



Foto: MGDboston.

Open research data

Onderzoeksinstituten zien meer en meer het belang in van open science. Open access promoten is al jarenlang een taak binnen de wetenschappelijke bibliotheek, de uitbreiding naar andere takken van open science is ondertussen ook volop aan de gang. Het beheer van data wordt daarbij steeds belangrijker. Er circuleert immers ontzettend veel data resulterend uit onderzoek.

Meestal blijft de toegang tot die data echter beperkt. Soms is dat uit noodzaak omdat de data nu eenmaal gevoelig zijn en persoonlijke gegevens bevatten. In andere gevallen is het echter om andere, minder relevante redenen, zoals “is te tijdrovend”, “wie begrijpt die data”, “het formaat is niet standaard”, etc.¹

Ter bevordering van de kwaliteit van de data, samenwerking met anderen, innovatie binnen en buiten onderzoeksinstituten en vele andere redenen is open data de weg die we moeten opgaan. Als we spreken over open data dan gebruiken we de open definition² om die te definiëren: “*Open means anyone can freely access, use, modify, and share for any purpose (subject, at most, to requirements that preserve provenance and openness).*”

Met andere woorden, data is open als iedereen er toegang toe heeft, ze kan gebruiken, aanpassen en delen, voor welk doel dan ook. De maximale beperking is refereren aan de bron en eisen ze te delen onder dezelfde openheid. Als je dit vertaalt naar creative commons licenties, dan is dat data met de licenties CC0, CC-BY en CC-BY-SA.³

Data openstellen is echter niet genoeg om ze herbruikbaar te maken. Gebruik van open standaarden die machineleesbaar zijn is even belangrijk, alsook een beschrijving met duidelijke metadata en het toekennen van een unieke ID.⁴ In dit artikel het verhaal van drie instellingen die een opendatabeleid implementeren en onderzoeksdata aanbieden aan geïnteresseerden. De koe bij de horens vatten en ervoor gaan, dat is hun boodschap.

Inge Van Nieuwerburgh

VDAB

Michael De Blauwe, Pieter Ronsse en Willem Vansina

De toekomstige verplichting om open data ter beschikking te stellen, was niet de eerste reden voor de VDAB om met open data te starten. Open data past binnen de bedrijfsfilosofie van de VDAB die focust op innovatie, wendbaarheid, transparantie en maximale zelfredzaamheid. Open data maakt het mogelijk dat applicaties voor nichemarkten of doelgroepen ontwikkeld worden door andere partijen. De VDAB kan en wil niet alles zelf doen op de arbeidsmarkt.

ARVASTAT

Met Arvastat heeft de VDAB een decennium van ervaring inzake openheid over onze data. Arvastat is een vrij toegankelijke toepassing waarmee allerhande cijfers over de werkzoekenden en de vacatures in Vlaanderen opgezocht kunnen worden. Wil je weten hoeveel werkzoekenden er zijn in je eigen gemeente of

“MET ARVASTAT HEEFT DE VDAB EEN DECENNIUM VAN ERVARING INZAKE OPENHEID OVER ONZE DATA.”

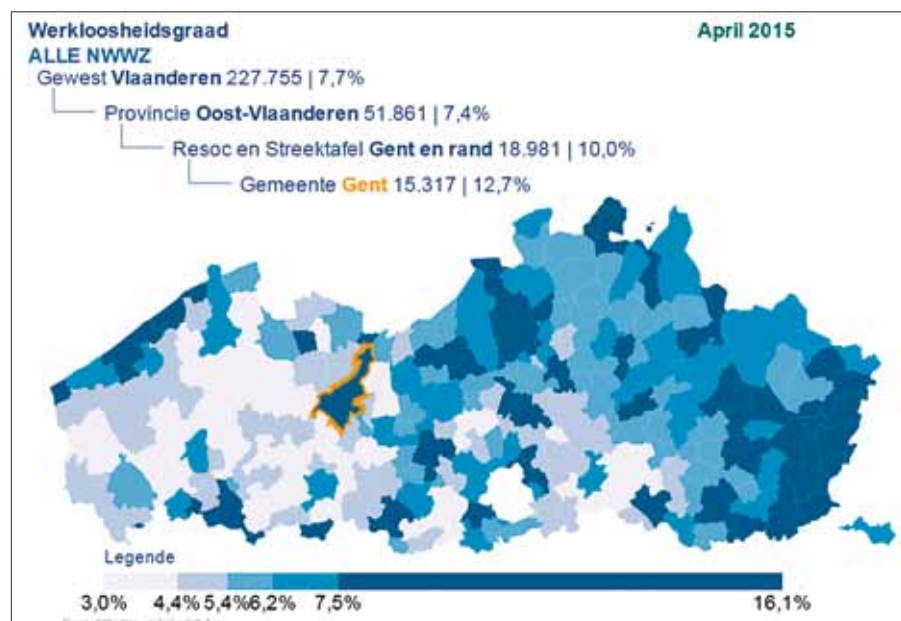
provincie? Wil je weten hoeveel van hen laaggeschoold zijn? Wil je weten hoeveel vacatures uit de bouwsector, waarvoor minimaal een hoger diploma en ervaring vereist is, ontvangen werden tussen februari 2001 en september 2010? Je kan het via enkele eenvoudige selecties allemaal zelf via Arvastat opzoeken.

