

BOMOS aanvullende modules: Linked data



Logius Handreiking

Werkversie 09 november 2022

Deze versie:

<https://logius-standaarden.github.io/BOMOS-LinkedData/>

Laatst gepubliceerde versie:

<https://gitdocumentatie.logius.nl/publicatie/bomos/linkedata/>

Laatste werkversie:

<https://logius-standaarden.github.io/BOMOS-LinkedData/>

Redacteurs:

Erwin Folmer ([Kadaster](#))

Edwin Wisse ([Logius](#))

Auteur:

Wouter van den Berg ([TNO](#))

Doe mee:

[GitHub Logius-standaarden/BOMOS-LinkedData](#)

[Dien een melding in](#)

[Revisiehistorie](#)

[Pull requests](#)

This document is also available in this non-normative format: [pdf](#)

This document is licensed under a [Creative Commons Attribution 4.0 License](#).

Samenvatting

BOMOS (Beheer- en OntwikkelModel voor Open Standaarden) is een hulpmiddel van en voor de standaardisatiewereld. Dit deel bevat een aanvullende module over linked data. Het is een verkenning van de rol die linked data speelt en kan gaan spelen in moderne semantische standaarden.

Het beheer van dit document ligt bij TNO (de oorspronkelijke auteur). Geïnteresseerden en experts worden van harte uitgenodigd om feedback te delen, het document aan te vullen of anderzijds te verbeteren.

Status van dit document

Dit is een werkversie die op elk moment kan worden gewijzigd, verwijderd of vervangen door andere documenten. Het is geen door het Technisch Overleg goedgekeurde consultatieversie.

Inhoudsopgave

Samenvatting

Status van dit document

1. LINKED DATA & ONTOLOGIËN

- 1.1 Opkomst van Linked Data
- 1.2 Publieke sector
- 1.3 Private sector
- 1.4 De FAIR-beweging
- 1.5 Waarom Linked Data?
- 1.6 Impact op standaardisatie en beheer
 - 1.6.1 Hergebruik
 - 1.6.2 Training en opleiding
- 1.7 Een (tussen)weg kiezen
- 1.8 Doorontwikkelen
- 1.9 Gebruiksanalyse
- 1.10 Hulpmiddelen
 - 1.10.1 NeoNeoLogism
 - 1.10.2 Pool Party
 - 1.10.3 Protégé
 - 1.10.4 Semantic Treehouse
 - 1.10.5 TopBraid Composer
 - 1.10.6 VocBench
 - 1.10.7 Matey

Documentbeheer

Datum	Versie	Auteur	Opmerkingen
13/11/2020	1.0	Wouter van den Berg	Oplevering module
01/06/2022	2.0	Wouter van den Berg	Aanvullende modules in eigen documenten ondergebracht

Colofon

Logius Servicecentrum:	Postbus 96810 2509 JE Den Haag tel. 0900 555 4555 (10 ct p/m) email servicecentrum@logius.nl
------------------------	---

§ 1. LINKED DATA & ONTOLOGIËN

Linked data is al langere tijd in opkomst. De standaarden en technologieën die hieronder vallen -- zoals RDF, SPARQL en OWL -- worden steeds meer in de praktijk geadopteerd. Dat begon in de academische wereld maar heeft zich de afgelopen jaren uitgebreid naar de publieke sector en het bedrijfsleven.

Linked Data brengt veranderingen in de semantische standaardisatiediscipline die BOMOS omschrijft. Dit hoofdstuk licht die veranderingen toe. Het gaat niet enkel om het gebruik van nieuwe technologieën en hulpmiddelen. Linked Data betekent een andere benadering tot semantische interoperabiliteit: de nadruk verschuift van het standaardiseren van berichten (of 'transacties') naar het standaardiseren van de onderliggende domeinbeschrijving.

