

Nota toekomst Nederlandse infrastructuur voor auteursidentifiers



DAI werkgroep

Magchiel Bijsterbosch (SURF)

Chris Baars (KNAW-DANS)

Marnix van Berchum (KNAW-DANS)

Gert Jan Bokdam (RUG)

Petra Dickhaut (UvT)

John Doove (SURF)

Frits van Latum (TUDelft)

Karin Lodder (KB, voorzitter)

Martin van Muyen (OCLC)

Maarten Steenhuis (EUR)



Dit werk valt onder een [Creative Commons Naamsvermelding 4.0 Internationaal-licentie](https://creativecommons.org/licenses/by/4.0/).

1 OVER DIT DOCUMENT

De Digital Author Identifier (DAI) is het unieke nummer voor Nederlandse wetenschappelijke auteurs. Dit nummer is in 2005 als onderdeel van het SURF DARE programma ontwikkeld en sindsdien in gebruik bij alle Nederlandse universiteiten en een aantal onderzoeksinstituten.

Door de Nederlandse focus van dit systeem is de DAI volgens velen van beperkte waarde binnen het internationaal karakter van de wetenschap. Anno 2014 is er een aantal ontwikkelingen dat van invloed is op keuzes in de ontwikkeling van deze nationale infrastructuur.

Dit document beschrijft de achtergrond van auteurs identifiers, de opzet van de Nederlandse DAI infrastructuur en de ontwikkeling van twee internationale standaarden. Tot slot worden enkele relevante ontwikkelingen in de nationale context waarna er een aantal conclusies en aanbevelingen voor vervolg gedaan wordt.

Deze nota is tot stand gekomen met inbreng van de DAI werkgroep. De DAI werkgroep kende in 2014/2015 een wisselende samenstelling. In de werkgroep namen deel: Magchiel Bijsterbosch, John Doove (SURF), Chris Baars en Marnix van Berchum (KNAW/DANS), Gert Jan Bokdam (RUG), Petra Dickhaut (UvT), Frits van Latum (TU Delft), Karin Lodder (KB, voorzitter), Martin van Muyen (OCLC), Maarten Steenhuis (EUR).

2 MANAGEMENTSAMENVATTING

Ten behoeve van eenduidige registratie van onderzoeksresultaten op persoonsniveau wordt in Nederland de Digital Author Identifier (DAI) gebruikt. Een Digital Author Identifier komt tot stand door het koppelen van een persoon binnen het onderzoeksinformatiesysteem bij de universiteit aan een record in het Nederlandse Thesaurus van Auteursnamen (NTA).

In de afgelopen jaren hebben belangrijke internationale ontwikkelingen plaatsgevonden. Met name de International Standard Name Identifier (ISNI) en de Open Researchers and Contributor ID (ORCID) dienen zich aan als twee internationale alternatieven voor de DAI. De ontwikkeling van ISNI is afkomstig uit de Rights Management organisaties. ISNI positioneert zich als 'bridge identifier' en richt zich voornamelijk op het integreren van bestaande databases en authority files zoals de NTA. Hiermee staat het dus dicht bij de bibliotheekwereld daar deze sector van oorsprong dergelijke kwalitatief hoogwaardig bronbestanden heeft.

De ORCID is een identificatienummer gericht op wetenschappelijke auteurs. ORCID richt zich direct op deze auteurs op basis van directe zelfregistratie (zogenaamde author-claim). Daarbij wordt de mogelijkheid geboden een profiel bij te houden met een overzicht van werken. Het beheer van dit profiel kan worden gedelegeerd aan derden, bijvoorbeeld de werkgever. De ORCID organisatie biedt zelf geen kwaliteitsprocessen. Wel wordt via het systeem voor leden de functionaliteit geboden gegevens te valideren.

De systemen hebben vanuit hun eigen achterban al een behoorlijke adoptie status. De facto zullen beide identifiers dan ook in de infrastructuur voor gaan komen. DAI's worden inmiddels reeds opgenomen in de ISNI database en wetenschappers claimen al regelmatig een ORCID. Het feit dat er meerdere (overlappende) systemen bestaan is wellicht in de organisatie en communicatie niet ideaal, maar vanuit een informatiekundig perspectief niet onoverkomelijk.

Beide (non-profit) organisaties hanteren ook een ander verdienmodel. Het model van ISNI is gebaseerd op relatief lage lidmaatschapskosten en variabele kosten voor de registratie van identifiers terwijl bij ORCID het individuele registreren gratis is maar wordt verdiend door hogere lidmaatschapskosten voor integratie zonder aanvullende variabele kosten.

In dit document worden op basis van bovenstaande ontwikkelingen een aantal adviezen geformuleerd voor de toekomst van de Nederlandse infrastructuur voor auteursidentifiers.

