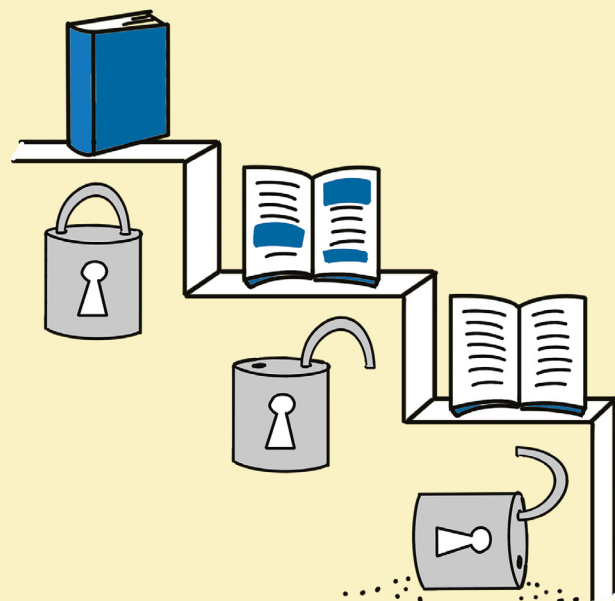
- ✓ Bevat de databron data die direct of indirect, zoals na het koppelen met andere databronnen, tot individuen is te herleiden? De data bevat persoonsgegevens: worden deze geanonimiseerd bij uitlevering?
- ✓ Worden gebruikers op de hoogte gesteld van de privacygevoelige aard? Wat gebeurt met persoonsgegevens die niet meer voor het doel noodzakelijk zijn? (Bewaard; geanonimiseerd; verwijderd?) Is het aantoonbaar dat het informatieverwerkingsproces conform de AVG wordt uitgevoerd?
- ✓ Is de databron aangemeld bij de RWS Privacy Coördinator?
- ✓ Is er een privacy assessment beschikbaar? (PIA/GEB)
- ✓ Is er een proces ingericht waarmee burgers en bedrijven het recht op inzage van hun data kunnen benutten?

Gerelateerd aan begrip

→ Privacy; Retentie

ISO

- 8000 -
- 25012 -
- 19157 -



RUBRICERING/CLASSIFICATIE

Rubricering / classificatie

De mate waarin gevoelige data op de juiste manier zijn geclassificeerd en dit geborgd is in het proces.

Waarom belangrijk

Data-analyses worden steeds breder toegepast. Het is daarom belangrijk om integraal te borgen dat gevoelige data als zodanig gemarkeerd en gecommuniceerd wordt. Het gaat naast persoonsgegevens ook om commercieel en/of strategisch gevoelige informatie.

Voorbeeld van controleregels

- ✓ Als de databron gevoelige data bevat, over welk onderwerp gaan de data? (commercieel; strategisch; persoonsgegevens; missiekritisch)
- ✓ Zijn alle onderdelen van de databron voorzien van een classificatie?
- ✓ Wordt daarbij de richtlijn van classificeren vanuit het securitycenter toegepast inclusief beslisboom?
- ✓ Is de classificatie geaccordeerd door de verantwoordelijke?
- ✓ Wordt de classificatie regelmatig getoetst / herzien?
- ✓ Mag de data openbaar gemaakt worden via een WOB-verzoek?
- ✓ Worden de gebruikers op de hoogte gesteld van de classificatie van de data?

Gerelateerd aan begrip

→ Vertrouwelijk; Legitimiteit voor gebruiksdoel

ISO

- | | |
|---------|-------------------------|
| → 8000 | Accessibility; Security |
| → 25012 | Confidentiality |
| → 19157 | - |

Beschikbaarheid



Bruikbaarheid

De mate waarin de juiste persoon op het juiste moment tegen de juiste kosten over de data kan beschikken.

