KB nationale bibliotheek

Bijlage 17 - Verslag marktconsultatie thesauri

Ten behoeve van het beheren van vier nationale thesauri:

1. **Nederlandse Thesaurus van Auteursnamen (NTA)**
2. **Nederlandse Thesaurus Corporaties (NTC)**
3. **Nederlandse Drukkersthesaurus (NTD)**
4. **Brinkmanthesaurus**

INHOUDSOPGAVE

INHOUDSOPGAVE **2**

1 INLEIDING **3**

2 DOEL VAN DE MARKTCONSULTATIE **3**

3 SAMENVATTING MARKTINPUT PER VRAAG **3**

3.1 TECHNIEK EN DIENSTVERLENING 3

3.2 DOORONTWIKKELING PRODUCT 7

3.3 ALGEMENE VRAGEN 7

1 Inleiding

De KB heeft een marktconsultatie uitgevoerd ter voorbereiding op de aanbesteding voor het beheren van nationale thesauri en terminologiebronnen. Middels een openbare marktconsultatie op TenderNed (TN 569146) zijn marktpartijen uitgenodigd inzicht te geven in mogelijke oplossingsrichtingen, technische en functionele mogelijkheden, ondersteuning, doorontwikkeling en aandachtspunten voor de aanbesteding en overeenkomst.

In totaal hebben acht marktpartijen antwoord gegeven op de vragen uit de marktconsultatie. Met vier marktpartijen heeft de KB verdiepende gesprekken gevoerd. Dit verslag bevat een samenvatting op hoofdlijnen van de ontvangen marktinput.

2 Doel van de marktconsultatie

Het doel van de marktconsultatie was om inzicht te verkrijgen in de mate waarin de markt oplossingen kan bieden voor het beheren, publiceren, raadplegen en integreren van meerdere thesauri en terminologiebronnen met verschillende datamodellen, governanceprocessen en gebruikersgroepen. Daarnaast is gevraagd naar aandachtspunten voor gunningscriteria, duurzaamheid, SROI, looptijd en contractuele inrichting.

3 Samenvatting marktinput per vraag

3.1 Techniek en dienstverlening

1.1.a Leveringsvorm

De markt laat een breed aanbod zien. SaaS is de meest genoemde leveringsvorm, veelal als beheerde EU-gehoste dienst. Daarnaast bieden meerdere partijen ook on-premise of private-cloud varianten aan. Voor on-premise hosting worden vooral containerisatie, Linux-omgevingen, Kubernetes/Docker, persistente opslag, back-upvoorzieningen, observability en afspraken over toegang voor beheer genoemd. Enkele partijen geven aan dat dezelfde release- en deploymentlogica in SaaS en on-premise kan worden gebruikt om overdraagbaarheid te vergroten.

1.1.b Technologiestack on-premise

Niet alle partijen beantwoorden deze vraag, omdat zij uitsluitend of primair SaaS leveren. Waar on-premise wordt ondersteund, worden onder meer Linux, Docker/Kubernetes, Java-, Python- en Rust-gebaseerde applicatiestacks, relationele databases, RDF-/graph-databases, zoekcomponenten, REST-, GraphQL-, SPARQL- en gRPC-koppelingen, CI/CD, monitoring en identity- en accessmanagement genoemd. De precieze stack verschilt sterk per oplossing.

1.1.c Verschillen in datamodel, beheerders en beheerproces

De markt geeft aan dat multi-thesaurusbeheer mogelijk is via afzonderlijke projecten, tenants, containers, profielen of omgevingen per thesaurus. Per thesaurus kunnen datamodellen, validatieregels, rollen, rechten, workflows en publicatieregels verschillen. Gezamenlijk beheer met netwerkpartners wordt doorgaans ondersteund met rolgebaseerde toegang, partner-specifieke rechten, audittrail en configureerbare goedkeuringsprocessen. Meerdere partijen benadrukken dat configureerbaarheid belangrijker is dan maatwerk.

1.1.d Standaard voor modellering

SKOS/RDF wordt door vrijwel alle partijen genoemd als belangrijke basis voor thesauri en linked data. Daarnaast worden onder meer ISO 25964, OWL, RDFS, SHACL, Dublin Core, schema.org, FOAF, BIBFRAME, MADS, DCAT, NL-SBB, LinkML en SSSOM genoemd. De rode draad is dat open standaarden en W3C Semantic Web-standaarden belangrijk worden geacht voor interoperabiliteit, duurzaamheid en beperking van vendor lock-in.

