



**Monitor Nederlandse
Onderzoek Repositories
2011**

Colofon

*Monitor Nederlandse Onderzoek Repositories 2011
Stand van zaken bij de aan NARCIS en aan de HBO Kennisbank toeleverende repositories*

SURFfoundation
PO Box 2290
NL-3500 GG Utrecht
T + 31 30 234 66 00
F + 31 30 233 29 60

info@surf.nl
www.surf.nl

Auteurs

Leo Waaijers – *Open Access Consultant*
Maurits van der Graaf – *Pleiade Management en Consultancy*

Eindredactie

Annemiek van der Kuil - *SURFfoundation*

SURF is de ICT-samenwerkingsorganisatie van het hoger onderwijs en onderzoek (www.surf.nl).
Deze publicatie is digitaal beschikbaar via de website van SURFfoundation:
www.surffoundation.nl/publicaties

© Stichting SURF
februari 2012

Deze publicatie verschijnt onder de Creative Commons licentie Naamsvermelding 3.0 Nederland.



Inhoudsopgave

1	Inleiding	5
1.1	Het netwerk van repositories in Nederland: infrastructuur voor Open Access tot kennis	5
1.2	Leeswijzer	5
2	Strategische synopsis	6
2.1	Strategische synopsis m.b.t. aan NARCIS toeleverende repositories	6
2.2	Strategische synopsis m.b.t. aan HBO Kennisbank toeleverende repositories	8
3	Methoden	11
4	Aan NARCIS toeleverende repositories	13
4.1	Onderzoeksvragen	13
4.2	Open Access beschikbaarstelling van Nederlandse wetenschappelijke output	13
4.3	Beheer en organisatie	14
4.4	Technische ontwikkelingen	16
4.5	Dienstverlening op basis van repositories	17
4.6	Wetenschappers als deponerende auteurs en als informatiezoekers	18
4.7	Naar een toekomstscenario voor de institutionele repositories	20
4.7.1	Doelstellingen voor de institutionele repositories	20
4.7.2	Realisatie van succesfactoren	20
4.7.3	Criteria voor het bereiken van de hoofddoelstellingen	21
4.7.4	Bouwstenen voor een toekomstscenario	22
4.7.5	Voorstel voor een toekomstscenario	23
5	Aan HBO Kennisbank toeleverende repositories	27
5.1	Onderzoeksvragen	27
5.2	Open Access beschikbaarstelling en content	27
5.3	Beheer en organisatie	29
5.4	Technische ontwikkelingen	29
5.5	Dienstverlening op basis van repositories	30
5.6	Naar een toekomstscenario voor de HBO repositories	30
5.6.1	Doelstellingen	30
5.6.2	Realisatie van succesfactoren	31
5.6.3	Criteria voor het bereiken van de hoofddoelstellingen	31
5.6.4	Bouwstenen voor een toekomstscenario	33
5.6.5	Voorstel voor een toekomstscenario	33
	Bijlage A - Aan NARCIS toeleverende repositories	37
	Bijlage B - Aan de HBO Kennisbank toeleverende repositories	39

1 Inleiding

1.1 Het netwerk van repositories in Nederland: infrastructuur voor Open Access tot kennis

Publieke toegang tot onderzoeksresultaten van de universiteiten was het doel van het hogeronderwijs in het door SURFfoundation opgezette DARE-programma (2003-2006). Daartoe werd een interoperabel netwerk van institutionele academische repositories opgezet met een gezamenlijke portal met zoek- en andere functionaliteit voor gebruikers van deze informatie. Uiteindelijk werden er twee zulke infrastructuren gecreëerd: een netwerk van repositories met universitair-wetenschappelijke informatie met als portal NARCIS (www.NARCIS.nl), en - enkele jaren later - een netwerk van HBO-repositories met als portal de HBO Kennisbank (www.HBO-Kennisbank.nl).

De opvolger van het DARE- programma - het SURFshare-programma (2007-2011) - bouwde voort op deze kennisinfrastructuur voor open toegang tot onderzoeksresultaten. Het programma richt zich onder andere op de verdere verbetering van deze infrastructuur en de mede op basis van deze infrastructuur mogelijk geworden uitbouw van de wetenschappelijke communicatie via verrijkte publicaties, samenwerkingsomgevingen voor onderzoekers (collaboratories) en toegang tot onderzoeksdata.

Deze studie fungeert als een effectmeting voor wat betreft de twee netwerken van repositories, brengt de huidige stand van zaken in kaart en verkent mogelijke toekomstscenario's. Deze studie kan deels gezien worden als een vervolg op de studie 'Nederlandse Academische Repositories, SURFshare Nulmeting' van maart 2010. Dit betrof een nulmeting onder de repositories van de Nederlandse universiteiten die uitgevoerd was uitgevoerd eind 2008 met aanvullend onderzoek in de eerste helft van 2009. Belangrijke getallen in het onderzoek betreffen het kalenderjaar 2007. Met andere woorden, ondanks het feit dat de resultaten publiek werden gemaakt in 2010, betrof het voor een deel al gegevens van enkele jaren oud, rond de start van het SURFshare-programma.

Dit onderzoek werd uitgevoerd door Leo Waaijers (Open Access consultant) en Maurits van der Graaf (Pleiade Management en Consultancy) in opdracht van SURFfoundation en vond plaats onder supervisie van Annemiek van der Kuil (SURFfoundation).

1.2 Leeswijzer

	Aan NARCIS toeleverende repositories	Aan HBO-Kennisbank toeleverende repositories
Strategische synopsis	Blz. 6 - 7	Blz. 8 - 9
Over het onderzoek	Blz. 13 - 14	
Toelichting	Blz. 15 - 27	Blz. 29 - 35
Onderliggende data: kwantitatieve analyse en enquête	Bijlage A	Bijlage B

2 Strategische synopsis

2.1 Strategische synopsis m.b.t. aan NARCIS toeleverende repositories

Analyse

De opbouw van een nationale infrastructuur voor open toegang tot de resultaten van wetenschappelijke onderzoek stagneert.

Kernobservaties over Open Access

- Het percentage publicaties dat in Open Access (OA) via de repositories jaarlijks ter beschikking komt (hierna dekkingsgraad genoemd), is niet significant gestegen sinds 2007 en schommelt in 2010 voor nagenoeg alle instellingen tussen 7% en 30%.
- Alleen voor de proefschriften is een ruime meerderheid van de universiteiten er in geslaagd om gedurende de afgelopen vier jaar (2007-2010) een dekkingsgraad van meer dan 80% te bereiken.
- De mate van OA succes (bij de proefschriften) of stagnatie (bij alle overige publicaties) is niet gecorreleerd met universitaire mandaten of personele capaciteit.
- De implementatie van zes van de negen ´standaard´ succesfactoren voor institutionele repositories bij meer dan de helft van de instellingen, heeft niet tot een hogere dekkingsgraad geleid. Op grond van de voorlopige ervaringen met de drie overige succesfactoren zijn ook hiervan geen doorbraken te verwachten.

Stagnerende cultuur

- Er is een grote verscheidenheid tussen de instellingen op bijna alle OA terreinen. Dit betreft het instellingsbeleid (hoewel wel alle universiteiten, alsmede de KNAW en NWO, de Berlin Declaration hebben ondertekend), de infrastructuur (software, persistent identifier, onderwerpsontsluiting, format metadata), het opnamebeleid, de werkprocessen en de personele inzet, de aanlevering aan NARCIS. Vergeleken met het eerdere onderzoek uit 2008 is deze verscheidenheid niet afgenomen. Klaarblijkelijk wordt de nationale infrastructuur onvoldoende als een gezamenlijke opgave beleefd.

Operationele effecten

- De DARE-afspraken uit 2007 voor de opzet van de nationale infrastructuur zijn gedateerd en worden deels niet meer nageleefd. Nieuwe afspraken zijn nog niet tot stand gekomen. Dit tast de robuustheid van de infrastructuur aan.
- De gezamenlijke NARCIS web site vertoont de nodige onvolkomenheden, mede doordat er onvoldoende systematische kwaliteitsbewaking is.
- Het streven naar een landelijk onderzoeksregistratiesysteem ter vervanging van het oude METIS systeem is gestaakt. Er gaat nu gestreefd worden naar een landelijk uniform datamodel voor onderzoeksregistratie. Vanuit de onderzoekregistratiesystemen worden de metadata voor de institutionele repositories en NARCIS aangeleverd.
- De duurzame opslag van de inhoud van de repositories in het e-Depot van de KB is weliswaar medio 2010 opnieuw gestart, maar de uitrol verloopt moeizaam door de eerder genoemde verscheidenheid tussen de repositories en door het ontbreken van collectieve werkafspraken.

Uitzondering

- Een uitzondering vormen de DAI's – Digital Author Identifiers. Hiervoor is een gezamenlijk systeem opgezet. Per eind 2011 is aan circa 80% van de wetenschappelijke onderzoekers een DAI toegekend. Dit fungeert als een belangrijk hulpmiddel bij het zoeken naar publicaties op naam van een auteur en als basis voor dienstverlening aan auteurs.

Aanbevelingen

De content van een repository is geen lokale collectie, maar een bijdrage aan een globaal kennisnetwerk. De eerste resultaten van het SURE2 project bevestigen dat vrijwel alle vraag naar de content van de repositories tot stand komt via Google en Google Scholar. En dat die vraag zeer internationaal is. Deze gedachte vormt de rode draad bij de hiernavolgende aanbevelingen.

Primair doel

Spreek gezamenlijk met de instellingen een primair doel af:

Primair doel: Vanaf eind 2015 is 70% van de jaarproductie van Nederlandse onderzoekspublicaties toegankelijk in Open Access ¹.

Gezamenlijkheid

- Maak het primaire doel de toetssteen voor alle gezamenlijke beslissingen en ontwikkelingen.

Meewerkende auteurs

- Het primaire doel is niet haalbaar zonder de medewerking van wetenschappelijke auteurs. Ontwikkel daartoe een serie apps die deponerende auteurs aantrekkelijke diensten bieden. Bied deze diensten aan op meerdere aggregatieniveaus (faculteiten, leerstoelgroepen, projecten, instellingen et cetera) en haak in op internationale successen, zoals Google Scholar Citations voor publicatielijsten, CiteULike/Mendeley/Zotero voor publicaties, Dataverse/DataCite voor datasets of Linked Open Data voor webdiensten.

Infrastructuur

- Maak NARCIS het logistieke hart voor toelevering aan Google (Scholar), het e-Depot en andere (disciplinaire) (inter)nationale dienstverleners. Beleg daartoe een duidelijke, niet-vrijblijvende regiefunctie bij een overkoepelende instelling.

Green én Gold

- Omdat alleen de *'Green road'* vrijwel zeker niet voldoende zal zijn om het primaire doel te halen, wordt aanbevolen ook het beleid ten aanzien van de *'Golden road'* te intensiveren. Het voorstel is om hiervoor een *task force* op te zetten, onder leiding van de voorzitter van de VSNU, die zich namens de universiteiten inzet voor de overgang van het Licentie model naar het OA model.

¹ Deze doelstelling dient duidelijk toetsbaar gemaakt te worden. Bijvoorbeeld: 70% van alle publicaties die zijn geregistreerd in het VSNU jaarverslag over 2014 (= wetenschappelijke publicaties inclusief proefschriften + vakpublicaties) zijn eind 2015 in Open Access toegankelijk via Google Scholar.

2.2 Strategische synopsis m.b.t. aan HBO Kennisbank toeleverende repositories

Analyse

Voor open toegang tot de resultaten van toegepast onderzoek is een sobere nationale infrastructuur opgebouwd; het gebruik er van is nog beperkt.

