

Guidelines voor de business laag

Collibra Data Catalog

Juni 2023 – Versie 1.0

Inhoudsopgave

Inleiding

Opbouw PostNL Data Catalogus

Business terms – basis principes, opvoeren, wijzigen, status

Vast te leggen metadata

Relaties

Inleiding

De PostNL Data Catalogus

Introductie

PostNL gebruikt het platform Collibra Data Governance Center (DGC) ter ondersteuning van haar data organisatie op het gebied van data kwaliteit, data governance en data analyse. Het helpt bij het vinden van data, biedt een inventarisatie van metadata en verstrekt informatie die nodig is om vast te stellen of de data geschikt is voor het beoogd gebruik.

De PostNL Data Catalogus is een instrument waarmee de data organisatie van PostNL de toegankelijkheid, juistheid en relevantie van de data bedrijfsbreed kan verbeteren. Hiermee wordt cruciale ondersteuning geboden bij:

- **Data gebruik:** metadata verhoogt de kennis over data. Hoe meer gebruikers weten over data hoe beter ze de bruikbaarheid en de grenzen daarvan kunnen bepalen.
- **Datamanagement:** de data catalogus biedt inzicht en een beter begrip van de data die PostNL bezit. Dit maakt de data bekend en beheersbaar en is een randvoorwaarde voor professionalisering van data management capabilities zoals data governance en data kwaliteit management.

De PostNL Data Catalogus is een van de solution building blocks binnen de data capability 'Meta Data Management'.

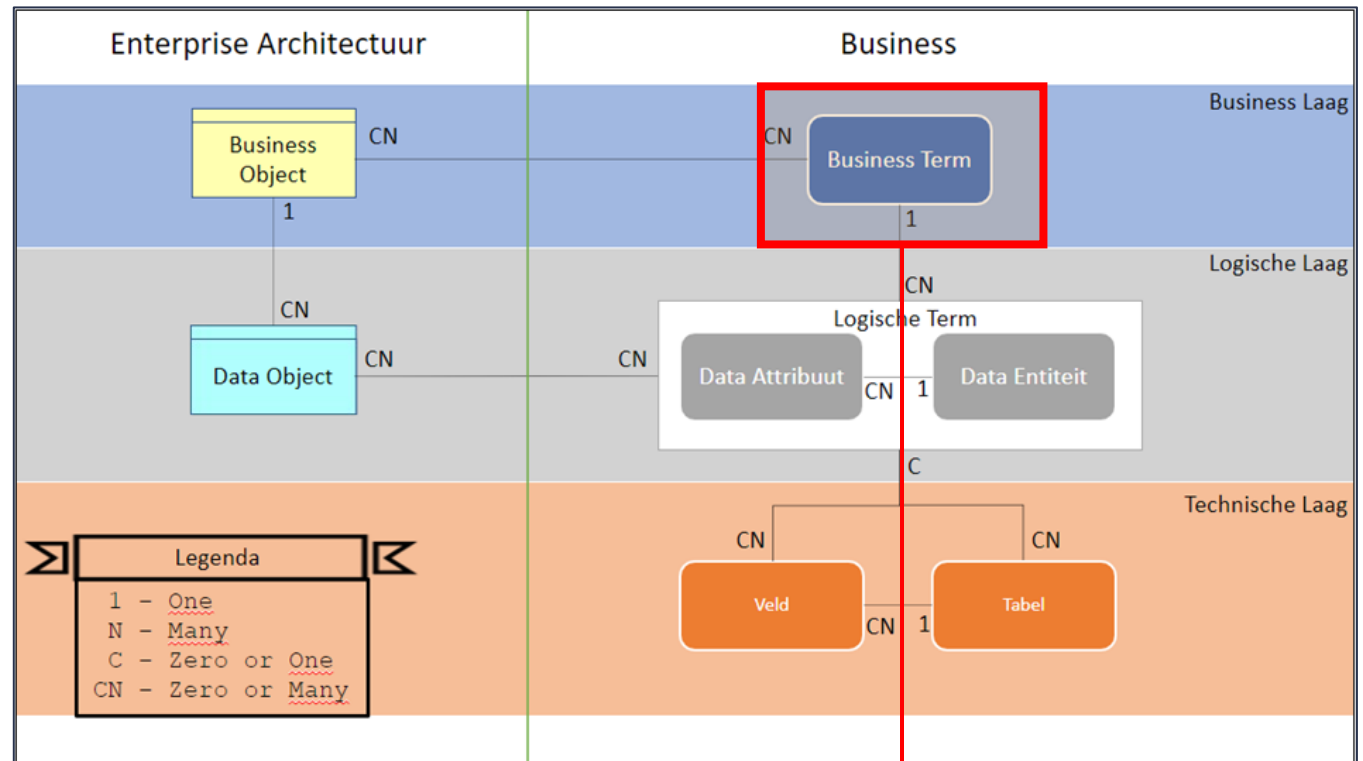
*“**Metadata Management** is de discipline die gaat over het verzamelen, onderhouden en standaardiseren van metadata (inclusief toegang en distributie)”.*