1.1.e Raadplegen door eindgebruikers

De markt noemt verschillende raadpleegkanalen: webinterfaces voor zoeken en bladeren, thesaurusviewers, hiërarchische boomstructuren, graphvisualisaties, publieke raadpleegportalen, REST API’s, SPARQL-endpoints, downloads en integratie met bestaande zoek- of catalogiseersystemen. Enkele partijen noemen daarnaast wijzigings- of suggestieformulieren voor niet-beherende gebruikers en conversationele of natuurlijke-taalinterfaces als aanvullende laag.

1.1.f Duurzame identifiers van buiten de oplossing

Partijen geven aan duurzame identifiers te kunnen ondersteunen via persistente HTTP-URI’s, externe URI’s, PID’s en identifierrelaties. Externe identifiers kunnen doorgaans worden

opgeslagen, gemapt of gekoppeld via relaties zoals sameAs, exactMatch of specifieke velden. Genoemde externe bronnen en identifiers zijn onder meer VIAF, ISNI, ORCID, GND, Wikidata en andere authority files. Content negotiation en dereferenceable URI's worden door enkele partijen expliciet genoemd.

1.1.g Verwijzingen tussen termen en datasets

De meeste partijen ondersteunen verwijzingen via SKOS- en OWL-relaties, zoals broader, narrower, related, exactMatch, closeMatch, relatedMatch, sameAs, seeAlso en custom properties. Voor echte naam en pseudoniem wordt zowel het modelleren als alternatieve labels binnen één entiteit genoemd als het beheren van afzonderlijke entiteiten met expliciete semantische koppeling. Enkele partijen adviseren om mapping- of linkset-artefacten met provenance te gebruiken voor beheer en kwaliteitscontrole.

1.1.h Import, export en metadatastandaarden

Veelgenoemde import- en exportformaten zijn RDF-serialisaties, CSV, Excel, XML, JSON-LD, MARC/MARCXML, MADS, BIBFRAME/RDF, ONIX, TEI/XML en dumps. Ook API's, SPARQL CONSTRUCT, OAI-PMH, Z39.50, SFTP, REST, GraphQL en incrementele event- of synchronisatiestromen worden genoemd. Een terugkerend aandachtspunt is volledige export in open formaat, mede ten behoeve van exit en dataportabiliteit.

1.1.i Integratie catalogiseren en terminologiebeheer

De markt ziet integratie met catalogiseren als essentieel. Genoemde mogelijkheden zijn lookup vanuit catalogiseersystemen, direct aanmaken van nieuwe entiteiten, conceptvoorstellen vanuit het catalogiseerproces, review- en goedkeuringsstappen, API-koppelingen, deeplinks, SSO en integratie met bibliotheek- of ILS-omgevingen. Voor NTA wordt de koppeling met het catalogiseerproces als bijzonder belangrijk gezien.

1.1.j Workflow en gezamenlijk beheer

Workflowfunctionaliteit varieert per oplossing. Veelgenoemd zijn stappen als voorstel, review, validatie, goedkeuring, publicatie, versiebeheer, intrekking en notificaties. Gezamenlijk beheer wordt ondersteund met gescheiden rollen, partner-specifieke rechten, wijzigingsvoorstellen, audittrail, provenance en traceerbaarheid. Eén partij geeft aan dat workflowbeheer nog niet standaard aanwezig is en in samenwerking ontwikkeld kan worden.

1.1.k Licentiemodel

De markt noemt vooral SaaS-abonnementen, jaarlijkse licenties, raam- of abonnementsmodellen en eenmalige implementatiekosten. Prijsbepalende factoren zijn onder meer aantal beheerders of gebruikers, aantal thesauri, omvang van datasets, implementatiemodel, rekencapaciteit, supportniveau en optionele modules. Ook worden open-sourceachtige of broncode-toegankelijke modellen genoemd. Voor de aanbesteding is het raadzaam transparantie te vragen over licentiegrondslag, meerwerk, indexatie, API-gebruik, dataverkeer en exitkosten.

1.1.l Beperkingen gebruikers en records

De meeste partijen geven aan geen harde beperkingen te hanteren voor het aantal gebruikers of thesaurusrecords. Wel wordt door meerdere partijen opgemerkt dat performance en schaalbaarheid afhankelijk zijn van infrastructuur, sizing, opslagcapaciteit en representatieve benchmarks. Aandachtspunt is dat schaalbaarheid aantoonbaar moet worden gemaakt, met name voor grote en frequent muterende datasets.

1.1.m Ondersteuning

Genoemd worden implementatiebegeleiding, migratieondersteuning, trainingen, train-the-trainer, trainings- en acceptatieomgevingen, supportdesk, SLA/DAP, ticketsystemen, telefonische ondersteuning, documentatie, workshops, releasecommunicatie en escalatiepaden. Meerdere partijen adviseren om servicelevels en rollen in de supportketen expliciet af te stemmen, inclusief de inrichting van eerste- en tweedelijnsondersteuning.