Arvastat biedt dus een binnen Europa unieke openheid met betrekking tot data over werkzoekenden en vacatures, ook al is het technisch geen ‘open data’. Toen we in 2005 Arvastat lanceerden was dit een revolutionair instrument. Omdat de toepassing ook regelmatig gebruikt wordt, zijn de consequenties die deze toepassing voor de VDAB heeft gehad sterk vergelijkbaar met die van open data.

ARVASTAT: LESSONS LEARNED

MINDER VRAGEN? THINK AGAIN!

Arvastat werd onder andere gelanceerd om het aantal voorspelbare vragen naar cijfers te doen verminderen. De toepassing zorgt er immers voor dat we niet meer elke maand cijfers voor allerhande gemeenten en provincies moeten maken en opsturen. Of het aantal vragen met de lancering van Arvastat effectief gedaald is, durven we echter niet met zekerheid stellen. Arvastat heeft weliswaar het aantal eenvoudige vragen naar aantallen werkzoekenden en vacatures sterk doen verminderen. Meer openheid inzake data genereert echter ook vragen. Enerzijds krijgen we meer vragen naar data die net niet in Arvastat beschikbaar zijn. Anderzijds verschuift de aard van de vragen. Waar men ons vroeger vroeg naar de cijfers, vraagt men ons nu naar een verklaring voor cijfers en evoluties die men zelf heeft opgezocht. Aangezien we via o.a. Arvastat tonen dat we heel wat over de arbeidsmarkt weten, krijgen we ook meer en meer vragen die helemaal niets met de VDAB-data te maken hebben. Dit was de doelstelling. We willen immers een



Kerncijfer werkzoekenden gemeente Gent.



kenniscentrum voor de gehele arbeidsmarkt zijn. Dit is echter ook een uitdaging. Nu hebben we mensen nodig die naast de eigen cijfers de hele arbeidsmarktcontext door en door kennen.

DATAKWALITEIT EN HET MANAGEN VAN VERWACHTINGEN

Gebruikers kennen de periodiciteit van de updates en reageren snel als een update later dan normaal is, als het systeem niet bereikbaar is of als ze een bug of fout ontdekt hebben. Zeker als de doelstelling is om op de eerste werkdag van de maand te rapporteren over de cijfers van de voorgaande maand, leg je jezelf heel wat discipline op. Er is echter nog een uitdaging aan deze openheid. Omdat alle cijfers open zijn en één samenhangend geheel vormen, heeft de kleinste fout grote consequenties. Stelt één gemeente bij het maandelijks raadplegen van de cijfers een zware fout vast, moeten we alle cijfers — ook die op Vlaams niveau — opnieuw berekenen. Aangezien de Vlaamse minister van Werk op de eerste werkdag van de maand de officiële Vlaamse werkzoekendencijfers al heeft gecommuniceerd, moeten we dergelijke correcties kost wat kost vermijden. Arvastat dwong ons om strakke werkprocessen op te zetten en een dagdagelijkse interne rapportering voor de werkzoekendencijfers op te zetten zodat we onmiddellijk kunnen optreden bij fouten en de vereiste persteksten kunnen schrijven voor het einde van de maand.

WAT IS 'CREATE YOUR OWN JOBADVISOR'?

Via Arvastat hebben we ervaring met openheid over onze data. Hoewel we nog maar pas structureel open data ter beschikking stellen, heeft de VDAB ook al ervaring met echte open data. Sinds schooljaar 2012-2013 organiseert de IT-afdeling van de VDAB 'Create your own Jobadvisor', een wedstrijd voor Vlaamse IT-studenten waarbij we hen uitdagen om zelf een prototype te ontwikkelen voor tools die mensen helpen bij het vinden

van werk. In functie van deze oefening krijgen de studenten toegang tot heel wat VDAB-data, waaronder de vacaturedata-bank van de VDAB. Naast het positioneren van de VDAB als innovatieve, open en dynamische organisatie met oog voor de jeugd, is dit een manier om jongeren te consulteren over welke instrumenten zij van ons verwachten.