1. *Kies ISNI als duurzame identifier voor Auteursnamen.*
 - De ISNI is sterk in duurzame toegang en kwaliteitsborging en kent een goede
2. *Gebruik ORCID als identifier binnen de wetenschappelijke communicatie.*
 - ORCID staat dicht bij de onderzoeker door steun van uitgevers en standaard integratie mogelijkheden in de onderzoeksinformatiesystemen en bibliografisch databases.
3. *Motiveer onderzoekers hun ORCID te claimen.*
 - ORCID hanteert het 'author-claim' principe dat de wetenschapper zelf eigenaarschap neemt van zijn of haar identifier. Tijdens het registratieproces is het van belang dat de onderzoeker zoveel mogelijk andere bekende identifiers koppelt aan het ORCID account zodat de koppeling met infrastructuren die (nog) gebruik maken van andere identifiersystemen zoals DAI en ISNI eenvoudiger wordt gemaakt.
4. *Registreer ORCID (en ISNI zie punt 6) in het onderzoeksinformatiesysteem.*
 - Bij voorkeur worden beide Identifiers in het onderzoeksregistratie systeem vastgelegd (zie 8.) zodat via het CRIS de benodigde matching tussen ISNI en ORCID zou kunnen plaatsvinden.
5. *Bij voldoende animo een nationale licentie laten onderhandelen voor ORCID.*
 - Een licentie biedt aanvullende mogelijkheden voor gedelegeerd beheer en automatische synchronisatie van gegevens tussen ORCID en onderzoeksinformatiesystemen. Een nationale licentie biedt de mogelijkheid voor kwantumkorting en is van strategische waarde bij het bepalen van de ontwikkeling van de dienst (zie ook 9.).
6. *Sluit onderzoeksinformatiesystemen aan op de ISNI infrastructuur.*
 - De CRIS leveranciers zijn actief in contact met ORCID. Als Nederlandse community kan druk worden gezet op de leveranciers om ook actief in contact te treden over koppeling met de ISNI infrastructuur.
7. *Richt een Nederlands Registration Agency voor ISNI op.*
 - Is door de KB al uitgevoerd. De KB kan voor Nederlandse auteurs ISNI's (laten) registreren.
8. *Universiteiten/UKB bibliotheken en KNAW/DANS-Narcis zouden ondersteund door OCLC een overgang van DAI naar ISNI moeten realiseren in 2015/2016.*
 - Geadviseerd wordt dit projectmatig in te steken, bijvoorbeeld onder regie van de werkgroep innovatie van de UKB ondersteund door OCLC.
9. *Een aantal universiteiten gaat samen met SURF in 2015/2016 in een pilotproject de mogelijkheden van ORCID voor versterking van de Nederlandse research infrastructuur onderzoeken.*
 - Geadviseerd wordt een pilot project onder coördinatie van SURF te starten waarin een aantal universiteiten een samenwerking (mogelijk met licentie) aangaat met ORCID. In de pilot kan een aantal relevante use cases worden uitgevoerd om de waarde van een ORCID licentie te kunnen inschatten en ervaring op te doen voor ORCID implementatie bij andere geïnteresseerde hoger onderwijsinstellingen.

3 IDENTIFICATIENUMMERS VOOR AUTEURS

Auteurs publiceren in hun werkzame leven één of meerdere bibliografische werken. Deze bronnen worden in bibliografische databases en/of repositories beschreven. Sinds de komst van het internet worden dergelijke bronnen steeds meer met elkaar verbonden. Hiermee wordt het creëren van complete overzichten van de productiviteit van personen steeds eenvoudiger. Deze overzichten kunnen gebruikt worden voor het vergroten van de zichtbaarheid van een auteur of zijn of haar affiliatie(s) en voor uiteenlopende evaluatiedoeleinden.

Wanneer lijsten voor dergelijke doeleinden gebruikt gaan worden, is het bij samenstelling van deze lijsten van belang dat zij bibliografisch volledig correct zijn. Bij het samenstellen en koppelen van lijsten van bibliografische werken aan personen bestaat het risico van persoonsverwisseling en persoonsdubbeling.

In het eerste geval gaat het om twee of meer natuurlijke personen met dezelfde naamsvariant waarbij werken abusievelijk aan de verkeerde natuurlijke persoon worden toegewezen. Dit komt vooral voor bij veelvoorkomende namen, zoals het bekende 'J. Jansen', 'H. Wang' of 'W. Li'. Op publicatielijsten van deze personen zullen vervolgens werken vermeld staan die niet van hun hand zijn of werken ontbreken omdat zij aan een ander worden toegeschreven.

Er is sprake van persoonsdubbeling wanneer één natuurlijke persoon onder meerdere naamsvarianten voorkomt en zijn of haar werken als van twee verschillende personen wordt herkend. Dit is het geval bij het publiceren onder het pseudoniem, maar kan ook abusievelijk ontstaan door naamsveranderingen als gevolg van een huwelijk, gebruik van verschillende initialen of roepnamen. Het gevolg is dat publicatielijsten van een persoon incompleet zijn. Het is de abusievelijke kwestie waar de auteursidentificatie zich op richt.

Bovendien garandeert het gebruik van een auteursidentificatie de uitwisselbaarheid van persoonsgerelateerde informatie tussen verschillende digitale systemen.

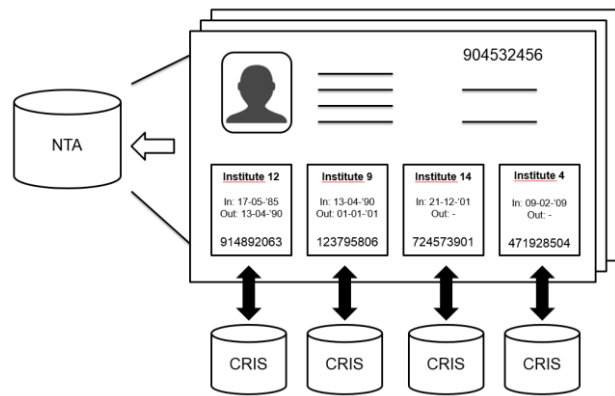
4 DIGITAL AUTHOR IDENTIFIER

De Digital Author Identifier (DAI) infrastructuur is het resultaat van het Digital Author Identification project dat eind 2005 en gedurende 2006 onder het DARE programma is uitgevoerd door de Rijksuniversiteit Groningen, Radboud Universiteit Nijmegen en OCLC. Deze infrastructuur moest auteurs binnen de DAREnet repository-infrastructuur voor Open Access wetenschappelijke publicaties – het huidige NARCIS – uniek identificeerbaar maken.

