



NORA
Nederlandse Overheid
Referentie Architectuur

KOMEN & VIEREN & DELEN & WERKEN & GAAN

Sturen op gegevens Gegevenskwaliteit – een raamwerk vanuit NORA

Door: Wim Stolk (voorzitter Expertgroep Gegevensmanagement NORA).

31 mei 2021

versie 1.0 na publieke consultatie

Jaap van den Berg, Gerald Groot Roessink, Danny Greefhorst, Kasper Kisjes, Wim Stolk

Het doel en de uitgangspunten

- Een gemeenschappelijk raamwerk voor gegevenskwaliteit voor de Nederlandse overheid:
 - Kwaliteitsdimensies, kwaliteitsattributen en voorbeelden
- Gebaseerd op beschikbare theorie, standaarden en best practices
- Uitgaan van eerder resultaat, met name Omgevingswet raamwerk
- Gericht op praktische toepasbaarheid

Basisbegrippen

Kwaliteit: Mate waarin een geheel van eigenschappen en kenmerken van een object voldoet aan eisen (ISO 9000:2015 DIS).

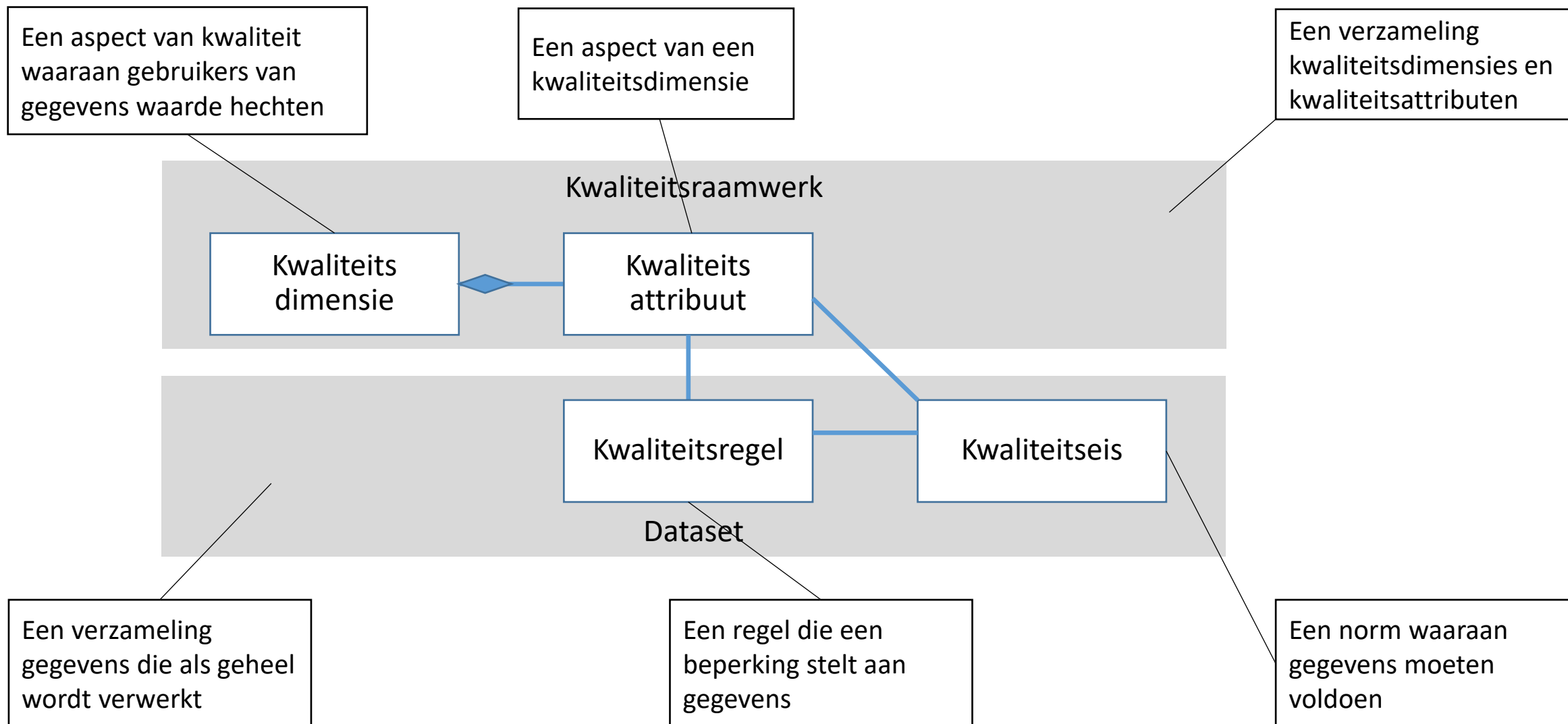
Gegevenskwaliteit: Mate waarin een geheel van eigenschappen en kenmerken van één of meer gegevens voldoet aan eisen.

