

Gebruik van OWL versus Named Graphs in Bouw Informatie Model (BIM) systemen

Bastiaan Bijl MSc

Sysunite B.V.
SYS-2014-01NL

Keywords:

BIM, RDF, OWL, Semantic Web, Named Graphs, closed-world, open-world

Abstract:

Het loont in een closed-world BIM-systeem om een objectenbibliotheek-structuur te hanteren die klasse-beschrijvingen bevat waaraan instantiedata getoetst kan worden voordat deze tot het informatiesysteem wordt toegelaten. Dit paper gaat in op de onnodige complexiteit die ontstaat als OWL gebruikt wordt in een closed-world.

De beste richting lijkt het ontwikkelen van een eigen structuur op basis van RDFS die gebruik kan maken van omgevingskenmerken en waarin zo min mogelijk reasoning nodig is. Hiermee verschilt het closed-world BIM-systeem van het Semantic Web (open-world).

Gebruik van OWL versus Named Graphs in Bouw Informatie Model (BIM) systemen

1. *Wat is een BIM-systeem?*

Een BIM-systeem is een centraal informatiesysteem waarmee alle betrokken partijen bij een bouwproject gegevens kunnen delen waarbij deze gegevens eenduidig, expliciet en compleet (in de context van de projectfase) gedefinieerd zijn. Elke partij heeft zijn eigen *view* op het informatiesysteem, wat betekent dat hij alleen de voor hem relevante gegevens inziet en kan wijzigen.

2. *Wat is RDF?*

RDF is een techniek om informatie te structureren aan de hand van statements. Initieel werd een statement beschouwd als een triple. Een triple in de context van RDF is een set van drie URI's. Later werd een vierde element aan de statement toegevoegd. Dit element, de naam van de triple, is ook een URI, en kan gebruikt worden om naar het statement als geheel te verwijzen. Het statement inclusief zijn naam wordt daarom ook een Named Graph genoemd.

3. *Waarom is RDF een bruikbare techniek voor een BIM-systeem?*

De RDF-techniek is bedacht om het Internet om te bouwen van een World Wide Web naar een **Semantic Web**. Dit zijn de belangrijkste eigenschappen van het Semantic Web:

- a) URI's worden gebruikt als **globaal unieke identifiers** van informatie-eenheden.
- b) De informatie is opgeslagen in een **netwerkstructuur**, met als grootste voordeel dat de modelstructuur van de informatie niet op voorhand vastgelegd hoeft te worden. Dit is b.v. wel het geval bij databases met een tabelstructuur (relationele databases).
- c) De **semantische betekenis** van URI's wordt op een eenduidige wijze vastgelegd met behulp van statements. Deze statements verwijzen naar objectbibliotheken waarin klassen van objecten zijn gedefinieerd. Deze definities zijn qua structuur hetzelfde opgebouwd als alle andere informatie op het Semantic Web.
- d) De informatie kan worden geleverd door **verschillende deelsystemen** die alle te benaderen zijn via het Internet.
- e) Het Semantic Web gaat uit van een **open-world**.

Een BIM-systeem lijkt wat de eerste vier punten betreft sterk op het Semantic Web. De open-world aanname is daarentegen een radicaal verschil met een BIM-systeem. Het verschil zit er in dat een BIM-systeem uitgaat van een dataset die eenduidig gedefinieerd is, consistent is en dat de bronnen die bijdragen aan deze dataset altijd beschikbaar zijn. Een query in een BIM-omgeving dient namelijk altijd een eenduidige, vooraf te voorspellen en volledig resultaat op te leveren en niet afhankelijk te zijn van het wel of niet beschikbaar zijn van bronnen.

4. Wat is de **open-world** aanname?

De open-world aanname voor het Semantic Web houdt in dat er geen conclusie getrokken mag worden uit de afwezigheid van data. Als bijvoorbeeld een query op het Semantic Web naar alle gebouwen van boven de 800 meter geen resultaten oplevert betekent dit niet dat deze gebouwen niet bestaan. Dit omdat op het Internet niet gegarandeerd kan worden dat alle data-services altijd online zijn. Kort gezegd is in de open-world een ja wel een ja, maar een nee geen absolute nee.

Behalve deze openheid m.b.t. de beschikbaarheid van data is het Semantic Web ook open om later nieuwe objectbibliotheken toe te voegen. Deze bevatten dan nieuwe klassen waaraan de al bestaande informatie gerelateerd kan worden. Dit hoeft niet expliciet te gebeuren. OWL voorziet in het later kunnen koppelen van oude informatie en nieuw toegevoegde objectbibliotheken.