Opbouw PostNL Data Catalogus

Drie lagen: business, logisch en technisch

De PostNL Data Catalogus is opgebouwd uit drie verschillende lagen:

- **Business laag**
De termen die voor 'de business' gemeengoed zijn en in het dagelijks gebruik worden gehanteerd.
- **Logische laag**
De benodigde termen (entiteiten en attributen) om de business termen te koppelen aan de technische laag.
- **Technische laag**
De velden en tabellen zoals die in de systemen en applicaties daadwerkelijk voorkomen.



Focus in dit document ligt specifiek op de 'Business laag'. Oftewel, wat is een business term en wat is de bijbehorende metadata?

Business terms

Basis principes

- “Een business term is een woord dat wordt gebruikt om iets of een concept uit te drukken in een bepaalde context”.
- In de Data Catalog, ontstaat een business term door middel van een business behoefte. Het gaat hier om de woorden en termen die gebruikt worden door de business. Het is daarom ook de business die de regie heeft over wat wel, en wat niet wordt opgenomen in de Data Catalog. De business wordt hierin vertegenwoordigd door de Data Eigenaar, Data Manager en Data Steward
- Een business term kan hetzelfde zijn als een business object of data object. Hiermee kan het zelfs zo zijn dat definities en andere metadata een-op-een gelijk zijn aan elkaar. Toch blijft, in ieder geval Architectuur eigenaar en verantwoordelijk voor business objecten en data objecten. Dit betreft dus ook de bijbehorende metadata.

Business Metadata

De Business bepaalt de **business definitie** van de te behandelen term. Daarnaast wordt metadata op business niveau toegevoegd (i.e. business ruling, synoniemen, homoniemen, etc.)

- Data Eigenaar
- Data Manager
- Data Steward

Business Term

Business terms

Opvoeren

Er is veel vrijheid om te bepalen wat gezien wordt als een business term – en wat niet. Het is niet de bedoeling om zodanig veel restricties op te leggen dat het invullen van de Data Catalog gezien wordt als een vervelende klus. Er wordt daarom vertrouwd op de expertise van de business om te bepalen wat wel op de business laag komt en wat niet.

- ‘Handheld’, ‘Branchenummer’ of ‘Postcode’ zijn allemaal duidelijk business terms en komen dus op de business laag.
- ‘HH__Code’, ‘Branch_number__c’ of ‘PC6_NL’ zijn technische velden en horen dus niet op de business laag.

Business terms dienen geschreven te worden zoals ze ook door de business gebruikt worden. De spelling volgt hierin hoe je het woord ook in het woordenboek zou vinden:

- De beginletter is altijd een hoofdletter, gevolgd door kleine letters.
- In het geval van meerdere woorden, worden er spaties gebruikt.

Business terms

Wijzigen

Om een business term toe te voegen of te wijzigen dien je onderdeel te zijn van de business Layer Curator groep, oftewel een 'Vuller'. Binnen deze groep zijn er **geen** restricties voor wat betreft het aanpassen van business terms. Dit betreft de business terms die door jezelf zijn aangemaakt, als ook de business terms die door een ander zijn aangemaakt.