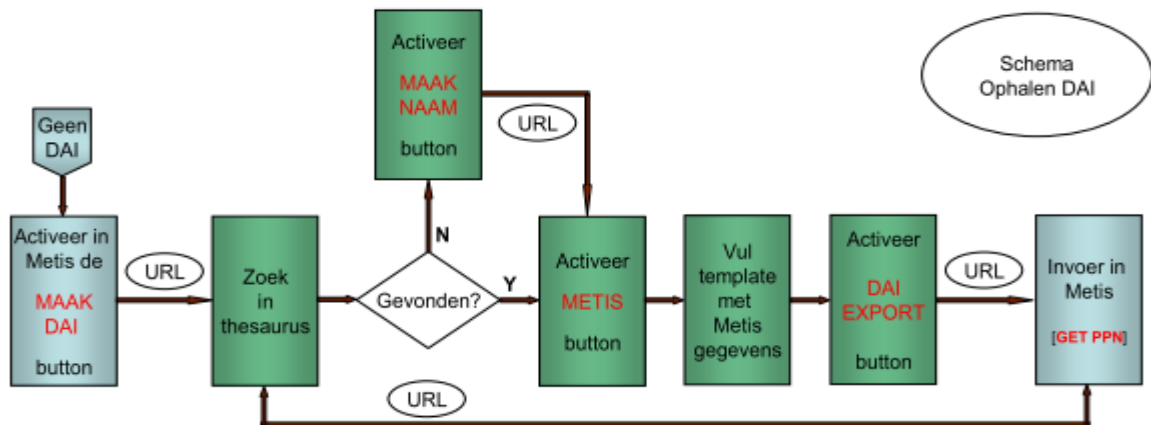
4.1 WERKPROCES

Een Digital Author Identifier komt tot stand door het koppelen van een persoon binnen het onderzoeksinformatiesysteem bij de universiteit aan een record in het Nederlandse Thesaurus van Auteursnamen (NTA). De NTA is onderdeel van de Gemeenschappelijk Geautomatiseerde Catalogus (GGC) en bevat alle publicerende Nederlandse auteurs. Het gaat in de NTA dus niet uitsluitend om wetenschappelijke auteurs. Op het moment van schrijven is alleen het METIS pakket gekoppeld, dat tot voor kort bij alle Nederlandse universiteiten in gebruik was als onderzoeksinformatiesysteem.

Het koppelen van een onderzoeker binnen METIS en een persoon in de NTA wordt handmatig gedaan door de METIS-beheerder. Op basis van enkele kenmerken van de persoon zoals bekend in METIS wordt een zoekopdracht gedaan in de NTA. Wanneer de beheerder een match vindt worden de kenmerken uit METIS in een zogenaamd 'onderzoeksblok' toegevoegd aan het persoonsrecord in de NTA (zie fig. 1). Als er geen match wordt gevonden, bijvoorbeeld wanneer het gaat om een promovendus welke nog niet eerder heeft gepubliceerd, kan de beheerder een nieuw persoonsrecord in de NTA aan maken. Het PICA Productie Nummer (PPN), de primaire sleutel in de NTA, wordt op dat moment 'gepromoveerd' tot DAI en teruggestuurd naar METIS. Vanuit METIS wordt het nummer verder doorgegeven aan de institutionele repository waar vanuit het nummer naar NARCIS wordt geharvest als onderdeel van de bibliografische metadata van publicaties.



Figuur 1: Visualisatie van NTA record met onderzoeksblokken van verschillende instellingen.



Figuur 2: Werkproces van het koppelen van een onderzoeker in METIS met een persoonsrecord in de NTA.

Hoewel het koppelen van personen een arbeidsintensief en foutgevoelig proces is, opent dit – eenmaal op orde – de deur voor legio toepassingen binnen bijvoorbeeld NARCIS zoals volledige publicatielijsten uit verschillende bronnen of (download)statistieken op het aggregatieniveau van auteurs. Maar bijvoorbeeld ook binnen de informatievoorziening rondom de toekenning van en verantwoording over subsidies bij financiers van wetenschappelijk onderzoek maakt een uniek nummer het mogelijk om op persoonsniveau in- en output meting te verrichten.

4.2 BESTURING EN FINANCIERING

De DAI infrastructuur wordt met gesloten beurzen onderhouden. Het strategisch eigenaarschap van de DAI was belegd bij SURF i.c. het platformbestuur ICT & Onderzoek welke hierbij werd geadviseerd door de Beleidsgroep Innovatie Kennisinfrastructuur (BIK). In de BIK namen namens de Colleges van Bestuur afgevaardigden uit bibliotheek en ICT diensten van universiteiten en KNAW deel.

Met de reorganisatie bij SURF is het platform ICT en Onderzoek en de daarmee het platformbestuur van het platform ICT en Onderzoek opgeheven. Ook de BIK bestaat niet langer als adviserend gremium. Auteursidentifiers zullen als thema vanaf 2015 binnen SURF in het innovatieprogramma

Open Access, dat getrokken wordt bij SURFmarket terugkomen. Er zijn contacten geweest met de Werkgroep Innovatie van de UKB om te bekijken hoe aan dit thema het beste invulling gegeven kan worden door de verschillende stakeholders.

De verschillende componenten van de infrastructuur hebben echter hun eigen besturing en beheerconstructies. OCLC is verantwoordelijk voor het technisch beheer van de DAI en stelt binnen de overeenkomst met het consortium Gemeenschappelijke Informatie Infrastructuur (GII) de NTA als onderdeel van de GGC beschikbaar. Er is een gereduceerd tarief waarbij organisaties die niet binnen deze overeenkomst vallen, zoals onderzoeksinstituten en hogescholen gebruik kunnen maken van de NTA webinterface.

Het functioneel beheer op de inhoud van de NTA wordt uitgevoerd door de Koninklijke Bibliotheek. Daar wordt ook een helpdesk bemand waar functionele en technische vragen rondom DAI kunnen worden aangemeld en afgehandeld, dan wel doorgezet naar de daarvoor aangewezen organisaties.

De individuele universiteiten zijn verantwoordelijk voor de onderzoeksinformatiesystemen en institutionele repositories. Het beheer van de NARCIS infrastructuur ligt bij DANS welke wordt geadviseerd door de NARCIS-adviesraad.

De coördinatie van de DAI infrastructuur vindt plaats in het DAI werkgroepoverleg. Deze werkgroep bestaat uit vertegenwoordiging van de reeds genoemde organisaties OCLC, KB, DANS en SURF alsmede vertegenwoordiging van METIS beheerders en deskundigen vanuit de RUG, UvT, TUDelft en EUR.

