



notitie

FORUM STANDAARDISATIE 16 december 2014
Agendapunt 5. Open standaarden, lijsten
Stuknummer 5B. Intake-advies SKOS

Advies

Het Forum Standaardisatie wordt geadviseerd om SKOS, een standaard voor het online delen en linken van systemen voor kennisrepresentaties (zoals thesauri, begrippenwoordenboeken en classificatielijsten) via het internet, in procedure te nemen op de 'pas toe of leg uit'-lijst.

Over de standaard:

Het publiceren van thesauri en taxonomieën door overheidsorganisaties gebeurt vaak in een vorm van documenten die niet bruikbaar zijn voor computerprogramma's. Daarnaast worden aan de begrippen die in thesauri en taxonomieën zijn opgenomen verschillende classificaties toegekend. Een goed voorbeeld hiervan is de SBI-lijst, een classificatielijst van economische activiteiten, die door de Kamer van Koophandel in de vorm van een downloadbaar document met een 'platte tabel' wordt gepubliceerd. Wanneer een andere organisatie met de SBI-codes wil werken moeten er eerst handmatig relaties worden gelegd met de in de eigen organisatie gehanteerde begrippen. Dit is onnodig en kan leiden tot interpretatieverschillen. Als daarbij nog gegevens worden gekoppeld uit een ander systeem is de kans op interpretatiefouten nog groter.

Simple Knowledge Organization System (SKOS) zorgt er voor dat gegevensmodellen van kennisrepresentaties (zoals thesauri, begrippenwoordenboeken en classificatielijsten) bruikbaar zijn voor computerprogramma's (machine readable) en uitgewisseld kunnen worden tussen computerprogramma's onderling. Ook kunnen kennisrepresentaties gedeeld worden op het internet. Gebruik van de standaard maakt de (familie)relaties tussen de verschillende definities van begrippen inzichtelijk.

Het gebruik van SKOS geeft actoren beter toegang tot begrippenapparaten binnen en buiten de organisatie, en ICT-systemen kunnen direct worden gevoed met referentiegegevens uit kennisrepresentaties in plaats van een eigen set referentiegegevens te moeten creëren. Hierdoor kunnen

'begrippenapparaten' geanalyseerd, vergeleken en vertaald worden naar de eigen organisatie. Dit zorgt voor tijdswinst omdat relevante informatie sneller gevonden kan worden en geen tijd hoeft worden verspild aan het creëren van eigen begrippenlijsten. Het geeft inzicht in de samenhang en (in)consistentie van begrippen (en bijbehorende definities), zorg voor toename in het gebruik van eigen overheidsinformatie.

Voorbeeld I

In de Basisregistratie Adressen en Gebouwen (BAG) zijn alle adressen in Nederland verzameld. Veel andere basisregistraties maken gebruik van de definitie 'adres', maar deze adressen hoeven niet per definitie geregistreerd te zijn in de BAG. Bijvoorbeeld wanneer de eigenaar van organisatie X in het buitenland woont, staat zijn adres wel in het Handelsregister, maar niet in de BAG.

Toch wordt in beide gevallen gesproken van een 'adres'. Het begrip 'adres' in het Handelsregister is een breder begrip dan het begrip 'adres' in de BAG. Met SKOS kunnen deze begrippen (ondanks dat zij niet exact hetzelfde zijn) toch met elkaar in verband worden gebracht.

SKOS bouwt voort op de Linked Data-standaarden RDF, RDFS en OWL. Linked data is een methode om gestructureerde gegevens te publiceren op het internet en met andere data te verbinden. Hierdoor worden verbanden tussen verschillende open databronnen zichtbaar. Deze samenhang maakt het makkelijk om relevante informatie te vinden, te structureren en te koppelen. Verwacht wordt dat naarmate meer (overheids)organisaties gebruik maken van Linked Data en op Linked Data-gebaseerde standaarden, de toepassingsmogelijkheden van SKOS zullen toenemen. Gebruikers van SKOS hoeven RDF en RDFS niet apart te implementeren voordat gebruik kan worden gemaakt van SKOS. Dit wordt door de tooling van SKOS automatisch ('onder water') geregeld.

