# NEN 3610 Linked Data Profiel (bijgewerkt 29112018)

## Doelstelling en resultaten – originele PvA

In het originele PvA wordt de volgende doelstelling en resultaten geformuleerd:

***Doel:*** *Ontwikkelen van een standaard voor linked data toepassing van NEN 3610. Concepttitel: Linked Data Profiel op NEN 3610.*

***Resultaat:*** *De standaard Linked data profiel op NEN 3610 is gepubliceerd. Mogelijk betreft het een voorloper hierop middels een handreiking. De standaard bevat een aantal onderdelen:*

* *Semantisch model van NEN 3610 omgezet naar een linked data ontologie. Dit kan bijvoorbeeld de vorm hebben van een gepubliceerde (OWL?) ontologie van het semantisch model (met de bekende klassen GeoObject, Weg, Water, Terrein enz) maar ook een ontologie voor het temporeel model en/of een ontologie voor identificatie enz. Deze ontologieën kunnen gebruikt worden om ontologieën die specifiek voor een sectormodel gemaakt worden, expliciet te koppelen aan NEN3610.*
* *Mapping (i.e. een set relaties tussen klassen/eigenschappen) van bovengenoemde ontologieën met bestaande internationale ontologieën die relevant zijn, bijvoorbeeld GeoSPARQL en OWL Time; en waar relevant ook met nationale ontologieën, voor zover aanwezig, bijvoorbeeld CB-NL.*
* *Regels, aanbevelingen, en/of handreikingen (modeling guide) voor omzetten van NEN 3610 modellen naar linked data (ontologiën).*
* *Visie op samenhang en governance van een stelsel van linked data ontologiën.*

## Aanscherping doel en resultaten na ongeveer een jaar ontwikkeling

**Doel:** Komen tot een standaard voor linked data toepassing van NEN 3610.Dit houdt in:

1. Een Linked Data implementatie van de NEN3610 standaard zelf;
2. Een transformatie-aanpak van UML modellen die conform NEN3610 zijn opgesteld naar het/de overeenkomstige model(len) in Linked Data.

Ondersteunend hieraan wordt:

1. Invulling geven aan het MDA metamodel ontwikkeld door de OMG om metamodellen naar elkaar te transformeren en uit te wisselen;
2. Een voorbeeld-uitwerking gemaakt van IMGolf, een voorbeeld casus van NEN3610

**Resultaten:**

*(1) Linked Data modellen:*

1. Een SKOS ConceptScheme van de *begrippen* zoals deze voorkomen in de NEN3610 standaard (met name hoofdstuk 3 uit de standaard). Hiermee leveren we een formalisatie van het begrippenkader en een gepubliceerde thesaurus voor NEN3610, waar anderen naar kunnen verwijzen bij het opstellen van hun eigen thesaurus (via eigenschappen als skos:broadMatch, skos:exactMatch, etc);
2. Een RDFS/OWL vocabulaire met de *termen* van klassen en eigenschappen zoals deze voorkomen in de NEN3610 standaard. Hiermee leveren we een formalisatie van de NEN3610 vocabulaire, zodat anderen bij leveren geometrische data als Linked Data kunnen verwijzen naar deze top-ontologie voor feature types (door middel van rdfs:subClassOf). Dit verhoogd de interoperabiliteit van geometrische datasets, en zorgt voor een geometrische *knowledge graph* met een potentiele dekking over heel Nederland;
3. Een SHACL vocabulaire voor de constraints die in de NEN3610 standaard zijn opgenomen. Hiermee wordt het mogelijk om te *valideren* of een specifiek Linked Data dataset voldoet aan NEN3610;
4. Een OWL ontologie met de (afleidings)regels die in de NEN3610 standaard zijn opgenomen. Hiermee wordt het mogelijk om afleidingen (AI) toe te passen op Linked Data sets die zich houden aan de NEN3610 standaard.

De modellen worden geleverd in Linked Data serialisaties (turtle, RDF/XML en JSON-LD) en als voor mensen leesbare website (HTML)

*(2) Beschrijving van de best-practice, het profiel*

Een ReSpec document dat beschrijft hoe het NEN3610 Linked Data profiel in elkaar zit. Dit document kent de volgende onderdelen:

Het LD model van NEN3610 willen we als standaard neerzetten: officieel vanuit Geonovum als beheerder van NEN3610 gepubliceerd. De transformatieregels zien we als een handreiking, best-practice, aanbeveling.