CREATE YOUR OWN JOBADVISOR: LESSONS LEARNED KEEP IT SIMPLE

Wanneer we als overheid digitale dienstverlening organiseren, werken we klassiek door eerst zo veel mogelijk informatie over de burger te verzamelen. Dit vanuit het idee dat, als we een zo compleet en exact mogelijke dossier hebben, we ook de beste dienstverlening kunnen aanbieden. De studenten uit de studentenwedstrijd draaien deze logica volledig om: zorg dat de gebruiker zo snel mogelijk kan starten met de toepassing, zelfs als de dienstverlening dan nog slecht is, en moedig hem of haar bij het gebruik van de toepassing aan om meer data te geven zodat ook de dienstverlening stapsgewijs beter wordt. Als het niet absoluut noodzakelijk is voor de dienstverlening, vraag het dan niet onmiddellijk. Haal de data ook waar ze al beschikbaar is. Als het al op mijn LinkedIn-pagina staat of opgevraagd is door een andere overheidsinstelling, haal het dan daar.

JUST SHOW ME

Jongeren willen daarnaast geen droge pagina's vol tekst met beschrijvingen van vacatures of werkgevers. Ze willen hun informatie visueel. De vacatures en de eigen woonplaats moeten op kaart uitgezet worden zodat de woon-werkafstand onmiddellijk zichtbaar is. Wil je vanuit hun toepassing een beeld krijgen van de werkgever, dan krijg je exact dat (via een link met bijvoorbeeld Google Street View). Als er bedrijfsfilmpjes op het net staan, gebruik die dan. Cijfers dienen ter beschikking gesteld via kaarten of grafieken zodat ze spreken. Uiteraard moet de

toepassing in zijn geheel er ook aantrekkelijk uitzien.

MOBILE, MOBILE, MOBILE

Toepassingen dienen gemaakt te worden op maat van de mobiele gebruiker. Toepassingen dienen voor een touchscreen ontwikkeld te worden en er mag niet al te veel informatie op één pagina staan zodat het ook via de smartphone nog hanteerbaar blijft.

De boodschap van deze IT-studenten is met andere woorden zeer duidelijk: VDAB, werk als de sociale media, niet als een overheid.

CONCLUSIE: EEN UITDAGENDE KANS

Zowel Arvastat als 'Create your own jobadvisor' dagen onze organisatie uit. De eerste met betrekking tot onze data, cijfers en informatiehuishouding, de tweede in de manier waarop wij onze digitale dienstverlening vorm geven.

Gemeenschappelijk aan zowel Arvastat als de studentenwedstrijd is dat we zelf actief en bewust met onze (open) data aan de slag gaan of er bewust voor zorgen dat anderen er mee aan de slag kunnen gaan. Dit is belangrijk. Pas als de data ook effectief worden gebruikt, worden we ten volle uitgedaagd en kunnen we het maximale uit onze open data halen.

Zijn we van deze acties efficiënter geworden? Dat durven we niet te beweren. Wel is het zeker dat we er beter van kunnen worden. Arvastat heeft ons gedwongen om beter te worden inzake informatie-management en kennis van de arbeidsmarkt. De Digistrijd toont ons hoe onze digitale dienstverlening beter kan worden.

De open data van VDAB is terug te vinden via opendata.vdab.be. Via www.vdab.be/digitalewedstrijd en kan je enkele van de producten van voorgaande wedstrijden bekijken. ■■

Het Instituut voor Natuur en Bosonderzoek

Bart Goossens, Dimitri Brosens, Peter Desmet, Daniel Du Seuil

Het Instituut voor Natuur- en Bosonderzoek (INBO) is het Vlaams onderzoeks- en kenniscentrum voor natuur en het duurzame beheer en gebruik ervan. Het INBO verricht onderzoek en levert kennis aan al wie het beleid voorbereidt, uitvoert of erin geïnteresseerd is. Als toonaangevende wetenschappelijke instelling werkt het INBO in de eerste plaats voor de Vlaamse overheid, maar het levert ook informatie voor internationale rapporteringen en gaat in op vragen van lokale besturen. Daarnaast ondersteunt het INBO onder meer organisaties voor natuurbeheer, bosbouw, landbouw, jacht en visserij. Het INBO maakt deel uit van nationale en Europese onderzoeksnetwerken.

WAAROM OPEN DATA?

In haar missie geeft het INBO aan dat het inzicht wil geven in natuur en bos via kwaliteitsvol onderzoek om zo het beleid voor een duurzame samenleving te onderbouwen. Deze missie vertaalt zich in de strategische doelstellingen ter ondersteuning van regeringsinitiatieven, partners en de implementatie van de Europese biodiversiteitsstrategie.