Enige basiskennis van de werking van Linked Data wordt in dit hoofdstuk verondersteld. Eerst komen de verschillen met conventionele semantische standaarden aan bod, zowel op het vlak van beheer en doorontwikkeling als in de toepassing door gebruikers. Tenslotte wordt er een overzicht geboden van digitale middelen die van nut kunnen zijn bij beheer en (door)ontwikkeling van met name ontologiën.

Toch eerst een toelichting op de betekenis van het woord 'ontologie'. Er is geen duidelijke scheiding tussen wat in de wereld van Linked Data wordt aangeduid als "vocabulaires" en "ontologiën". [Volgens de W3C](#) is het de trend om het woord "ontologie" te gebruiken voor complexere en formele verzamelingen van termen, terwijl "vocabulaire" wordt gebruikt voor simpelere varianten of waar minder strikt formalisme wordt gebruikt. In dit hoofdstuk wordt met het woord 'ontologie' het hele spectrum bedoeld.

Aan de basis van het hoofdstuk staat verder het volgende onderscheid tussen de twee soorten semantische standaarden (zie ook figuur 1):

1. Conventionele semantische standaarden. Deze standaarden lossen het interoperabiliteitsprobleem op door middel van gezamenlijke afspraken over de inhoudelijke structuur van berichten of transacties.
2. Semantische standaarden op basis van Linked Data. Deze standaarden lossen het interoperabiliteitsprobleem op door een gezamenlijke taal voor het beschrijven van dingen of

gebeurtenissen in een domein vast te leggen (in een ontologie). Dankzij die gezamenlijke taal is de gemeenschap niet meer afhankelijk van vaste berichtstructuren voor eenduidige interpretatie van de gedeelde data. (Het uitwisselen van gestandaardiseerde berichten blijft overigens gewoon mogelijk als daar binnen de gemeenschap behoefte toe is.)

§ 1.1 Opkomst van Linked Data

In de afgelopen jaren is in zowel de publieke als private sector steeds vaker de Linked Data benadering toegepast. Dit hoofdstuk beschrijft initiatieven in de publieke sector en de private sector, plus een die daar tussenin zit: de FAIR-beweging.

§ 1.2 Publieke sector

Binnen de publieke sector heeft de cultuursector altijd vooropgelopen. Het Europeana project (al gestart in 2008) was baanbrekend: Europeana Collections biedt inmiddels toegang tot meer dan 50 miljoen gedigitaliseerde items - boeken, muziek, kunst en meer - met geavanceerde zoek- en filterfuncties om het zoeken te vergemakkelijken. Net als in de andere sectoren begon het ook hier met de ontwikkeling van de modellen om de metadata van de collecties in te beschrijven. Ook andere culture instellingen hebben verschillende modellen in Linked Data uitgedrukt, bijvoorbeeld de Koninklijke Bibliotheek met de Nederlandse Thesaurus van Auteursnamen en de Rijksdienst Cultureel Erfgoed met onder andere het Archeologisch Basisregister.

Op het gebied van bouw en infrastructuur zijn zogenaamde Object Type Libraries (OTL) een bekend fenomeen. Een OTL is een soort semantische bibliotheek binnen de eigen organisatie. ProRail en Rijkswaterstaat zijn bekende organisaties die dit toepassen. Daarnaast is daar COINS, een andere belangrijke standaard die op Linked Data is gebaseerd, bedoeld voor de uitwisseling van BIM-modellen. Organisaties in de sector van geografische informatie waren er ook vroeg bij, met Geonovum als trekker van het Platform Linked Data Nederland. Het Kadaster had al snel de eerste modellen van de Basisregistraties Adressen en Gebouwen gepubliceerd, en vervolgens ook andere datasets. Na de publicatie van de modellen zijn de basisregistraties zelf als Linked Data gepubliceerd, en dan ook op 'productie'-niveau, dat wil zeggen met persistente URIs, actueel en gemonitord. Daarnaast wordt op labs.kadaster.nl ter inspiratie een groot aantal voorbeelden gegeven van de mogelijkheden die Linked Data biedt. Een belangrijke vervolgstap is dat de NEN3610 (het semantische 'moedermodel' van Nederlandse geo-informatiemodellen) inmiddels ook als Linked Data beschikbaar is.