Gegeven: Een formeel vastgelegd feit, begrip of aanwijzing (NORA)

Het streven is dat gegevenskwaliteit “fit-for-purpose” is; dat deze aansluit bij het gebruik

Kwaliteit heeft betrekking op de mate waarin wordt voldaan aan verwachtingen

Gerelateerde begrippen



Waarom is gegevenskwaliteit belangrijk?

- **Snellere en betere besluitvorming** – betere gegevens leiden tot minder discussie en meer betrouwbare besluiten
- **Meer efficiënte en voorspelbare processen** – betere gegevens voorkomen inefficiënte handelingen en onverwachte verrassingen
- **Compliance** – door betere gegevens kun je beter, efficiënter en met meer vertrouwen rapporteren over het voldoen aan wet- en regelgeving

Afbakening: kwaliteit van de gegevens zelf

Waarde van gegevens
in gebruik (ook wel:
externe kwaliteit)



gegevenskwaliteit

De input

Open standaarden

- ISO/IEC 25012/25024, NEN/ISO 19157, ISO 5725, ISO 8000

De facto standaarden

- DAMA-NL Code voor Informatiekwaliteit, DAMA-NL Dimensions of Data Quality, DAMA DMBok, Conformed Dimensions of Data Quality

Onderzoek

- R.W. Wang, C. Batini, D. McGilvary, Y. Lee, D. Loshin, G. Tejay, C. Besselink, P. van Oort

Zelf-ontwikkelde raamwerken

- Omgevingswet, Rijkswaterstaat, IPO

Ervaringen bij de Nederlandse overheid

- DSO, Kadaster, RIVM, Rijkswaterstaat, RCE, DUO, IPO, Waterschapshuis, VNG, InspectieViews

Basisraamwerk

Juistheid	De mate waarin gegevens de echte waarde goed weergeven.
Compleetheid	De mate waarin gegevens aanwezig zijn.
Validiteit	De mate waarin gegevens voldoen aan de verwachte structuur en opslagvorm.
Consistentie	De mate waarin gegevens vrij van tegenspraak zijn en samenhang vertonen met andere gegevens.
Actualiteit	De mate waarin gegevens recent genoeg zijn.
Precisie	De mate waarin gegevens exact of onderscheidend genoeg zijn.
Plausibiliteit	De mate waarin gegevens worden beschouwd als waar en geloofwaardig door gebruikers.
Traceerbaarheid	De mate waarin de totstandkoming en het gebruik van gegevens zijn vastgelegd.
Begrijpelijkheid	De mate waarin gegevens eenvoudig gelezen en geïnterpreteerd kunnen worden door gebruikers.

Basisraamwerk

Juistheid	Thematische juistheid	Classificatie juistheid				
Compleetheid	Dataset compleetheid	Attribuut compleetheid	Over compleetheid	Historie compleetheid	Metadata compleetheid	Identificeerbaarheid
Validiteit	Formaat validiteit	Domein validiteit				
Consistentie	Logische consistentie	Referentiële integriteit	Homogeniteit			
Actualiteit	Update frequentie	Versheid				
Precisie	Opslagprecisie					
Plausibiliteit	Authenticiteit	Reputatie	Bewijsbaarheid	Waarschijnlijkheid		
Traceerbaarheid	Herleidbaarheid	Reproduceerbaarheid	Gebruiks inzicht			
Begrijpelijkheid	Duidelijkheid					

