

Toelichting op de werkwijze rondom de Monitor Makers 2016

Rapportage monitor en selectie algemene publieksboeken

Besloten is dat we de monitor voor de makers op twee niveaus zullen rapporteren. Bij de presentatie van de cijfers zal telkens vermeld worden op welk niveau de gepresenteerde gegevens betrekking hebben. Hieronder lichten we de twee niveaus toe.

Op **totaalniveau** – als referentiekader – waarbij geen onderscheid wordt gemaakt tussen boeksoort AWSO. Hier gaat het om de totale titelproductie en de auteurs en vertalers die daar verantwoordelijk voor zijn. Deze analyse zal op basis van Nederlandse papieren titels gemaakt worden.

Omdat we toch wat willen zeggen over auteurs en vertalers van algemene boeken (A+O) versus studie- en wetenschappelijke boeken (S+W), zal er een selectie gemaakt worden op die titels waarbij de boeksoort bekend is.

Er is uit het totale titelbestand van de KB (met 193.527 ISBNs in de periode 2008 t/m 2016) een selectie gemaakt van titels waarvan vaststaat dat ze binnen de categorie ‘**algemeen publieksboek**’ vallen. Deze selectie betreft 55% van het totaal door KB geregistreerde nieuwe Nederlandstalige papieren titels (135.827). 16% van het totale bestand zijn titels met een wetenschappelijk en/of studie karakter. Daarnaast zijn er nog titels waarvan niet bekend is in welke categorie ze vallen (29%). Deze titels zijn bij het bespreken van de algemene markt buiten beschouwing gelaten, net als de wetenschappelijke- en studieboeken.

Binnen de context van de Monitor Makers 2017 is de markt voor **algemene publieksboeken dus de selectie titels die gedistribueerd wordt door de grootste boekendistributeur van Nederland CB**. CB – voorheen Centraal Boekhuis – is de logistieke dienstverlener van de boekensector en bestrijkt het grootste gedeelte van Nederland en Vlaanderen. Er zijn redenen voor uitgevers om niet via CB te distribueren: het kan voor een uitgever niet rendabel zijn om via CB te leveren (kosten/baten), er kan gebruik worden gemaakt van andere distributiekanaal (denk aan aldi/ betapress), of omdat een uitgave heel gericht te koop wordt aangeboden bij een specifieke selectie winkels (denk aan museumwinkels).

Op de volgende pagina wordt in detail beschreven op welke wijze het titelbestand en de selectie van algemene publieksboeken tot stand is gekomen.

Titelbestand en bestand met auteurs en vertalers

Op basis van de KB-titelbestanden heeft Ingressus een titelbestand opgeleverd met 193.527 ISBNs. Daarnaast is er een bestand opgeleverd waar de auteurs en vertalers betrokken bij de ISBNs vermeld zijn. Uit analyse blijkt dat het aantal (unieke) ISBNs in de twee bestanden verschilt.

Vanuit deze selectie titels is bekeken welke auteurs en vertalers Nederlandse titels op hun naam hebben staan. Indien er een auteur of vertaler door KB is geregistreerd, is deze

opgenomen in het auteurs/vertalers bestand. Indien meerdere auteurs of vertalers gewerkt hebben aan een titel, zijn die ook opgenomen.

Het verschil is te verklaren door het feit dat niet alle titels een auteur of vertaler hebben: bloemlezingen, bundels samengesteld werk, bundels onder redactie en fotoboeken bijvoorbeeld. Fotografen, samenstellers en redacteuren worden niet als auteur gecatalogiseerd, maar vallen voor deze monitor in de categorie 'titels van overige makers'.

Databestand

Het originele bronbestand dat, net als in 2016, vanuit de Koninklijke Bibliotheek (KB) naar onderzoeksbureau Ingressus is verstuurd bevat records van titels (ISBNs) die diverse eigenschappen hebben. Bijvoorbeeld een NUR waarmee het boek ingedeeld wordt in een genre en een AWSO-codering die gebruikt wordt voor de indeling in het soort boek: algemeen, wetenschappelijk, studie en overig.

Uit de analyse van het bestand dat gebruikt is voor de monitor editie 2016 bleek dat veel velden van NUR en AWSO niet gevuld zijn in de catalogus van de KB, omdat dit niet standaard data zijn die KB aangeleverd krijgt. Is de informatie er, dan neemt het KB het op in de titelinformatie, is het er niet, dan niet.

KVB Boekwerk wil graag specifiek rapporteren over titels die behoren tot het Algemene publieksboek, dit zijn titels met een A of O codering. Daarnaast neemt het literair-culturele segment een belangrijk aandeel in de markt in. Dit marktsegment wordt geselecteerd op basis van de NUR (een lijst met NUR-codes is hier te vinden). Het is daarom van belang dat deze data bij zoveel mogelijk titels aanwezig is in de dataset.

Om deze reden is voor de Monitor 2017 onderzocht of de dataset van de KB verrijkt kan worden met data vanuit een andere bron. Die bron is gevonden, namelijk het CB.