Zo zijn er nog vele andere initiatieven in de publieke sector te noemen, zoals die in de landbouw en de vreemdelingenketen. Maar ook de Onderwijssector heeft al vele stappen richting Linked Data

gezet op basis van de ROSA, waarbij wederom de semantische modellen het startpunt zijn. Het voert te ver om die in dit hoofdstuk allemaal te belichten.

§ 1.3 Private sector

In de private sector komen Linked Data-ontwikkelingen in alle lagen voor: op het niveau van de branche en ketenprocessen, in individuele bedrijven, en in specifieke bedrijfsprocessen.

In het bedrijfsleven zijn er branches en gemeenschappen waar de rol van Linked Data steeds groter wordt, al is dat meestal in een verkennende fase. Het bedrijfsleven lijkt vooral nieuwsgierig naar de enorme hoeveelheid werk die in de academische wereld is verricht naar semantische technologieën en wat dit concreet kan betekenen om hun problemen te adresseren.

Een voorbeeld is te vinden in de uitzendbranche. Partijen uit die branche hebben in samenwerking met TNO het meerjarige Flexible IT (FIT) with Ontologies-project opgezet. Met dit project testen zij de voordelen van een overstap van uitwisseling middels de SETU-berichtstandaarden naar flexibelere uitwisseling middels de SETU-ontologie.

Daarnaast zijn er individuele bedrijven die met Linked Data aan de slag gaan. Bijvoorbeeld firma NXP Semiconductors, die in 2015 met hun Enterprise Datahub in de prijzen vielen in Nederlandse en Europese Linked Data competities. Dit systeem zorgt ervoor dat data hun vele producten (meer dan 20.000) uit verschillende databases geïntegreerd kunnen worden. Waar mogelijk heeft NXP gebruik gemaakt van bestaande standaard ontologieën als Dublin Core, SKOS, FOAF en Schema.org. Omdat er een aantal voor de sector benodigde termen en concepten ontbraken, is ook een NXP-vocabulaire ontwikkeld en aan de eerdergenoemde set gekoppeld. Op deze manier kan NXP Semiconductors de data rijk genoeg beschrijven en kunnen externe gebruikers de betekenis van begrippen gemakkelijker nagaan.

Tot slot worden Linked Data technologieën steeds meer toegepast in de processen in het bedrijfsleven. Een bekend voorbeeld is te vinden in de online marketing, specifiek in zoekmachineoptimalisatie (Engels: search engine optimization, of SEO). Met SEO streven bedrijven ernaar om een webpagina zo hoog mogelijk te laten scoren in de zoekresultaten van een zoekmachine. Tegenwoordig past men hiervoor JSON-LD toe. JSON-LD, een methode om Linked Data over te dragen via het bekende JSON-formaat, wordt gebruikt door zoekmachines om informatie over websites te verzamelen om zo de gewenste zoekresultaten te bepalen. Google, eigenaar van 's werelds meest gebruikte zoekmachine, raadt de toepassing van JSON-LD aan. De opkomst van JSON-LD is ook een voorbeeld van Linked Data die verbinding vindt met veelgebruikte data-formaten.

§ 1.4 De FAIR-beweging

De hoeveelheden data waar wetenschappers en andere data-analysten mee te maken hebben blijven groeien. Daarmee groeit ook de vraag naar meer machine-ondersteuning bij verwerking en analyse van die data. Maar hier bestond simpelweg de benodigde data-infrastructuur niet voor. De FAIR Data Principles is een reactie op dit gemis. FAIR staat voor een set richtlijnen om data meer vindbaar (**F**indable), toegankelijk (**A**ccesible), Interoperabel en herbruikbaar (**R**eusable) te maken. De nadruk van FAIR ligt op het vergroten van het vermogen van machines om data automatisch te vinden en te gebruiken.

