

# Handleiding opstellen dataset

EENHEID VAN TAAL

Betere zorg  
door betere informatie



<b>Datum</b> 28-09-2011	<b>ID Nummer</b> 11101		
<b>Auteur(s)</b> Claartje Hülsmann Elze de Groot Henk Hutink Ron van Holland			

**Wanneer een zorgverlener of patiënt digitaal informatie wil vastleggen of uitwisselen, is het van belang dat ondubbelzinnig is vastgelegd om welke gegevens het precies gaat, en hoe deze gegevens worden geregistreerd of uitgewisseld. Een dataset biedt hiervoor een eerste stap. In een dataset zijn definities van de gegevens vastgelegd die in een bepaald zorgproces, of binnen een specifiek vakgebied worden geregistreerd of uitgewisseld. Deze definities zijn niet-technisch van aard en worden bepaald door de gebruikers van ICT-systemen, door zorgverleners en patiënten dus.**

**Deze handleiding biedt handvatten voor hoe een dataset op uniforme wijze opgesteld kan worden. De bijbehorende template is een spreadsheet met voorgedefinieerde kolommen. De handleiding en template zijn bestemd voor informatieanalisten die gebruikers helpen bij het opstellen van een dataset.**

## Samenvatting

Uitwisseling en registratie van medische gegevens begint bij het definiëren van een dataset. Een dataset bevat functionele definities van data-elementen die geregistreerd of uitgewisseld kunnen worden. Deze definities vormen de basis voor technische specificaties van ICT-systemen. Tot nu toe worden datasets op verschillende manieren gedefinieerd. Dit leidt tot veel discussies waardoor de totstandkoming van datasets vaak moeizaam verloopt. Deze handleiding geeft kaders voor hoe een dataset op een uniforme en gestandaardiseerde manier opgesteld kan worden en hoe de dataset inhoudelijk gevuld moet worden. Daarnaast zijn een aantal best practices geformuleerd.

## Doelgroep

Een dataset slaat een brug tussen zorginhoudelijke begrippen en technische specificaties. Aanbevolen wordt om de zorginhoudelijke begrippen te laten vaststellen en invullen door zorgverleners. Zorgverleners vullen een aantal kolommen in. Deze kolommen zijn in deze handleiding als 'verplicht' gemarkeerd.

## Vaststellen van uitgangspunten

Bij het opstellen van een dataset zijn de volgende uitgangspunten van belang:

- Is de dataset bedoeld voor registratie of uitwisseling van medische gegevens in het kader van een specifieke use case?
- Is de dataset bedoeld om gegevens die binnen een specifiek vakgebied of met betrekking tot een bepaald onderwerp vaak voorkomen te benoemen en definiëren?
- Is de dataset bedoeld als één op één vervanging van de huidige werkwijze of zijn er extra doelstellingen bepaald waardoor er ook nieuwe data-elementen gedefinieerd moeten worden?
- Welke kaders en richtlijnen zijn afgesproken om besluiten te nemen?
- Welk zorgproces of welke zorgstandaard wordt als uitgangspunt genomen?

Bij het opstellen van een dataset is het menselijk eigen standpunten te verdedigen. Indien bovenstaande uitgangspunten niet gedefinieerd worden, bestaat het risico dat in een latere fase intensieve discussies ontstaan.

*Best practices:*

- *Implementatie van een dataset staat of valt met de acceptatie van de gebruikers en koepelorganisaties. Parallel aan de ontwikkeling van de dataset is het raadzaam afspraken te maken met deze partijen over het beheer en over de implementatiestrategie.*
- *Bij het opstellen van de dataset is het essentieel om zorgverleners te consulteren.*

## Terminologie en begrippen

Begrip	Definitie
Dataset	Een dataset is een verzameling van definities van elementen en attributen die gebruikt wordt in het kader van één gebruiksdoel of een aantal nauw aan elkaar gerelateerde gebruiksdoelen.
Attribuut	Een attribuut is een gegeven waaraan een waarde kan worden toegekend. Een voorbeeld van een attribuut is 'straatnaam' of 'frequentie'.
Element	Een element is een verzameling van attributen of andere elementen.

## Toelichting op de template voor datasets

De template bestaat uit drie tabbladen die in onderstaande tabel worden beschreven. De inhoud van de tabbladen wordt in dit hoofdstuk verder uitgewerkt.