In de operationele werking van het instituut worden deze strategische doelstellingen nagestreefd door wetenschappelijk onderzoek en monitoring uit te voeren die hun output kennen in wetenschappelijke artikels, rapporten en adviezen die beleid en maatschappij ondersteunen in beheers- en beleidskeuzes. Maar, voor zowel interne als externe spelers is niet enkel het uiteindelijke advies, rapport of wetenschappelijke artikel van belang, maar ook de data waarop deze zijn gebaseerd. Het openstellen van de data draagt niet enkel bij tot de transparantie van de werking van het INBO, maar heeft ook

HOEWEL OPEN DATA DOOR IEDEREEN VRIJ GEBRUIKT KUNNEN WORDEN, IS HET NUTTIG OM ALS INSTITUUT ONZE WENSEN BETREFFENDE HET DATAGEBRUIK TE COMMUNICEREN, VOORAL VOOR GEBRUIK DOOR DE WETENSCHAPPELIJKE GEMEENSCHAP.

een belangrijke maatschappelijke, wetenschappelijke en soms commerciële meerwaarde voor derden. De roep om deze data openbaar, toegankelijk, gedocumenteerd en actueel ter beschikking te stellen wordt steeds groter voor het INBO.

WAT DOET HET INBO VANDAAG AL?

- Het INBO is al enkele jaren geregistreerd als GBIF⁵ data publisher en heeft met de ondertekening van de Bouchout declaratie⁶ nog eens bekrachtigd zich in te zetten voor open biodiversiteitsdata.

- Het INBO publiceert momenteel 17 gestandaardiseerde en gedocumenteerde datasets⁷ via GBIF, goed voor meer dan 5,2 miljoen waarnemingen. Het publiceren van een dergelijke dataset kan gecombineerd worden met het schrijven van een datapaper⁸. Dit is een peer reviewed publicatie die een dataset beschrijft en correct hergebruik en datacitatie stimuleert. Uitgevers⁹ van datapapers vereisen dat de data toegankelijk zijn via GBIF of in andere open repository. Het INBO heeft reeds drie datapapers gepubliceerd en er zijn meerdere in uitvoering.



Visualisatie op basis van 'bird tracking data' gepubliceerd via GBIF.

- Sinds november 2014 biedt het INBO verscheidene geospatiale datasets aan via Geopunt (AGIV) onder een Open Data licentie.
- Alle data verzameld binnen de LifeWatch¹⁰ projecten worden gepubliceerd als Open Data (bijv. de meeuwen-tracking data¹¹).
- Het INBO opteert voor Dryad¹² als platform voor het ontsluiten¹³ van ondersteunende data voor A1-publicaties.

RICHTLIJNEN

Een Open Data instituut heeft een duidelijk opendatabeleid nodig. Het beleid van het INBO werd daartoe opgebouwd rond deze 10 richtlijnen:

- 1 Alle wetenschappelijke data waar het INBO (mede)eigenaar van is, vallen onder dit beleid en kunnen als Open Data ontsloten worden, met toepassing van de wettelijke uitzonderingsgronden.**
Wetenschappelijke data waar het INBO (mede)eigenaar van is, vallen onder dit beleid. Externe data waar het INBO toegang toe heeft, maar geen (mede)eigenaar van is, vallen enkel onder dit beleid indien ontsluiting als Open Data contractueel is vastgelegd of als de andere partij hiermee schriftelijk akkoord gaat. Voor nieuwe opdrachten wordt contractueel vastgelegd dat het INBO (mede) eigenaar wordt van de bekomen data en zij het recht heeft deze te ontsluiten volgens de eigen richtlijnen.