De FAIR-principes zijn in 2016 in een wetenschappelijk artikel geïntroduceerd. De FAIR-beweging is daarna in korte tijd flink in kracht toegenomen. Zo spraken de leiders van de G20 in hetzelfde jaar hun steun uit voor FAIR, en richtten Duitsland, Nederland en Frankrijk in 2017 samen het GO FAIR initiatief op.

De *I* in FAIR onderstreept het belang van geautomatiseerde data-interoperabiliteit. Het doel is het bereiken en bewaken van interoperabiliteit, zelfs wanneer gegevens in een heel ander formaat, in verschillende talen en in verschillende onderzoeksdomeinen zijn gemaakt. De link tussen FAIR en Linked Data is hier heel sterk. Dat blijkt uit de drie FAIR-principes die onder de *I* van *Interoperability* vallen:

1. (meta)data gebruiken een formele, toegankelijke, gedeelde en breed toepasbare taal voor kennisrepresentatie.
2. (meta)data gebruiken vocabulaires die FAIR-principes volgen.
3. (meta)data bevatten gekwalificeerde verwijzingen naar andere (meta)data.

Voor het implementeren van de Interoperability principes van FAIR wordt veel gebruik gemaakt van ontologiën (voor punt 1) en andere Linked Data standaarden (voor punten 2 en 3).

Links:

- [The FAIR Guiding Principles for scientific data management and stewardship](#), Nature, 2016, het wetenschappelijk artikel over FAIR
- [FAIR Principles](#), GO FAIR Initiative, de FAIR-principes

§ 1.5 Waarom Linked Data?

Linked Data maakt het mogelijk om rijke datasets te delen en gecontroleerd open te stellen zonder dat de die dataset de bron hoeft te verlaten. De Europeana Collections zijn hier een voorbeeld van. Niet alleen de data binnen zo een dataset is gelinkt, ook de datasets onderling zijn dat. Met een query-taal als SPARQL wordt het vervolgens mogelijk om de hele set van datasets te bevragen. Dit is waar de visie van het Semantische Web over gaat. Hoe meer Linked Data, hoe krachtiger de query's die mogelijk worden.

Met name overheden gebruiken Linked Data steeds meer om data openbaar te maken voor het publiek, het liefst op een manier waarmee geïnteresseerden de data makkelijk automatisch kunnen opvragen en verwerken. Hier is het bekende vijf-sterren model op van toepassing:

- 1 ster:** de data is beschikbaar via internet; in welk formaat maakt niet uit.
- 2 sterren:** de beschikbare data heeft een machine-leesbaar structuur (geen scan bijvoorbeeld).
- 3 sterren:** de data is beschikbaar in een niet-gepatenteerd formaat (zoals .csv en niet MS Excel).
- 4 sterren:** Linked Data gepubliceerd met behulp van W3C standaarden (RDF en SPARQL).
- 5 sterren:** al het bovenstaande en de dataset is gelinkt aan andere Linked Open Data.

Linked Data wordt ook ingezet door digitaliserende organisaties om hun data en systemen verder te integreren, zoals te zien bij Rijkswaterstaat en ProRail met hun OTLs. De scope van dit soort inspanningen is vaak eerst intern en breidt zich vervolgens uit naar grote leveranciers of andere vaste partnerorganisaties. Door die uitbreiding wordt er met relatief weinig inspanning extra voordeel uit de stap naar Linked Data behaald.

Linked Data maakt het verder mogelijk voor machines om automatisch de betekenis (semantiek) van data in te lezen en zo zelf nieuwe verbindingen te leggen, verbindingen die normaal gesproken alleen door mensen gelegd konden worden. Linked Data is een essentieel component van nieuwe generatie Kunstmatige Intelligentie (A.I.) technieken, waarbij enorme hoeveelheden ruwe data wordt gecombineerd met geformaliseerde domeinbeschrijvingen om steeds slimmere machines te bouwen.