Kernobservaties over Open Access

- Alle grotere hogescholen, totaal 21 samen meer dan 90% van de studenten omvattend, nemen deel in de HBO Kennisbank. Er is een zekere diversiteit op het gebied van software en definities, maar dit lijkt geen storende factor, mogelijk omdat de Kennisbank nog klein is. Deponeren van publicaties in de repository maakt in het algemeen geen deel uit van de werkprocessen in de instelling.
- Er bestaat een verscheidenheid aan belemmeringen voor het open toegankelijk maken van afstudeer- of bachelorscripties: er is expliciete toestemming van de student als auteursrechthebbende voor nodig, stagebedrijven vrezen voor het uitlekken van bedrijfsgeheimen en zwakke scripties zijn schadelijk voor het imago van de instelling. In het algemeen lijken zowel de studenten als de instellingen niet trots op dit product. Slechts een fractie van de scripties komt beschikbaar via de Kennisbank.
- Lectorenpublicaties lopen uiteen van ingezonden brieven tot bijdragen aan vakbladen en wetenschappelijke artikelen. Ook wordt er veel en uiteenlopend audiovisueel materiaal geproduceerd. Van een breed gedragen kwaliteitscultuur is nog geen sprake. Instellingen zijn nog op zoek naar hun onderzoeksprofiel. Over de onderzoeksproductie in termen van aantallen en type publicaties is niets bekend, zodat ook over het dekkingspercentage geen uitspraken kunnen worden gedaan. Gerelateerd aan het aantal lectoren en (onderzoek-)docenten lijkt het laag.
- De HBO Kennisbank wordt per maand 40.000 keer bezocht. Over de aard van dit bezoek (eenmalig, regelmatig, personen of crawlers) en het gebruik van het materiaal in de repositories is weinig bekend: leidt het tot downloads, tot citaties, is er sprake van valorisatie? Langs welke weg wordt de content van de repositories benaderd: via de Kennisbank of bijvoorbeeld via Google (zoals bij NARCIS voor het overgrote deel het geval is)?
- Gemiddeld stellen de instellingen 0,4 fte beschikbaar voor hun repository.

Aanbevelingen

Praktijkgericht onderzoek als structurele opdracht is bij het HBO pas een decennium oud en nog volop in ontwikkeling, inclusief de bijbehorende publicatiecultuur. De kwaliteit en toegankelijkheid van de onderzoekspublicaties maken amper deel uit van het profiel van de instelling en de status van lectoren en (onderzoek-)docenten. De HBO Kennisbank kan een katalysator zijn in dit proces. De onderstaande aanbevelingen stellen deze rol van de HBO Kennisbank centraal.

Primair doel

Spreek gezamenlijk met de instellingen een primair doel af:

Primair doel: maak van de HBO Kennisbank een kwaliteitsvenster op de kennisproducten van het HBO door:

- vanaf 2015 de publicaties voortkomende uit het praktijkgericht onderzoek door lectoren, onderzoekers en docenten aan de HBO instellingen openbaar toegankelijk te maken, tenzij de auteurs motiveren waarom openbaarheid niet mogelijk is.
- vanaf 2015 tenminste 50% van de met een hoog cijfer beoordeelde studentenscripties openbaar beschikbaar te stellen.

Output van lectoren, onderzoekers en docenten

- *Publicatielijst in jaarverslag:* Iedere instelling neemt vanaf 2013 in zijn jaarverslag een (door het onderzoeksregistratiesysteem gegenereerde) bibliografie op van de uit het praktijkgericht onderzoek voortkomende publicaties door lectoren, docenten en onderzoekers, op basis van

gemeenschappelijke (door de HBO-Raad ontwikkelde) definities. De HBO-Raad publiceert een totaal overzicht (à la de VSNU) van de onderzoeksproductie.

- *Opname in de repository:* Deze publicaties worden met bijbehorende full text opgenomen in het repository van de instelling en zoveel mogelijk Open Access beschikbaar gesteld via de HBO Kennisbank.
- *Diensten voor deponerende lectoren/docenten/onderzoekers:* Deponerende auteurs dienen voordelen geboden te worden in de vorm van aantrekkelijke diensten, zoals automatisch gegenereerde publicatielijsten, de mogelijkheid om het aantal downloads van hun publicaties te bekijken en dergelijke.
- *Stimuleren van Open Access:* Open Access tot de resultaten van praktijkgericht onderzoek maakt optimale benutting hiervan door de samenleving mogelijk. Daarom dienen de Hogescholen en de HBO-raad een actief beleid te voeren om de Open Access toegankelijkheid van de uitkomsten van praktijkgericht onderzoek te stimuleren door het bevorderen van het opzetten van Open Access tijdschriften voor professionals, zowel via nieuw op te zetten tijdschriften als via migratie van bestaande tijdschriften.

Afstudeerscripties

- *Alle afstudeerscripties archiveren:* iedere instelling neemt alle afstudeerscripties op in de repository van de instelling voor archivering ten behoeve van accreditatie.
- *Alleen afstudeerscripties met een 8 of hoger* komen in aanmerking voor Open Access beschikbaarstelling via de HBO Kennisbank.

Vraagkant

- *Inzet van de HBO Kennisbank bij het bereiken van kennisvalorisatie:* een diepgaand kwantitatief en kwalitatief onderzoek naar gebruikers en het gebruik van de HBO Kennisbank biedt de mogelijkheid de rol van de HBO Kennisbank bij kennisvalorisatie nader te preciseren en verder te intensiveren.

Investerings

- Het in omloop brengen van de onderzoeksresultaten is een inherent deel van het onderzoek als zodanig. Een (bescheiden) deel van het onderzoeksbudget dient daarvoor te worden uitgetrokken. Om de bovenstaande voorstellen uit te voeren, is een gemiddelde personele inzet van 0,4 fte per instelling zonder meer te weinig. Daarnaast wordt ook personele inzet van de HBO-raad noodzakelijk geacht om de bovenstaande voorstellen mede vorm te geven en te implementeren

3 Methoden

Het onderzoek werd in drie fasen uitgevoerd (zie ook tabel hieronder):

- **Fase 1:** Opzet van het onderzoek: in deze fase werden de repository managers van zowel aan NARCIS toeleverende repositories als aan de HBO Kennisbank toeleverende repositories geconsulteerd over de opzet van de vragenlijst.
- **Fase 2:** Uitvoeren van een enquête onder de repository managers van de beide groepen repositories in de periode november/december 2011 m.b.v. Pleiadesurvey (zie <http://www.pleiade.nl/ict.php>).
In aanvulling op de enquête werden enkele interviews gehouden met sleutelfiguren en relevante documenten en literatuur geanalyseerd.
Ook is een kwantitatieve analyse uitgevoerd op basis van gegevens beschikbaar in NARCIS en de HBO Kennisbank.
- **Fase 3:** Analyse en presentatie van de gegevens en bespreking daarvan op het overleg van de managers van de aan NARCIS toeleverende repositories (WRM overleg 13 december 2011), en op het overleg van de contactpersonen van de HBO's (17 januari 2012) en de repository managers (18 januari).

Respondenten

De NARCIS enquête is uitgestuurd naar 28 repository managers van alle aan NARCIS toeleverende repositories. Er zijn 21 vragenlijsten ingevuld, waaronder alle bij de VSNU aangesloten universiteiten.

De HBO Kennisbank enquête is uitgestuurd naar alle 21 aan de HBO Kennisbank deelnemende hogescholen. 19 vragenlijsten zijn ingevuld.

4 Aan NARCIS toeleverende repositories

4.1 Onderzoeksvragen

Bij deze studie stonden de volgende onderzoeksvragen centraal:

1. In hoeverre zijn de aan NARCIS toeleverende onderzoekrepositories succesvol bij het Open Access beschikbaar stellen van de Nederlandse wetenschappelijke output?
2. Wat is de stand van zaken wat betreft de organisatie en het beheer van de repositories?
3. Wat is de stand van zaken wat betreft de technische ontwikkelingen van de repositories?
4. Wat is de stand van zaken wat betreft de dienstverlening op basis van de repositories?
5. Wat is er bekend over de gebruikers van repositories: zowel de deponerende wetenschapper als de informatiezoeker?
6. Wat is een mogelijk toekomstscenario voor de institutionele repositories?

Hieronder worden de voornaamste resultaten en conclusies betreffende deze onderzoeksvragen besproken.

4.2 Open Access beschikbaarstelling van Nederlandse wetenschappelijke output

In hoeverre zijn de aan NARCIS toeleverende onderzoekrepositories succesvol bij het Open Access beschikbaar stellen van de Nederlandse wetenschappelijke output?

Uit de kwantitatieve analyse blijkt het volgende:

- Het percentage publicaties dat in Open Access via de repositories jaarlijks ter beschikking komt, is niet significant gestegen sinds 2007: dit percentage schommelt in de periode 2007-2010 voor de bij de VSNU aangesloten instellingen (n=14) tussen 7% en 30% met een gemiddelde van 20% tot 21%.
- Een ruime meerderheid van de universiteiten slaagt er in om gedurende de afgelopen vier jaar (2007-2010) meer dan 80% van hun proefschriften open toegankelijk te maken op internet.
- Voor de KNAW is ook een vergelijking mogelijk op basis van de gepresenteerde output in hun jaarverslag. Ook daar zijn de percentages OA publicaties laag ten opzichte van de totale output. Voor de overige niet bij de VSNU aangesloten onderzoeksinstellingen met een repository (n=14) is zo'n vergelijking niet mogelijk.

Uit de resultaten van de enquête blijkt het volgende:

- Het content beleid van de repositories is divers.

Vrijwel alle institutionele repositories slaan wetenschappelijke publicaties op in de vorm van boeken/boek hoofdstukken, rapporten, tijdschriftartikelen, proefschriften/dissertaties en/of proceedings/conference papers. Een meerderheid slaat ook working papers en ongepubliceerde (wetenschappelijke) materialen op. Masterscripties en bachelorscripties worden door een minderheid van de institutionele repositories opgeslagen. Tenslotte slaat een minderheid van de institutionele repositories ook niet-tekstuele bestanden op (multimedia, software, datasets) en zijn er twee repositories die ook onderwijsmaterialen opnemen.

- De opnamecriteria verschillen.

De meeste repositories hanteren als criterium wetenschappelijke output, zoals door de VSNU gedefinieerd (wetenschappelijke publicaties; vakpublicaties) en zoals in het onderzoeksregistratiesysteem (METIS) wordt vastgelegd. Een minderheid hanteert geen selectiecriteria en neemt op wat aangeboden wordt. Wat betreft het type record, neemt ruim 28% van de repositories uitsluitend full text records op, de overige nemen ook metadata records zonder full text op. Wat betreft de versie van tijdschriftartikelen, overheerst de

uitgeversversie bij de meeste repositories, hoewel een meerderheid ook postprintversie (ook wel auteursversie genoemd) van tijdschriftartikelen opneemt. Bij circa éénderde van de repositories is in veel gevallen niet duidelijk voor de gebruiker welke versie van het tijdschriftartikel getoond wordt. Tenslotte, wat betreft het zogenaamde 'remote Open Access', geeft 57% aan dat de full text van het artikel uit het Open Access tijdschrift ook in de repository wordt opgenomen, terwijl 10% alleen de metadata opneemt en een link naar de full text in het Open Access tijdschrift. De overige 33% heeft nog geen beleid hierin geformuleerd.

- De resultaten wat betreft de wijze van beschikbaarheid van de repository materialen wijzen op een tweedeling.