Het INBO hanteert voor het ontsluiten van data de wettelijke uitzonderingsgronden voor openbaarmaking van milieu-informatie (decreet van 26 maart 2004, artikel 15), zoals de bescherming van milieu, persoonsgegevens en vertrouwelijke gegevens. Deze uitzonderingen worden toegepast op specifieke attributen van een dataset (e.g. geografische coördinaten) en bij voorkeur met anonimiseren en verminderde resolutie.
- 2 Data komen ten vroegste 12 maanden na het verzamelen in aanmerking voor ontsluiting.**
Dit embargo geldt op recordniveau: voor langlopende projecten kan een subset van een groeiende dataset worden. Het in aanmerking komen voor ontsluiting betekent niet automatisch dat de data zullen ontsloten worden: dit hangt af van prioritering. Onderzoekers of projecten (zoals LifeWatch) kunnen actief beslissen om de data al vroeger te ontsluiten.
- 3 Ruwe data worden ontsloten.**
Ruwe data zijn verzamelde data in bruikbare vorm (bijv. soortwaarnemingen en metingen). Ze hebben het potentieel voor verscheidene onderzoeksvragen of doeleinden gebruikt te kunnen worden. Ruwe data komen in aanmerking om ontsloten te worden indien er 1) vraag naar is en/of 2) een infrastructuur voor is, zoals GBIF voor waarnemings- en taxonomische data, GenBank voor genetische data, en GeoPunt¹⁴ voor belangrijke GIS-lagen. De ontsluiting van de ruwe data wordt jaarlijks geprioriteerd en ingepland.
- 4 Resultaatdata die geassocieerd zijn met een wetenschappelijk artikel, rapport of advies worden ontsloten bij publicatie.**
Geassocieerde resultaatdata zijn data die gebruikt worden voor en/of besproken worden in een publicatie. In tegenstelling tot ruwe data hebben ze vaak een aantal transformaties (bijv. aggregaties, analyses) ondergaan. Bij de publicatie van een wetenschappelijk artikel, rapport of advies worden deze data mee ontsloten, zelfs al zou het embargo (zie punt 2) nog van toepassing zijn. De data worden in principe ontsloten via een data repository en er wordt vanuit de publicatie gerefereerd naar de publieke data. In bepaalde gevallen kan ook naar reeds gepubliceerde ruwe data verwezen worden.
- 5 Data worden ontsloten als Open Data, zoals gedefinieerd in de Open Definition, onder Creative Commons Zero.**
Alle data worden ontsloten volgens de Open Definition, die onder andere stelt deze voor iedereen vrij toegankelijk, te verspreiden en (her)bruikbaar zijn. De data worden ontsloten onder Creative Commons Zero, waarmee duidelijk wordt gemaakt dat geen auteursrechten golden en gelden op de data en deze in het publiek domein geplaatst worden. Deze waiver is juridisch de meest geschikte voor het type data dat het INBO hanteert.
- 6 Er wordt in gepubliceerde datasets steeds verwezen naar de INBO-normen voor datagebruik.**
Daar geen auteursrechten bestaan op feitelijke data, kunnen geen gebruiksvereisten afgedwongen worden op dergelijke publieke data. Wel maakt het INBO haar wensen betreft citatie en datagebruik — vooral dan in een wetenschappelijke context — kenbaar via normen (zie hieronder). Voor alle gepubliceerde datasets wordt naar deze normen verwezen.
- 7 Ontsloten data worden steeds voldoende gedocumenteerd met metadata.**
Metadata beschrijven een dataset en vermelden onder andere contactpersonen, methodologie, en taxonomische, temporale en geografische scope. Ze helpen de gebruikers van de data deze beter te begrijpen. Alle ontsloten data worden steeds gedocumenteerd met metadata. Het formaat en de uitgebreidheid van de metadata hangen af van de vereisten van het platform (GBIF, GenBank, Dryad, etc.) waarop deze wordt gepubliceerd.
- 8 Voor alle onderzoeksprojecten wordt een data management plan (DMP) opgesteld.**
De onderzoeker dient steeds vaker op verzoek van de subsidieverstrekker een data management plan (DMP)

op te stellen waarin aangegeven wordt op welke manier de onderzoeksdata beheerd, gedocumenteerd en gedeeld gaan worden. Het data management plan dient opgesteld te worden in de beginfase van een onderzoeksproject door de projectleider.

9 Onderzoekers passen het opendatabeleid toe.

Het opendatabeleid wordt opgenomen als doelstelling voor onderzoekers in hun jaarlijkse planning. Via de jaarlijkse evaluaties wordt toegezien op het naleven van de vastgelegde afspraken.

10 Het IDC ondersteunt de onderzoekers in de toepassing van het opendatabeleid.

Het INBO Informatie & Datacenter (IDC) implementeert het opendatabeleid en 1) zorgt voor de nodige ondersteuning en tools om Open Data te publiceren, 2) faciliteert het schrijven van datapapers, en 3) helpt in het standaardiseren van data en metadata, gebruik makend van internationaal aanvaarde standaarden.

NORMEN VOOR DATAGEBRUIK

Hoewel Open Data door iedereen vrij gebruikt kunnen worden, is het nuttig om als instituut onze wensen betreffende het datagebruik te communiceren, vooral voor gebruik door de wetenschappelijke gemeenschap. Hiertoe publiceert het INBO, net als enkele andere organisaties (zoals Canadensys¹⁵ en VertNet¹⁶), normen voor datagebruik, waarnaar telkens verwezen wordt in iedere gepubliceerde dataset. Deze normen zijn geen juridisch document, maar worden als ethische normen beschouwd:

- **Bronvermelding:** hoewel bronvermelding niet wettelijk verplicht is, is het wel de norm binnen de wetenschappelijke gemeenschap. Voor wetenschappelijke artikels adviseren we de richtlijnen van het tijdschrift te volgen voor het citeren van datasets, voor ander gebruik raden we aan een link naar de oorspronkelijke dataset toe te voegen.