INHOUD		BEHEER	GEBRUIK
CONFORMITEIT	(OVER) COMPLEETHEID	BEHEER-BAARHEID	BESCHIK-BAARHEID
Formaat-consistentie	Compleetheid per object	Wijzigbaarheid	Vindbaarheid
Domein-consistentie	Object-aanwezigheid	Herstelbaarheid	Benaderbaarheid
Temporele validiteit	Historie & frequentie	Controleerbaarheid	Bestendigheid
Geometrische correctheid	Ruimtelijke dekking	Herleidbaarheid tot bronnen	Tijdigheid (van het proces)
Opslagprecisie	Metadata compleetheid	Reproduceerbaarheid	Betaalbaarheid
Leesbaarheid		Documentatie	Retentie
		Rolinvulling	
SAMENHANG	JUISTHEID	VERTROUWELIJKHEID	BRUIKBAARHEID
Referentiele integriteit	Temporele juistheid	Afscherming van onbevoegden	Reputatie van de bron (leverancier)
Conceptuele consistentie	Positionele juistheid	Bescherming persoonsgegevens	Legitimiteit voor gebruiksdoel
Temporele consistentie	Kwantitatieve juistheid	Rubricering/classificatie	Gebruiksgemak
Topologische consistentie	Actuele juistheid		Meet-nauwkeurigheid
Statistische consistentie			Representativiteit
Uniciteit			Koppelbaarheid
			Inzicht in gebruik



Vindbaarheid

De mate waarin gebruikers data met passend gemak kunnen lokaliseren.

Waarom belangrijk

Vindbaarheid is vooral belangrijk om zoveel mogelijk waarde uit de data te halen. Data kan goed op orde zijn, maar als de data moeilijk vindbaar zijn zal het gebruik beperkt blijven. Dit speelt vooral in de wereld van open data, maar ook intern RWS.

Voorbeeld van controleregels

- ✓ Is er voor elk gegeven centraal geregistreerd wat de officiële bron is van deze data-assets (waar de echte waarheid zou moeten worden vastgelegd)?
- ✓ Zijn de data uit de databron vindbaar via Dataportaal/DMC/ServiceDesk/www.dataoverheid.nl data?
- ✓ Zijn de data op een andere manier vindbaar?

Gerelateerd aan begrip

→ -

ISO

- 8000 -
- 25012 Availability
- 19157 -



BENADERBAARHEID

Benaderbaarheid

De mate waarin gebruikers, na lokalisering, met passend gemak toegang kunnen verkrijgen tot de data.

Waarom belangrijk

Het spreekt voor zich dat gebruikers niets aan de data hebben als ze er niet daadwerkelijk bij kunnen. Hoe beter de gebruiker wordt gefaciliteerd in het benaderen of ontsluiten van data, des te meer waarde er uit de data kan worden gehaald en worden verspillingen voorkomen.

Voorbeeld van controleregels

- ✓ Als een gebruiker de databron heeft gelokaliseerd, is dan duidelijk welke stappen hij moet doorlopen om daadwerkelijk toegang te krijgen?
- ✓ Ervaren de gebruikers problemen bij het benaderen van de data (DMC, blijven niet hangen, up tijd, geen lange laadtijden, etc)? Ook bij piekbelasting?

Gerelateerd aan begrip

→ Toegankelijkheid

ISO

- 8000 -
- 25012 Accessibility
- 19157 -



Bestendigheid

De mate waarin techniek en organisatie zijn voorbereid op incidenten en calamiteiten.

Waarom belangrijk

De organisatie moet ervan uit kunnen gaan dat de data ook beschikbaar blijven (of kunnen worden gemaakt) na een incident of calamiteit. Bestaande processen en best practices spelen hierin een rol, maar ook de invulling van rollen en opleiding van betrokkenen. Risicobeheersing is hierbij een belangrijk onderwerp.

Voorbeeld van controleregels

- ✓ Zijn de grootste risico's in kaart gebracht en maatregelen getroffen?
- ✓ Zijn er processen ingericht voor probleem- en incidentmanagement? O.b.v. ITIL?
- ✓ Wordt Topdesk gebruikt voor het aanmelden van deze data issues?
- ✓ Is er een proces ingericht om datalekken tijdig te kunnen melden?
- ✓ Kennen de rolhouders hun taken, verantwoordelijkheden en bevoegdheden in geval van incidenten en calamiteiten?