14 van de 21 repositories hebben meer dan 75% van hun materialen Open Access beschikbaar, ofwel direct ofwel na een embargoperiode. Zeven van de 21 repositories hebben minder dan 70% van hun materialen Open Access beschikbaar. Dit verschil is waarschijnlijk het resultaat van een ander beleid: in het eerste geval staat de doelstelling Open Access (etalagefunctie) voorop, in het tweede geval staat de dienstverlening richting onderzoekers van de instelling voorop.

- Opslag en beschikbaarstelling van onderzoekdatasets en verrijkte publicaties staan nog in de kinderschoenen.

Een relatief recente ontwikkeling is de nadruk op het archiveren en beschikbaar stellen van onderzoekdatasets. Ruim de helft van de instellingen heeft (nog) geen mogelijkheden ontwikkeld voor opslag van datasets van de eigen onderzoekers. Van de overige instellingen geeft een groot deel aan dat onderzoekdatasets of in een apart repository of in hetzelfde repository kunnen worden opgeslagen, terwijl enkele instellingen doorverwijzen naar data-archieven en datacenters elders. Het beeld wat betreft verrijkte publicaties is vergelijkbaar: meer dan de helft geeft aan hiervoor nog geen mogelijkheden ontwikkeld te hebben, een deel stelt dat verrijkte publicaties in hetzelfde repository kunnen worden opgenomen, terwijl enkele aangeven hiervoor een apart repository te hebben.

4.3 Beheer en organisatie

Nationale repository infrastructuur

In de afgelopen jaren heeft elke instelling een repository ontwikkeld, waarin de wetenschappelijke productie van de instellingen (al dan niet in Open Access) gedeponeerd kan worden. In deze repositories worden vooral artikelen en proefschriften verzameld. Een goede aansluiting op het onderzoeksregistratiesysteem is van cruciaal belang voor het gemak waarmee auteurs hun full text content kunnen aanleveren. Landelijke invoering van een nieuw onderzoeksregistratiesysteem bij alle HO-instellingen vanaf 2012 (project NLRIS) is dan ook een kans voor bevordering van de 'green road' tot Open Access. UKB-leden participeren in het overleg dat Surf hiervoor organiseert (BIK).

Doelstellingen voor de komende periode zijn:

- optimale aansluiting op het nieuwe onderzoek registratie systeem
- ontwikkeling van een gezamenlijk repository systeem, in aansluiting op de landelijke samenwerking m.b.t. het NLRIS, met als doel verhoging van de efficiency in de processen rondom deze diensten in de bibliotheken.

Uit: UKB beleidsplan 2011-2015: De wetenschappelijke bibliotheek op weg naar 'the cloud'

Wat is de stand van zaken wat betreft de organisatie en het beheer van de repositories? Deze vraag is extra relevant gezien de beleidsvoornemens van het UKB (beleidsplan 2011-2015; zie tekstkader). De resultaten van de enquête laten het volgende zien:

- De inspanningen rond het beheer en organisatie van de institutionele repositories nemen bijna 40 FTE in beslag.

Bij de negentien repositories, waarvan een opgave werd ontvangen, zijn in het totaal 38 FTE betrokken. Bij repositories van de dertien universiteiten zijn in het totaal 34,5 FTE betrokken.

- Grote verschillen wat betreft inzet van menskracht tussen de universitaire repositories.

Bij vier universitaire repositories zijn 4 FTE of meer betrokken, terwijl bij twee andere repositories slechts 1 FTE of minder betrokken is. De overigen nemen tussenposities in: bij een universitair repository is 3 FTE betrokken, bij zes universitaire repositories zijn (ruim) 2 FTE betrokken.

- Administratieve en bibliotheektechnische werkzaamheden nemen meer dan de helft van de inspanningen in.

Vijftien repositories gaven ook een uitsplitsing van hun FTE naar verschillende taken. Uit deze gegevens blijkt dat circa 60% van de FTE inspanning gaat naar administratieve en bibliotheektechnische werkzaamheden, terwijl circa 20% aan managementtaken en circa 20% aan technische/ICT werkzaamheden wordt besteed. In de kwantitatieve analyse (zie bijlage A) werd geen correlatie gevonden tussen de personele inzet en het OA percentage.

- De werkprocessen wat betreft het deponeren van metadata van materialen in de repositories vertonen een grote diversiteit.

Bij 28% van de repositories wordt de invoer van metadata verzorgd door de medewerkers van de repository zelf, bij 19% van de repositories gebeurt de invoer van de metadata door de auteurs. Bij één repository (4,8%) worden de metadata overgenomen uit andere bestanden. 33% van de repositories meldt een combinatie van deze methoden en 14% zet andere methoden in. In vergelijking met de nulmeting in 2008 is er een verschuiving opgetreden richting combinatie van methoden.

- Het percentage repositories met een verplichtstelling is in de afgelopen jaren fors gestegen zodat nu bij een meerderheid van de repositories de instelling een of andere vorm van verplichtstelling heeft geïntroduceerd voor het deponeren van publicaties.

Bij meer dan 70% van de repositories geldt een verplichtstelling tot deponeren: bij 38% voor alle onderzoekpublicaties, bij 33% uitsluitend voor dissertaties/proefschriften. De overige 28% van de institutionele repositories hebben geen verplichtstelling. In 2008 waren er slechts enkele repositories met een verplichtstelling voor dissertaties/proefschriften. In de kwantitatieve analyse (zie bijlage A) werd geen correlatie gevonden tussen deze verplichtstelling en het OA percentage.

- Overwegingen van duurzaamheid (energiegebruik, groene ICT) spelen een ondergeschikte rol bij de (toekomstige) opzet en inrichting van repositories.

In de enquête geeft 38% van de respondenten aan dat dergelijke overwegingen een rol zouden moeten spelen. In de voorbereidingen van de vragenlijst werd echter door meerdere repository managers gesteld dat dergelijke overwegingen in de praktijk geen rol spelen.

4.4 Technische ontwikkelingen

Wat is de stand van zaken wat betreft de technische ontwikkelingen van de repositories?

Uit de enquête en de kwantitatieve analyse komt het volgende naar voren:

- De diversiteit in softwarepakketten is onveranderd groot gebleven.

Bij de 21 repositories zijn vijf standaard softwarepakketten in gebruik, terwijl vier repositories lokaal ontwikkelde software in gebruik hebben voor hun institutionele repository. Dus in het totaal negen verschillende softwarepakketten. Dit is - ondanks enkele veranderingen - onveranderd gebleven ten opzichte van 2008.

- Een grote meerderheid wetenschappelijke auteurs heeft een Digital Author Identifier (DAI).

Ten opzichte van de nulmeting van 2008 is het aantal repositories dat een procedure heeft voor het toekennen van een digital author identifier sterk toegenomen. Een meerderheid (ruim 60%) heeft zo'n procedure, van hen geeft 77% aan dat zij inmiddels bij meer dan 50% van de auteurs van hun instelling een DAI hebben toegekend. De toename van de DAI's blijkt ook uit de kwantitatieve analyse: in NARCIS hebben 19.500 personen een DAI. Uitgaande van circa 25.000 personen die in aanmerking komen voor een DAI, bedraagt de DAI-dekkingsgraad voor personen 80%. Dit is een grote verandering ten opzichte van 2008, toen de DAI nog bij bijna de helft van de repositories niet was ingevoerd.

- SURE2 - een dashboard met gebruiksstatistieken - zal bij tweederde van de repositories worden geïmplementeerd.

In het SURFshare project SURE2 zijn gebruikstatistieken van objecten in repositories zichtbaar gemaakt in NARCIS. Daardoor kan een gebruiker o.a. zien hoe vaak een bepaalde publicatie is gedownload. De verwachting is dat dit voor deponerende auteurs interessante informatie is die het deponeren van andere publicaties zal stimuleren. Tweederde van de repositories doet aan dit project mee.

- De digitale duurzaamheid van de content van de repositories is voor een belangrijk deel (nog) niet gewaarborgd.

Ruim de helft van de repositories geeft aan dat de digitale duurzaamheid van de materialen gewaarborgd is. Het overgrote deel daarvan geeft aan dat dit gebeurt door de Koninklijke Bibliotheek. In de nulmeting van 2008 gaf 100% aan dat de digitale duurzaamheid gewaarborgd was. De terugloop is waarschijnlijk te danken aan de toegenomen kennis rond de problematiek van digitale duurzaamheid en de problemen rond het harvesten van de content van de repositories door de Koninklijke Bibliotheek ten behoeve van hun e-depot. Uit het interview met de respondent van de Koninklijke bibliotheek hierover blijkt het volgende:

- Na een eerdere mislukking, is sinds 2010 een nieuwe, sterk verbeterde infrastructuur voor de harvesting van de repositories ten behoeve van het e-depot opgezet. Op dit moment zijn vier of vijf repositories geharvest. Men stuit daarbij op diverse problemen. De verschillen in de gebruikte softwarepakketten spelen daarbij een belangrijke rol.
- Er spelen diverse kwesties rond het harvesten van full text, die niet publiek beschikbaar is, hoe om te gaan met publicaties die teruggetrokken zijn en dergelijke, waarvoor de huidige afspraken tekortschieten.
- In de werkafspraken met de repositories is afgesproken dat er een set van circa dertig bestandsformaten zal worden geharvest. Echter, de bestandsformaten ontwikkelen zich verder en worden steeds complexer door onder andere insluiting van niet-tekstuele componenten (zoals bij verrijkte publicaties). Om de problematiek rond de digitale duurzaamheid van deze complexe bestandsformaten aan te pakken zullen de instituten die voor deze materiaalvormen de digitale duurzaamheid in Nederland organiseren nauw gaan samenwerken: het betreft KB, DANS, 3TU Datacentrum en Beeld en Geluid.

- Wat betreft onderwerpsontsluiting is er een tweedeling tussen de repositories.

Iets minder dan de helft van de repositories geven aan geen onderwerpsontsluiting te hebben. De overige repositories hebben wel verschillende vormen van onderwerpsontsluiting. De VSNU indeling in HOOP-gebieden wordt overigens door niemand gebruikt. Enkele repositories leggen relaties aan tussen objecten, zoals 'part of'-relatie en preprint-postprint relatie.

- Veranderingen in enkele overige technische aspecten zijn gering en/of volgens verwachting.

De toepassing van persistente identifiers in repositories is nauwelijks veranderd, wat betreft de toepassing van metadata standaarden is er een verschuiving te zien richting DIDL en MODS.

4.5 Dienstverlening op basis van repositories

Wat is de stand van zaken wat betreft de dienstverlening op basis van de repositories? Hieronder de voornaamste uitkomsten uit de enquête:

- De meeste repositories zijn geregistreerd bij Open DOAR en de content van de meeste repositories is ontsloten via Google, Google Scholar, OAIster en DRIVER.

Het overgrote deel van de repositories heeft zich geregistreerd bij Open DOAR, slechts een minderheid bij andere internationale registers (ROARMAP, The University of Illinois OAI-PMH Data Provider Registry, Open Archives registered data providers). Wat betreft ontsluiting via OAI gebaseerde harvesters is de meerderheid ontsloten via OAIster en DRIVER, een minderheid ook via BASE. De content van vrijwel alle repositories is eveneens vindbaar via algemene, internationale zoekmachines als Google, Yahoo en MSN. Wat betreft wetenschappelijke zoekmachines scoort Google Scholar heel hoog. De content van de meeste repositories is echter niet vindbaar via andere, minder bekende wetenschappelijke zoekmachines op het internet (Scirus en Scientific Commons). Voor ontsluiting via Google of Google Scholar dient een repository speciale sitemaps aan te maken: slechts een kleine minderheid van de repositories heeft dit gedaan. Dit betekent dat de ontsluiting vooral indirect via NARCIS gebeurt, dat wel dergelijke sitemaps aanmaakt.

- De content van een kleine meerderheid van de repositories is ontsloten via de eigen digitale bibliotheek.