Ten slotte vormt Linked Data de kern van een nieuwe aanpak tot het vergroten van interoperabiliteit tussen partijen in een branche of waardeketen. De semantische standaard die men hiervoor vastlegt, bevat dan een of meer ontologiën. Net als bij andere soorten semantische standaarden is het doel om het voor de gebruikers zo makkelijk mogelijk te maken om onderling koppelvlakken in te richten die aan alle wensen voldoen. Het Flexible IT with Ontologies-project van de uitzendbranche is een voorbeeld van een gebruikersgemeenschap die de overstap verkent naar informatie-uitwisseling middels een gestandaardiseerde ontologie.

Het gebruik van Linked Data is onderhevig aan het bekende netwerk-effect. Voor elk van bovenstaande toepassingen geldt: hoe meer wijdverspreid het gebruik van Linked Data wordt, des te groter de voordelen voor gebruikers.

Ga naar de website van Platform Linked Data Nederland (www.pldn.nl) voor meer informatie over hoe Linked Data in zijn werk gaat

§ 1.6 Impact op standaardisatie en beheer

Oorspronkelijk werd in BOMOS geen nader onderscheid gemaakt tussen de soorten semantische standaarden. Het is bij zowel de conventionele aanpak als bij de Linked Data-benadering van toepassing. Maar de precieze invulling van de activiteiten verschilt wel degelijk tussen de twee. Invulling geven aan het BOMOS-model was overigens altijd al een opgave, omdat elke context maatwerk behoeft. Maar bij de Linked Data-benadering tot semantische standaarden komen nieuwe uitdagingen kijken.

Hieronder volgen wat uitdagingen en lessen uit het FIT with Ontologies-innovatietraject van de uitzendbranche. De volgorde is grofweg van vroege ontwikkelingsfase naar beheerfase.

§ 1.6.1 Hergebruik

Ontologies ontstaan zelden in een **green-field** scenario. Juist digitale waardeketens waar al vergevorderde uitwisselingsafspraken gemaakt zijn maken sneller de stap naar een oplossing op basis van Linked Data. De documentatie van deze standaarden bevat vaak veel informatie die van waarde is voor de nieuwe oplossing, zoals gedetailleerde beschrijvingen van ketenprocessen en bedrijfsregels (*business rules* in het Engels). Het is belangrijk dat die kennis en informatie zoveel mogelijk wordt hergebruikt. Hergebruik zorgt niet alleen voor tijdsbesparing in het ontwikkelproces van de nieuwe oplossing, maar ook voor meer vertrouwen in het eindresultaat en dus voor betere adoptie door de gebruikersgemeenschap.

§ 1.6.2 Training en opleiding

De overstap naar gestandaardiseerde uitwisseling op basis van Linked Data betekent vrijwel altijd dat er een andere technology stack gebruik wordt (bijvoorbeeld met o.a. OWL, RDF en/of SPARQL). Dat vraagt om andere kennis en vaardigheden, zowel van de betrokkenen bij ontwikkeling en beheer van de standaard als bij de gebruikers.

De beheerorganisatie dient voldoende expertise over deze technologieën in huis te hebben om haar

beheertaken uit te kunnen voeren, zoals het inschatten van impact van verandervoorstellen op de standaard of het beantwoorden van technische vragen. Dat spreekt voor zich. Maar ook aan de kant van de gebruikersorganisaties zal men waarschijnlijk het een en ander moeten bijleren. Uit ervaring blijkt dat de meeste softwareontwikkelaars weinig of geen ervaring hebben met deze technologieën. Om de actieve deelname van gebruikers in het ontwikkelproces te borgen is het belangrijk dat er voldoende training en opleiding beschikbaar is, en dat gebruikers de tijd krijgen om de technologie zich eigen te maken. De beheerdersorganisatie kan daarin op allerlei manieren voorzien, zoals het schrijven van gidsen, het geven van demo's, of zelfs het organiseren van een Linked Data hackathon.