De standaard is aangemeld door het Platform Linked Data Nederland en de aanmelding wordt ondersteund door de Rijksdienst voor het Cultureel Erfgoed van het Ministerie van Onderwijs Cultuur en Wetenschappen (OCW). In Nederland wordt de standaard momenteel door een aantal (semi-) overheidsorganisaties gebruikt zoals de Rijksdienst voor Cultureel Erfgoed, Brandweer Nederland, Stichting Bibliotheek.nl en het Nederlands Instituut voor Beeld en Geluid. Binnen het Platform Linked Data zijn organisaties als Belastingdienst, Kadaster, Rijkswaterstaat, Geonovum en Ministerie BZK betrokken. (zie <http://www.pilod.nl/wiki/Categorie:Organisaties>).

Over het toepassingsgebied

Geadviseerd wordt om bij de expertsessie stil te staan bij een goede definiëring van het functioneel toepassingsgebied van de standaard. Een voorstel is om het functioneel toepassingsgebied van de standaard als volgt te definiëren:

Het online beschikbaar en deelbaar maken van kennisrepresentaties (zoals begrippenwoordenboeken, classificatielijsten en thesauri).

Toelichting

1. Aanmelding, intakegesprek en toetsingsprocedure

Op 30 oktober 2014 is door Jos Voskuil namens het Platform Linked Data Nederland een melding ingediend, betreffende de aanmelding van SKOS voor de lijst voor 'pas toe of leg uit'.

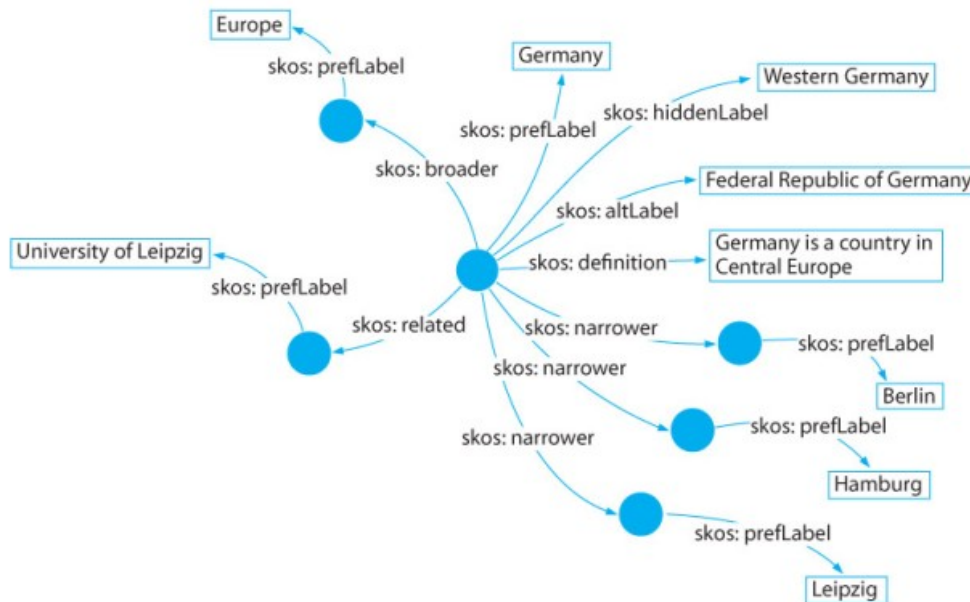
Op 11 november 2014 heeft een intakegesprek plaatsgevonden met de aanmelder. Hierbij is gekeken of alle basisinformatie aanwezig is en of de standaard voldoet aan de criteria voor inbehandelname. Daarnaast is vooruitgeblikt op de procedure.

2. Korte beschrijving standaard

Waar gaat de standaard over?

Simple Knowledge Organization System (SKOS) is een gegevensmodel voor het delen van kennisrepresentaties (bijvoorbeeld via het internet) en het koppelen van systemen voor kennisrepresentatie. Kennisrepresentatie is de wijze waarop en de structuur waarin kennis is vastgelegd. Systemen voor kennisrepresentatie worden ook wel Knowledge Organisation Systems (KOS) genoemd.