Ondernemingen om de besturing en financiering van deze infrastructuur op een centraal niveau te borgen hebben vooralsnog niets opgeleverd. Voor de hand liggende partijen zouden zijn bijvoorbeeld UKB of GII zijn. In het verleden heeft de UKB te kennen geven de voorkeur te hebben voor borging binnen het GII gezien haar rol in de besturing van de GGC – waar de NTA onderdeel van uitmaakt. Binnen het GII sluit dit bij de openbare bibliotheken op bezwaren daar de DAI gericht is op wetenschappelijke auteurs en in haar optiek een zaak is voor UKB.

Inmiddels zijn de UKB bibliotheken samen met OCLC een traject ingegaan, waarin de functie van de GGC als nationale infrastructuur voor gemeenschappelijk catalogiseren zal worden overgenomen door het nieuwe *WorldShare platform* van OCLC.

Vrijwel alle universiteiten zijn begonnen aan de vervanging van hun METIS systeem door een nieuw CRIS, of hebben dit traject inmiddels afgerond. Hiermee komt op afzienbare termijn een einde aan de bestaande technische infrastructuur die tot nu toe de basis vormde voor het onderhoud van de DAI.

5 INTERNATIONALE STANDAARDEN

Een terugkerend punt van kritiek op de Nederlandse DAI infrastructuur in de afgelopen jaren was dat het hier een nationale infrastructuur betreft terwijl wetenschappelijk onderzoek een internationale aangelegenheid is. Een tweede punt van kritiek is dat de DAI infrastructuur ver weg staat van de dagelijkse praktijk van de onderzoeker.

In de afgelopen jaren hebben er drie belangrijke internationale ontwikkelingen plaatsgevonden: de Virtual International Authority File (VIAF), de International Standard Name Identifier (ISNI) en de Open Researchers and Contributor ID (ORCID).

Hieronder worden de belangrijkste kenmerken van deze identificersystemen beschreven. Samen kunnen deze systemen de basis vormen waarop een nieuwe nationale infrastructuur voor auteursidentifiers kan worden gebouwd.

Bijlage A bevat een vergelijkingsmatrix voor de systemen en een uitsplitsing van de kosten die deelname/gebruik van deze systemen voor belanghebbende organisaties met zich meebrengen.

5.1 VIRTUAL INTERNATIONAL AUTHORITY FILE

De VIAF is een gezamenlijk project van zeven nationale bibliotheken waarin verschillende nationale bronbestanden worden geïntegreerd is één virtueel bronbestand (<http://www.oclc.org/viaf.en.html>)

5.2 INTERNATIONAL STANDARD NAME IDENTIFIER (ISO 27729)

De International Standard Name Identifier is een ISO standaard voor de identificatie van bijdragers aan creatieve werken en de verspreiding daarvan. Als voorbeelden worden genoemd wetenschappers, schrijvers, componisten, musici en kunstenaars, maar ook producenten en uitgevers.

De ontwikkeling van ISNI is afkomstig uit de Rights Management organisaties.

ISNI positioneert zich als 'bridge identifier' en richt zich voornamelijk op het integreren van bestaande databases en authority files zoals de NTA. Hiermee staat het dus dicht bij de bibliotheekwereld daar deze sector van oorsprong dergelijke kwalitatief hoogwaardig bronbestanden heeft. Het ISNI proces voorziet in een centrale kwaliteitscontrole die is belegd bij de Bibliothèque National de France en de British Library.

De ISNI database is gelanceerd in 2011 en bevat (dd. 20-02-2015) 8,6 miljoen records van personen met een ISNI waarvan 2,25 miljoen wetenschappers¹. De database draait op hetzelfde platform als de GGC: het door OCLC ontwikkelde CBS. De database is benaderbaar via een webinterface voor eindgebruikers en op open standaarden gebaseerde API voor integratiedoeleinden (i.c. SRU en AtomPub).

Het verdienmodel van ISNI is gebaseerd op lidmaatschapsbijdragen en additionele kosten voor de registraties van identifiers. Het ISNI lidmaatschapsmodel maakt onderscheid tussen regulier lidmaatschap en Registration Agencies (RA).

Het reguliere lidmaatschap kost €800 p.a. Met het lidmaatschap verkrijgt men het gebruiksrecht op de toegang tot de volledige database exclusief vertrouwelijke onderdelen, additionele indices, de mogelijkheid bestanden te verrijken en te linken. Organisatorisch hebben de gezamenlijke leden het recht één bestuurslid te nomineren om hen te vertegenwoordigen.

Ieder lid heeft recht op één batch load. De kosten voor een dergelijk batch load zijn onderverdeeld in staffels op basis van het aantal records. De hiermee gemoeide kosten lopen uiteen van K€5 tot K€10. Het verder bijhouden verloopt vervolgens via een RA. De hiermee gemoeide kosten zullen per RA verschillen.

¹ Bron: <http://www.isni.org/> (dd. 20-02-2015)

Registration agencies hebben naast het reguliere lidmaatschap de mogelijkheid om namens anderen records aan te maken. De vaste kosten voor een RA bestaan uit eenmalige kosten variërend van K€2,5 tot K€7,5 en verder een jaarlijks bedrag van K€1 of K€2, afhankelijk van de jaarlijkse omzet van de betreffende organisatie. Anders de reguliere leden hebben RA's het recht op ongelimiteerde batch loads. De kosten zijn gelijk met die van reguliere leden. Daarnaast kan een RA ook per identifier aanvragen. De kosten per identifier variëren tussen de €0,10 en €0,50.

5.3 OPEN RESEARCH AND CONTRIBUTOR ID

De Open Researcher and Contributor Identifier (ORCID) is een identificatienummer gericht op wetenschappelijke auteurs. Het nummer heeft hetzelfde format als ISNI en ISNI heeft een blok gereserveerd in hun range die zij niet uitdelen waardoor het niet mogelijk is dat ISNI en ORCID hetzelfde nummer uitdelen aan andere personen. Merk op dat het wel nog steeds mogelijk blijft dat één persoon zowel een ORCID als een ISNI heeft.