Training en opleiding is ook van belang wanneer de gebruikersgemeenschap al standaarden voor informatie-uitwisseling toepaste voordat men de overstap naar Linked Data maakte. De gebruikers en zelfs de beheerders zijn dan gewend te denken in termen van de oude technology stack en de gebruikte methoden van uitwisseling. In die gevallen wordt de stap naar de Linked Data aanpak vaak behandeld als een wisseling van syntax, terwijl er in werkelijkheid ook een andere denkwijze nodig is. Dat geldt voor zowel de ontwikkelingsfase als voor de daadwerkelijke toepassing van de semantische standaard. Bij de ontwikkeling loop je anders het risico dat de status quo een te grote invloed heeft op hoe de ontologie gemodelleerd wordt, met verarming van semantiek van het model tot gevolg. Bij de toepassing kan de oude denkwijze leiden tot omslachtige manieren van data opvragen of uitwisselen, wat leidt tot verarming in de semantiek van de data. Zo observeerde de beheerders van de Basisregistratie Adressen en Gebouwen (BAG) dat veel gebruikers bij het opvragen van data onbedoeld ook de hele historie van de data meevroegen.

§ 1.7 Een (tussen)weg kiezen

Het is overigens niet zo dat de keuze enkel bestaat uit ofwel een volledige overstap maken naar Linked Data, of alles te houden bij het oude. Er zijn gradaties en tussenstappen denkbaar.

Een technologie die hierin een essentiële rol speelt is RML (RDF Mapping Language), ontwikkeld door een team van de Universiteit van Gent. Met RML is het mogelijk om semi-gestructureerde data als JSON, XML en CSV te transformeren naar Linked Data. Dat doe je door eerst de RML regels die voor de betreffende data moeten gelden op te schrijven. Daarna voed je deze RML regels plus de oorspronkelijke data aan een zogenaamde RMLMapper, die er Linked Data triples van maakt.

Het toepassen van RML zou wel eens voor veel gemeenschappen met standaarden het beste alternatief balans kunnen zijn op volledig overstappen op Linked Data of niets doen. In principe volstaat het om door middel van RML mappings een tussenlaag aan te brengen die de vertaling maakt tussen de relevante interne systemen van de gebruiker en de gemeenschappelijke ontologie. Dankzij die mappings kan men vervolgens door middel van RESTful APIs toch Linked Data uitwisselen, terwijl de huidige systemen in stand kunnen blijven.

VOORBEELD 1

Deze strategie werd ook in het FIT with Ontologies-project gehanteerd. Een belangrijk doel in dat project was het bewaken van backwards compatibility met de berichtuitwisseling volgens de conventionele SETU-standaarden. Met andere woorden, het moest met de nieuwe techniek mogelijk blijven om SETU-berichten te ontvangen en verwerken. Men heeft daarom gewerkt aan de FIT engine: een softwarecomponent dat RML toepast om SETU-berichten om te zetten naar Linked Data. Met de FIT engine kunnen de partijen die de overstap naar uitwisseling van Linked Data gemaakt hadden alle vruchten hiervan plukken, zonder dat hun bestaande koppelingen met klanten daarvoor op de schop hoeven.

§ 1.8 Doorontwikkelen

Stel: de eerste versie van de nieuwe semantische standaard op basis van Linked Data is af. De gebruikersgemeenschap is naartig aan de slag met implementeren. Waar wordt het beheer en doorontwikkeling vanaf nu door gekenmerkt?

Ten eerste is hetgeen dat in beheer genomen wordt, de nieuwe standaardontologie, meer modulair van aard dan je als beheerder van een conventionele standaard gewend bent. Doorontwikkeling is nog meer gericht op samenstelling van bestaande ontologiën en andere bouwblokken dan bij de conventionele aanpak het geval was. Dit is bijvoorbeeld te zien in de Geo-standaarden. Daar hebben nieuwe ontologiën relatief weinig toegevoegd. Er zijn veel bestaande vocabulaires geassembleerd. De gezamenlijke afspraken gaan dan over hoe je die zaken compileert. Dat was met de conventionele aanpak voor semantische standaarden al in mindere mate zo, bijvoorbeeld met XML Schema, maar met ontologieën is dat meer en voor grotere blokken.