Document blijft een standaard, het stuk over de transformatie wordt non-normatief en we zeggen dat dit stuk een handreiking betreft.

De NEN 3610 – ontologie/vocabulaire wordt op een aparte pagina gepubliceerd. Die pagina wordt de officieele bron van de de ontologie. In de inleidende pagina daarvan wordt verwezen naar het NEN 3610-LD respec doc.

is gerealiseerd.

**Algemeen/structuur**

Huidige status: nog niet alle formele elementen zijn benoemd. Redacteur niet (zou Paul zijn), Auteurs nog niet benoemd (alle mensen die een commit hebben gedaan)

Actie: Technische punten toevoegen aan de html, conform ReSpec. [Linda]

Actie: herstructuring en hernummering oppakken conform dit document [Marco]

**Samenvatting**

Huidige status: nog niet opgeschreven

Actie: voorlopig nog even open laten staan tot dat document 90% af is [Paul]

**1. Inleiding**

Huidige status: eerste concept staat. Op zich redelijk goed, mag nog wat scherper om doelstelling en wat-heb-ik-hier-aan duidelijk neer te zetten;

Actie: beter beschrijven naar de doelstelling en resultaten [Paul]

**2. Nederlandse LOD-Cloud voor GeoDatasets**

Huidige status: beoogd doel gaat hier best ver: een eigen LOD-Cloud. Dit lijkt bijna wel als een afzonderlijke project, en als we dat echt willen, dan moeten we em ook eigenlijk beheren.

Actie: in het document de status op moment van schrijven opnemen. Dit expliciet IN het document zetten. En verwijzen naar de LOD-cloud spin-off. [Marco]

5/10: Nog een aantal LOD sets toevoegen. Bijvoorbeeld CBNL referentie, GWSW en andere. (iedereen)

25/10: Marco heeft een klein LOD-nl cloud gemaakt die clickable is naar de individuele datasets. Er komt in github een folder (C:\Users\pjanssen\Documents\GitHub\NEN3610-Linkeddata\dataset) waarin conform DCAT de datasets beschreven kunnen worden.

An de hand van ingevulde sets wordt de LD-NL cloud gevult. Marco richt dat in. Iedereen vult het in.

**3 (MISSING) Use cases**

Huidige status: ontbreekt. Het “waarom doen we dit” hoort bovenin het document.

Actie: toevoegen van het “waarom”: waarom stellen we een NEN3610 LD model op, en waarom beschrijven we een transformatie vanuit een UML model? [Paul, Joep]

**4. Review beschikbare standaarden**

Huidige status: er is een kort lijstje gemaakt, en voor de meeste is ook wel een aardige beschrijving. Nog onvoldoende is beschreven wat de “pijn” is de in betreffende transformatieregels zit.

Actie: voor nu stellen we vast dat dit de relevante standaarden zijn. Afschrijven van de teksten. [Paul]

5/10: toevoegen LD proxy.

**5. Basisbeginselen Linked Data**

Huidige status: een initiele tekst is opgeleverd. Nog niet scherp wat het doel en diepgang zou moeten zijn.

Actie: besluit: zie hieronder voor essentie. Actie is uitschrijven conform beschrijving [Linda,Pano] hieronder

Opdelen in drie stukjes:

* Linked Data 101, met vooral verwijzingen naar andere plekken. Belangrijk puntje is dat LD gebruik maakt van vocabulaires (zoals bv SKOS, FOAF, DCAT, etc..), om interoperabiliteit ;
* Beschrijven wat we als LD model nu eigenlijk neerzet (begrippen, regels, structuur)
* Beschrijven wat essentiele verschillen zijn tussen OO en LD m.b.t. de transformatie

Vooral ook veel verwijzen naar bestaande bronnen.

5/10: Aanpassen aan de hand van resultaat discussie over punten Hans en Marinus.

Een paar punten:

UNA wordt nog steeds niet goed omschreven. Setorientatie nog niet goed omschreven. Het is eigenlijk allemaal set georienteerd maar op een verschillende manier. Gebruik dit om de verschillen uit te leggen.

RDF: Instantie ontstaat voor de klasse

OWL: Klasse is er voor de instantie.

Normalisatie is geen LD concept of doel. Het komt uit de database wereld. Dit klopt maar het kan nog steeds gebruik worden als concept om de verschillende methoden van vertaling van werkelijkheid naar model te vergelijken.