ORCID richt zich direct op de onderzoeker op basis van directe zelfregistratie (zogenaamde author-claim). Daarbij wordt de mogelijkheid geboden een profiel bij te houden met een overzicht van werken. Het beheer van dit profiel kan worden gedelegeerd aan derden, bijvoorbeeld de werkgever. De ORCID organisatie biedt zelf geen kwaliteitsprocessen. Wel wordt via het systeem voor leden de functionaliteit geboden gegevens te valideren.

De ORCID database is gelanceerd op 16 oktober 2012 en telt 1.172.865 (dd. 20-02-2015)² records. De onderliggende software is afkomstig van Thomson Reuters. De toegang wordt geboden via een webinterface voor eindgebruikers en RESTful API voor doorzoeken, aanmaken en bijhouden van records. ORCID biedt twee versies aan: een publieke en één specifiek voor leden. Via deze leden-API is het mogelijk om niet alleen publieke gegevens, maar gegevens die staan gemarkeerd als 'limited access' in te zien. Verder is het via de leden-API mogelijk records aan te maken en bij te werken. Tot slot wordt er jaarlijks een publiek databestand onder CC-zero licentie gepubliceerd. Voor leden is dat twee keer per jaar.

Het registreren en beheren van een ORCID-profiel is gratis voor individuele gebruikers. Het verdienmodel bestaat uit een basislidmaatschap van K\$5 voor organisaties rechten geeft op de besturing van ORCID, bredere toegang tot de data en verbeterde integratiemogelijkheden via de leden-API. Voor groepen groter dan vijf instellingen geldt een korting van 10%.

Naast het basislidmaatschap is er een premiumlidmaatschap van K\$10 of K\$25, afhankelijk van de jaarlijkse omzet of ontvangen onderzoeksubsidies. Deze vorm biedt meer integratiemogelijkheden dan het basislidmaatschap zoals geautomatiseerde synchronisatie met lokale systemen, een groter aantal verzoeken per dag via de API en meerdere gebruikers voor de API. Daarnaast krijgen premium leden maandelijks een publiek databestand en statistieken van de integratiewebsite.

Lidmaatschappen kunnen op individueel organisatieniveau, via consortia of nationaal worden afgesloten. Voor een consortium gelden sinds december 2014 de volgende prijsstellingen:

5-9 members	US\$ 6.000 per member
10-19 members	US\$ 5.000 per member
20-29 members	US\$ 4.000 per member

² <http://orcid.org/> (dd. 20-02-2015)

30-99 members US\$ 135.000 per consortium

100-250 members US\$ 200.000 per consortium

Daarnaast bestaan er kortingen voor non-profit organisaties en kleine organisaties. Deze kortingen kunnen oplopen tot 75% van het lidmaatschapsbedrag.

Contact met ORCID heeft uitgewezen dat de prijsstellingen verder non-negotiable zijn maar dat ze in de dienstverlening (support bij implementatie) tegemoetkomingen kunnen doen.

5.4 ADOPTIE

Het ISNI consortium kende op 20-02-2015 164 leden. Omdat ISNI bestaande databases aggregereert en registraties via Registration Agencies verlopen, staat ISNI als entiteit op afstand van de eindgebruiker. Adoptie van ISNI is daarmee in belangrijke mate afhankelijk van de mate van adoptie van de delen van de som. Deze delen zijn op dit moment afkomstig uit de bibliotheekwereld.

Op hetzelfde moment telde ORCID volgens haar website 196 leden in verschillende stadia van adoptie.³ ORCID zoekt 'outreach meetings' op verschillende plekken in de wereld direct contact met lokale wetenschappelijke communities. ORCID onderhoud nauwe banden met de grote uitgevers, niet verbazingwekkend gezien de inbreng van de Thomson-Reuters software als fundament voor de ORCID infrastructuur. Dit betekent ook dat ORCID snel in diensten van uitgevers wordt opgenomen het geen bijdraagt aan de bekendheid en het vertrouwen dat wetenschappers hebben in ORCID als 'canonieke identifieer' voor de wetenschap. Daarnaast richt ORCID zich op de ontwikkelaars van onderzoeksinformatiesystemen om standaard integratie van deze pakketten met de ORCID infrastructuur aan te bieden.

In het Verenigd Koninkrijk is in een Jisc rapport uit juli 2012 ORCID voorgesteld als de preferente identifieer voor wetenschappers.⁴ Voornaamste argument hierbij is het 'bottom-up' karakter waardoor ORCID dicht bij de wetenschapper staat.

In Denemarken is per augustus een consortium agreement aangegaan voor een (basic) lidmaatschap (lead organisatie: Deense Technische Universiteit), waarna per 1 september ORCID implementatie projecten zijn gestart⁵.

6 NATIONALE ONTWIKKELINGEN

6.1 ONDERZOEKSINFORMATIESYSTEMEN BIJ UNIVERSITEITEN

Het Nederlandse veld van onderzoeksinformatie is drastisch in beweging. Na een periode van ruim 20 jaar waarin uitsluitend het METIS in gebruik was als onderzoeksinformatiesysteem bij de Nederlandse universiteiten, zijn op het moment van schrijven vrijwel alle universiteiten bezig met het implementeren van een systeem ter vervanging van METIS.

Daar de koppeling tussen een auteur in de NTA en een onderzoeker in onderzoeksinformatiesysteem bij de universiteit de basis is voor de uitgifte van een DAI, heeft deze

³ <http://orcid.org/organizations/integrators/integration-chart>

⁴ http://www.jisc.ac.uk/whatwedo/programmes/di_researchmanagement/researchinformation/orcid.aspx

⁵ <http://orcid.dk/>

ontwikkeling een majeure impact op de organisatie en de infrastructuur. Er zal in de toekomst ten minste rekening gehouden moet worden met de pakketten PURE (Elsevier) en CONVERIS (Thomson Reuters). Ontwikkelingen zullen moeten worden afgestemd met meerdere pakketleveranciers het geen nieuwe eisen stelt aan de regie op de infrastructuur.