Hiervoor gelden de volgende procesregels:

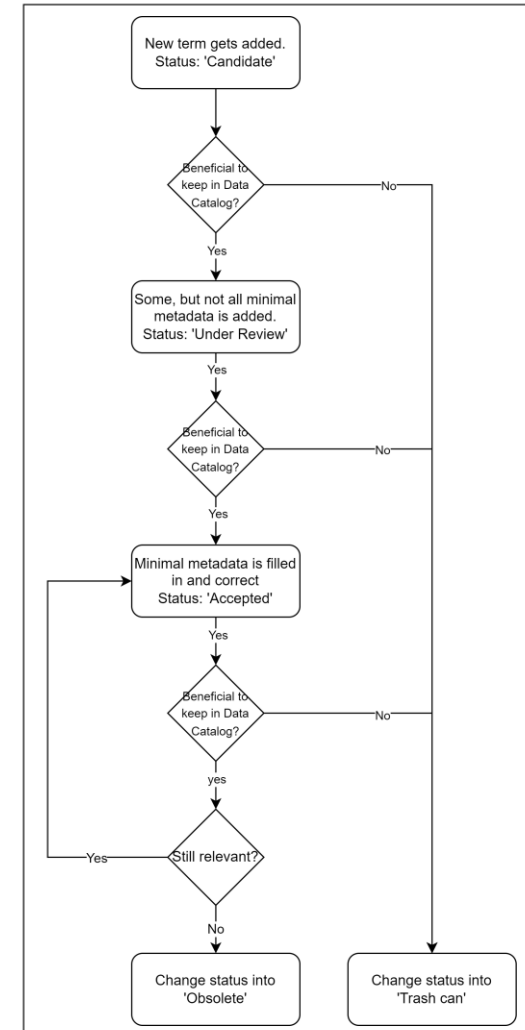
- Je bent **vrij** om je **eigen** business term (dus zelf aangemaakt) te wijzigen.
- Je bent **niet vrij** om zomaar de business term te wijzigen die door een **ander** is aangemaakt. Wanneer de metadata van een andere business term incorrect of incompleet is, dan dien je contact op te nemen met de eigenaar van die andere business term om tot een onderlinge oplossing te komen. Wijkt de context waarin de business term gebruikt wordt teveel af van de bestaande business term; voeg dan de business term toe in jouw specifieke context.

Business terms

Status

- Een term kan in verschillende fases verkeren. Met de status wordt aangegeven of het bijvoorbeeld een geaccepteerde term met metadata is of dat het een term is die 'Under Review' is.

Status	Betekenis
Candidate	Initiële status van een term. Dit betekent dat de term is aangemaakt of geïmporteerd. Er wordt hier nog niet gekeken naar eventuele metadata.
Under Review	Stakeholders reviewen de Asset. Dit betekent dat er metadata is toegevoegd, maar dat nog niet alle verplichte velden zijn ingevuld of kloppend zijn.
Accepted	De term én de verplichte metadata zijn volledig en correct ingevuld. De betekenis wordt gedragen door de belangrijkste stakeholders.
Obsolete	De term is achterhaald. De metadata van deze term blijft in dit geval wel beschikbaar voor informatie. Na bepaalde periodes wordt er gekeken of hierop een schoning nodig is.
Trash can	Wanneer een term foutief is opgevoerd krijgt deze de status van Trash can. Termen met deze status worden periodiek, met goedkeuring van de eigenaar, verwijderd
Invalid	Buiten scope
In Progress	Buiten scope
Approval Pending	Buiten scope



Vast te leggen metadata

Afgestemd met de Data Management Organisaties

1	Definition
2	Definition (English)
3	Explanation
4	Example
5	Data Manager
6	Data Domain
7	Language
8	System Of Record
9	System of Entry
10	Data Labelling
11	Personally Identifiable Information
12	PII Type
13	CIA Classification
14	Confidentiality
15	Integrity
16	Availability
17	Business Ruling
18	Standard Value Format
19	Abbreviation
20	Date Created
21	Date Changed
22	Made/Changed by
23	Status
24	Source (origin metadata)
25	Synonym
26	Translation
27	Rel Architecture - Objects
28	Rel Logical Layer – Data Entity
29	Rel Logical Layer – Data Attribute

In samenspraak met de Data Management Organisaties binnen PostNL is vastgesteld dat de hier links genoemde zgn. metadata velden in de PostNL Data Catalogus dienen te worden vastgelegd.

- De rood gekleurde velden zijn verplicht.
- De zwart gekleurde velden zijn optioneel.
- De blauw gekleurde velden laten de relatie zien tussen de drie lagen en/of business terms onderling.
- De velden 14 t/m 16 zijn vooralsnog buiten scope.
- De velden 20 t/m 22 worden automatisch door Colibra zelf gegenereerd (standaard functionaliteit).

Op de volgende slides wordt elk van deze velden nader toegelicht.