- **Wees verantwoordelijk:** (mis)bruik de data niet op een onwettelijke, schadelijke of misleidende manier. Informeer jezelf over de omvang en het oorspronkelijk doel van de data aan de hand van de metadata, en besef dat deze onderhevig kunnen zijn aan verandering en fouten.
- **Respecteer openheid:** verzwijg de open status van de data niet.
- **Contacteer ons:** we zijn steeds geïnteresseerd om te zien hoe de data gebruikt of gevisualiseerd wordt, of om meer informatie te verkrijgen over de data. Het kan de start zijn van een boeiende samenwerking.

BESLUIT

Aan de ontwikkeling en de goedkeuring van het INBO opendatabeleid is heel wat werk vooraf gegaan. De volgende stappen zijn het toepassen van dit beleid in de praktijk en het overtuigen van de onderzoekers van het belang van Open Data in relatie tot hun onderzoek.

Om tot resultaat te komen staat in 2015 naast de uitwerking en implementatie van een aantal processen (bijv. opstellen van een data management plan, selecteren van ruwe data, etc) en het voeren van een bewustmakingscampagne ook nog de uitrol van het dataportaal gepland zodat alle informatie tegen eind dit jaar via één portaal ter beschikking zal worden gesteld.

Het Onderzoeksinstituut voor Natuur en Bos hoopt met dit opendatabeleid het wetenschappelijk onderzoek in Vlaanderen inzake natuur en biodiversiteit transparanter en toegankelijker te maken. Dit kan alleen maar voordelen opleveren voor iedereen die met natuurbeleid te maken heeft. ■■

VLIZ DATA - OPENDATABELEID

Het VLIZ is voorstander van een vrije uitwisseling van data. Data dienen zo veel mogelijk beschikbaar gesteld voor wetenschappelijk onderzoek op zowel nationaal als internationaal niveau. Het VLIZ ziet het als zijn kerntaak om die dataontsluiting te faciliteren. Data waar het VLIZ de primaire bron van is worden steeds publiek ontsloten. Metadata die het bestaan van een dataset illustreren worden steeds publiek ontsloten, tenzij expliciet aangevraagd werd om dit niet te doen.

In vele gevallen zijn er argumenten waardoor geen of slechts een beperkte toegang verleend kan worden. Het VLIZ respecteert deze beperking en zal in samenspraak met de data-eigenaar een op maat uitgewerkt databeleid opzetten. Het VLIZ zal de ontsluiting van data volgens dit beleid verder opvolgen, ook na het verloop van de context waarin de datapolicy werd opgesteld. Bijvoorbeeld, in het kader van specifieke projecten wordt vaak door de partners een moratoriumperiode gevraagd tot de gegevens of resultaten van het project gepubliceerd zijn.

Gebruik van gegevens moet steeds erkend worden of het nu om publieke of aangevraagde data gaat. Op elk moment van download of transfer vermeldt het VLIZ de erkenning van de databron. Indien een substantieel deel van de gegevens gebruikt wordt spoort het VLIZ de gebruiker aan om contact op te nemen met de oorspronkelijke bron van de gegevens. Het VLIZ organiseert dit door de inbouw van disclaimers in data-systemen en door opname van citaties in het metarecord van de dataset. (actuele versie te vinden op <http://www.vliz.be/nl/databeleid>)

(uitgebreid beschreven in Haspeslagh & Vanden Berghe, *Bibliotheek- & archief-gids* 79/2). Het oorspronkelijke concept van IMIS werd in de afgelopen vijftien jaar verder ontwikkeld en verfijnd om vooral de Belgische mariene wetenschapswereld optimaal te kunnen documenteren. Hierbij werd sterk de focus gelegd op het gedetailleerd behandelen en ontsluiten van de wetenschappelijke output: publicaties, opgebouwde expertise, projectwerk en — last but not least — datasets. Ondertussen werd een eerste open informatieproduct gepubliceerd, het Compendium voor Kust en Zee (zie kaderstuk p. 16), waarin onder andere de geïntegreerde open data en informatie uit IMIS gebruikt wordt.