Dat neemt niet weg dat een deel van de domeinbeschrijving in een eigen ontologie gemodelleerd is. Die eigen ontologie moet doorontwikkeld worden om met de tijd mee te gaan. De uitdaging hierbij is om dit zo zorgvuldig en precies te blijven doen. De geavanceerdere toepassingen van ontologiën, zoals complexe query's of automatisch redeneren, zijn namelijk gevoeliger voor fouten. Meer nog dan bij andere semantische standaarden is dus de nauwkeurigheid en coherentie van de ontologie van kritisch belang.

Niet alle wijzigingsverzoeken leiden tot veranderingen in de gemeenschappelijke ontologie. Soms is het beter om een uitbreiding te ontwikkelen, die vervolgens in gebruik wordt genomen door het deel van de gebruikersgemeenschap dat hier profijt van heeft. Ontologiën en andere Linked Data standaarden zijn hiervoor ontworpen. Als beheerder is het zaak om uitbreidingen voor de gebruikers expliciet en inzichtelijk te maken. Hier kan je vervolgens controle en analyses op uitvoeren.

§ 1.9 Gebruiksanalyse

In het FIT with Ontologies-project maakt men gebruik van een webapplicatie om, uit de SETU-ontologie, zelf berichtdefinities bij elkaar te klikken. Deze specifieke berichtdefinities zien er anders uit, maar voldoen allemaal aan de ontologie en zijn eenduidig te vertalen naar feiten uit die ontologie. Doordat deze verschillende berichtdefinities centraal opgeslagen staan, wordt gedetailleerde gebruiksanalyse mogelijk, vergelijkbaar met het code coverage concept uit de software engineering. Deze term duidt op het uitzoeken welke delen (e.g. concepten, relaties) van een ontologie door wie in welke berichten daadwerkelijk worden gebruikt, alsmede de trends daarin. Bevindingen uit code coverage testen bieden belangrijke aanwijzingen voor doorontwikkeling en kunnen discussies over de standaard voeden met echte gebruiksdata.

§ 1.10 Hulpmiddelen

Hier zijn een aantal hulpmiddelen die voor beheerders van gemeenschappelijke ontologiën van nut kunnen zijn. Pool Party, Protégé en TopBraid Composer hebben aanzienlijke overlap; met alle drie is het mogelijk om ontologiën te modelleren. Semantic Treehouse is niet gericht op modelleren maar op het beheer- en doorontwikkelingsproces, en is de enige van de vier met een directe link met BOMOS. Matey (en de gerelateerde RML en YARRRML talen) zijn gericht op de transformatie van traditionele dataformaten zoals JSON, XML en CVS naar Linked Data die voldoet aan een specifieke ontologie.

Middel	Ontwikkelaar	Omschrijving
NeoNeoLogism	Netage B.V.	Een met de webbrowser te gebruiken RDF Schema vocabulaire-editor en publicatieomgeving.
Pool Party	Semantic Web Company GmbH	Knowledge management, data analytics & content organization
Protégé	Stanford University School of Medicine	Ontology editor (gratis en open source)
Semantic Treehouse	TNO	Publicatie- en beheeromgeving voor gezamenlijke ontologiën en andere datamodellen
TopBraid Composer	TopQuadrant	Integrated Development Environment (IDE) voor Semantic Web, inclusief ontology editor
VocBench	EU Publicatiebureau	Open-source publicatie- en beheeromgeving van meertalige vocabulaires

Middel	Ontwikkelaar	Omschrijving
Matey	Universiteit Gent	Matey is een browsergebaseerde editor waarmee je YARRRML-regels kunt schrijven.