Extensies voor geografie en statistiek

Juistheid

Compleetheid

Validiteit

Consistentie

Actualiteit

Precisie

Plausibiliteit

Traceerbaarheid

Begrijpelijkheid

Positionele
juistheid

Ruimtelijke
dekking

Geometrische
validiteit

Topologische
consistentie

Geometrische
precisie

Geografie

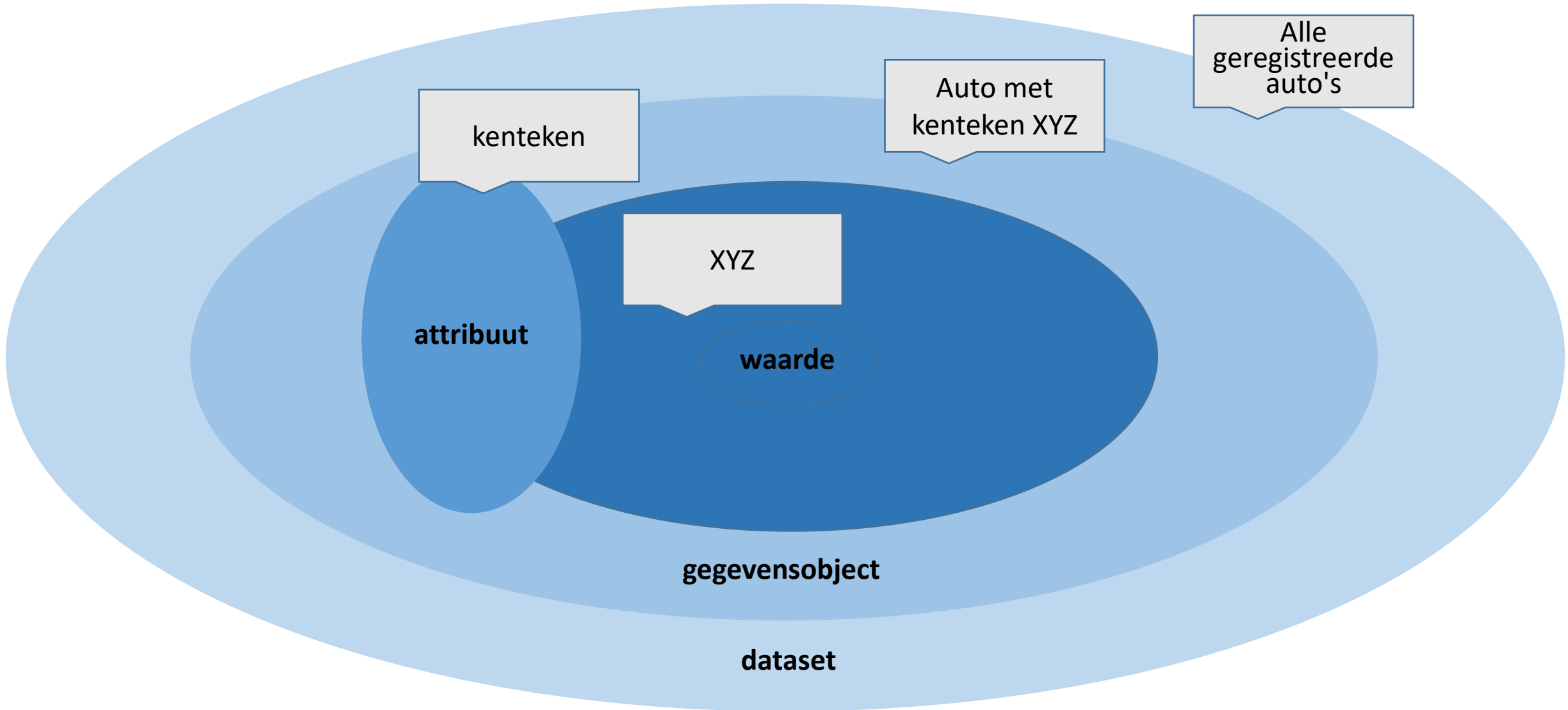
Kwantitatieve
juistheid

Statistische
precisie

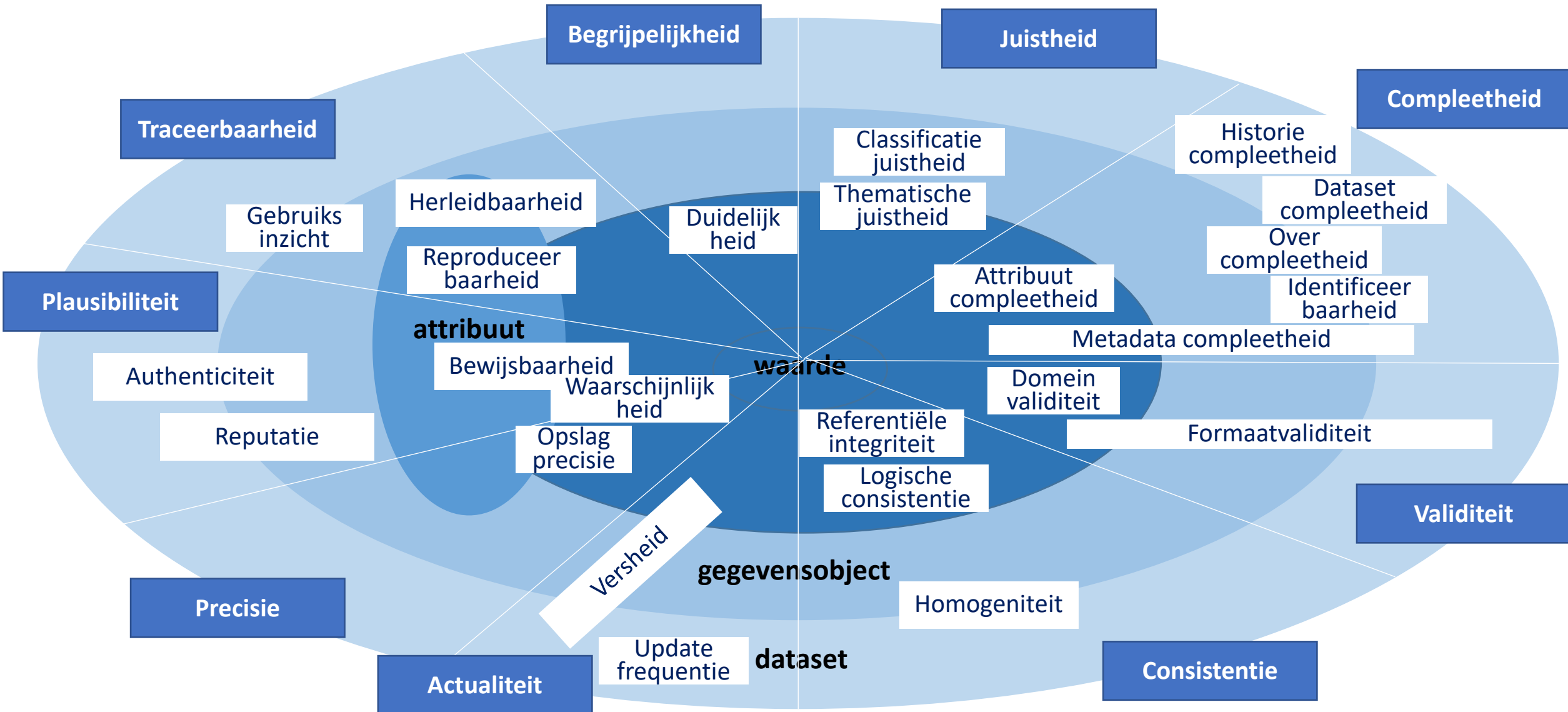
Repre-
sentativiteit

Statistiek

Toepassingsniveaus



Toepassingsniveaus



Definities kwaliteitsattributen basis (1)