NIUWE DATAPRODUCTEN

Naast het beheren van MDA, is het VLIZ-datacentrum ook actief betrokken in het creëren van nieuwe en afgeleide producten die gebaseerd zijn op open data van diverse oorsprong. Het MDA-archief zelf wordt hiervoor gebruikt, maar er wordt ook nauw samengewerkt met de data

creators. VLIZ voorziet in platformen die wetenschappers toelaten datasets aan te leveren en te documenteren in systemen die deze data onmiddellijk integreren en verspreiden in open, online data- en informatieproducten. Daarnaast wordt het ook mogelijk om deze gegevens te publiceren en zo datacitaties te genereren. We illustreren dit proces met voorbeelden uit MDA en uit het European Marine Observation and Data Network (EMODNET).

VAN ARTIKEL NAAR DATAPUBLICATIE EN -CITATIE

Een onderzoeksgroep publiceert een peer-reviewed artikel in een openaccess-tijdschrift (BRAECKMAN et al., *PLoS ONE* 9(10): e108153). De onderliggende dataset wordt opgeslagen in het MDA, waarbij de onderzoekers uitgebreide metadata toevoegen. De auteurs stellen deze data vrij ter beschikking van andere onderzoekers via een Creative Commons CC-BY-NC licentie. Gebruikers kunnen deze set dus via een simpele downloadknop uit MDA

The screenshot shows the IMIS interface. The top part displays a scientific article titled "Empirical evidence reveals seasonally dependent reduction in nitrification in coastal sediments subjected to near future ocean acidification" by Braeckman, U.; Van Colen, C.; Gilleki, K.; Van Gasbeke, D.; Soetaert, K.; Vincu, M.; Vanaverbeke, J. (2014). Below the article, there is a section for "Bijbehorende data" (Associated data) with a link to a DOI: <https://doi.org/10.14284/1>. A blue arrow points from this link down to the dataset page. The dataset page shows the title "Ocean acidification effects on nitrification in natural sediment communities from Belgian part of the North Sea" by Braeckman, U.; Marine Biology Research Group - Uperil, Belgium (2014). At the bottom right of the dataset page, there is a blue arrow pointing to the "MDA" logo.

Datapublicatie: de dataset die gebruikt werd voor de bovenliggende publicatie, krijgt een eigen DOI-link en referentie. Deze data zijn onder een CC-BY-NC licentie vrij beschikbaar vanuit het MDA via de downloadknop.



Het Compendium voor Kust en Zee integreert alle informatie over de Vlaams/Belgische mariene onderzoekswereld. ©VLIZ

COMPENDIUM VOOR KUST EN ZEE

Er bestaat een groeiende vraag in de mariene en maritieme wereld naar wetenschappelijk onderbouwde en geactualiseerde informatie als antwoord op maatschappelijke vragen en probleemstellingen. Deze vragen vloeien veelal voort uit de toenemende nood aan een geïntegreerde benadering waarbij zowel milieu- en natuuronderzoek in het mariene systeem als socio-culturele, economische en institutionele aspecten aan bod komen. De informatie is deels voorhanden maar dikwijls sterk gefragmenteerd, sectorgebonden of weinig toegankelijk. Het 'Compendium voor Kust en Zee' wil deze disperse informatie en gegevens uit het Vlaams en Belgisch marien en maritiem onderzoek bundelen. Al het thematisch onderzoek behorend tot diverse wetenschapsdisciplines uitgevoerd in mariene en kustgebonden ecosystemen en estuaria, inclusief het beleidsondersteunend onderzoek en het onderzoek gericht op gebruikstoepassingen en sectoren, komt hierbij aan bod. Het geïntegreerde en grensoverschrijdende karakter van het Compendium, inclusief de land-zee grens, draagt bij tot een verhoogde communicatie binnen het netwerk van zeewetenschappers en deskundigen die professioneel betrokken zijn met kust en zee, en verhoogt ook de zichtbaarheid en toegankelijkheid van het marien onderzoek.

halen en zelf verder gebruiken in hun onderzoek.

De dataset heeft een DOI meegekregen, en wordt dus een citeerbare datapublicatie. Deze citatie kan dan opgenomen worden in bijvoorbeeld de Data Citation Index (een recente aanvulling van de Web of Knowledge databankgroep) en draagt zo bij tot de wetenschappelijke reputatie van de onderzoeksgroep.

HERVERDELING EN NIEUWE DATAPRODUCTEN

EMODNET (<http://www.emodnet.eu/>) is een Europees project waarbij de partners Europese mariene data, dataproducten en metadata van verschillende oorsprong bijeenbrengen op een uniforme manier. Het hoofddoel is om de

gefragmenteerde en verborgen data te ontsluiten, geïntegreerd aan te bieden en te herverdelen. Op die manier faciliteert de Europese Commissie meer duurzaam onderzoek via kwalitatieve, gestandaardiseerde en vrije toegang tot mariene data. VLIZ voorziet een aantal technische platformen die aan de onderzoekers toelaten om hun data, vanuit MDA of van andere bronnen, te koppelen aan het centrale EMODNET-portaal. Dit portaal integreert de data met een aantal visualisatietools en brengt vooral gegevens van verschillende oorsprong samen in één open omgeving van waaruit onderzoekers nieuwe combinaties met, en analyses op mariene data kunnen maken. Dit leidt tot nieuwe wetenschappelijke datasets, producten, publicaties en inzichten.