57% van de repositories geeft aan dat de materialen van de repository ontsloten zijn in de eigen digitale bibliotheek, ofwel via de catalogus ofwel via een uniform discovery tool.

- Geen vooruitgang in overige diensten, zoals persoonlijke diensten aan auteurs of printing on demand services.

27% van de repositories geeft aan een printing on demand service aan te bieden. 47% biedt persoonlijke diensten aan de auteurs aan (zoals automatisch gegenereerde publicatielijsten). Beide diensten zijn vrijwel onveranderd gebleven ten opzichte van 2008.

4.6 Wetenschappers als deponerende auteurs en als informatiezoekers

Wat is er bekend over de gebruikers van repositories: zowel de deponerende wetenschapper als de informatiezoeker? Onder Nederlandse wetenschappers zijn voor zover ons bekend geen onderzoeken hiernaar verricht. Wel is er een omvangrijke studie gedaan onder Europese onderzoekers door het PEER Observatory. Dit onderzoek bestond uit een enquête in 2009² onder Europese onderzoekers (3139 respondenten, waarvan 156 uit Nederland) en een tweede enquête onder Europese auteurs van tijdschriftartikelen, dat op een aantal zaken wat dieper ingaat³. Dit laatste survey werd uitgevoerd in 2010/2011 en werd door 1427 respondenten ingevuld, waarvan 92 uit Nederland. Dit survey werd in 2010/2011 uitgevoerd. De belangrijkste resultaten worden hieronder kort samengevat. Het betreft de resultaten uit de eerste enquête, tenzij anders vermeld.

- Bekendheid van onderzoekers met repositories is matig.

32,2% van de onderzoekers geeft aan dat het eigen instituut een repository heeft; 37,6% geeft aan een subject-based repository voor het eigen onderzoek te kennen. Daarnaast geeft nog 13,1% aan andere geschikte repositories te kennen.

- Meer dan de helft van de onderzoekers deponeert wel eens.

52,7% van de respondenten (waarvan 98,8 % auteurs) geeft aan een versie van een eigen artikel, verschenen in een peer-reviewed tijdschrift in een repository te hebben gedeponerd: 23,6% in een institutioneel repository, 24,4% in een subject-based repository (11,9% elders/weet niet).

- Bewegredenen om te deponeren zijn ideëel, eigenbelang en in mindere mate onder invloed van de omgeving.

Gevraagd naar de belangrijkheid van verschillende driving factoren om tijdschriftartikelen in een repository te deponeren zijn er drie clusters te onderscheiden (weergegeven percentages zijn van respondenten die de betreffende factor (heel) belangrijk vonden; clustering door de opstellers van het rapport):

- Ideële factoren in relatie met Open Access:
 - 89%: wijde beschikbaarheid van mijn onderzoeksresultaten
 - 86%: het principe van vrije toegang voor allen
 - 81%: beschikbaarheid voor onderzoekers met beperkte toegang tot tijdschriften
 - 53%: hoge kosten tijdschrift abonnementen
- Eigenbelang:
 - 70%: snelheid van disseminatie
 - 68%: mogelijkheid van verhoogd aantal citaties
 - 43%: ten behoeve van de eigen carrière
- Druk omgeving:
 - 34%: vereisten onderzoeksfinanciers
 - 29%: mandaat eigen instelling
 - 29%: aandrang collega's

² PEER Behavioural Research: Authors and Users vis-a-vis Journals and Repositories; baseline report; J. Fry, C. Oppenheim, S. Proberts, C. Creaser, H. Greenwoord, V. Spezi, S. White; September 2009

³ PEER Behavioural Research: Authors and Users vis-a-vis Journals and Repositories; final report; J. Fry, S. Proberts, C. Creaser, H. Greenwoord, V. Spezi, S. White; August 2011

- De remmende factoren wat betreft deponeren betreffen voornamelijk (1) zorgen over copyright en dergelijke en (2) weinig interesse en kennis.

Gevraagd naar de belangrijkheid van de verschillende remmende factoren om tijdschriftartikelen in het Repository te deponeren zijn er twee clusters remmende factoren te onderscheiden:

- zorgen over copyright en dergelijke:
 - 48%: bezorgd over overtreden copyrightregels van uitgevers.
 - 48%: terughoudend om eigen onderzoekspublicaties in de repository te deponeren tussen andere niet- peer reviewed publicaties.
 - 41%: bezorgd over overtreden regels van uitgevers m.b.t. embargo periode.
 - 35%: vindt het niet prettig een peer reviewed versie van mijn artikel te deponeren dat niet geredigeerd is door de uitgever.
 - 31%: bezorgd over plagiaat/hoe eigen materiaal gebruikt zal worden door lezers.
- weinig interesse/kennis:
 - 37%: kent geen geschikt repository
 - 30%: onvoldoende bewust van wat een Open Access repository is.
 - 30%: weet niet hoe materiaal te deponeren.
 - 17%: vindt het deponeren te gecompliceerd/te tijdrovend.
 - 15%: ziet dit niet als eigen verantwoordelijkheid.
 - 14%: vergeet het meestal.
 - 12%: ziet Open Access niet als voordelig voor eigen onderzoek.

- Veel wetenschappers hebben een voorkeur voor een subject-based repository.

Gevraagd naar hun voorkeur, geeft 46,1% van de respondenten aan een voorkeur te hebben voor subject-based repository en 22,1% voor een institutioneel repository.

- Wat betreft zoekgedrag komt het regelmatig voor dat men geen toegang heeft tot een tijdschrift(-artikel) en dat een Open Access bron dan uitkomst biedt.
 - 34,4% van de respondenten geeft aan regelmatig geen toegang te hebben tot een tijdschriftartikel van een peer reviewed tijdschrift. 42,8% geeft aan dit soms mee te maken.
 - In geval van geen directe toegang tot het tijdschriftartikel geeft 50,2% aan een Open Access bron te raadplegen (26,4% vraagt een kopie aan via IBL, 21,7% neemt contact op met de auteur; 22,9% vraagt een collega; 19,5% vraagt de bibliotheek; 3,7% betaalt voor het downloaden van het artikel en 11,1% gebruikt het artikel in dat geval niet).
 - 74% vindt het belangrijk te weten welke versie van het tijdschriftartikel men heeft gevonden. 39% stelt dat het meestal wel duidelijk is welke versie men heeft gevonden.
- De meesten citeren de gepubliceerde versie bij gebruik van een tijdschriftartikel uit een repository.
 - 87%: probeert de gepubliceerde versie te vinden voordat het artikel geciteerd wordt.
 - 81%: geeft de voorkeur eraan om uitsluitend de gepubliceerde versie van het artikel te citeren
 - 71%: citeert meestal de versie van de publicatie die is gelezen
 - 61%: citeert in het algemeen alleen de auteursversie als deze is peer reviewed en geaccepteerd voor publicatie
 - 43%: citeert eventueel een preprintversie
 - 20%: citeert eventueel elke versie
- Twee derde van de onderzoekers ziet een rol voor repositories in het wetenschappelijke communicatieproces en driekwart vindt de voordelen van het Open Access beschikbaar stellen van hun artikel opwegen tegen de extra inspanningen.

- o 66% van de respondenten ziet een rol voor repositories in het wetenschappelijke communicatieproces. Slechts 4% ziet geen rol voor repositories terwijl 30% geen mening heeft.
- o 75% denkt dat de voordelen van het Open Access beschikbaar stellen van het artikel opwogen tegen de extra inspanning (7% niet; 18% weet niet; resultaat tweede enquête).

4.7 Naar een toekomstscenario voor de institutionele repositories

4.7.1 Doelstellingen voor de institutionele repositories

Wat is een mogelijk toekomstscenario voor de institutionele repositories die aan NARCIS toeleveren? Om deze vraag goed te kunnen beantwoorden, dient eerst naar de doelstellingen van de institutionele repositories gekeken te worden. Uit de enquête blijkt dat de repositories één primaire doelstelling hebben: de etalagefunctie (Open Access) (100% van de respondenten onderschrijven deze doelstelling, waarvan slechts 14% met 'enigszins mee eens'). Een andere belangrijke doelstelling is de servicefunctie voor de wetenschappers (95% onderschrijft deze doelstelling, waarvan 33% met 'enigszins mee eens'). De archieffunctie van de repositories wordt eveneens breed onderschreven (90%), in tegenstelling tot de registratie- en rapportagefunctie (respectievelijk 67% en 43%)⁴. Overigens blijkt dat dit vooral de visie van de respondenten weergeeft (48%), slechts bij een derde is dit ook vastgesteld beleid.

Al eerder is gesignaleerd dat de doelstellingen 'etalagefunctie' en 'servicefunctie' kunnen leiden tot verschillende uitkomsten: zo hebben 14 van de 21 repositories meer dan 75% van hun materialen Open Access beschikbaar, terwijl 7 repositories juist een belangrijk deel van hun content via Campus Access beschikbaar. Dit verschil is waarschijnlijk het resultaat van een ander beleid: in het eerste geval staat de doelstelling Open Access (etalagefunctie) voorop, in het tweede geval staat de dienstverlening richting onderzoekers van de instelling voorop. Niettemin ondersteunen de twee hoofddoelstellingen in het algemeen elkaar en zullen slechts in uitzonderingsgevallen tot tegenstrijdige beleidsbeslissingen leiden.

4.7.2 Realisatie van succesfactoren

Erkende succesfactoren zijn voor een belangrijk deel gerealiseerd.

In de enquête zijn negen - algemeen geaccepteerde - succesfactoren voor repositories voorgelegd aan de respondenten met de vraag (1) in hoeverre deze gerealiseerd zijn, (2) in hoeverre ze bijgedragen hebben tot het succes van de repository en (3) - indien niet gerealiseerd - welke bijdrage men verwacht. Uit de resultaten blijkt dat zes van de negen succesfactoren bij meer dan 50% van de repositories gerealiseerd zijn (zie tabel; lichtgroen). Vijf daarvan hebben volgens de respondenten ook daadwerkelijk bijgedragen aan het succes van de repository. Eén van de veronderstelde succesfactoren blijkt in de praktijk weinig bij te dragen aan het succes van de repository (long term archivering). Van de overgebleven drie succesfactoren, die nog niet bij 50% van de repositories gerealiseerd zijn (namelijk: betrokkenheid van beslissers van uw instelling; bekendheid van onderzoekers met vereisten van financiers m.b.t. deponeren; verplichting tot deponeren; zie ook tabel, lichtrood) wordt door de meeste respondenten een bijdrage verwacht aan het succes van de repository.

Uit een vergelijking van deze drie gemeten waarden in deze enquête kan worden afgeleid dat er bij het realiseren van twee succesfactoren nog kansen liggen voor de NARCIS repositories als geheel. Dit is het geval voor bekendheid van onderzoekers met vereisten financiers met betrekking tot deponeren en integratie en/of koppeling met andere systemen in de instelling. Op grond van deze

⁴ De registratie en rapportagefunctie wordt bij universiteiten voornamelijk toegeschreven aan METIS, het onderzoeksregistratiesysteem.

analyse kan voor twee andere, nog niet bij meer dan 50% van de repositories gerealiseerde succesfactoren gesteld worden dat het effect ervan in de praktijk tegenvalt: dit geldt voor de betrokkenheid van beslissers van de instelling zelf en voor verplichting tot deponeren.