Naast dat deze pakketten geschikt moeten zijn voor een dergelijke koppeling, zal ook de NTA geschikt moeten worden gemaakt voor koppelingen met andere pakketten dan METIS. De huidige koppeling laat daarvoor het gebied van architectuur en stabiliteit te wensen over. Een ontwerp voor een nieuwe op webservices gebaseerde koppeling is al beschikbaar, maar niet geïmplementeerd voor de systemen Pure en Converis. Het is niet waarschijnlijk dat deze koppeling alsnog zal worden geïmplementeerd.

6.2 TOEKOMST NATIONALE INFRASTRUCTUUR VOOR CATALOGISEREN

Rondom het catalogiseren zijn er ontwikkelingen in de nationale infrastructuur voor de UKB. Het scenario is dat er direct gecatalogiseerd gaat worden op het nieuwe Worldshare platform van OCLC. In principe komt hiermee de GGC – waar de NTA onderdeel van is – voor de UKB te vervallen en rust nog slechts op gebruik door de Openbare Bibliotheken.

Tegelijkertijd blijven door de architectuur van het WorldShare platform lokale authority files nog een belangrijke rol vervullen; met andere woorden de functionaliteit van bijvoorbeeld een NTA als nationale authority file zal mogelijk blijven bestaan.

Het tijdspad van de migratie van de universitaire bibliotheken naar WorldShare vertaalt zich in een horizon van eind 2015 voor het uitfasen van een nationale auteursidentificer. Dat betekent dat voor deze tijd een keuze moet zijn gemaakt voor en gemigreerd naar een alternatieve internationale auteursidentificer.

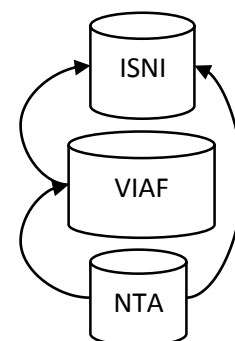
6.3 AANSLUITING BIJ INTERNATIONALE ONTWIKKELINGEN

Zowel met ISNI als ORCID is gesproken over mogelijke integratie met of migratie van de DAI infrastructuur.

Tijdens contacten met ORCID in 2012 is gesproken over een zoekinterface bij ORCID die op het moment van registratie de eindgebruiker haar DAI laat koppelen aan het ORCID profiel. Hier is verder geen opvolging aan gegeven.

In samenwerking met OCLC is gewerkt aan de koppeling van DAI aan ISNI. Hiervoor worden twee routes gebruikt. Ten eerste is de gehele NTA geïntegreerd in de VIAF. De VIAF bevat alleen de publieke metadata. Via de VIAF komen zo alle NTA records in de ISNI database en – wanneer zij met voldoen aan de criteria – krijgen zij een identifier toegewezen. Daarnaast is de subset van DAI records – dus NTA records met een affiliatie – geëxporteerd en direct geladen in de ISNI database. Deze export bevat ook vertrouwelijke gegevens waarmee de betrouwbaarheid van de matching wordt verbeterd.

Het resultaat van deze aanpak in 2014 is een totaal van 72.771 DAI records in de ISNI database zijn, waarvan 54.673 toegewezen en 18.098 met een voorlopige. De opname van bestaande en nieuwe DAI records in de ISNI database wordt nu door OCLC op reguliere basis gedaan. Deze werkwijze moet garanderen, dat voor elke DAI identifiers nu of straks een geldige ISNI identifier beschikbaar komt.



Figuur 3: Inlezen van NTA in ISNI via VIAF.

Het verder terug leveren van ISNI's aan onderzoeksinformatiesystemen en NARCIS kan daarom de nieuwe basis gaan vormen voor de nationale infrastructuur. Wel zijn er punten van aandacht. Registratie van de DAI of ISNI is nog niet mogelijk in alle versies/implementaties van Pure of Converis. Een tweede punt van aandacht is de vertraging die door de tweemaandelijks (eenmaal voor VIAF en eenmaal voor ISNI) batchlading ontstaat in de periode tot aan de uitfasering van de DAI. Hierdoor ontstaat een ingewikkeld asynchroon proces van het koppelen van een onderzoeker in een onderzoeksinformatiesysteem aan de uitgifte van een ISNI welke kan leiden tot onduidelijkheid.

7 CONCLUSIES

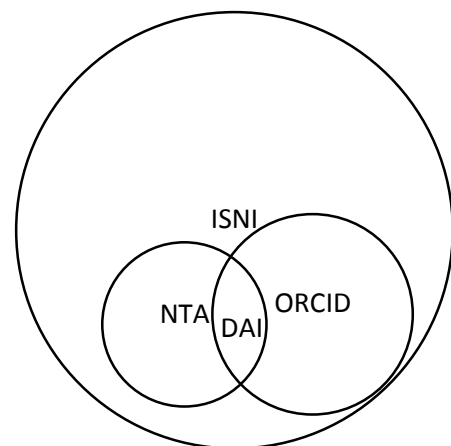
De ontwikkeling van een internationale auteursidentificer biedt grote voordelen op het gebied van adoptie en de mogelijkheden voor de ontwikkeling van diensten voor wetenschappers op basis van deze identificers. Het feit dat er meerdere (overlappende) systemen bestaan is vanuit een informatiekundig perspectief niet onoverkomelijk, maar zowel in de organisatie en communicatie niet ideaal.

7.1 VERHOUDINGEN TUSSEN IDENTIFIERSYSTEMEN

Over de verhouding tussen ISNI en ORCID is geen consensus. In hun communicatie positioneren zij zich als 'complementaire systemen die op verschillende wijze een gezamenlijk doel nastreven'. De integratie van beide systemen wordt echter afgewenteld op de gebruikers en afnemers.

Zij gebruiken dezelfde structuur voor de identificers die zij uitdelen en ISNI heeft een blok identificers gereserveerd voor exclusief gebruik door ORCID. Daarnaast heeft ORCID in haar registratieproces opgenomen dat er in de ISNI database gezocht wordt. Het is echter onduidelijk wat de procedure is wanneer een persoon zowel een ORCID als een ISNI nummer heeft. En hoe een eventuele koppeling tussen beiden tot stand kan komen. Verder beperkt ORCID zich tot wetenschappers terwijl ISNI zich ook nadrukkelijk richt op cultureel erfgoed en andere actoren die betrokken zijn bij de productie van creatieve werken. Tot slot is een belangrijk verschil tussen deze twee systemen dat ORCID zich baseert op zelfregistratie of author-claim, terwijl ISNI uitgaat van bestaande databases en registration agencies met achterliggende kwaliteitsprocessen. Een belangrijke consequentie van deze zelfregistratie is overigens dat ORCID niet toepasbaar is voor niet inactieve of overleden auteurs.