1.1.n Gebruikersorganisatie

De markt beschrijft de gebruikersorganisatie vooral als rol- en rechtenmodel met beheerders, redacteuren, validators, reviewers, approvers, catalogiseerders, technische beheerders, netwerkpartners en eventueel publieke verzoekers. SSO, identity providers, groepen, projectrechten en fijnmazige autorisatie worden veel genoemd. Enkele partijen bieden ondersteuning bij beleid, governance, kwaliteitsbeheer en inrichting van beheerprocessen.

3.2 Doorontwikkeling product

1.2.a Ontwikkelroadmap en toekomstvisie

Vrijwel alle partijen geven aan te beschikken over een roadmap of toekomstvisie. Terugkerende thema's zijn linked data, open standaarden, schaalbaarheid, AI-ondersteuning, knowledge graphs, semantische interoperabiliteit, toegankelijkheid, beveiliging, configureerbaarheid, performance en uitbreiding naar toekomstige terminologiebronnen.

1.2.b Totstandkoming roadmap

Roadmaps komen volgens de markt tot stand via een combinatie van interne productstrategie, klantfeedback, gebruikersoverleg, sectorontwikkelingen, standaardisatieontwikkelingen, research, klantadviesraden, roadmap-sessies en periodieke prioritering. Meerdere partijen noemen co-creatie en structurele afstemming als manier om functionaliteit relevant en toekomstbestendig te houden.

1.2.c Invloed KB

Partijen geven aan dat de KB invloed kan uitoefenen via roadmap-overleg, gebruikersoverleg, klantadviesraden, co-creatietrajecten, ontwikkelbudgetten, change-controlprocessen en gezamenlijke prioritering. Voor contractrelevante onderwerpen zoals URI-beleid, exportformaten, validatieregels, mappings, publicatiecontracten en integraties is formele governance aan te bevelen.

3.3 Algemene vragen

2.1.a Passende gunningscriteria

De markt noemt als onderscheidende onderwerpen vooral open standaarden, linked-data-architectuur, schaalbaarheid en performance, migratieaanpak, integratie met bestaande systemen, governance en workflowondersteuning, datakwaliteit, security, privacy, toegankelijkheid, gebruikersvriendelijkheid, support en kennisoverdracht. Ook worden dataportabiliteit, beperking van vendor lock-in, aantoonbare ervaring met thesauri/authority data en proof-of-concept of demonstratie als relevante criteria genoemd.

2.1.b Duurzaamheid en SROI

Duurzaamheid kan volgens de markt worden ingevuld via duurzame hosting, energie-efficiënte software, beperking van onnodige dataverwerking, open standaarden en dataportabiliteit, kennisdeling, opleiding, digitale toegankelijkheid, inclusie, CSR/ESG-beleid, circulaire IT en maatschappelijke bijdragen. SROI wordt in deze specialistische IT-context

vooral gezien in kennisoverdracht, opleidingen, inclusieve arbeidspraktijken, open kennisdeling en ontwikkel- of leertrajecten.

2.1.c Looptijd raamovereenkomst

De adviezen lopen uiteen. Genoemd worden onder meer 2 jaar met 2 maal 1 jaar verlenging, 3 jaar met 1 jaar verlenging, 3 tot 4 jaar initieel, en een initiële looptijd van 4 jaar met aanvullende verlengopties. Gegronde redenen voor een langere looptijd zijn migratiecomplexiteit, inrichting van governance, terugverdienen van implementatie-investeringen, continuïteit van persistente URI's, stabiliteit van nationale infrastructuur en beperking van heraanbestedingslasten.

2.1.c Overige contractuele aandachtspunten

Terugkerende aandachtspunten zijn prijsindexering, duidelijke exitregeling, volledige data-export in open formaat, overdraagbare documentatie, afspraken over eigendom van data, schema's en mappings, servicelevels, beveiliging, change management, versiebeheer, deprecationbeleid, portability-tests, acceptatiecriteria, benchmarks, governance rond identifiers en expliciete afspraken over doorontwikkeling.

2.2 Overige tips en overwegingen

De markt adviseert onder meer om interoperabiliteit centraal te stellen, migratie zwaarder te wegen dan gebruikelijk, een gefaseerde aanpak of proof of concept te overwegen, performance-eisen expliciet te maken, NL-SBB/DERA/NDE-compatibiliteit toetsbaar te formuleren, duurzame URI-strategie contractueel te borgen, toekomstige uitbreiding mogelijk te maken, roadmapoverleg vast te leggen en leveranciers te laten aangeven welke functionaliteit reeds beschikbaar is en welke nog specifiek ontwikkeld moet worden.