Vast te leggen metadata

Toelichting op de velden (1/6)

Metadata	Description	Guideline	Mandatory
Definition	Een duidelijke beschrijving van de business term in één à twee zinnen.	• Tekstveld • Definitie bevat níét de opgevoerde term zelf • Volgt de meest gebruikelijke spelling en leestekens • Is uniek voor de context waarin de term beschreven wordt	Verplicht
Definition (English)	De Engelse vertaling van de origineel opgevoerde definitie.	• Tekstveld • Definitie bevat niet de opgevoerde term zelf • Volgt de meest gebruikelijke spelling en leestekens • Dient een sluitende vertaling te zijn van de origineel opgevoerde definitie	Optioneel
Explanation	Een aanvulling op de definitie die duidelijk de term beschrijft.	• Tekstveld • Uitleg bevat niet de opgevoerde term zelf • Volgt de meest gebruikelijke spelling en leestekens • Uniek voor de context waarin de term beschreven wordt	Optioneel
Example	Voorbeelden van de term.	• Tekstveld • Volgt de meest gebruikelijke spelling en leestekens • Uniek voor de context waarin de term beschreven wordt	Optioneel

Vast te leggen metadata

Toelichting op de velden (2/6)

Metadata	Description	Guideline	Mandatory
Data Manager	Verantwoordelijke Data Manager voor deze specifieke term.	- Selectielijst met daarin de specifieke Data Manager die verantwoordelijk is voor deze specifieke term - Laat alleen de verantwoordelijke functie zien; niet de naam Actueel overzicht	Verplicht
Data Domain	Verantwoordelijke domein voor deze specifieke term.	- Selectielijst met daarin het specifieke Data Domein dat verantwoordelijk is voor deze specifieke term Actueel overzicht	Verplicht
Language	De taal waarin een term betekenis heeft.	Selectielijst met daarin de waarden: NL, EN, DE en FR	Verplicht
System of Record	Applicatie waarin dé waarheid wordt vastgelegd en van waaruit distributie naar ontvangende systemen en applicaties plaatsvindt.	- Selectielijst met daarin alle systemen en applicaties die gebruikt worden binnen PostNL die als Golden Record zijn gedefinieerd - Kan dezelfde waarde bevatten als de SoE. - Als de term een lineage heeft tot en met de technische laag, dan is dit veld verplicht	Optioneel
System of Entry	Applicatie waarin de gegevens initieel worden opgevoerd c.q. gegenereerd.	- Selectielijst met daarin alle systemen en applicaties die gebruikt worden binnen PostNL voor data opvoer - Kan dezelfde waarde bevatten als de SoR. - Als de term een lineage heeft tot en met de technische laag, dan is dit veld verplicht	Optioneel

Vast te leggen metadata

Toelichting op de velden (3/6)

Metadata	Description	Guideline	Mandatory
Data Labelling	Duidt het label aan dat van toepassing is op gegevensbescherming	- Selectielijst met waarden: Public: Komt overeen met gegevens met de vertrouwelijkheidswaarde 'None'. Internal: Komt overeen met gegevens met de vertrouwelijkheidswaarde 'Low'. Confidential**: Komt overeen met gegevens met de vertrouwelijkheidswaarde 'Medium'. Secret: Komt overeen met gegevens met de vertrouwelijkheidswaarde 'High'. - Moet afgestemd zijn met en voldoen aan de regels en principes van PostNL Cybersecurity Office. ** Als de kolom PII-gegevens bevat, dan dient deze minimaal het label van "Confidential" te krijgen.	Optioneel
Personally Identifiable Information	Geeft aan of het persoonsgegevens bevat.	- Selectielijst met daarin de waarden: - Yes, does contain Personally Identifiable Information - No, does not contain Personally Identifiable Information - Uncertain if it contains Personally Identifiable Information - Unknown	Verplicht
PII Type	Geeft aan wat voor soort persoonsgegevens het bevat.	- Meerkeuze selectielijst met daarin de waarden: - Business Partner - Consumer - Employee - Kan in de toekomst verder uitgebreid worden. - Indien 'Personal Identifiable Information' is ingevuld, dan is dit veld verplicht	Optioneel

Vast te leggen metadata

Toelichting op de velden (4/6)