Systemen voor kennisrepresentatie zijn veelal gegroundvest op eenzelfde conceptueel karakter, wat het mogelijk maakt om deze representaties met elkaar te vergelijken. SKOS maakt de overeenkomstige structuurelementen uit deze kennisrepresentaties, bijvoorbeeld thesauri, taxonomieën, begrippenwoordenboeken, classificatieschema's expliciet. Hierdoor zijn de kennisrepresentaties bruikbaar voor computerprogramma's ('machine readable') en kunnen deze uitgewisseld worden tussen applicaties onderling en op het internet. Dit laatste draagt bij aan verbeterde semantische interoperabiliteit bij gegevensuitwisseling, omdat de (familie)relaties tussen de verschillende definities inzichtelijk worden gemaakt. Het gaat hierbij om de relaties bij begripsdefinities als 'gelijk aan', 'breder dan', 'smaller dan' en 'gerelateerd aan'.



Figuur 1: voorbeeld van SKOS-vocabulaire.

Door het gebruik van SKOS krijgen gebruikers beter toegang tot begrippenapparaten binnen en buiten de eigen organisatie (kennismanagement) en kunnen ICT-systemen direct worden gevoed met referentiegegevens uit thesauri in plaats van een eigen set van referentiegegevens te moeten creëren (verbeterde ICT-ondersteuning).

Voorbeeld II

In de wettelijke regelingen die uitvoeringsorganisaties zoals SVB, UWV, Belastingdienst, DUO en gemeenten uitvoeren wordt het begrip loon en/of inkomen gehanteerd. De definitie van het begrip komt niet altijd met elkaar overeen.

Het gebruik van SKOS maakt het mogelijk om gegevensuitwisseling plaats te laten vinden, met behoud van de verschillende definities van het begrip. Hierdoor hoeven definitiekwesties niet eerst beslecht te worden voordat er gegevensuitwisseling tussen deze organisaties kan plaatsvinden.

SKOS maakt het dus mogelijk om definities naast elkaar te gebruiken en toe te passen. Door het gebruik van SKOS worden verschillende definities onder één noemer gestructureerd door de duiding van relaties tussen de definities. Intersectorale gegevensuitwisseling en internationale gegevensuitwisseling worden hierdoor beter mogelijk gemaakt. De toegevoegde waarde van deze standaard wordt alleen maar groter doordat steeds meer verschillende organisaties en verschillende disciplines in ketenverband met elkaar samenwerken. (zie ook: 7. Functionele use case)

Wie beheert de standaard?

De standaard wordt beheerd door World Wide Web Consortium (W3C).

Waarom is de standaard aangemeld voor 'pas toe of leg uit'?

SKOS bouwt voort op de Linked Data standaarden RDF, RDFS en OWL. Gebruikers van SKOS hoeven RDF en RDFS niet apart te implementeren voordat gebruik kan worden gemaakt van SKOS. Dit wordt door de tooling van SKOS automatisch ('onder water') geregeld.

Naarmate Linked Data standaarden breder worden toegepast en de hoeveelheid op Linked Data-gebaseerde standaarden toeneemt, ontstaan er steeds meer toepassingsmogelijkheden voor SKOS-gebaseerde vocabulaires. SKOS-vocabulaires kunnen ingezet worden voor de bevordering van semantische interoperabiliteit.

3. Criteria voor inbehandelname

Om een standaard in behandeling te nemen moet de standaard vallen binnen de scope van de lijsten. Hiervoor gelden drie criteria:

1. Is de standaard toepasbaar voor elektronische gegevensuitwisseling tussen (semi-)overheidsorganisaties en bedrijven, tussen (semi-)overheidsorganisaties en burgers of tussen (semi-)overheidsorganisaties onderling?

Ja. De standaard is specifiek bedoeld om elektronische gegevensuitwisseling te verbeteren voor met name gegevenssets waarbij het belangrijk is dat er onderling afspraken gemaakt moeten worden over waardenlijsten.