Tabblad	Omschrijving
Dataset	De structuur en de definitie van de daadwerkelijke functionele dataset met zijn data-elementen en attributen.
Valuesets	Specificatie van de valuesets die worden gebruikt in de dataset. Valueset bevatten mogelijke waarden die een bepaald attribuut kan aannemen, bijvoorbeeld 'man', 'vrouw'.
Datatypes	Specificatie van de gebruikte datatypes in een dataset. Datatypes specificeren het type van een bepaald attribuut, bijvoorbeeld 'alfanumeriek'.

## Dataset een naam geven

Het is van belang uniformiteit te hanteren in de naamgeving van de datasets die worden uitgegeven, zodat datasets herkenbaar zijn.

De naam van de dataset wordt als volgt opgebouwd:

- Dataset <onderwerp>

Voorbeelden zijn:

- Dataset huisartswaarneming
- Dataset kern dossier ziekenhuizen
- Dataset perinatologie

### Opbouw van een dataset

Een dataset is opgebouwd uit attributen en elementen. Een element is opgebouwd uit één of meerdere elementen en/of één of meerdere attributen. Aan attributen kan een waarde worden toegekend, die geregistreerd en/of uitgewisseld kan worden. Deze geneste structuur van elementen en attributen wordt in 'Rubriek' van de template beschreven. Vervolgens worden verschillende aspecten van de elementen en attributen gespecificeerd. Dit wordt gedaan in een aantal additionele kolommen. Deze kolommen worden later in dit document uitgewerkt.

*Gebruiksregels:*

1. *Elementen en attributen worden zo veel mogelijk hergebruikt.*
2. *Elementen en attributen worden op een dusdanige manier benoemd en gedefinieerd zodat meerdere datasets deze kunnen gebruiken.*

### Elementen en attributen een naam geven

Elementen krijgen een naam die logisch en begrijpelijk voor de zorgverlener is. De namen worden met een hoofdletter geschreven. In geval van een naam die uit meerdere woorden bestaat, worden de woorden door een spatie van elkaar gescheiden. De tweede en volgende woorden beginnen met een kleine letter.

Attributen krijgen ook een betekenisvolle naam. Namen van attributen worden met een kleine letter geschreven, ook voor attributen geldt dat wanneer zij uit meer dan één woord bestaan, tweede en volgende woorden met een kleine letter geschreven worden en dat alle woorden gescheiden worden door spaties.

### Verplichte kolommen van het tabblad 'dataset'

Een element of attribuut wordt gespecificeerd met de verplichte kolommen 'Omschrijving', 'Eenheid',

'Voorbeeld', 'Conformance' en 'Conditie'.

### Omschrijving

Deze kolom bevat bij voorkeur een nauwkeurige specificatie van de bron en de locatie binnen deze bron waar het element of attribuut wordt beschreven. Indien niet naar een andere bron verwezen kan worden, dan wordt in deze kolom een tekstuele omschrijving van het element opgenomen. In dit geval dient als bron de eigen organisatie of werkgroep te worden vermeld.

*Gebruiksregel:*

1. *Aansluiten bij gangbare en landelijk geldende bronnen.*

### Eenheid

Specificatie van de eenheid van een gemeten waarde. Voor attributen waarop een eenheid van toepassing is, is het verplicht deze kolom te vullen. Er zijn legio meeteenheden in te vullen, een aantal voorbeelden zijn:

- mm (millimeter)
- µmol (micromol)
- µg (microgram)

### Voorbeeld

In deze kolom wordt altijd een voorbeeld van een mogelijke waarde voor een attribuut in.

*Gebruiksregel:*

1. *Geef altijd een voorbeeld invulling, ook als het nog zo logisch lijkt.*

### Conformance

De conformance van een element/attribuut geeft aan of en wanneer het betreffende element/attribuut een waarde moet hebben. Indien de dataset wordt gebruikt in het kader van uitwisseling van gegevens, worden door de conformance impliciet verplichtingen voor de verzender en ontvanger vastgelegd. De kolom conformance moet worden gevuld.

Bepaling van de conformance kan vanuit verschillende perspectieven worden gedaan, bijvoorbeeld:

- Eis van zorgverlener: dient het element of attribuut door een zorgverlener te worden geregistreerd of aangeleverd?
- Eis aan ICT-systeem: dient het ICT-systeem de registratie of uitwisseling van het element of attribuut te ondersteunen?