Zo beschouwd zou de conclusie getrokken kunnen worden dat ORCID niet meer dan een zelfservice registration agency is voor ISNI met meer autonomie maar minder hoog niveau in vertrouwen in het kwaliteitsniveau, waardoor ORCID voor een kwalitatief hoogwaardige infrastructuur in beginsel minder geschikt lijkt. Tegelijkertijd heeft ORCID een sterke band met de uitgevers en gepaard met een goede marketing en communicatie daarmee de aandacht van de wetenschappelijke community.



Figuur 4: Overlap van de verschillende identificersystemen c.q. thesauri.

7.2 GESCHIKTHEID VOOR TOEPASSING IN DE NEDERLANDSE CONTEXT

Wanneer wordt gekeken naar de verhoudingen in het dekkingsgebied tussen ISNI en ORCID en de huidige Nederlandse situatie (zie fig. 4) wordt duidelijk dat voor het toepassingsgebied van DAI – het identificeren van wetenschappelijke auteurs over organisaties heen – ORCID toereikend is.

De huidige structuur van besturing, financiering en beheer kan echter bestaan omdat de DAI voor een belangrijk deel ‘meelift’ op een infrastructuur die een bredere toepassing kent, zoals het beheer van de NTA door OCLC en de KB als onderdeel van de GGC. Een beperking tot wetenschappelijke auteurs maakt de opdracht te smal om een separate activiteit te rechtvaardigen zonder dat hier een explicitering van kosten tegenover staat. Daarbij komt de ontwikkeling richting WorldCat waarbij er gezocht moet worden naar oplossingen welke ook geschikt is voor niet-wetenschappelijke auteurs. Zo beschouwd past een keuze voor ISNI beter in het model waarbij er een brede, internationale en kwalitatief hoogwaardig thesaurus voor auteurs wordt nagestreefd.

De keuze voor een systeem wordt hiermee ook een bestuurlijke en financiële kwestie. Het tegengestelde verdienmodel van beide oplossingen – waarbij het model van ISNI is gebaseerd op relatief lage lidmaatschapskosten en variabele kosten voor de registratie van identifiers terwijl bij ORCID het individuele registreren gratis is maar wordt verdiend door hogere lidmaatschapskosten voor integratie zonder aanvullende variabele kosten – speelt hier een belangrijke rol. Beide organisaties zijn overigens non-profit organisaties.

7.3 SCENARIO'S

Overwegende het bovenstaande wordt er voor elk van de twee oplossingen een scenario voorgesteld. Belangrijk om hierbij op te merken is dat deze scenario's elkaar niet uitsluiten. Tegelijkertijd lijkt het vanuit kosten oogpunt niet opportuun om beide scenario's integraal te implementeren. In het derde scenario wordt een meer synergetische aanpak voorgesteld waarbij de verschillende systemen elk voor hun verschillende doelen worden geïmplementeerd.

7.3.1 Scenario 1: Nationale ORCID overeenkomst

Het eerste scenario gaat uit van een oplossing op basis van ORCID. In dit scenario wordt in opdracht van de Nederlandse universiteiten door SURFmarket een nationale licentie onderhandeld. In een implementatieproject worden voor alle wetenschappers met een DAI ORCID records aangemaakt en gekoppeld met hun ISNI die is ontstaan uit het proof-of-concept project via de VIAF.

Wetenschappers worden door hun universiteit geïnformeerd en het beheer van de profielen wordt door hen gedelegeerd aan hun eigen organisatie. Wijzigingen in onderzoeksinformatiesystemen worden via een standaardkoppeling automatisch gesynchroniseerd met de ORCID database.

Voor historische werken en andere domeinen zoals cultureel erfgoed wordt daarnaast een aparte oplossing op basis van ISNI ontwikkeld.

Voor de realisatie van deze oplossing is een zogeheten basic ORCID membership van de deelnemende instellingen nodig (US\$5000 per jaar) met een groep van groter dan 5 zou uitkomen op 4500 dollar per jaar per instelling.

Een premium; consortium membership (alleen voor not for profit instellingen) is in december 2014 geïntroduceerd en stelt een prijs van 5000 dollar per instelling per jaar voor een groep tussen de 10-19 deelnemers. Dit premium lidmaatschap maakt het ook mogelijk om ook de registratie van publicaties te synchroniseren tussen de CRIS systemen en de ORCID databases.

Hierbij zal nog moeten worden gekeken hoe andere instellingen (hogescholen en bijvoorbeeld het CWI, die ook DAI's registreren) mee kunnen in deze ontwikkelingen. Eerste contact met ORCID heeft uitgewezen dat de Nederlandse hogescholen samen als één deelnemer mee zouden mogen gaan.

Het voordeel van het scenario op basis van ORCID is dat er geen aparte nationale infrastructuur nodig is. Daarnaast beschikken de onderzoeksinformatiesystemen PURE en CONVERIS over standaardkoppelingen met de ORCID diensten. Er hoeven dus geen kosten betaald te worden voor het onderhoud van een dergelijke infrastructuur of de voor ontwikkeling van koppelingen met de onderzoeksinformatiesystemen.

Het nadeel is dat de lidmaatschapskosten behoorlijk kunnen oplopen. Ook zal er nog steeds een oplossing voor historische wetenschappelijk werk en cultureel erfgoed gezocht moeten worden. Tot slot is het probleem van zelfassertie en betrouwbaarheid van de data hiermee nog niet opgelost.

7.3.2 Scenario 2: ISNI registration agency

In het tweede scenario wordt een nationaal ISNI Registration Agency opgezet.