**6. De NEN3610 ontologie**

De ontologie is klaar. Review op begrip en inhoud. [Iedereen]

## **7. Transformatie van een NEN3610 UML model naar Linked Data**

GeoSPARQL toevoegen als vocabulaire.

Verwijzen naar INSPIRE RDF guidelines als standaard.

7.3 Handmatige aanpassing van transformatie.

Huidige status: al een vrij uitgebreide beschrijving, maar er zitten nog issues en keuzes in. Hier zouden we een aantal definitieve keuzen in moeten maken. Hier ook uitleggen waarom we het zo doen, en niet direct van UML naar een “juist” model

Actie: transformatieregels beschrijven die geautomatiseerd uitgevoerd kunnen worden [Marco]

Actie: transformatieregels om tot een juist model te komen voor de DSO [Pano]

Actie: transformatieregels om tot een juist model te komen voor COINS [Niels]

Actie: transformatieregels om tot een juist model te komen voor Orox [Marinus]

5/10

De tekst is nu geschreven op het totaal proces van omzetten van een UML naar LD. Deze tekst behouden maar ergens anders opnemen. Hier moeten meer algemene regels komen die specifiek het onderwerp van de handmatige aanpassing en overwegingen daarin adresseren. [Marco] Pano Niels Marinus

**8. Ontwerp van de transformatie**

8.2.1 Huidige status: een vrij ingewikkelde tekst. De matrix uit het verslag ontbreekt nog.

Actie: tekst redigeren, duidelijker uitleggen (dit nog even zo laten). Matrix toevoegen en onderscheid maken tussen enerzijds de toepassingen (COINS, Orox, DSO) en anderzijds de componenten (SKOS, OWL, RDF/RDFS, SHACL). [Paul]

## **9. Voorbeeld: UML-Golfbaan naar ontologie en vocabulaire**

Huidige status: is nu hoofdstuk 9, en kent nog geen invulling. Het document gaat in hoofdstuk 5 verder met de transformatie van het huidige model. Mijn gevoel is dat voor veel mensen het logischer zou zijn om nu het werkelijke profiel te leren kennen. De aanpak *hoe* je tot zo’n profiel komt (of tot een vertaling van jou eigen model) is voor mijn gevoel iets dat hierna kan komen. Eigenlijk beschrijven we het LD profiel nu helemaal niet!

Actie: hoofdstuk beschrijven, kan vrij kort. Gaat om de beschrijving van het model en eventueel de keuzen die we daarbij gemaakt hebben. We verwijzen dan naar de werkelijke plaats waar de modellen gepubliceerd staan. [Marco]

Actie: URI strategie vaststellen voor de LD resources die we gaan beschrijven [Marco]

Actie: Beschrijven van de LD resources (dwz: maken van de modellen) [Marco]

**8. Voorbeelduitwerking**

Huidige status: de voorbeelduitwerking kent nu een aantal plaatjes, maar de achterliggende consequenties, ontwerpbeslissingen en ideeen komen amper terug

Actie: wellicht goed om een bijlage te hebben voor elke voorbeelduitwerking, en hier alleen hoofdlijnen bespreken. (voorlopig even laten)

Actie: verwijzingen opnemen naar de voorbeelduitwerkingen, en deze ook gepubliceerd beschikbaar maken c.q. houden. (voorlopig even laten)

We hebben bij de ontwikkeling van de 3 voorbeelduitwerkingen nooit expliciet stap 1 uitgewerkt, maar direct tot de drie juiste modellen gekomen. Wat we nu nog moeten doen is:

1. Transformatiescript maken obv de transformatieregels die een NEN3610 UML model (source) technisch omzet naar de 4 componenten (SKOS, OWL, RDF/RDFS, SHACL);
2. Handmatige regels om te komen tot een “juist” model in de betreffende stijl (DSO, COINS, OROX) (target)

**10. (EXTRA)**

Huidige status: verborgen in de IMGolf subpagina: we beschrijven ook hoe we verschillende Linked Data datasets met elkaar kunnen gebruiken, door een model te “genereren” uit de data, en deze modellen vervolgens op eenzelfde manier te representeren.

Actie: dit uitwerken in een hoofdstukje. [Marco]

**A: Glossery**

Nog verder invullen. Kernbegrippen hier opnemen. Ook UNA, AAA, Setorientatie e.d.