Metadata	Description	Guideline	Mandatory
CIA Classification	De classificatie van data en de gerelateerde systemen is bepaald door wat de impact voor PostNL is wanneer de gestelde eisen aan Confidentiality, Integrity en Availability niet gehaald worden.	• Selectielijst met daarin de waarden Baseline en Above Baseline • Wanneer de impact van een van de drie categorieën (Confidentiality, Integrity or Availability) hoog is, dan is de CIA Classificatie Above Baseline. In alle andere gevallen is Baseline voldoende. • Dient afgestemd te zijn en overeen te komen met de regels en principes van PostNL Security Office.	Optioneel
Confidentiality	Wat is de impact wanneer er sprake is van onbevoegde verstrekking van informatie.	• Selectielijst met daarin de waardes: low, medium en high: • Low: Het verlies van Confidentiality, Integrity of Availability heeft naar verwachting een <u>beperkt nadelig effect</u> op bedrijfsprocessen, bedrijfsmiddelen of individuen. • Medium: Het verlies van Confidentiality, Integrity of Availability heeft naar verwachting een <u>nadelig effect</u> op bedrijfsprocessen, bedrijfsmiddelen of individuen. • High: Het verlies van Confidentiality, Integrity of Availability heeft naar verwachting een <u>ernstig of catastrofaal nadelig effect</u> op bedrijfsprocessen, bedrijfsmiddelen of individuen. • Dient afgestemd te zijn en overeen te komen met de regels en principes van PostNL Security Office.	Optioneel
Integrity	Wat is de impact wanneer er sprake is van onbevoegde aanpassing of vernietiging van informatie.		Optioneel
Availability	Wat is de impact wanneer er sprake is van een verstoring in de toegang voor het gebruik van informatie of informatiesystemen.		Optioneel

Vast te leggen metadata

Toelichting op de velden (5/6)

Metadata	Description	Guideline	Mandatory
Business Ruling	De business regels die gelden voor een term.	• Tekstveld • Voorbeeld: Nederlandse Postcode: Bestaat vier cijfers en twee letters. Het eerste getal mag niet met 0 beginnen. De lettercombinaties 'SS', 'SD' en 'SA' worden niet gebruikt.	Optioneel
Standard Value Format	De standaard waarde opmaak die geldt voor een term.	• Tekstveld • Geeft aan volgens welk format de term ingevuld dient te worden. • Kan ingevuld worden als reguliere expressie • Voorbeeld: Nederlandse Postcode: /^[1-9][0-9][0-9][0-9][A-Z][A-Z]\$/gm	Optioneel
Abbreviation	Afkorting van de term.	• Tekstveld • Indien een afkorting gebruikt wordt als een term, dan dient deze ook als een nieuwe term opgevoerd te worden	Optioneel
Source	Geeft aan wat of wie (functie/rol) de bron is van de ingevulde metadata.	• Tekstveld • Kan een of meerdere bronnen bevatten • Kan verwijzen naar functies, rollen en/of systemen.	Optioneel

Vast te leggen metadata

Toelichting op de velden (6/6)

Metadata	Description	Guideline	Mandatory
Synonym	De relatie met terms die eenzelfde of soortgelijke definitie hebben binnen hetzelfde Data Domein.	• Heeft een relatie met andere term in Collibra. Dit vereist dat de term waarmee een relatie moet worden gelegd bestaat in Collibra. • Kan alleen bestaan met andere terms op dezelfde laag • Relatie dient zowel bij term A als term B aangegeven te worden	Optioneel
Translation*	De relatie met terms die opgevoerd zijn in een andere taal binnen hetzelfde data domein.	• Heeft een relatie met andere term in Collibra. Dit vereist dat de term waarmee een relatie moet worden gelegd bestaat in Collibra • Kan alleen bestaan met andere terms op dezelfde laag • Is alleen van toepassing op de business laag • Relatie dient zowel bij term A als term B aangegeven te worden	Optioneel
Relaties naar andere lagen	De relaties naar andere termen worden op aparte dia's uitgelegd.	• Heeft een relatie met andere term in Collibra. Dit vereist dat de term waarmee een relatie moet worden gelegd bestaat in Collibra.	Optioneel

* Naar aanleiding van het DGB besluit is het initiële aantal velden uitgebreid met twee optioneel in te vullen velden:

- Definition_EN: hierin komt de Engelstalige vertaling van de definitie te staan.
- Relatie 'Translation': hierin wordt bij de vastlegging van een NL business term de relatie gelegd met de Engelstalige variant.

Daarnaast wordt het huidige veld 'Synoniem' NIET meer gebruikt voor het leggen van de relatie tussen de NL en de EN business term. Hiervoor in de plaats komt de 'Relation'.