Gerelateerd aan begrip

→ -

ISO

- 8000 -
- 25012 -
- 19157 -

Tijdigheid (van het proces)

De mate waarin doorlooptijden en frequentie van de dataverwerking voldoen aan de eisen vanuit het gebruik.

Waarom belangrijk

Doorlooptijden en stiptheid zijn typisch zaken die veel bijdragen aan de waarde van data in een proces. Repeterende metingen: moeten altijd op tijd en vaak genoeg volgens de wens van de gebruiker. Statische aspecten: up to date blijven met de werkelijkheid (processen om de data tijdig bij te werken in het geval van veranderingen).

Voorbeeld van controleregels

- ✓ In het geval van repeterende metingen: in hoeverre is de gebruiker tevreden met de tijdigheid van de geleverde data? En met de frequentie?
- ✓ In het geval van statische aspecten: in hoeverre worden de wijzigingen van de werkelijkheid binnen acceptabele tijd verwerkt in de data?

Gerelateerd aan begrip

→ Frequentie (vaak genoeg?)

ISO

→ 8000	Timeliness
→ 25012	-
→ 19157	-



TIJDIGHEID (VAN HET PROCES)



Betaalbaarheid

De mate waarin de kosten om de data te kunnen gebruiken acceptabel zijn voor de gebruiker of eigenaar.

Waarom belangrijk

Ook de kosten voor de gebruiker/eigenaar zijn bepalend voor de 'geschiktheid voor gebruik' waar datakwaliteit uiteindelijk om draait. Als de kosten voor veel gebruikers een te hoge drempel blijken, kun je je afvragen of dit in overeenstemming is met de gestelde doelen. Anderzijds kan inzicht in de kosten helpen om processen verder te stroomlijnen en zo voor iedereen de kosten omlaag te brengen.

Voorbeeld van controleregels

- ✓ Is het bekend wat het kost om deze data in te winnen, te beheren en te ontsluiten?
- ✓ Zijn de kosten om de gegevens te verkrijgen acceptabel voor de eigenaar?
- ✓ In hoeverre zie je verspillingen in het inwinproces (bijv: is de aanlevering uniform?)?

Gerelateerd aan begrip

→ -

ISO

- | | |
|---------|---------------|
| → 8000 | Cost; Benefit |
| → 25012 | - |
| → 19157 | - |



Retentie

De mate waarin data na registratie bewaard blijven.

Waarom belangrijk

In allerlei gevallen is het handig om terug te kunnen gaan naar historische data: registraties van lang geleden, zelfs als die geen betrekking meer hebben op de huidige werkelijkheid. Hier zijn echter wel (steeds meer) wettelijke beperkingen aan gesteld: bepaalde data mag niet langer worden opgeslagen dan strikt nodig vanuit het doel waarmee de data zijn verzameld. Zowel technische als procesmatige oplossingen zijn denkbaar om hier invulling aan te geven.

Voorbeeld van controleregels

- ✓ Is er beleid met betrekking tot retentie opgesteld? Zo ja, wordt dit beleid nageleefd?
- ✓ Tot hoever terug (in jaren) gaat de historie van data?
- ✓ Indien van toepassing (persoonsgegevens): wordt dan aantoonbaar voldaan aan de wettelijke kaders van bewaartermijnen?
- ✓ Geeft het systeem een signaal als de termijn verstreken is?

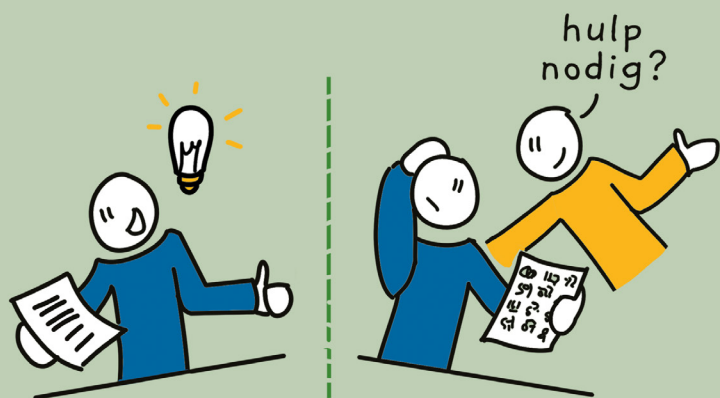
Gerelateerd aan begrip

→ -

ISO

- 8000 -
- 25012 -
- 19157 -

Bruikbaarheid



Bruikbaarheid

Worden gebruikers voldoende ondersteund om de data te gebruiken en op waarde te schatten?