Proces vanuit Ingressus

- Ingressus heeft bij CB een volledig titelbestand opgevraagd en gekoppeld aan de KB-data. Het CB- bestand bevatte de NUR-data die de KB-data sterk verrijkte, maar dit bestand bevatte geen AWSO-informatie.
- Om toch over AWSO-informatie te kunnen beschikken, heeft Ingressus namens KVB Boekwerk bij CB ook een uitvoer van AWSO-informatie gekocht.
- Ingressus heeft beide CB-bestanden gekoppeld en bewerkt, zodat er alleen ISBNs in een nieuw bestand overbleven waarvan de NUR en/of AWSO-informatie beschikbaar is.
- ISBN+NUR: 1.640.162 records
- ISBN+AWSO: 317.036 records
- De aangeleverde XML-bestanden van de KB zijn tot slot gematcht op ISBN met het hiervoor beschreven nieuwe bestand, waardoor de NUR en AWSO-informatie is toegevoegd aan de KB-data. Hierdoor was het mogelijk de procedure voor de monitor 2016 te herhalen.

Resultaat verrijking voor NUR

De verrijking is voor NUR goed gelukt. De NUR is opgenomen in het titelbank bestand en door verrijking zijn er nu nog 14.296 records zonder NUR in de KB-bestanden 2008 t/m 2016. Zo'n 7,4% van het totaal aantal records dat door Boekwerk geanalyseerd wordt ten behoeve van de Monitor Makers 2017, waarvan het merendeel (10.786) ISBNs betreft die buitenlands zijn en überhaupt niet meegenomen worden in de analyse.

177.605 titels in het KB-bestand zijn verrijkt met een NUR uit CB. 1755 titels zijn nog aanwezig met een NUR uit KB. Doordat CB leidend is gemaakt op basis van het daar aanwezig ISBN is het in 76 gevallen zo dat KB wel een NUR Veld heeft en CB niet, en dat daardoor de NUR als onbekend is aangemerkt. Het aantal ISBNs dat wel in KB met een NUR-code aanwezig is en niet in CB is zodanig klein (76 ISBNs op 194.000 in totaal) dat is besloten dat we deze niet toevoegen aan de dataset.

Resultaat verrijking voor AWSO

De verrijking voor het veld AWSO bleek niet vanuit Titelbank te kunnen. Ingressus heeft een apart bestand opgevraagd bij CB. Door navraag bleek dat bestand alleen data van ISBNs te bevatten die door CB gedistribueerd worden. De toekenning van een AWSO-code wordt uitgevoerd door CB in samenwerking met de eigenaren van de titels. De verrijking is uitgevoerd op basis van dat bestand. Uit analyse na de verrijking bleek dat het aantal titels zonder AWSO-codering nog aanzienlijk groot was. Ter illustratie: in 2015 ging het om 6.000 titels op een totaal van ruim 20.000 titels.

Deze 'restgroep' wilden we verder onderverdelen en zo klein mogelijk maken. Vandaar dat we gekeken hebben of we de AWSO op een andere wijze konden toekennen. Dit is niet gelukt. Er is geen methode gevonden op basis waarvan we een getrouwe inschatting kunnen maken. Adviesbureau Berenschot is gevraagd mee te lopen in dit proces, wat ook gebeurd is.

Na doorlopen van dit proces kunnen we concluderen dat alleen titels die via het CB gedistribueerd worden een AWSO-codering krijgen. Bij titels die buiten het CB om gedistribueerd worden ontbreekt het veld AWSO. Een uitzondering hierop zijn de titels die bij KB aangemeld worden met een AWSO, maar vervolgens niet door CB zijn gedistribueerd.

Er is dus een aanzienlijk aantal titels in het databestand waarvan de AWSO-code ontbreekt. Hieronder een overzicht van de onderverdeling.

Vergelijking KB en CB AWSO veld	
Totaal records	193.527
AWSO bekend	125.620
AWSO source CB	124.353
AWSO source KB	1.267
Totaal AWSO onbekend (X)	67.907

Bij het verrijken van het databestand is het veld vanuit CB leidend gemaakt, voor zowel NUR als AWSO. Indien het veld bij CB gevuld was is CB leidend gemaakt en is NUR of AWSO gevuld met de gegevens vanuit CB. Bij nadere analyse bleek dat er nog een restgroep was waarbij AWSO wel gevuld was bij KB, maar door CB overschreven werd, terwijl de waarde

leeg was. Omdat we een zo groot mogelijk aantal ISBNs met een AWSO-code mee willen nemen in de analyse voor Boekwerk is onderzocht of we deze ISBNs vanuit de KB konden worden toegevoegd. Om na te gaan of dit methodologisch verantwoord is, is er een aparte analyse uitgevoerd. Voor een set aan titels waarvoor zowel bij CB als bij de KB een AWSO AWSO-codering beschikbaar is, is nagegaan of deze coderingen van elkaar afwijken en in welke mate. Uit deze analyse bleek dat voor 2.348 titels de AWSO-code – in (hoofd)categorie – afwijkt tussen CB en KB. Op het totaal van de in de analyse meegenomen titels (92.440) is dit 2,5%.

In overleg met adviesbureau Berenschot is besloten dat deze afwijking van 2,5% verwaarloosbaar klein is om de resultaten te beïnvloeden. Voor de analyse op AWSO-code kan gewerkt worden met zowel het AWSO-veld uit KB als uit CB. We hebben de 4.106 records van KB ook gebruikt voor toekennen AWSO en een nieuwe kolom AWSO v2 gemaakt en gevuld. Deze kolom is gebruikt voor de analyses in de Monitor Makers 2017.