5. Hoe werkt OWL?

Met behulp van OWL kan een klasse worden beschreven door eigenschappen te identificeren waaraan instanties van de klasse moeten voldoen. De klasse *fiets* kan bijvoorbeeld beschreven worden als een object met twee wielen en een fietsbel. Vervolgens kan automatisch worden uitgerekend welke objecten aan deze eisen voldoen (vanwege het feit dat deze twee wielen en fietsbel bezitten). Er wordt dan afgeleid dat sommige objecten nu ook lid zijn van de klasse fiets.

- a) OWL is bedoeld om vooraf of achteraf restricties op te leggen aan welke eisen een object moet voldoen om tot een bepaalde klasse te behoren.
- b) In OWL bestaan complexe begrippen die gebruikt kunnen worden om precies te beschrijven hoe klassen zich tot elkaar verhouden.
- c) Veel uitspraken die in deze OWL-terminologie gedaan kunnen worden over objecten (het simpelste begrip is het lidmaatschap van een instantie tot een klasse) zijn niet bij voorbaat deel van het informatiesysteem, maar moeten afgeleid worden wanneer er vraag naar is.
- d) In OWL is het opstellen van klasse-beschrijvingen niet de grootste moeilijkheid, maar het afleiden van missende en mogelijk conflicterende relaties tussen de aanwezige objecten en de klasse-beschrijvingen uit beschikbare objectenbibliotheek. Dit proces is softwarematig ingewikkeld en kost veel performance.
- e) Deze opzet in OWL is nodig om het **open-world** karakter van het Semantic Web vorm te geven.
- f) OWL en met name de beschikbare tools en software libraries hebben nog een hoog academisch gehalte. Er is geen goede .NET-ondersteuning.
- g) OWL is ontwikkeld in de tijd dat statements nog uit drie URI's bestonden. Een koppeling maken van OWL naar de Named Graphs-structuur is niet eenvoudig. Een eigen klasse-beschrijving structuur op basis Named Graphs is wel gemakkelijk aan OWL-klassen te koppelen.

6. Wat is nodig in een closed-world zoals het BIM-systeem?

Omdat in een closed-world wel gegarandeerd kan worden dat alle informatiebonnen beschikbaar zijn en een meer rigide controle gevoerd kan worden op welke data wordt toegelaten tot het informatiesysteem is het mogelijk een structuur te hanteren waarin

later nauwelijks of geen informatie hoeft te worden afgeleid. OWL biedt in zo'n omgeving derhalve geen toegevoegde waarde.

Voordelen van een specifieke eenduidige klasse-beschrijving structuur in plaats van OWL:

- a) Als wordt geëist dat alle informatie die naar een BIM-systeem wordt verstuurd expliciete verwijzingen bevat naar elementen uit de objectenbibliotheek hoeft het lidmaatschap van een object tot een klasse nooit meer afgeleid te worden, maar kan dit direct worden opgevraagd. Dit scheelt aanzienlijk in performance en evenzo in de complexiteit van de software.
- b) De benodigde ondersteuning van software libraries wordt daarmee minimaal.
- c) Voor een specifiek gebruik kan het meest efficiënte gebruik van de Named Graphs methodologie gekozen worden.
- d) Alle objecten kunnen vooraf van restricties en mogelijk default-waarden en/of voorgedefinieerde template-waarden worden voorzien.
- e) De relaties kunnen in een eigen structuur (los van de objecten) van restricties worden voorzien. Hiermee wordt de objectenbibliotheek fijnmaziger omdat klassen van objecten en klassen van relaties los kunnen worden beschreven.

Conclusie

OWL heeft de potentie veel te gaan betekenen in de Internet-wereld (open-world) daar waar het gaat om het exploreren van de vele aanwezige informatiebronnen op het Internet ten einde enerzijds een maximaal resultaat met anderzijds een focus op relevante antwoorden op een gestelde vraag. In het socio-georiënteerde WEB waar goede richting gevende informatie vooralsnog belangrijker is dan exacte, eenduidige informatie, is dit een terechte benadering.

In een BIM omgeving is echter de beschikbare data evenals relaties tussen de data gelimiteerd tot de daadwerkelijk aanwezige en bewust gelegde relaties. Een vraag aan een BIM-omgeving leidt daarom tot een van te voren gedefinieerd antwoord: er is maar één antwoord mogelijk. Het reasoning principe van OWL is hier vooralsnog niet relevant en maakt een BIM omgeving nodeloos complex. RDFS (wat de basis vormt van OWL) en specifiek de uitbreiding Named Graph past heel goed in de behoefte van BIM-omgevingen om de grote rijkheid van verschillende bronnen van informatie in een gebouwde omgeving te integreren, te ontsluiten en of uit te wisselen.