Voor de meeste BU's betekent dit dat ze alle termen alleen in het Nederlands opvoeren én het veld 'Definition_EN' vullen met de Engelse vertaling. [hierin heeft men nog de 'escape' dat dit feitelijk alleen voor 'nieuwe' termen hoeft]

Voor E-Commerce Operations en de BU's die veel met Engelstaligen te maken heeft, betekent het dat alle business terms wél dubbel worden opgevoerd; dat wil zeggen dat elke NL business term ook een EN business term kent die d.m.v. de relatie 'Translation' aan elkaar zijn gekoppeld. [de verantwoordelijkheid qua onderhoud ligt bij de betreffende DMO]

Het is daarbij verplicht dat de metadata in dezelfde taal wordt opgevoerd als de business term. Dit betekent dat een Engelse term, Engelse metadata krijgt en een Nederlandse term, Nederlandse metadata krijgt.

Vast te leggen metadata

Een voorbeeld

- Een aantal voorbeelden van ingevulde metadata van de business term 'Medewerker'.

The screenshot shows the 'Business Layer Glossary' interface for the business term 'Medewerker'. The header includes the 'BT' logo, the term name 'Medewerker', a star rating of five stars with '(0)' reviews, and metadata: 'Type: Business Term', 'Status: Candidate', and buttons for 'Edit', 'Move', 'Delete', and 'Auto hyperlinks'. A green 'Add characteristic' button is on the left. The main content area is divided into sections: 'Definition' (Een natuurlijk persoon in dienst van een bedrijf, die werkzaamheden verricht voor PostNL en daartoe vanuit PostNL individueel wordt aangestuurd.), 'Definition (English)' (No value has been given yet. Double click or use the edit button.), 'Explanation' (Er zijn personen die werkzaamheden verrichten voor Post NL, maar hier niet bij in dienst zijn. Dit zijn geen medewerkers.), 'Example' (No value has been given yet. Double click or use the edit button.), 'Data Manager' (HR), 'Data Domain' (HR - NL), and 'Language' (NL). A left sidebar contains navigation links: Overview, Tags, Ratings, Comments, Diagram, Pictures, Responsibilities, References, History, and Files.

Relaties

Relatie met entiteiten op de logische laag (1)

- Wanneer een business term ook voorkomt op applicatie en systeem niveau (technische laag) dient dit te worden aangegeven d.m.v. een 'relatie'.
 - Voorbeeld: De zakelijke klant die te vinden is in Salesforce of de Retaillocatie die in BLS te vinden is.
- Het is dan in de meeste gevallen niet mogelijk om een business term direct te koppelen aan een technische term (tabellen en velden). Hiervoor is de logische laag nodig. Hierin staan logische terms die als tussenstap dienen om de relatie te laten zien. Daarnaast kunnen er op de technische laag systeemvelden voorkomen die moeilijk te snappen zijn. In deze gevallen biedt de logische term de uitkomst.
- Een business term kan op de logische laag gekoppeld worden aan een Data Entiteit of een Data Attribuut.
 - Data Entiteit is de logische tegenhanger van een tabel
 - Data Attribuut is de logische tegenhanger van een veld

Business term

Relatie met entiteiten op de logische laag (2)

- In overleg met een Data Architect, Domein Architect of Data Informatie Analist bepaalt een Business Layer Curator wat benodigd is om de volledige lineage te kunnen weergeven.
 - Er kan een relatie gelegd worden met de Data Entiteiten en Data Attributen die vanuit Aris zijn ingeladen.
 - Er kan een relatie gelegd worden met nieuwe Data Entiteiten en Data Attributen die niet in Aris staan.
- Het kan voorkomen dat er *geen* nieuwe logische termen nodig zijn om de business laag te koppelen aan de technische laag.
 - In dit geval krijgt de business term een relatie met een Data Entiteit op logische laag die dezelfde naam en metadata krijgt als de business term

Business term

Relatie met enterprise objecten

- Zoals eerder beschreven kan een business term hetzelfde zijn als een business object of een data object.
 - In deze gevallen krijgt de business term dezelfde naam en/of definitie, maar blijft de verantwoordelijkheid ervoor bij de business.
 - In deze gevallen kan er vanuit de business term gerefereerd worden naar het object waar deze op gebaseerd is.
 - Dit doe je door een relatie te leggen tussen de business term en het object.
- Ook wanneer de business terms niet exact overeenkomen met de objecten, dan kan er alsnog gerefereerd worden naar een object.