INHOUD		BEHEER	GEBRUIK
CONFORMITEIT	(OVER) COMPLEETHEID	BEHEER-BAARHEID	BESCHIK-BAARHEID
Formaat-consistentie	Compleetheid per object	Wijzigbaarheid	Vindbaarheid
Domein-consistentie	Object-aanwezigheid	Herstelbaarheid	Benaderbaarheid
Temporele validiteit	Historie & frequentie	Controleerbaarheid	Bestendigheid
Geometrische correctheid	Ruimtelijke dekking	Herleidbaarheid tot bronnen	Tijdsigheid (van het proces)
Opslagprecisie	Metadata compleetheid	Reproduceerbaarheid	Betaalbaarheid
Leesbaarheid		Documentatie	Retentie
		Rolinvulling	
SAMENHANG	JUISTHEID	VERTROUWELIJKHEID	BRUIKBAARHEID
Referentiele integriteit	Temporele juistheid	Afscherming van onbevoegden	Reputatie van de bron (leverancier)
Conceptuele consistentie	Positionele juistheid	Bescherming persoonsgegevens	Legitimiteit voor gebruiksdoel
Temporele consistentie	Kwantitatieve juistheid	Rubricering/classificatie	Gebruiksgemak
Topologische consistentie	Actuele juistheid		Meet-nauwkeurigheid
Statistische consistentie			Representativiteit
Uniciteit			Koppelbaarheid
			Inzicht in gebruik

Reputatie van de bron (leverancier)

De mate waarin de herkomst van de data vertrouwen geeft in de juistheid van de inhoud.

Waarom belangrijk

Slechte dataleveranciers kunnen dan in de toekomst worden uitgesloten van het uitvoeren van werkzaamheden voor RWS. Dit kan op langere termijn veel datakwaliteitsproblemen voorkomen.

Voorbeeld van controleregels

- ✓ Is de leverancier gekoppeld aan de door hem/haar geleverde data?
- ✓ Zijn er leveranciers die structureel data leveren van onvoldoende kwaliteit (bijv: externe opdrachtnemer, of intern afdeling of IPM team)?
- ✓ In geval van interne leveringen: vormt deze info input voor een verbetercyclus?

Gerelateerd aan begrip

- Vertrouwen
- Geloofwaardig

ISO

- 8000 -
- 25012 Credibility
- 19157 -



Legitimiteit voor gebruikdoel

De mate waarin het doel van inwinning en de beperkingen in het gebruik duidelijk worden gecommuniceerd.

Waarom belangrijk

Doelbinding is een belangrijk concept om op integere wijze met data om te kunnen gaan. Het feit dat data binnen een organisatie beschikbaar zijn, betekent niet dat deze altijd voor elk doel kunnen worden ingezet. Er kan imagoschade of persoonlijke schade optreden als blijkt dat de data voor allerlei andere doeleinden worden ingezet dan het doel dat is gecommuniceerd bij de inwinning. Eigenaarschap van data is hier een belangrijk aspect: de eigenaar moet zeggenschap hebben waar de data wel of niet.

Voorbeeld van controleregels

- ✓ Is in de bron gedocumenteerd met welk doel en in welke context de data zijn ingewonnen of worden uitgewisseld?
- ✓ Worden de beperkingen gecommuniceerd naar alle gebruikers?
- ✓ Is het inzichtelijk voor welke doelen de data worden gebruikt?
- ✓ Heeft de eigenaar (binnen RWS) zeggenschap over de doelen waarvoor de data wel of niet mogen worden gebruikt?
- ✓ Bestaan er inmiddels beter geïnformeerde bronnen?
- ✓ Worden gegevens uit deze databron gebruikt die eigenlijk uit basisregistraties gehaald moeten worden?