Er moet nog veel water naar de zee vloeien voor alle mariene data in open

DateAcquired	CatalogNumber	ScientificName	Year	Month	Day	Lat	Lon	Pressure (m)	WaterTemp (m)	MaxDepth (m)	Sea	InCont	SampleSize	MethodCode	Release	QC
2005-12-19	410276	Leptocottus armatus				57.83	40.19	3	4					HarDFP	679	1
2005-12-19	410280	Leptocottus armatus				57.83	40.19	3	4					HarDFP	679	1
2005-12-19	410281	Leptocottus armatus				57.83	40.19	3	4					HarDFP	679	1
2005-12-19	410282	Callinectes sapidus				57.83	40.19	3	4					HarDFP	679	1
2005-12-19	410283	Leptocottus armatus				57.83	40.19	3	4					HarDFP	679	1
2005-12-19	410284	Other species				57.83	40.19	3	4					HarDFP	679	1
2005-12-19	410285	Other species				57.83	40.19	3	4					HarDFP	679	1
2005-12-19	410286	Acipenserina (fish sp.)				57.83	40.19	3	4					HarDFP	679	1
2005-12-19	410287	Acipenserina (fish sp.)				57.83	40.19	3	4					HarDFP	679	1

Het EMODNET-portaal integreert en visualiseert mariene datasets. Deze interactieve kaart toont waar Zeekevers (een zeeslakkensoort) gevonden werden. De code '647' verwijst rechtstreeks naar de volledige dataset in het MDA.

repositories beschikbaar zullen worden. Overheden — nog steeds de grootste geldschietters voor wetenschappelijk onderzoek — hebben er alle belang bij om op internationale schaal het wettelijk kader hiervoor aan te reiken. Ondertussen wordt de mariene onderzoekswereld gemotiveerd om nu reeds hun gegevens vrij aan te bieden door hen zo veel mogelijk meerwaarde te geven: veilige en permanente opslag, centraal institutioneel beheer, blijvende controle over datagebruik, datapublicatie en -citatie, integratie, en het creëren van nieuwe dataproducten die voor iedereen beschikbaar zijn. ■■

Referenties

- BRAECKMAN, U. et al. (2014). Empirical evidence reveals seasonally dependent reduction in nitrification in coastal sediments subjected to near future ocean acidification. *PLoS One* 9(10): e108153. dx.doi.org/10.1371/journal.pone.0108153 - Bijhorende data: dx.doi.org/10.14284/1
 - HASPELAGH, J.; VANDEN BERGHE, E. (2003). IMIS, de eerste Vlaamse mariene informatiedatabank. *Bibliotheek- & archiefgids* 79(2): 12-16
 - LESCRAUWAET, A.-K. et al. (Ed.) (2013). Compendium voor Kust en Zee 2013: Een geïntegreerd kennisdocument over de socio-economische, ecologische en institutionele aspecten van de kust en zee in Vlaanderen en België. Vlaams Instituut voor de Zee (VLIZ): Oostende. ISBN 978-90-820731-5-7. 342 pp. www.compendiumkustenzee.be
- 1 Bekijk ook de "50 shades of no": <http://opendataforum.info/index.php/news2/open-data-nieuws>
 - 2 [Opendefinition.org](http://opendefinition.org)
 - 3 Zie creativecommons.be
 - 4 Zie ook <http://5stardata.info/> voor meer
 - 5 www.gbif.org
 - 6 www.bouchoutdeclaration.org/
 - 7 www.gbif.org/publisher/1cd669d0-80ea-11dea9d0-f1765f95f18b/datasets
 - 8 vb datapaper: <http://phytokeys.pensoft.net/articles.php?id=1417>
 - 9 www.pensoft.net/
 - 10 lifewatch.inbo.be
 - 11 lifewatch.inbo.be/blog/posts/bird-tracking-data-published.html
 - 12 <http://datadryad.org/>
 - 13 We gebruiken in dit artikel de termen "publicatie" en "ontsluiting". Beide betekenen het "online beschikbaar stellen van data (als Open Data)."
 - 14 www.geopunt.be
 - 15 www.canadensys.net/norms
 - 16 www.vertnet.org/resources/norms.html