2. Is het beoogde functioneel toepassingsgebied en het organisatorisch werkingsgebied van de standaard, voldoende breed om substantieel bij te dragen aan de interoperabiliteit van de (semi-)overheid?

Ja. De standaard is algemeen toepasbaar, ook binnen het werkgebied van de (semi-)overheid.

Bij de behandeling van de standaard door de experts dient nagedacht te worden over een goede definiëring van het toepassingsgebied. Een voorstel is om het functioneel toepassingsgebied van de standaard als volgt te definiëren:

Het online beschikbaar en deelbaar maken van kennisrepresentaties (zoals begrippenwoordenboeken, classificatielijsten en thesauri).

3. Is het zinvol de standaard op te nemen, gezien het feit dat deze niet al wettelijk verplicht is voor het beoogde functioneel toepassingsgebied en organisatorisch werkingsgebied?

Ja. De standaard is niet wettelijk verplicht en opname van de standaard op de 'pas toe of leg uit'-lijst kan helpen om de adoptie van de standaard verder te bevorderen.

Conclusie

De standaard voldoet aan de criteria voor inbehandelname.

4. Toetsing kansrijkheid procedure

Het Forum Standaardisatie wil voorkomen dat er standaarden in procedure worden genomen, waarvan bij voorbaat al bekend is dat deze in de expertronde of consultatieronde zullen stranden op één van de inhoudelijke criteria. Daarom heeft de procedurebegeleider de beantwoording van de criteriavragen nagelopen, waar mogelijk zelf aangevuld en vervolgens besproken met de indiener.

1. Open standaardisatieproces

De ontwikkeling en het beheer van de standaard moeten op een open, onafhankelijke, toegankelijke, inzichtelijke, zorgvuldige en duurzame wijze zijn ingericht.

De standaard wordt beheerd door World Wide Web Consortium (W3C), een internationaal samenwerkingsverband dat als doel heeft open standaarden te ontwikkelen om de groei van het internet op lange termijn te verzekeren.

Het specificatiedocument en documentatie over het ontwikkel- en beheerproces zijn kosteloos en voor een ieder te downloaden van de website van W3C (www.w3.org/TR/skos-reference). Deelname aan het standaardisatieproces is kosteloos en voor een ieder toegankelijk. Gebruikers kunnen via de website van W3C doorlopend issues, wensen en behoeften kenbaar maken op het gebied van ontwerp en de toepassing van SKOS (<http://www.w3.org/2001/sw/wiki/SKOS/Issues>). Er zijn geen aanvullende standaardisatieprofielen noodzakelijk voor het inzetten van SKOS.

Om deel te kunnen nemen aan de ontwikkeling en strategievorming van de standaard kunnen organisaties lid worden van W3C. Organisaties die een partnership met W3C aangaan kunnen deelnemen aan diverse werkgroepen op het gebied van ontwikkeling en strategievorming. Het lidmaatschap kost voor (semi-)overheidsorganisaties jaarlijks € 7.800,00. Logius is per 1 juli 2014 als eerste overheidsorganisatie in de Benelux toegetreden als lid van W3C.

Door het niet-verplichte karakter van het lidmaatschap en de beperkte kosten voor een eventueel lidmaatschap wordt dit niet gezien als een belemmering voor het gebruik van de standaard.

2. Toegevoegde waarde

De interoperabiliteitswinst en andere voordelen van adoptie van de standaard wegen overheidsbreed en maatschappelijk op tegen de kosten, de risico's en nadelen. Voor elk van de te onderscheiden stakeholders (overheid, bedrijven en burgers) afzonderlijk zouden de baten voor de informatievoorziening en de bedrijfsvoering op moeten wegen tegen de kosten. Verder moeten de risico's aan overheidsbrede adoptie van de standaard (beveiliging, privacy) acceptabel zijn. Het publiceren van thesauri en taxonomieën door overheidsorganisaties gebeurt nu vaak in een vorm van documenten die niet bruikbaar is voor

computerprogramma's. Een goed voorbeeld hiervan is de SBI-lijst, een classificatielijst van economische activiteiten, die door de Kamer van Koophandel in de vorm van een downloadbaar document met een 'platte tabel' wordt gepubliceerd. Wanneer een andere organisatie met de hiërarchisch geordende SBI-codes wil werken moeten deze relaties eerst gelegd worden. Dit is onnodig, leidt mogelijk tot interpretatieverschillen en daardoor ook tot fouten. SBI-codes zouden met SKOS eenduidig beschikbaar kunnen worden gemaakt, waardoor de hiërarchie tussen verschillende SBI-codes behouden blijft.