Gerelateerd aan begrip

- Doelbinding; Relevant; Geschikt; Toepasbaar; Passend;
- Proportioneel; Transparant; Aanspreekbaar; Rechtmatig;
- Wettelijke grondslag; Ethiek

ISO

- | | |
|---------|-----------|
| → 8000 | Relevance |
| → 25012 | - |
| → 19157 | - |

LEGITIMITEIT VOOR GEBRUIKSDOEL

Gebruiksgemak

De mate waarin gebruikers met de data aan de slag kunnen zonder diepe kennis van de inhoud of de techniek.

Waarom belangrijk

Data, zeker in ruwe formaten, spreken lang niet altijd voor zich. Om succesvol gebruik te faciliteren, is het allereerst belangrijk dat de data zelf al zo begrijpelijk mogelijk zijn. Daarnaast helpt het natuurlijk om de gebruiker uitleg te geven van wat hij ziet (dit kan op verschillende manieren), en bij verdere vragen gebruikersondersteuning paraat te hebben staan. Het gaat hier zowel om de aspecten "bij de databron zelf" (leesbaarheid, metadata) als om aspecten verderop in de keten van ontsluiting (formats, services).

Voorbeeld van controleregels

In het geval dat de gebruikers rechtstreeks bij de databron data benaderen en/of in het geval dat data wordt ontsloten via DMC/dataportalen:

- ✓ Is er een goed bruikbaar gebruikersinterface?
- ✓ Is er gebruikersondersteuning ingericht?
- ✓ In hoeverre zijn de metadata beschikbaar voor gebruikers?

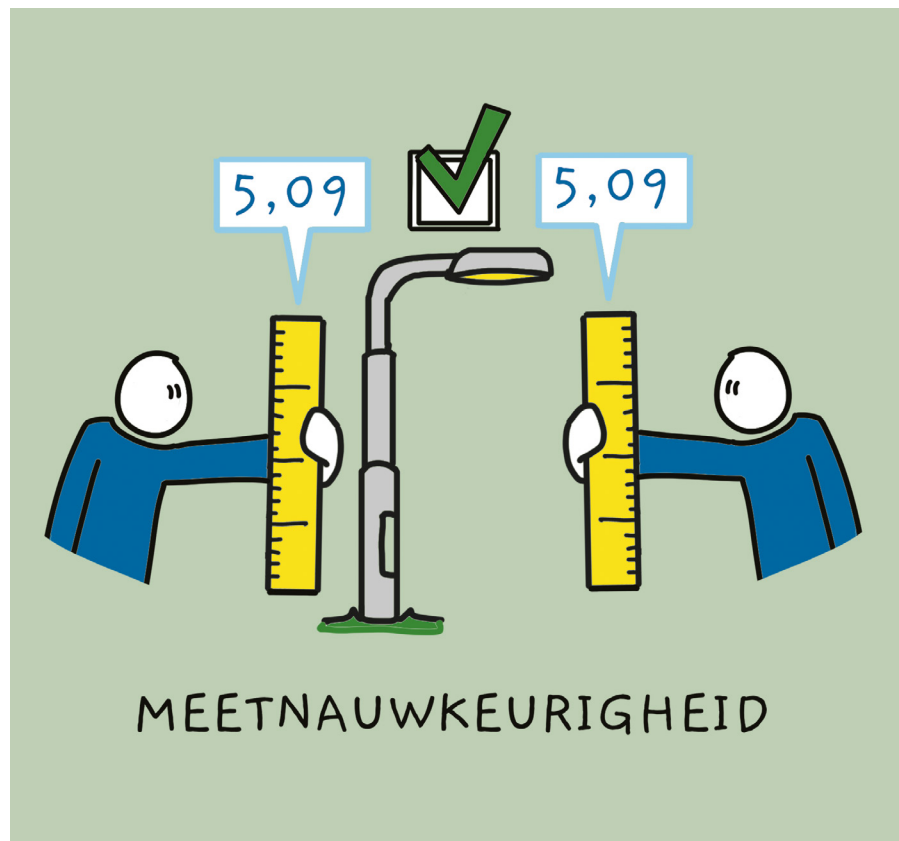
Gerelateerd aan begrip

→ Toegankelijkheid

ISO

- 8000 -
- 25012 Accessibility; Efficiency
- 19157 -





Meetnaauwkeurigheid

De mate waarin metingen of berekeningen bij herhaling dezelfde waarde opleveren.