Succesfactoren	gerealiseerd	draagt bij (indien gerealiseerd)	zal bijdragen (indien niet gerealiseerd)
Eenvoudig en gebruikersvriendelijk deponeren	71,4	62,5	63,7
Voorlichtingscampagnes over uw repository	66,7	68,7	42,9
Integratie en/of koppeling met andere systemen in uw instelling	61,9	80,0	70,0
Bekendheid onderzoeker met verhoogde citatiescores voor OA publicaties	57,1	64,3	63,7
Informatievoorziening over copyright in relatie met het deponeren	57,1	60,0	50,0
De long-term archivering	55,0	42,8	30,0
Betrokkenheid van beslissers van uw instelling	47,6	46,2	83,3
Verplichting tot deponeren	28,6	40,0	64,3
Bekendheid van onderzoekers met vereisten financiers deponeren	14,3	50,0	57,1

4.7.3 Criteria voor het bereiken van de hoofddoelstellingen

Voor het adequaat realiseren van de etalagefunctie is het huidige percentage van Open Access beschikbaarheid van de wetenschappelijke output - gemiddeld circa 20% van de repositories - volstrekt onvoldoende. Uit de bovenstaande analyse met betrekking tot de succesfactoren blijkt dat er bij het realiseren van succesfactoren nog wel winst te behalen valt, maar dat de verwachtingen hiervan niet te hoog gespannen moeten zijn. Is het niet nodig om te komen tot een volstrekt andere aanpak? Deze vraag werd voorgelegd aan de werkgroep repository managers (WRM) in hun overleg van 13 december.

Deze discussie leidde tot de volgende resultaten:

- Er is consensus over de doelstelling dat in 2015 elke instelling circa 70% van zijn wetenschappelijke output via de repository Open Access ter beschikking dient te stellen: Gesteld werd dat de achilleshiel van de Open Access beweging is dat iedereen mee moet doen om een aanzienlijk percentage van de wetenschappelijke output te realiseren dat daadwerkelijk Open Access beschikbaar is. Wanneer dit percentage niet ver boven de 20% komt, zouden de repositories een belangrijk deel van hun functie kwijtraken. Het bereiken van een zekere kritische massa hierbij essentieel is. De verwachting is dat wanneer iedere instelling circa 70% van zijn wetenschappelijke output Open Access ter beschikking stelt via de repository, het totale percentage Open Access van de Nederlandse wetenschappelijke output in de buurt van 100% zal komen. Dan is deze doelstelling bereikt en kunnen de repositories als succesvol worden aangeduid.
- Voor wat betreft de doelstelling 'service functie voor eigen onderzoekers' dient in 2015 tenminste de helft van de onderzoekers verbonden aan de eigen instelling gebruik te maken van de diensten rondom het institutionele repository: uit de discussie kwam naar voren dat één succes criterium het aantal downloads vanuit de repository zal zijn. Dit wordt binnenkort zichtbaar gemaakt met gebruiksstatistieken in NARCIS als resultaat van het SURE2 project. Wanneer deze aantallen downloads voldoende hoog zullen blijken, wordt verwacht dat dit een belangrijke stimulans zal zijn voor wetenschappers om hun publicaties te deponeren. De eerste

gegevens wijzen overigens daarop.

Een tweede aspect van de dienstverleningsdoelstelling is de DAI. Eén deelnemer stelt dat 'wij onze professor O. het beste kennen'. Deze filosofie volgend besteden zij veel aandacht aan het implementeren van de DAI en het opschonen van publicaties die wel of niet bij deze DAI horen. Met een opgeschoonde DAI kunnen zij weer uitgevers en anderen informeren over welke publicaties bij welke auteur (DAI) horen, daarmee de citatiescores van de betreffende auteur bij andere uitgevers/databanken betrouwbaarder makend. In de groepsdiscussie werd geen criterium genoemd voor het bereiken van deze doelstelling. Het lijkt echter redelijk te veronderstellen dat deze doelstelling bereikt is wanneer meer dan de helft van de interne onderzoekers gebruik maakt van de dienstverlening rond de repository.

4.7.4 Bouwstenen voor een toekomstscenario

Wat zijn de bouwstenen voor een toekomstscenario voor de Nederlandse onderzoekrepositories? De voor deze studie verzamelde gegevens leidde tot het volgende observaties/conclusies, die als bouwstenen fungeren voor het opstellen van een toekomstscenario:

- De aan NARCIS toeleverende repositories kennen één primaire doelstelling: het zo breed mogelijk beschikbaar stellen van wetenschappelijke output (etalagefunctie; Open Access). In het verlengde daarvan ligt een andere belangrijke doelstelling: een service functie voor de onderzoekers van de eigen instelling. Deze doelstellingen versterken elkaar.
- De aan NARCIS toeleverende repositories kunnen worden opgesplitst in twee categorieën:
 - Vijftien repositories van de grotere academische instellingen, aangesloten bij de VSNU, en de KNAW en NWO. Vanuit deze instellingen wordt het overgrote deel van de Nederlandse academische wetenschappelijke output verzorgd en met een aantal van ruim 80.000 publicaties jaarlijks. Er is overigens een opvallende lacune in het landschap van institutionele repositories van deze instellingen: afgezien van het CWI ontbreken institutionele repositories van de overige onderzoeksinstituten van NWO.
 - Dertien repositories van kleinere academische instellingen, onderzoeksinstituten en verwante instellingen. Deze hebben een jaarlijks totaal van ruim 1000 publicaties.
- Sinds 2005 hebben alle Nederlandse universiteiten, de HBO-raad namens alle hogescholen alsmede de KNAW en NWO (en de KB en SURF) de [Berlin Declaration](#)⁵ ondertekend en ondersteunen dus in principe Open Access tot wetenschappelijke output. De twee wegen om Open Access te bereiken zijn de 'green road' en de 'golden road'. De 'green road' is met behulp van institutionele repositories, zoals hier beschreven, De 'golden road' is het publiceren in Open Access journals. Uit een inventarisatie van de UKB werkgroep OA onder de dertien grote universiteiten en de KNAW blijkt dat het beleid ten aanzien van de 'golden road' nauwelijks van de grond is gekomen: er zijn twee instellingen met een Open Access fonds (ter ondersteuning van het publiceren in OA tijdschriften), er zijn vier instellingen met een OA membership (waardoor auteurs kunnen publiceren zonder de 'publication fee' in bijvoorbeeld BioMed Central of PLoS tijdschriften) en er zijn acht instellingen die zelf het uitgeven van OA tijdschriften faciliteren (dit betreft ruim dertig tijdschrifttitels)⁶. Hieruit kan worden geconcludeerd dat de Nederlandse onderzoeksinstituten hoofdzakelijk de 'green road' tot Open Access – de route via de institutionele repositories – bewandelen en de 'golden road' – de route via de uitgevers - tot nu toe grotendeels links laten liggen.
- Van de negen erkende 'standaard' succesfactoren van de institutionele repositories blijken er zes bij meer dan de helft van de repositories gerealiseerd. Van de drie succesfactoren, die nog niet bij meer dan de helft van de repositories gerealiseerd zijn, kan op grond van de ervaringen van die repositories, waar dit wel gebeurd is, geen doorbraken worden verwacht in de dekkingsgraad van de repositories.

⁵ <http://oa.mpg.de/openaccess-berlin/signatories.html>

⁶ D.d. 5 april 2011; met dank aan Saskia Franken, Igitur, Universiteit Utrecht.

- De dekkingsgraad voor proefschriften bedraagt meer dan 80%. Een belangrijke reden hiervoor is waarschijnlijk dat de opname van het proefschrift in de repository in de workflow van het promoveren is opgenomen.
- De dekkingsgraad voor de overige wetenschappelijke output bedraagt circa 20%¹ en is niet gestegen in de periode 2008-2010. Een belangrijke reden hiervoor is dat het deponeren in een repository los staat van de workflow voor het publiceren en als zodanig niet als aantrekkelijk wordt ervaren.
- Als succes criterium voor de primaire doelstelling wordt gesteld dat de dekkingsgraad van de repositories in 2015 tenminste 70% dient te bedragen. En in aanvulling daarop dat de helft of meer van de wetenschappers van de eigen instelling gebruik maakt van de diensten rondom de repository.
- Uit dit onderzoek blijkt een grote verscheidenheid tussen de repositories wat betreft content beleid, werkprocessen, technische opzet en de dienstverlening rondom de repositories. Dit maakt het gezamenlijke beheer en beleid van de gehele repository infrastructuur complex en de voortgang ervan langzaam. Daarbij komt dat de gekozen overlegvorm (Werkgroep Repository Managers) een vrijblijvend karakter heeft, waardoor bestaande werkafspraken onvoldoende worden nageleefd en onvoldoende geactualiseerd. Dit leidt ertoe dat de periodes van tekentafel tot implementatie voor nieuwe technische ontwikkelingen en diensten, zoals de Digital Author Identifier en SURE2, enkele jaren in beslag neemt. Ook dan is er zelden een volledige implementatie bij alle repositories.
- De rol van NARCIS als nationale portal is vooral van belang voor nationale metadata opslag, met filters voor het doen van specifieke tellingen en site maps voor het aanleveren aan Google en Google Scholar. Uit de voorlopige gegevens van het aantal downloads van publicaties in repositories blijkt dat de vraag vrijwel uitsluitend komt via Google of Google Scholar. Daaruit blijkt dat de betekenis van NARCIS als zoekingang beperkt is.
- Naast NARCIS harvest ook de Koninklijke Bibliotheek de content van de repositories om deze duurzaam op te slaan in het e-depot. De Koninklijke Bibliotheek ondervindt daarbij moeilijkheden om de content op een adequate en consistente manier binnen te halen.

4.7.5 Voorstel voor een toekomstscenario

Op basis van bovengenoemde bouwstenen, wordt een toekomstscenario voorgesteld. Dit toekomstscenario richt zich in eerste instantie op de vijftien institutionele repositories van de grote onderzoeksinstituten (de bij de VSNU aangesloten instellingen, KNAW en NWO).

Doelstelling van het toekomstscenario is dat in 2015 meer dan 70% van de jaarlijkse wetenschappelijke output van deze instellingen in Open Access beschikbaar wordt gesteld.

De 70% doelstelling zal in 2015 nog niet gehaald kunnen worden door middel van de uitgeversversie van het tijdschriftartikel; de auteursversie van het tijdschriftartikel is dus nodig en kan alleen met medewerking van de wetenschappelijke auteur verkregen worden.

Om medewerking van de auteurs te verkrijgen zal een dienstenpakket rond de repositories ontwikkeld worden waarmee enerzijds het deponeren onderdeel wordt van de workflow van het publiceren en anderzijds deponerende auteurs beloond worden met voor hen aantrekkelijke diensten.

Het voorgestelde toekomstscenario omvat vier thema's:

1. **Gezamenlijkheid:** Maak het primaire doel de toetssteen voor alle gezamenlijke beslissingen en ontwikkelingen. Stel een gezamenlijk plan op hoofdlijnen op voor de periode 2012-2015 met afgesproken mijlpalen.

2. **Meewerkende auteurs:** Ontwikkel een serie apps die het deponeren makkelijker maken via aantrekkelijke diensten voor de auteurs. Bied deze diensten aan op meerdere aggregatieniveaus (faculteiten, leerstoelgroepen, projecten, instellingen et cetera).