Een dataset bevat functionele eisen met betrekking tot de informatievoorziening die worden gesteld door zorgverleners aan ICT-systemen. Conformance geeft informatie van de zorgverlener aan de ICT-leverancier. De kolom conformance geeft aan welke elementen en/of attributen geregistreerd of uitgewisseld zullen worden en welke geregistreerd of uitgewisseld kunnen worden.

De volgende waarden zijn mogelijk:

Conformance	Betekenis
V	Verplicht: Het element of attribuut bevat een waarde. Verplicht verwerken door het ontvangend systeem in geval van uitwisseling.
O	Optioneel: Het element of attribuut kan aanwezig zijn en mag een waarde bevatten. Verplicht verwerken door het ontvangend systeem in geval van uitwisseling.
C	Conditioneel: Het element of attribuut moet alleen verplicht een waarde bevatten indien aan de conditie wordt voldaan. Verplicht verwerken het ontvangend systeem in geval van uitwisseling.

Verwerken door het ontvangend systeem houdt in dat het ICT-systeem van de ontvanger in staat moet zijn om het desbetreffende element of attribuut aan de gebruiker te tonen en/of in het eigen systeem op te slaan.

### Conditie

Indien van toepassing wordt in deze kolom de betreffende conditie die geldt voor een element of attribuut genoemd. Een voorbeeld van een conditie is:

- Het attribuut 'triageklasse' dient gevuld te zijn als het

attribuut 'gewondennummer' is gevuld. Het element of attribuut moet verplicht gevuld worden indien een ander element of attribuut is gevuld.

De kolom 'conditie' geeft aan onder welke condities een zorgverlener het betreffende element of attribuut zal vastleggen of uitwisselen.

*Best practice:*

- *Nooit verwijzen naar een regelnummer, rijnummer of een cel dat/die door Excel zijn toegekend. De regelnummers veranderen automatisch doordat elementen tussengevoegd worden. Verwijs naar namen van elementen of attributen met unieke paden vanaf het eerste element in de dataset.*

### Optionele kolommen van het tabblad 'dataset'

Naast de verplichte kolommen bevat het tabblad dataset ook optionele kolommen:

#### Cardinaliteit

Met een cardinaliteit wordt aangegeven hoe vaak een element of attribuut met een waarde gevuld kan worden. Hieronder worden de verschillende mogelijkheden beschreven:

Cardinaliteit	Omschrijving
0..1	Het element of attribuut hoeft niet voor te komen, maar mag maximaal één keer voorkomen.
0..*	Het element of attribuut hoeft niet voor te komen, maar mag meerdere keren tot een onbeperkt aantal keer voorkomen.
1..1	Het element of attribuut moet exact één keer voorkomen.
1..*	Het element of attribuut moet minimaal één keer voorkomen en mag meerdere keren tot een onbeperkt aantal keer voorkomen.

## Type

In deze kolom wordt een type meegegeven van de waarde van het element of attribuut. De volgende mogelijkheden zijn van toepassing:

- De attribuutwaarde is een keuze uit verschillende waarden die zijn gedefinieerd in een valueset. Het type is dan de naam van een valueset.
- Het type is een datatype. Gebruikte datatypes zijn in het template gedefinieerd onder tabblad 'datatypes': Voorbeelden van datatypes die gebruikt kunnen worden zijn:
  - String (betreft een reeks karakters, vrije tekst)
  - Boolean (betreft een ja of nee waarde)
  - Numeriek (getal)
  - Datum (een datum zonder een tijd)
  - DatumTijd (een datum met een tijd)
  - Tijd (een tijdstip)
  - Eigen datatype.

## Vaste waarde

Indien van toepassing wordt in deze kolom een vaste waarde van het attribuut gespecificeerd.

## Lengte

Indien gewenst en van toepassing wordt in deze kolom de maximale lengte in aantal karakterposities opgenomen.

## Code

Indien beschikbaar wordt in deze kolom een code opgenomen, zoals die voor dit element of attribuut bekend is binnen een bestaand codestelsel. Een code mag slechts worden ingevuld indien sprake is van een exacte match tussen de betekenis van het element of attribuut en het betreffende concept uit het codestelsel.

## Codestelsel

Indien een code is gespecificeerd, dan dient in deze kolom het bijbehorende codestelsel te worden opgenomen. Voorbeelden van codestelsels zijn LOINC en SNOMED CT.