Een ander voorbeeld is de Justitiethesaurus van het Wetenschappelijk Onderzoek- en Documentatiecentrum (WODC). Deze thesaurus wordt uitgewisseld met de Politieacademie en de Raad voor de Kinderbescherming. Het uitwisselingsformaat van deze thesaurus is een specifiek formaat van een Content Management System leverancier (CMS-leverancier). Daardoor kunnen alleen organisaties gebruik maken van deze thesaurus als zij een CMS gebruiken van de desbetreffende leverancier. Het toepassen van SKOS doorbreekt deze leveranciersafhankelijkheid.

Het belang in inzicht in de onderliggende waardenlijsten neemt toe naarmate er in de toekomst (in ketenverband) meer samengewerkt wordt tussen verschillende (semi-) overheidsorganisaties en tussen verschillende disciplines.

De invoering van SKOS is laagdrempelig. Er zijn meerdere SKOS-tools beschikbaar, ook via Open Source (http://en.wikipedia.org/wiki/Simple_Knowledge_Organization_System).

3. Draagvlak

Aanbieders en gebruikers moeten voldoende ervaring hebben met de implementatie en het gebruik van de standaard.

De standaard wordt in Nederland momenteel door een aantal (semi-) overheidsorganisaties gebruikt zoals de Rijksdienst voor Cultureel Erfgoed, Brandweer Nederland, Stichting Bibliotheek.nl en het Nederlands Instituut voor Beeld en Geluid.

De standaard wordt in Nederland ook in de private sector gebruikt en kent een breder draagvlak in Europa en daarbuiten. SKOS wordt ook door de Britse Cabinet Office aanbevolen als standaard voor de Britse overheid.

4. Opname bevordert adoptie

De opname op de lijst moet een geschikt middel zijn om de adoptie van de standaard te bevorderen.

SKOS is aangemeld voor opname op de 'pas toe of leg uit'-lijst. Gebruik van de standaard is groeiende, maar kent op dit moment nog niet de omvang die nodig is om te kunnen worden beschouwd als gangbare standaard. Opname van de standaard op de 'pas toe of leg uit'-lijst kan zodoende helpen om de adoptie van de standaard verder te bevorderen.

Conclusie

Er zijn op voorhand geen struikelblokken te verwachten in de procedure.

5. Samenhang

Forumstandaardisatie wil weten of de aangemelde standaard samenhangt met standaarden die reeds op de 'pas toe of leg uit' -lijst en gangbare lijst zijn opgenomen, of standaarden die voor toetsing in aanmerking komen. Uit de intake moet duidelijk worden of dit gevolgen heeft voor de toetsing en eventuele opname van de aangemelde standaard.