De aangeboden diensten dienen de wetenschappelijke auteurs te ondersteunen bij het distribueren van hun publicatie (zoals: via het institutionele repository, NARCIS en verder; doorsluizen naar relevante subject repositories), bij het verhogen van de zichtbaarheid en de herkenbaarheid van de betreffende wetenschappelijke auteurs (harmoniseren/opschonen van door wetenschappelijke databanken zoals Web of Science of Scopus gebruikte author identifiers, bijv. via Linked Data technieken), en bij het bieden van persoonlijke diensten (automatisch gegenereerde publicatielijsten, dashboard van het aantal downloads van eigen publicaties en dergelijke). De Digital Author Identifier en de SURE2 technieken leveren een technische basis voor deze diensten. Het is aanbevelenswaardig deze diensten te ontwikkelen in de vorm van 'apps', die op individuele en gezamenlijke onderzoekomgevingen gebruikt kunnen worden. Ook verdient het aanbeveling om deze diensten zo te ontwikkelen, dat de wetenschappelijke auteurs wijzigingen kunnen aanbrengen, die vervolgens weer gebruikt kunnen worden ter verbetering van de content van de repository (voorbeeld: Google Scholar Citations). Tenslotte wordt benadrukt om bij de ontwikkeling van deze diensten zoveel mogelijk aan te sluiten bij internationale successen, zoals Google Scholar Citations voor publicatielijsten, CiteULike/Mendeley/Zotero voor publicaties, Dataverse/DataCite voor datasets of Linked Open Data voor webdiensten.

Ontwikkel een op het bovenstaande gebaseerd gezamenlijk, uniform dienstenpakket op instellingsniveau, faculteits-niveau en leerstoelgroepsniveau: Deze diensten dienen te leiden tot het gebruik van de content van de repositories om de wetenschappelijke output op het niveau van groepen te kunnen etaleren, monitoren en de impact daarvan te meten. Het gaat bijvoorbeeld om dashboard-achtige diensten (waarbij instellingen bijvoorbeeld kunnen monitoren hoeveel publicaties er zijn gepubliceerd, hoeveel daarvan Open Access beschikbaar zijn en hoeveel downloads daarvan plaatsvinden) en om diensten die de etalagefunctie versterken (automatisch gegenereerde publicatielijsten en dergelijke).

3. **Infrastructuur:** Maak NARCIS het logistieke hart voor toelevering aan Google Scholar, het e-Depot en andere (disciplinaire) (inter)nationale dienstverleners. Beleg daartoe een duidelijke, niet-vrijblijvende regiefunctie bij een overkoepelende instelling.

Er bestaat nu een wirwar van afspraken met betrekking tot het harvesten ten behoeve van NARCIS (beheerd door DANS) en ten behoeve van het e-Depot (beheerd door de Koninklijke Bibliotheek), waarbij een deel van de technische inspanningen dubbel geschiedt. Een bundeling van krachten tussen DANS en de KB, een vereenvoudiging en stroomlijning van de afspraken tussen de institutionele repositories enerzijds en DANS en de KB anderzijds zal tot een efficiëntere inzet van middelen en een beter resultaat leiden.

Harmoniseer stapsgewijs de repository infrastructuur: voor wat betreft het content beleid wordt een afstemming met de definities van de VSNU voor wetenschappelijke output bepleit. Verder dienen er een eenduidig content beleid en eenduidige werkprocessen te worden opgesteld en geïmplementeerd, die helder te communiceren zijn richting deponerende wetenschappers in Nederland. Een harmonisatie van de technische infrastructuur van de repositories is van belang. Dit kan geschieden door bij elk ontwikkelstap convergentie af te dwingen en (verdere) divergentie tegen te gaan.

4. **Green én gold:** Omdat alleen de green road" vrijwel zeker niet voldoende zal zijn om het primaire doel te halen, wordt aanbevolen ook het beleid ten aanzien van de 'golden road' te intensiveren. Het voorstel is om hiervoor een *taskforce* op te zetten, onder leiding van de voorzitter van de VSNU, die zich namens de universiteiten inzet voor de overgang van het licentiemodel naar het OA-model.

¹ (noot bij blz. 25) In december 2011 is door Wouter Gerritsma van Wageningen Universiteit en Researchcentrum een onderzoek gedaan naar de Open Access dekkingsgraad in NARCIS. (A census of Open Access in the Netherlands, 20120202, to be blogged.) Het betreft een ad-hoc totaalopname, geen meerjarige trendanalyse. Ook zijn alle typen publicaties meegenomen, dus niet alleen de wetenschappelijke en vakpublicaties die door de VSNU worden geregistreerd, maar ook technische documentatie, working papers, preprints e.d.. Zowel het collectiebeleid als het aanbiedingsbeleid aan NARCIS van de onderscheiden instellingen verschilt sterk op deze punten. In dit licht moet de volgende uitspraak worden geplaatst: "It can easily be argued that Wageningen UR, of all repositories in the Netherlands excels at disseminating grey literature by means of their open access repository Wageningen Yield." De overige conclusies uit de studie van Gerritsma vallen grosso modo binnen de bevindingen van deze Monitor.

5 Aan HBO Kennisbank toeleverende repositories

5.1 Onderzoeksvragen

De infrastructuur van repositories voor het HBO is van recentere datum en er heeft geen eerdere nulmeting plaatsgevonden zoals bij de aan NARCIS toeleverende repositories. De onderzoeksvragen voor dit onderdeel van de studie waren dan ook voornamelijk inventariserend van aard.

Bij deze studie stonden de volgende onderzoeksvragen centraal:

1. Wat is de stand van zaken wat betreft Open Access beschikbaarstelling en content?
2. Wat is de stand van zaken met betrekking tot het beheer en organisatie rondom de repositories?
3. Wat is de stand van zaken met betrekking tot de technische ontwikkeling van de repositories?
4. Wat is de stand van zaken met betrekking tot de dienstverlening gebaseerd op de repositories?
5. Wat is een mogelijk toekomstscenario?

Hieronder worden de voornaamste resultaten en conclusies betreffende deze onderzoeksvragen besproken.

5.2 Open Access beschikbaarstelling en content

De resultaten van de kwantitatieve analyse en de enquête wat betreft deze onderzoeksvraag zijn als volgt:

De HBO Kennisbank oogst en toont de inhoud van 21 HBO repositories. De site biedt voornamelijk uitsluitend publicaties aan, geen projecten, personen of andere onderzoeksinformatie uit de HBO-wereld. In totaal gaat het per december 2011 om bijna 20.000 publicaties, waarvan 15.000 Open Access.

De HBO Kennisbank is daarmee de voornaamste 'outlet' voor deze informatie, aangezien een belangrijk deel van de aan de HBO Kennisbank toeleverende repositories zelf niet direct toegankelijk zijn via internet.

Ongeveer driekwart van het aantal publicaties in de HBO Kennisbank is een scriptie van een student (bachelor thesis; master thesis, student thesis, student report). De overige publicaties bestaan voor ruim 10% uit artikelen en andere bijdragen aan tijdschriften en voor het overige deel uit verschillende publicatievormen, zoals boeken, boekhoofdstukken, conferentiebijdragen en rapporten. De publicaties van de lectoren in de HBO Kennisbank worden tevens getoond op de lectorensite www.lectoren.nl.

De HBO Kennisbank geeft inzicht in het aantal publicaties per instelling en per jaar dat zich in de repository van die instelling bevindt, evenals van het OA aandeel daarin. Hoe zich dat verhoudt tot het totaal aantal publicaties van een instelling is niet direct na te gaan omdat de HBO-raad noch de instellingsjaarverslagen deze cijfers bevatten. Niettemin lijkt de dekkingsgraad ten opzichte van de totale output van de HBO's beperkt:

- Het aantal in de HBO Kennisbank opgenomen lectoren en docenten bedraagt in 2009, 2010 en 2011 respectievelijk 146, 66, 46 (lectoren) en 139, 98, 85 (docenten). Bij de betrokken hogescholen werken 438 lectoren en 2078 (docent)-onderzoekers per november 2011 (tabel 1 en 2 in bijlage B). Het lijkt er daarom op dat maar een beperkt deel van de onderzoekspublicaties in de HBO Kennisbank terecht komt.
- Van de studenten worden kennelijk alleen OA publicaties in de repositories opgenomen. De aantallen zijn blijven stijgen over de afgelopen jaren van bijdragen van net geen duizend studenten in 2007 tot bijna tweeduizend studenten in 2010. Overigens is dit aantal maar een fractie van het totaal aantal afgestudeerden aan de hogescholen: in 2009 ruim 67.000⁸. Hoewel niet alle afstuderende studenten een individuele scriptie schrijven (bij sommige studierichtingen wordt er een afstudeerverslag door een groep studenten gemaakt, bij andere studierichtingen is er in het geheel geen sprake van een afstudeerscriptie), wordt het uit deze getallen duidelijk dat slechts een zeer beperkt deel van de afstudeerscripties in de HBO Kennisbank terechtkomt.
- Alle repositories bieden de mogelijkheid om hun materialen Open Access ter beschikking te stellen, bij ruim de helft van de repositories is dat ook meer dan 80% van het materiaal. Bij 58% van de repositories is er ook de mogelijkheid om de materialen via Open Access na een embargo beschikbaar te stellen. Tenslotte hebben ruim 30% van de repositories ook de mogelijkheid om materialen via Campus Access ter beschikking te stellen.
 - Bij tien repositories is meer dan 80% van het materiaal Open Access beschikbaar.
 - Bij vier repositories is tussen de 48% en 70% van het materiaal Open Access beschikbaar.
 - Bij drie repositories is minder dan 40% van het materiaal Open Access beschikbaar.
 - Elf repositories bieden de mogelijkheid om materialen ter beschikking te stellen via Open Access na een embargo. Het percentage materiaal dat op die manier beschikbaar wordt gesteld varieert van 2% tot 20%.
 - Zes repositories bieden de mogelijkheid om materialen via campus Access ter beschikking te stellen. Het percentage materiaal dat op die manier beschikbaar wordt gesteld varieert van 1% tot 40%.

Een ruime meerderheid van de repositories hebben selectiecriteria met betrekking tot het op te nemen materiaal.

- 42% van de repositories blijkt selectiecriteria te hebben voor materialen van studenten. 21% van de repositories blijkt selectiecriteria te hebben voor zowel materialen van studenten als materialen van onderzoekers. 26% heeft geen selectiecriteria, terwijl 10% 'weet niet/geen beleid in deze' heeft aangevinkt.
- Wat betreft de selectiecriteria voor het materiaal van studenten blijkt het hier meestal om een grens gaan wat betreft de beoordeling ervan: de meeste respondenten noemen een cijfer van zeven of hoger.
- De selectiecriteria voor wat betreft materiaal van onderzoekers blijken meestal de verantwoordelijkheid van de onderzoekers zelf te benadrukken.

De overgrote meerderheid van de repositories hebben uitsluitend full text records. Slechts drie repositories hebben ook metadata records zonder full text.

- Uit de toelichtingen blijkt echter dat het beleid recentelijk gewijzigd is en dat sommige mogelijk ook metadata records gaan opnemen. Dit geldt ook voor het beleid voor wat betreft het doorsluizen naar de HBO Kennisbank

⁷ Gegevens uit www.hbo-raad.nl/hbo-raad/publicaties/cat_view/43-publicaties/138-2011

⁸ Idem.

5.3 Beheer en organisatie

- De inspanningen rond het beheer en organisatie van de aan de HBO Kennisbank toeleverende repositories kan geschat worden op in het totaal 8 tot 9 FTE. Dit is gebaseerd op de opgave die ontvangen werd van dertien hogescholen die in het totaal 5,5 FTE hiervoor opgaven. De gemiddelde inzet is circa 0,4 FTE.
 - Drie hogescholen zetten ongeveer één FTE in voor de repository.
 - Er zijn daarnaast ook vier hogescholen die een inzet van 0,25-0,5 FTE voor de repository hebben.
 - Vier hogescholen hebben een inzet van 0,2 FTE en twee hogescholen 0,1 FTE of minder.
- De verschillende taakgebieden nemen het volgende deel van de inspanning in beslag: 35% management, 44% administratief/bibliotheek technisch en 25% technisch/ICT.
- Bij een meerderheid van de repositories (58%) wordt de full text aangeleverd door de auteurs en verzorgen de medewerkers van de repository het deponeren.
 - Bij 10,5% van de repositories deponeren de auteurs zelf direct in de repository ('self-deposit').
 - Bij 15,8% van de repositories worden verschillende werkprocessen naast elkaar gebruikt.
- Bij ruim één derde van de repositories geldt een verplichting voor het deponeren van (een deel van) het materiaal. Bij bijna 60% geldt er geen verplichting vanuit de instelling.
 - Bij 31,6% van de repositories geldt een verplichting voor een deel van het materiaal: dit heeft meestal betrekking heeft op scripties van studenten. Bij een repository (5,3%) geldt een verplichting voor alle studenten en medewerkers.
- Overwegingen van duurzaamheid (energiegebruik, groene ICT) zouden volgens een meerderheid van de respondenten een rol moeten spelen bij de toekomstige opzet en inrichting van repositories.