Waarom belangrijk

Deze dimensie gaat erover hoe zeker we weten dat een meting of waarneming de juiste waarde genereert. Een meetwaarde van 0,74632 zegt bijvoorbeeld niets als er bij de volgende meting 0,75230 wordt gemeten: maar in beide gevallen is het 0,75. Handmatige (subjectieve) metingen leiden ook snel tot foutjes: bijvoorbeeld een lantaarnpaal die in werkelijkheid een meter verderop staat. Uiteindelijk is het vooral belangrijk dat de betrouwbaarheid van een meting volstaat voor de wensen van de gebruiker, en dat deze voor de gebruiker bekend is.

Voorbeeld van controleregels

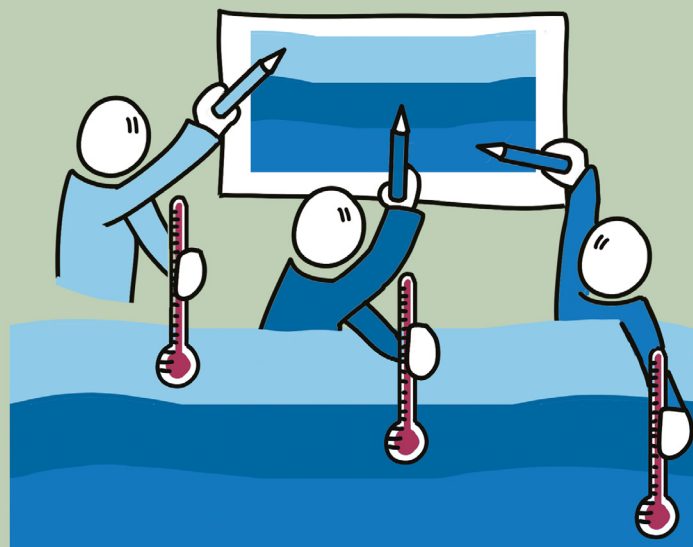
- ✓ Is er een betrouwbaarheidsinterval bekend voor de gebruiker?
- ✓ In hoeverre komen foute meetwaarden voor in de bron en heeft dit impact op gebruikers?
- ✓ Zitten er onnodige menselijke stappen in het inwinnen en bewerken van de data?

Gerelateerd aan begrip

→ Herhaalbaarheid; Betrouwbaarheid

ISO

- 8000 -
- 25012 -
- 19157 -



REPRESENTATIVITEIT

Representativiteit

De mate waarin de databron een goede weergave geeft voor de situatie (Omgevingswet).

Waarom belangrijk

Deze dimensie speelt een rol bij een databron waarin data is opgeslagen die op basis van relatief schaarse of mogelijk gekleurde waarnemingen is opgenomen. Zoals bijvoorbeeld waarnemingen van diersoorten of metingen van waterstanden die worden geëxtrapoleerd naar de gehele rivierloop. In zulke gevallen is het goed te bepalen of de data voldoende representatief zijn en te bepalen dat meer of andersoortige metingen nodig zijn voor het beoogde doel.

Voorbeeld van controleregels

- ✓ Voor zover bekend: zouden de gebruikers vanuit het oogpunt van representativiteit willen dat er op meer plekken of frequenter gemeten gaat worden?
- ✓ Indien metingen: Bestaat er statistische (of andere) onderbouwing voor op hoeveel plekken en hoe vaak gemeten wordt?

Gerelateerd aan begrip

→ Granulariteit

ISO

- 8000 -
- 25012 -
- 19157 -

Koppelbaarheid

De mate waarin de kenmerken van de databron het mogelijk maken om data te koppelen aan andere data.