1. *Bestaat er samenhang tussen de aangemelde standaard en de standaarden die reeds op de 'pas toe of leg uit' -lijst zijn opgenomen en wat betekent dit voor de toetsing en eventuele opname van de standaard?*
Er zijn geen standaarden gevonden die samenhang kennen met SKOS. De toepassing van SKOS heeft wel een ondersteunende werking voor standaarden die waardenlijsten bevatten, zoals het Semantisch Model e-Factuur (SMeF) of XBRL, maar conflicteert niet met deze standaarden.
2. *Bestaat er samenhang tussen de aangemelde standaard en de standaarden die reeds op de gangbare lijst zijn opgenomen en wat betekent dit voor de toetsing en eventuele opname van de standaard?*
SKOS bouwt voort op de Linked Data-standaard RDF en maakt gebruik van de standaard XML. SKOS conflicteert niet met deze standaarden. Toename in het gebruik van Linked Data-standaarden zoals RDF leidt tot meer toepassingsmogelijkheden voor SKOS-gebaseerde vocabulaires.
3. *Bestaat er samenhang tussen de aangemelde standaard en standaarden die in aanmerking komen voor opname op één van de lijsten en wat betekent dit voor de toetsing van de standaard(en)? (Denk bijvoorbeeld ook aan een gezamenlijke toetsing met (een deel van) deze aanvullende standaarden).*
SKOS kent samenhang met de standaarden OWL en ISO 25964. OWL is een door W3C beheerde standaard voor het definiëren van webontologieën. SKOS onderscheidt zich van OWL doordat de vocabulaires in OWL onderdeel zijn van een geformaliseerd kennissysteem, terwijl SKOS van toepassing is op niet-geformaliseerde kennissystemen. OWL en SKOS sluiten naadloos op elkaar aan.

Daarnaast kent de standaard samenhang met SPARQL Query Language for RDF en Turtle. SPARQL is een standaard die gebruikt wordt om RDF-gebaseerde data door middel van zoekopdrachten te bevragen (zoekfunctionaliteit). Turtle (Terse RDF Triple Language) is een universele taal voor het weergeven van informatie op het internet.

SKOS kent enige overlap met de standaard ISO 25964, een standaard die thesauri ondersteunt bij informatieontsluiting, en de keuze voor begrippen die gebruikt worden bij indexeren, taggen en

zoekopdrachten. Gebruik van deze ISO-standaard leidt niet tot een voor computers bruikbare kennisrepresentatie zoals SKOS dit doet.

SKOS conflicteert niet met deze standaarden.

6. Sponsorschap

De aanmelding van standaarden voor de lijsten van Forum en College dient ondersteund of gesponsord te worden door overheids- en/of (semi)publieke organisaties die de standaard reeds in gebruik hebben (of voornemens zijn dit te doen) en die de beoogde opname op de lijsten ondersteunen. Dit draagt bij aan het draagvlak voor de standaard, geeft zich op de functionele usecase voor de overheid en helpt bovendien om tijdens de toetsing de juiste experts te benaderen.

1. Welke overheden en/of (semi) publieke organisaties ondersteunen de aanmelding van de standaard?
De standaard wordt in Nederland momenteel door een aantal (semi-)overheidsorganisaties gebruikt zoals de Rijksdienst voor Cultureel Erfgoed, Brandweer Nederland, Stichting Bibliotheek.nl en het Nederlands Instituut voor Beeld en Geluid. Verwacht wordt dat deze organisaties de aanmelding van de standaard ondersteunen.

Daarnaast is de standaard ingediend namens het Platform Linked Data. Verwacht wordt dat de organisatie, zoals de Belastingdienst, Kamer van Koophandel en de provincie Overijssel, die betrokken zijn bij dit platform de aanmelding steunen, dit wordt nog voorgelegd (zie [http://www.pilod.nl/wiki/Categorie: Organisations](http://www.pilod.nl/wiki/Categorie:Organisaties)).

2. Hebben deze organisaties de standaard geïmplementeerd?
(zie ook punt 7 voor een uitwerking)
Ja. Een aantal van deze organisaties hebben de standaard geïmplementeerd. Zie 6.1.

7. Functionele use case

Voor de standaard dient een duidelijke use case beschikbaar te zijn op basis waarvan overheden en/of instellingen uit de (semi) publieke kunnen bepalen of de aangemelde standaard voor hen relevant is en wie eventueel moet deelnemen aan de experttoetsing van de standaard.

In 2012 heeft het programma CATCHPlus van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) in samenwerking met de Rijksdienst voor Cultureel Erfgoed en een aantal softwareleveranciers het OpenSKOS-platform opgezet. Op dit platform kunnen door de toepassing van SKOS zowel thesauri worden geraadpleegd als uitgewisseld. Het doel van dit platform is om samenwerking tussen de verschillende instituten in de cultureel erfgoedsector te bevorderen, en zo tijd en kosten te besparen (<http://openskos.org/>).