5.4 Technische ontwikkelingen

- Er zijn vijf verschillende softwarepakketten in gebruik bij de HBO repositories. De meeste HBO repositories delen een gemeenschappelijke ICT infrastructuur: aan Sharekit nemen tien hogescholen deel, vijf hogescholen delen hun infrastructuur met de Universiteit van Tilburg, twee met twee andere universiteiten en twee met één gezamenlijk systeem.
- Bijna 80% van de repositories kennen een onderwerpsontsluiting toe aan de objecten in hun repository. Ruim 20% geeft aan geen onderwerpsontsluiting te hebben.
 - Verreweg de meeste repositories (78,9%) hebben vrijelijk toegevoegde trefwoorden in het Nederlands.
 - 47,4% van de repositories hebben (ook) vrijelijk toegevoegde trefwoorden in het Engels.
 - 36,1% van de repositories hebben een ontsluitingssysteem volgens een thesaurus/classificatiesysteem in het Nederlands; van deze heeft 5,3% ook dit ontsluitingssysteem in het Engels.
- MODS en DIDL zijn de metadata standaarden die bij een ruime meerderheid van de repositories in gebruik zijn.

Een aantal vragen in de enquête met betrekking tot de technische aspecten lijkt (nog) niet relevant voor de HBO-repositories. Dit lijkt de reden te zijn dat sommige vragen over technische aspecten door een groot deel van de respondenten met 'weet niet' zijn beantwoord. Hieronder een aantal resultaten uit de enquête waar dit voor geldt:

- Slechts 16% van de repositories zou een persistent identifier toekennen aan de objecten (meer dan 40% antwoordt met 'weet niet'). Bij de zeven hogescholen, die de repository infrastructuur delen met een universiteit, wordt echter wel een persistente identifier toegekend.
- Ruim 20% geeft aan dat de digitale duurzaamheid van de materialen in de repository zijn gewaarborgd. 16% stelt hiertoe concrete plannen te hebben. 37% zegt dat de digitale duurzaamheid niet is gegarandeerd, terwijl 26% 'weet niet' antwoordt. In de praktijk blijkt dat de digitale duurzaamheid door opname in het e-depot van de Koninklijke Bibliotheek nog niet gerealiseerd te zijn.
- Op de vraag of een digital author identifier (DAI) wordt toegekend, antwoordt één repository positief. De DAI hangt nauw samen met de implementatie van het onderzoeksregistratiesysteem voor de vastlegging van metadata. Wanneer dit systeem wordt geïmplementeerd, kan de DAI systematisch toegekend worden aan de medewerkers van de hogeschool.

5.5 Dienstverlening op basis van repositories

- De content van een minderheid van de repositories (42%) is ontsloten via de digitale bibliotheek van de instelling.
 - Bij 21,1% van de repositories bestaan daarvoor concrete plannen.
- Eén repository heeft een printing on demand service.
- Bij drie repositories (16%) worden er persoonlijke diensten aan de deponerende auteurs aangeboden.
 - Bij een repository (5,3%) bestaan hier toe concrete plannen.
 - Bij de overige repositories (78,9%) zijn er geen persoonlijke diensten aan de deponerende auteurs.
- Drie repositories geven aan dat er andere diensten in ontwikkeling zijn. Men noemt onder andere een digitale brievenbus.

5.6 Naar een toekomstscenario voor de HBO repositories

5.6.1 Doelstellingen

Wat is een mogelijk toekomstscenario voor de aan de HBO Kennisbank toeleverende repositories? Om deze vraag goed te kunnen beantwoorden, dient eerst naar de doelstellingen van de HBO repositories gekeken te worden.

Uit de enquête blijkt dat de repositories één duidelijke, algemeen onderschreven primaire doelstelling hebben: de etalagefunctie (Open Access). Deze etalagefunctie wordt door vrijwel alle respondenten onderschreven (95%).

In het verlengde daarvan wordt ook de doelstelling 'servicefunctie voor onderzoek en onderwijs' breed onderschreven (79%). Een archieffunctie voor de repository wordt minder ondersteund: hoewel 73% deze functie (enigszins) onderschrijft, geeft ruim 20% aan deze niet te

onderschrijven. Overigens blijkt dat deze uitkomsten vooral de visie van de respondenten weergegeven (53%), bij slechts 26% van de repositories/instellingen is dit ook vastgesteld beleid.

5.6.2 Realisatie van succesfactoren

Van acht erkende succesfactoren zijn er twee gerealiseerd.

In de enquête zijn acht - algemeen geaccepteerde – succesfactoren⁹ voor repositories voorgelegd aan de respondenten met de vraag (1) in hoeverre deze gerealiseerd zijn, (2) in hoeverre ze bijgedragen hebben tot het succes van de repository en (3) - indien niet gerealiseerd - welke bijdrage men verwacht. De resultaten staan vermeld in de tabel hieronder.

Vergelijking aspecten succesfactoren	Gerealiseerd	Draagt bij	Zal bijdragen
Voorlichtingscampagnes onder de studenten	57,9	60,0	63,3
Informatievoorziening over copyright in relatie met deponeren	52,7	38,5	45,5
Betrokkenheid van beslissers van uw instelling	47,4	46,2	80,0
Voorlichtingscampagnes onder de lectoren/onderzoekers/auteurs	42,1	56,3	66,6
Eenvoudig en gebruikersvriendelijk deponeren	31,6	50,0	53,3
Verplichting tot deponeren	31,6	45,5	69,3
Integratie en/of koppeling van uw repository met andere systemen	26,4	55,5	50,0
Long-term archivering	15,8	10,0	26,7

Twee van de acht succesfactoren blijken bij een meerderheid van de repositories zijn gerealiseerd: voorlichtingscampagnes onder de studenten en informatievoorziening over copyright in relatie met deponeren.

De overige zes succesfactoren zijn dus (nog) niet bij een meerderheid van de repositories gerealiseerd. Vijf van deze succesfactoren verdienen een hoge prioriteit om deze te realiseren:

- verplichting tot deponeren
- voorlichtingscampagnes onder de lectoren/onderzoeker/auteurs
- de betrokkenheid van beslissers van uw instelling
- eenvoudig en gebruikersvriendelijk deponeren
- integratie en/of koppelingen van het repository met andere systemen binnen de instelling.

Deze succesfactoren dragen bij aan het succes van de repository, zoals blijkt uit de antwoorden van de respondenten van die repositories, waar deze al gerealiseerd zijn.

Van één nog niet gerealiseerde 'standaard' succesfactor (long term archivering) kan op grond van de enquêteresultaten gesteld worden dat deze in de praktijk weinig lijkt bij te dragen aan het succes van de repository daar waar het gerealiseerd is. Tenslotte kan vanuit de gegevens worden opgemaakt dat de succesfactor 'betrokkenheid van beslissers van uw instelling' indien gerealiseerd wél een bijdrage levert aan het succes van de repository, maar veel minder dan verwacht wordt door degenen die dit nog niet hebben gerealiseerd.

5.6.3 Criteria voor het bereiken van de hoofddoelstellingen

Op bijeenkomsten van de BKH en van de werkgroep repository managers (op respectievelijk 17 en 18 januari 2012) werden de voornaamste resultaten uit de enquête besproken met het oog op de

⁹ In de enquête voor de aan NARCIS toeleverende repositories waren 9 succesfactoren voorgelegd: de succesfactor 'bekendheid van onderzoekers met vereisten financiers met betrekking tot deponeren' leek niet/nauwelijks relevant voor de HBO's en is daarom weggelaten.

ontwikkeling van een toekomstscenario voor de HBO repositories. Met betrekking tot de hoofddoelstelling van de repository 'etalagefunctie voor de output en de kennis van de HBO instelling' kwamen uit de discussie drie criteria voor het bereiken hiervan naar voren:

- **Dekkingsgraad output lectoren en onderzoekers:** de deelnemers aan de discussies stelden dat de dekkingsgraad van de output van de lectoren en onderzoekers aan de hogescholen een belangrijk criterium voor het succes van de repositories is. Veel hbo-instellingen zijn bezig met het opstellen van kwaliteitsindicatoren, de output aan publicaties van lectoren zou hiervan er een kunnen zijn. Meerdere deelnemers stelden daarom ook een verplichtstelling van het deponeren van publicaties van lectoren in het repository van de instelling voor. Opgemerkt werd dat de lectoren praktijkgericht onderzoek doen en als gevolg daarvan naast wetenschappelijke publicaties vooral vakgerichte publicaties produceren. Veel wetenschappelijke kwaliteitsindicatoren voor publicaties (peer review tijdschriften; citaties) zijn daarom minder van toepassing op de output van lectoren en onderzoekers aan hogescholen. Door het maken van goede afspraken binnen de instelling, die onderbouwd worden door efficiënte werkprocessen, zou het volgens meerdere deelnemers mogelijk moeten zijn om in 2015 een dekkingsgraad na te streven van 80-90%. Wel werd daarbij opgemerkt dat het lastig is om de totale output van lectoren en onderzoekers aan hbo-instelling vast te stellen, omdat de HBO-raad - in tegenstelling tot bijvoorbeeld de VSNU – niet dergelijke gegevens verzamelt. Eén deelnemer merkte daarbij op dat zij wél gebruik maken van de VSNU definities voor de verschillende soorten output en dat deze goed toepasbaar bleken te zijn voor de output van lectoren en onderzoekers aan hogescholen.
- **Dekkingsgraad studentenscripties:** Ook de dekkingsgraad van studentenscripties (bachelorscripties, afstudeerverslagen e.d.) werd als een criterium genoemd om het succes van de repositories af te meten. Echter, de deelnemers noemden veel belemmeringen om deze dekkingsgraad significant op te hogen:
 - Er is geen cultuur van publiceren en openbaar maken in de hogescholen, noch bij docenten noch bij studenten.
 - Er is bij vele instellingen een grote aarzeling om kwalitatief mindere studentenscripties openbaar te maken.
 - Er bestaan hoge drempels met betrekking tot openbaarmaking bij opdrachtgevers, die stageplekken aanbieden.
 - Het is in de praktijk lastig om de werkprocessen met betrekking tot het deponeren van afstudeerverslagen/bachelor scripties op te zetten en in stand houden.
 - Een 'opt-in' procedure, waarbij de student expliciet toestemming geeft voor openbaarmaking van zijn/haar afstudeerscriptie, wordt noodzakelijk geacht en maakt een eventuele verplichting tot deponeren moeilijk.

Veel hogescholen of onderdelen van deze instellingen leggen selectiecriteria aan en wensen uitsluitend scripties met een cijfer van (bijvoorbeeld) acht of hoger openbaar te maken.