Waarom belangrijk

De organisatie legt vaak op meerdere plekken data vast over dezelfde objecten. Als achteraf relaties tussen die data gemaakt moeten worden, helpt het enorm als gebruik is gemaakt van gedeelde identificatiecodes of gestandaardiseerde namen, die ook buiten dit systeem gebruikt worden, en zodoende aan andere data gekoppeld kunnen worden. Ook het (kunnen) vastleggen van gegevens over ruimte (georefereerbaarheid) en tijd (tijdfereerbaarheid) maken het mogelijk om gegevens aan elkaar te koppelen.

Voorbeeld van controleregels

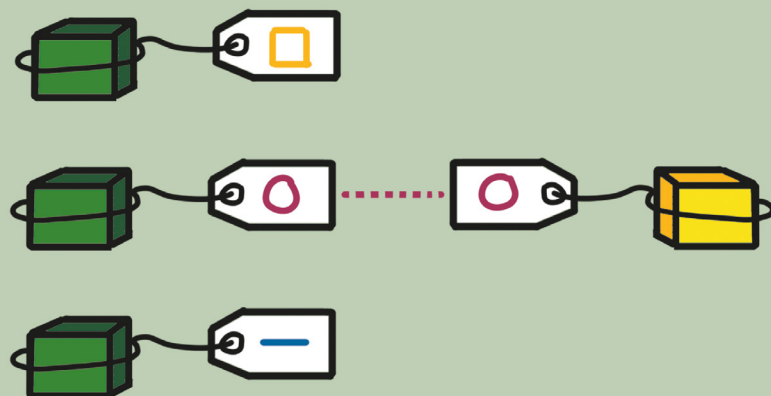
- ✓ Liggen er mogelijkheden voor verbetering om de data in verdere mate te voorzien van identificatiecodes of gestandaardiseerde namen?
- ✓ Is van elk gegeven in de bron, voor zover van toepassing, informatie vast te leggen over de locatie?
- ✓ Is van elk gegeven in de bron, voor zover van toepassing, informatie vast te leggen over de tijd van meting, vastlegging en mutatie?
- ✓ Wordt er voldaan aan de richtlijnen op het gebied van coördinatensystemen?

Gerelateerd aan begrip

→ Master data; Georefereerbaar; Tijdfereerbaar

ISO

- 8000 -
- 25012 -
- 19157 -



KOPPELBAARHEID



Inzicht in gebruik

De mate waarin bekend is welke personen, activiteiten en processen gebruik maken van de data en welke eisen zij stellen.

Waarom belangrijk

Je kunt alleen gericht tot bruikbare data komen als je weet wat de wensen zijn van je gebruikers. Vaak genoeg weten mensen aan het begin van de keten niet hoe hun data allemaal worden hergebruikt in latere processen. Binnen deze dimensie vallen ook standaarden voor gebruik: immers, als je niet precies weet wie je data allemaal gebruiken kun je op zijn minst zorgen dat de data aan (open) standaarden voldoen.

Voorbeeld van controleregels

- ✓ Is bekend welke (hoofd en sub)processen afhankelijk zijn van de data?
- ✓ Is bekend welke gebruikersgroepen de data daadwerkelijk gebruiken?
- ✓ In hoeverre zijn de datakwaliteitseisen vanuit alle stakeholders bekend?
- ✓ Is bekend welke standaarden, kaders, wetten, regelgeving van toepassing zijn?

Gerelateerd aan begrip

→ Verspreiding

ISO

- 8000 -
- 25012 -
- 19157 -

Datakwaliteitsstelsel

Nummer:	6189
Versienummer standaard:	1.0
Versienummer document:	1
Status:	In beheer
Type:	Kader
Inhoudelijk beheerder:	Ronald Jautze
Verantwoordelijke afdeling:	Afd. Gisanalyse en Procesadvisering Data
Netwerken:	Hoofdvaarwegennet, Hoofdwatersysteem, Hoofdwegennet
Rollen:	
Fase:	Planuitwerking, Verkenning, Realisatie, Onderhoud
Proceseigenaar:	Proceseigenaar Informatievoorziening
Link om te reageren:	Link