§ 1.10.1 NeoNeoLogism

NeoNeologisme is een eenvoudig hulpmiddel voor het schrijven en publiceren van vocabulaires in RDF(S), niets zoals TopBraid, Protégé, VocBench of PoolParty. Het is bedoeld om eenvoudig maar expressief genoeg te zijn om subklassen en subattributen en verschillende vormen van relaties te ondersteunen.

§ 1.10.2 Pool Party

De PoolParty Semantic Suite biedt verschillende functies waarmee organisaties meer waarde uit hun eigen gegevens kunnen halen. Het product commercialiseert Semantic Web-technologieën. In 2009 kwam de eerste release van de PoolParty Semantic Software op de markt. Sindsdien is het product geëvolueerd van een taxonomiebeheertool naar een functierijk semantisch softwareplatform waarmee bedrijven ontologiën kunnen inzetten om gestructureerde en ongestructureerde gegevens te integreren.

§ 1.10.3 Protégé

Protégé is een gratis, open-source platform dat een reeks tools aanbiedt om domeinmodellen en applicaties met ontologiën te bouwen. Er bestaat een webversie en een desktopversie. WebProtégé is een ontologie-ontwikkelomgeving voor het web om ontologiën te creëren, uploaden, wijzigen en delen. Gebruikers kunnen deze gezamenlijk bekijken en bewerken. Protégé Desktop bevat daarnaast extra functionaliteit zoals 'semantic reasoners' en visualisatietools.

§ 1.10.4 Semantic Treehouse

Semantic Treehouse is het communityplatform dat het beheer van ontologiën en andere soorten datamodellen ondersteunt. Wat Semantic Treehouse uniek maakt, is dat het de publicatie, het onderhoud en het beheer van uw datamodellen op één plek combineert. Semantic Treehouse is

gebaseerd op meer dan 10 jaar ervaring met het ontwikkelen, onderhouden en delen van datastandaarden volgens BOMOS. Semantic Treehouse valt onder de categorie Software as a Service, en afnemers hebben de mogelijkheid om een eigen huisstijl in te richten voor de gebruikersinterface voor een herkenbare gebruikerservaring.

§ 1.10.5 TopBraid Composer

TopBraid Composer is een desktoptool om ontologiemodellen mee te maken en gegevensbronnen te integreren. Het ondersteunt visuele bewerking van ontologiën en kan verbinden met database-backends zoals Oracle en verschillende populaire triplestores. De uitgebreide editie is een IDE die de mogelijkheid toevoegt om snel semantische webapplicaties te ontwikkelen.

§ 1.10.6 VocBench

VocBench 3 is gefinancierd door het ISA²-programma van de Europese Commissie. Het is een gratis platform dat gezamenlijke bewerking en beheer van codelijsten, taxonomieën, thesauri en ontologiën mogelijk maakt. Naast community management (beheer van meerdere gebruikers, roldefinitie, toegangscontrole), biedt VocBench 3 een mechanisme waarmee gebruikers hun model kunnen koppelen aan Jira, een soort issue en bug tracking platform voor software.

§ 1.10.7 Matey

Het idee van Matey is dat het RML toegankelijker maakt.

De RDF Mapping Language (RML) is een generieke mappingtaal om regels uit te drukken die gegevens in heterogene structuren en serialisaties (o.a. JSON, XML, CSV) toewijzen aan het RDF-gegevensmodel. Zie <https://rml.io/docs/rml/introduction/> voor meer informatie over RML.

YARRRML is een voor mensen leesbare op tekst gebaseerde weergave voor declaratieve regels voor het genereren van Linked Data. Het is een subset van YAML, een veelgebruikte data-serialisatie-taal die is ontworpen om mensvriendelijk te zijn. YARRRML kan worden gebruikt om RML-regels weer te geven. Zie <https://rml.io/yarrml/> voor meer informatie over YARRRML.