- **Gebruik van de HBO Kennisbank:** het gebruik van de HBO Kennisbank wordt eveneens als een criterium voor succes van de repositories genoemd. Op dit moment trekt de HBO Kennisbank circa 40.000 bezoekers per maand. Naast kwantitatieve cijfers over het gebruik, wordt gesteld dat het ook zeer belangrijk is om in kwalitatieve zin het gebruik van de HBO Kennisbank te meten: er is weinig bekend over wie de gebruikers zijn en wat zij doen met de informatie die ze vinden. In hoeverre vindt **kennisvalorisatie** plaats door het gebruik van informatie uit de HBO Kennisbank? Ook het woord **talentbank** valt in dit verband: in hoeverre wordt de HBO Kennisbank gebruikt door het bedrijfsleven en maatschappelijke organisaties om talenten op te sporen onder de afgestudeerden? In dit verband wordt opgemerkt dat profilering van hbo-instellingen toenemende aandacht opeist en dat de informatie uit de HBO Kennisbank hiervoor als bouwsteen kan dienen (bijv. bij het presenteren van de lectoraten, bij het opzetten van kennispleinen en dergelijke).

5.6.4 Bouwstenen voor een toekomstscenario

- De aan de HBO Kennisbank toeleverende repositories kennen een primaire doelstelling: het zo wijd mogelijk beschikbaar stellen van de kennis-output van de instelling (etalage functie). In het verlengde daarvan ligt een andere belangrijke doelstelling: een servicefunctie voor het onderwijs en onderzoek binnen de instelling.
- Eenentwintig van de veertig Nederlandse hogescholen hebben een repository dat aan de HBO Kennisbank toelevert. Deze eenentwintig instellingen vertegenwoordigen meer dan 90% van de HBO studenten. De hbo-instellingen zonder repository zijn in het algemeen klein, en in de meeste gevallen PABO's of kunstonderwijsinstellingen, waarvoor een repository minder relevant is.
- In het totaal bevatte HBO Kennisbank per december 2011 bijna 20.000 publicaties, waarvan 15.000 Open Access.
- De dekkingsgraad van wat er van de jaarlijkse output van de deelnemende hogescholen in de repositories opgenomen wordt is niet precies vast te stellen door het ontbreken van overkoepelende, uniforme gegevens over deze output. Niettemin kan vastgesteld worden dat de dekkingsgraad op dit moment zeer beperkt is: slechts een klein deel van de bij de HBO's werkende lectoren en (docent)-onderzoekers is vertegenwoordigd in de HBO Kennisbank. Het jaarlijkse aantal opgenomen afstudeerscripties bedraagt circa 2000, het geen schril afsteekt bij de meer dan 67.000 afgestudeerden per jaar.
- De meeste repositories hanteren selectiecriteria voor opname: dit geldt met name voor afstudeerscripties, waar het criterium het toegekende cijfer is (meestal acht of hoger).
- De gemiddelde personele inzet ten behoeve van het repository bij de hogescholen bedraagt 0,4 FTE. Naast enkele hogescholen met de inzet van 1 FTE, zijn er ook hogescholen met een inzet van 0,1 of 0,2 FTE.
- Van de acht erkende 'standaard' succesfactoren zijn er twee bij een meerderheid van de repositories gerealiseerd. De realisatie van vijf - nog niet bij een meerderheid gerealiseerde - succesfactoren verdient hoge prioriteit. Drie zijn voornamelijk organisatorische aard: verplichting tot deponeren; voorlichtingscampagnes onder de lectoren/onderzoeker/auteurs; de betrokkenheid van beslissers van de instelling. Twee zijn deels van technische aard: eenvoudig en gebruikersvriendelijk deponeren; integratie en/of koppelingen van het repository met andere systemen binnen de instelling.
- Als succescriteria zijn voorgesteld:
 - Dekkingsgraad output lectoren en onderzoekers – in 2015 is het streven 80% van de output van datzelfde jaar.
 - Dekkingsgraad afstudeerscripties – een streefgetal voor 2015 kan op basis van de twee groepsdiscussies niet vastgesteld worden. Wel is gesteld dat er een cultuuromslag plaatsgevonden dient te hebben, waarin studenten en docenten het als een voordeel zien om de afstudeerscripties van goede kwaliteit openbaar te maken.
 - Gebruik van de HBO Kennisbank: zowel de kwantiteit van het gebruik (aantal downloads) van de HBO Kennisbank als het soort gebruik (In hoeverre wordt de HBO Kennisbank gebruikt als talentenbank? In hoeverre wordt de HBO Kennisbank gebruikt om kennis te valoriseren?) is als een succes criterium genoemd.

5.6.5 Voorstel voor een toekomstscenario

Dit voorgestelde toekomstscenario richt zich in eerste instantie op de 21 institutionele repositories van de grotere hogescholen, die momenteel al aan de HBO Kennisbank deelnemen. Dit toekomstscenario heeft als overkoepelende doelstelling dat de HBO Kennisbank een kwaliteitsvenster biedt van de kennisproducten van het HBO. Operationele doelstellingen zijn voor 2015:

- Vanaf 2015 zijn de publicaties voortkomende uit het praktijkgericht onderzoek door lectoren, onderzoekers en docenten aan de hogescholen openbaar toegankelijk, tenzij de auteurs motiveren waarom openbaarheid niet mogelijk is.
- Vanaf 2015 is tenminste 50% van de met een hoog cijfer beoordeelde studentenscripties openbaar.

Teneinde deze doelstellingen te bereiken, bevat het toekomstscenario de volgende vier thema's:

1. Output van lectoren, onderzoekers en docenten

- **Publicatielijst in jaarverslag:** Iedere instelling neemt vanaf 2013 in zijn jaarverslag een bibliografie op van de uit het praktijkgericht onderzoek voortkomende publicaties door lectoren, docenten en onderzoekers, op basis van gemeenschappelijke (door de HBO-Raad ontwikkelde) criteria. Daardoor wordt de output van het praktijkgericht onderzoek zichtbaar, vergelijkbaar en meetbaar. De gegevens over de publicaties voor het jaarverslag worden gegenereerd vanuit het onderzoeksregistratiesysteem. De HBO-Raad publiceert een totaal overzicht (à la de VSNU) van de onderzoeksproductie.
- **Opname in de repository:** De publicaties met bijbehorende full text worden opgenomen in het repository van de instelling en zoveel mogelijk Open Access beschikbaar gesteld via de HBO Kennisbank. Indien Open Access niet mogelijk is, wordt de full text van de betreffende publicaties toch opgenomen in het repository met Campus Access of No Access als toegangsopties, terwijl de auteurs de redenen aan dienen te geven waarom OA beschikbaarstelling niet mogelijk is. Systematische gegevens over deze redenen zullen leiden tot nieuwe inzichten in de belemmeringen voor Open Access beschikbaarstelling en tot een aanpak om deze belemmeringen weg te nemen.
- **Diensten voor deponerende lectoren/docenten/onderzoekers:** het bovenstaande gaat uit van een verplichting tot deponeren van de resultaten van praktijkgericht onderzoek voor lectoren, docenten en onderzoekers. Deze verplichting zal beter nageleefd worden wanneer de deponerende auteurs ook voordelen geboden worden in de vorm van aantrekkelijke diensten, zoals automatisch gegenereerde publicatielijsten, de mogelijkheid om het aantal downloads van hun publicaties te bekijken (via een SURE2 implementatie) en het aantal citaties (d.m.v. doorsluizen van de wetenschappelijke publicaties naar Google Scholar en Google Scholar Citations).
- **Stimuleren van Open Access:** voor optimale benutting van de uitkomsten van het praktijkgericht onderzoek door de samenleving (bijv. het bedrijfsleven, maatschappelijke organisaties e.d.) is Open Access beschikbaarstelling in vrijwel alle gevallen een voorwaarde. Daarom dienen de hbo-instellingen en de HBO-raad een actief beleid te voeren om de Open Access toegankelijkheid van de uitkomsten van praktijkgericht onderzoek te stimuleren door het bevorderen van het opzetten van Open Access tijdschriften voor professionals, zowel via nieuw op te zetten tijdschriften als via migratie van bestaande tijdschriften.

2. Afstudeerscripties

- **Alle afstudeerscripties archiveren:** het voorstel is om alle afstudeerscripties op te slaan in de repository van de instelling voor doeleinden van archivering ten behoeve van accreditatie en dergelijke.
- **Alleen afstudeerscripties met een 8 of hoger Open Access in de HBO Kennisbank:** het voorstel is om alleen afstudeerscripties met een 8 of hoger in aanmerking te laten komen voor Open Access beschikbaarstelling via de HBO Kennisbank. Toestemming van de student is hiervoor vereist, maar door het kwaliteitskeurmerk (alleen cijfer acht of hoger) zal dit voor de betreffende student een aantrekkelijke optie zijn. De toekenning van een persistente identifier aan de Open Access toegankelijke afstudeerscriptie in de HBO

Kennisbank geeft de student de mogelijkheid om de link naar de afstudeerscriptie op te nemen in zijn/haar CV en presentaties op het internet (Linked-In, Facebook et cetera). In het verlengde hiervan dient een handvat te worden ontwikkeld om eventuele belemmeringen van de zijde van opdrachtgevers voor de Open Access beschikbaarstelling van de afstudeerscriptie weg te nemen (zoals bedrijfsgevoelige gegevens in vertrouwelijke bijlage et cetera). De verwachting is dat deze maatregelen mede een cultuuromslag wat betreft publicatiegedrag zullen bewerkstelligen binnen de hogescholen.

3. Vraagkant

- **HBO Kennisbank en kennisvalorisatie:** Hoe kan de HBO Kennisbank beter worden ingezet bij de valorisatie van kennis? Welke doelgroepen gebruiken de HBO Kennisbank? Voor welke doeleinden gebruiken zij de informatie uit de HBO Kennisbank? Wat zijn hun wensen en behoeften ten aanzien van de HBO Kennisbank en hoe kunnen deze (beter) ingevuld worden? Zowel in kwantitatieve als in kwalitatieve zin is er weinig bekend over het gebruik en de gebruikers van de HBO Kennisbank. Voorgesteld wordt om de vraagkant door middel van een diepgaand gebruikersonderzoek beter in kaart te brengen. Een dergelijk onderzoek zou kunnen bestaan uit een kwantitatieve analyse van het gebruik van de HBO Kennisbank en een kwalitatief onderzoek onder gebruikers. Daarnaast wordt een permanente monitoring van de gebruiker statistieken, aangevuld door mogelijkheden waarin gebruikers feedback kunnen geven aanbevolen. Resultaten van deze onderzoeken zullen leiden tot een beter begrip van de rol van de HBO Kennisbank bij kennisvalorisatie en inzicht in hoe deze rol verder geïntensiveerd kan worden. Daarmee wordt de functie van de HBO Kennisbank zowel voor de hogescholen als voor de deponerende auteurs duidelijker. Overigens, het huidige gemiddelde van 40.000 bezoekers per maand laat geen ruimte voor twijfel aan het nut van de HBO Kennisbank.

4. Investerings

- Het in omloop brengen van de onderzoeksresultaten is een inherent deel van het onderzoek als zodanig. Een (bescheiden) deel van het onderzoeksbudget dient daarvoor te worden uitgetrokken. Om de bovenstaande voorstellen uit te voeren, is een gemiddelde personele inzet van 0,4 fte per instelling zonder meer te weinig. Daarnaast wordt ook personele inzet van de HBO-raad noodzakelijk geacht om de bovenstaande voorstellen mede vorm te geven en te implementeren.

Bijlage A - Aan NARCIS toeleverende repositories

Resultaten kwantitatieve analyse en enquête 2011; inclusief vergelijkingen resultaten enquête 2008.

De bijlage is aan te vragen via info@surf.nl.

Bijlage B - Aan de HBO Kennisbank toeleverende repositories

Resultaten van de kwantitatieve analyse en enquête 2011.

De bijlage is aan te vragen via info@surf.nl.