Kwaliteitsattribuut	Definitie
Thematische juistheid	De mate waarin kwalitatief beschrijvende gegevens overeenkomen met de werkelijkheid
Classificatie juistheid	De mate waarin gegevensobjecten zijn geïdentificeerd als het juiste objecttype.
Dataset compleetheid	De mate waarin objecten waarvan het bestaan bekend is aanwezig zijn.
Attribuut compleetheid	De mate waarin bij gegevensobjecten waarden aanwezig zijn voor een attribuut.
Overcompleetheid	De mate waarin gegevensobjecten niet onterecht aanwezig zijn.
Historie compleetheid	De mate waarin historische gegevens aanwezig en temporeel dekkend zijn.
Metadata compleetheid	De mate waarin metadata aanwezig zijn.
Identificeerbaarheid	De mate waarin gegevens zijn voorzien van een unieke en stabiele identificatie die ook buiten de directe gebruikscontext betekenis heeft.
Formaatvaliditeit	De mate waarin gegevens syntactisch correct zijn.
Domeinvaliditeit	De mate waarin de inhoud van waarden consistent zijn met hun domein.
Logische consistentie	De mate waarin de combinaties van waarden logisch samenhangend zijn.
Referentiële integriteit	De mate waarin verwijzingen in gegevens verwijzen naar bestaande gegevens.

Definities kwaliteitsattributen basis (2)

Kwaliteitsattribuut	Definitie
Homogeniteit	De mate waarin de gegevens in een dataset gelijksoortig zijn.
Updatefrequentie	De snelheid waarmee gegevens worden vernieuwd.
Versheid	De mate waarin gegevens tijdig overeenkomen met de werkelijkheid.
Opslagprecisie	De mate van detail waarmee gegevens zijn geregistreerd.
Authenticiteit	De mate waarin de identiteit van de bron van de gegevens aantoonbaar is.
Reputatie	De mate waarin de gegevens worden vertrouwd vanwege de bron.
Bewijsbaarheid	De mate waarin de juistheid van een gegeven kan worden aangetoond.
Waarschijnlijkheid	De mate waarin gegevens waarschijnlijk zijn voor de situatie.
Herleidbaarheid	De mate waarin is vastgelegd wie of wat het gegeven waar, wanneer en op welke manier heeft ingewonnen of bewerkt.
Reproduceerbaarheid	De mate waarin de transformatieregels die zijn gebruikt om tot de gegevens te komen bekend zijn.
Gebruiksinzicht	De mate waarin er inzicht is in de gebruikers en hun gebruik.
Duidelijkheid	De mate waarin teksten voor de doelgroep begrijpelijk zijn geformuleerd.

Definities kwaliteitsattributen geografie

Kwaliteitsattribuut	Definitie
Positionele juistheid	De mate waarin locatiegegevens overeenkomen met de werkelijkheid.
Ruimtelijke dekking	De mate waarin de gegevens ruimtelijk dekkend zijn.
Topologische consistentie	De mate waarin ruimtelijke gegevens zich op de juiste wijze tot elkaar verhouden.
Geometrische validiteit	De mate waarin iedere geometrie voldoet aan geometrische regels.
Geometrische precisie	De mate van detail waarmee ruimtelijke gegevens worden ingewonnen.

Definities kwaliteitsattributen statistiek

Kwaliteitsattribuut	Definitie
Kwantitatieve juistheid	De mate waarin kwantitatieve gegevens overeenkomen met de werkelijkheid.
Statistische precisie	De mate waarin metingen of berekeningen bij herhaling dezelfde waarde opleveren.
Representativiteit	De mate waarin een dataset een goede weergave geeft van het geheel.

Adoptie

- Stelsel van Basisregistraties
- Loonaangifteketen

Vragen en bijdrage?

- Adoptie?
- Hebben jullie voorbeelden die de genoemde kwaliteitscriteria verduidelijken?
- Wat is er nog nodig om het Raamwerk te gaan gebruiken? Zijn er plaatsen of groepen waar we het verder onder de aandacht moeten brengen?
- Waar loop jij tegen aan als je het Raamwerk in je eigen praktijk gaat toepassen?
- Gerichtte feedback?

- Vragen?

w.stolk@berenschot.nl

Of

nora@ictu.nl