## Scope

Deze kolom geeft aan wat de geldigheidscope is van de gehele definitie (dus zoals beschreven in de som van alle kolommen van het element of attribuut) van een element of attribuut. Mogelijke waarden zijn:

- Globaal: de definitie is geüniformeerd en is ook beschikbaar voor andere datasets
- Dataset: de definitie wordt slechts toegepast binnen de dataset zelf

## Kolommen van het tabblad 'Valuesets'

In het tabblad 'Valuesets' worden de valuesets gespecificeerd die binnen de dataset worden gebruikt. Valuesets kunnen verschillende 'eigenaren' hebben die verantwoordelijk zijn voor de inhoud. Bij gebruik van valuesets is het van belang om zo veel mogelijk aan te sluiten bij bestaande valuesets.

## Valueset

In de kolom 'Valueset' worden de naam en de waarden ingevuld. Valuesets krijgen een naam die logisch en begrijpelijk voor de zorgverlener is. De namen worden met een hoofdletter geschreven, in geval van een naam van meer dan één woord worden de woorden door een spatie van elkaar gescheiden. De tweede en volgende woorden beginnen met een kleine letter.

Waarden (values) krijgen ook een betekenisvolle naam. Namen van values worden met een kleine letter geschreven. Ook voor values geldt dat wanneer zij uit meer dan één woord bestaan, tweede en volgende woorden met een kleine letter geschreven worden en dat alle woorden gescheiden worden door spaties.

## Omschrijving

Hiervoor geldt hetzelfde als voor de kolom 'Omschrijving' van het tabblad 'Dataset'.

## Code

Hiervoor geldt hetzelfde als voor de kolom 'Code' van het tabblad 'Dataset'.

## Codestelsel

Hiervoor geldt hetzelfde als voor de 'Codestelsel' kolom van het tabblad 'Dataset'.

## Gebruiksregels:

1. Indien een valueset in zijn geheel gebruikt wordt zoals elders gespecificeerd, wordt verwezen naar deze valueset.
2. Indien een valueset in zijn geheel gebruikt wordt en er bovendien een toevoeging of wijziging aan deze

*valueset gespecificeerd moet worden, wordt bij de beheerder van de valueset een aanvraag gedaan voor de wijziging van de set. Daarna kan men naar de set verwijzen met behulp van een naam en een code.*

### Scope

Deze kolom geeft aan wat de geldigheidscope is van de gehele definitie (dus zoals beschreven in de som van alle kolommen van het element of attribuut) van een element of attribuut. Mogelijke waarden zijn:

- Globaal: de definitie is geüniformeerd en is ook beschikbaar voor andere datasets
- Dataset: de definitie wordt slechts toegepast binnen de dataset zelf

### Kolommen van het tabblad 'Datatypes'

In het tabblad 'Datatypes' worden de datatypes gespecificeerd die binnen de dataset worden gebruikt.

### Datatype

De eerste kolom bevat de naam van het datatype. Datatypes krijgen een naam die logisch en begrijpelijk voor een zorgverlener is. De namen worden met een hoofdletter geschreven en bestaan altijd uit één woord.

### Omschrijving

Hiervoor geldt hetzelfde als voor de Omschrijving kolom 'Omschrijving' van het tabblad 'Dataset'.

### Voorbeeld

In deze kolom worden altijd één of meerdere voorbeelden ingevuld van mogelijke waarden die voldoen aan het betreffende datatype.

### Scope

Deze kolom geeft aan wat de geldigheidscope is van de gehele definitie (dus zoals beschreven in de som van alle kolommen van het element of attribuut) van een element of attribuut. Mogelijke waarden zijn:

- Globaal: de definitie is geüniformeerd en is ook beschikbaar voor andere datasets
- Dataset: de definitie wordt slechts toegepast binnen de dataset zelf

Nictiz is het landelijke expertisecentrum dat ontwikkeling van ICT in de zorg faciliteert. Met en voor de zorgsector voorziet Nictiz in mogelijkheden en randvoorwaarden voor elektronische informatie-uitwisseling voor en rondom de patiënt. Wij doen dit ter bevordering van de kwaliteit en doelmatigheid in de gezondheidszorg.

### Nictiz

Postbus 19121  
2500 CC Den Haag  
Oude Middenweg 55  
2491 AC Den Haag

T 070 - 317 34 50  
info@nictiz.nl  
www.nictiz.nl