Op basis van de huidige NTA –infrastructuur kan een dienst worden ontwikkeld waarmee onderzoeksinstellingen nieuwe ISNI's voor hun (nieuwe) wetenschappelijke auteurs (handmatig) kunnen aanmaken. Het nationale Registration Agency zorgt niet alleen voor de registratie van wetenschappelijke auteurs, maar ook voor auteurs en andere “producenten” van Nederlands cultureel erfgoed. Daarnaast is de ISNI geschikt voor het ontwikkelen van een identifieer voor organisaties. Kanttekening daarbij is dat ook daar verschillende identifieer opties voor zijn (zie bijvoorbeeld: <http://repository.jisc.ac.uk/5853/1/Review-of-orgIDs-usecases-clax.pdf>)

De kosten bestaan uit een initiële investering voor het opzetten van een Registration agency door de KB (maximaal K€7,5). Voor de gemiddeld 200 nieuwe onderzoekers per jaar per instelling wordt aangeraden de ISNI's handmatig te registreren in het CRIS. Er zal per identifieer of batch moeten worden afgerekend. Uitgaande van €0,50 per identifieer gaat het hierbij om een zeer gering bedrag van ca. K€1 a K€1,5 totaal. Voor geautomatiseerde registratie kunnen eventueel een centraal koppelvlak en adapters bij de universiteiten worden ontwikkeld (kosten onbekend). De verdere jaarlijkse kosten bestaan uit het lidmaatschap van ISNI-International Agency (maximaal K€2 p.a.), het eventueel onderhouden van een centrale database en koppelvlakken en bemensing van een helpdesk. Het (opnieuw) inlezen van een bestand ter grootte van alle in NARCIS bekende personen (23312 personen) kost daarnaast K€5 (<50K records).

Het voordeel van deze oplossing is dat door de achterliggende kwaliteitsprocessen een betrouwbare auteursdatabase wordt gecreëerd. Verder betreft het een breed toepasbare oplossing waardoor kosten met een breder publiek gedeeld kunnen worden.

Het nadeel is, dat er een gescheiden nationale infrastructuur onderhouden en bekostigd moet worden. Daarnaast staat deze oplossing verder van de wetenschapper af en is de verwachting dat ORCID in de toekomst een rol van betekenis zal spelen in de wetenschappelijke informatie infrastructuur. Het is dus zaak hierop aangesloten te blijven.

7.3.3 Scenario 3: Hybride model

Het hybride model gaat uit van het gegeven dat er meerdere identifiëersystemen zullen bestaan die elk hun functie en bestaansrecht hebben, i.c. ORCID voor actuele wetenschappelijke informatievoorziening en ISNI voor de duurzame disambiguïteit vanuit rechtenbeheer, de muziekindustrie en (nationale) bibliotheken. Uitgangspunt hierbij is dat de beide systemen elkaar complementeren en de administratieve werklast wordt geminimaliseerd: er wordt dus gestreefd

naar synergie en niet naar redundantie. In feite is dat al de bestaande situatie. De ISNI wordt reeds gematched met de DAI en onderzoekers claimen op individuele basis (in Nederland nog zonder licentie) hun ORCID's.

Zoals eerder beschreven biedt ORCID geen soelaas voor inactieve wetenschappers en de niet-wetenschappelijke doelgroep. Voor het reguliere catalogiseerproces, waarbij ook historische of niet-wetenschappelijke werken kunnen worden gethesaureerd, wordt daarom ISNI als duurzame en volledige oplossing gepositioneerd. Volgens het ISNI scenario zal de KB, al dan niet als vertegenwoordiger van de UKB als *'federated registration agency'*, toetreden tot ISNI (is reeds geschied). De onderzoeksinformatiesystemen publiceren via een standaardkoppeling de ORCID, bibliografische gegevens en affiliatie gegevens wat de wetenschapper aan de NTA. Via geautomatiseerde matchingalgoritmen en nauwe integratie met de ISNI infrastructuur wordt vervolgens een ISNI toegekend. Wanneer de KB migreert naar een ander platform dan de GGC en de daarmee zou komen NTA te vervallen, zal deze functionaliteit met de KB meeverhuizen.

Daarnaast registreren universiteiten in het hybride model volgens het ORCID principe. Dat kan eventueel zonder licentie waarbij hun onderzoekers stimuleren hun ORCID aan te vragen en deze te registreren in het onderzoeksinformatiesysteem. Volledige dekking door een dergelijk author self claim is op deze manier echter zeer moeilijk te bewerkstelligen. Gedurende de aanvraagprocedure wordt een onderzoeker door ORCID gevraagd andere bekende identifiers zoals een ISNI of een ResearcherID te koppelen aan zijn of haar ORCID. Daar via de eerder beschreven pilot van OCLC al een groot aantal records uit de NTA is ingelezen in de ISNI database, is er een gerede kans dat er een ISNI aan de onderzoeker is toegekend. Een dergelijke koppeling maakt het vervolgens eenvoudiger om de verschillende identifiersystemen (i.c. DAI, ORCID en ISNI) te ontdebellen. Hierbij dient wel te worden aangetekend dat deze procedure zoals aangeboden door ORCID nog van vrij primitieve aard is en lang niet altijd de andere identifier (zoals ResearcherID) kan terugvinden wanneer een auteur die wel heeft.

Een variant is dat de betreffende universiteit of onderzoeksinstituut een ORCID licentie afneemt waarbij de universiteit als gedelegeerd beheerder via een geautomatiseerde koppeling de ORCID in het onderzoeksinformatiesysteem koppelt. Naast dat deze variant een minder afhankelijk is van de proactieve houding van de onderzoeker, heeft dit als bijkomend voordeel dat ook bibliografische gegevens uit het onderzoeksinformatiesysteem aan ORCID wordt aangeleverd. Ook dit helpt bij eventuele ontdebelling tussen ORCID en ISNI. De verwachting is verder dat de aanwezigheid van een ORCID de koppeling met externe bibliografische databases eenvoudiger maakt waardoor meer gegevens geautomatiseerd het onderzoeksinformatiesysteem binnen kunnen komen en op de juiste wijze worden gekoppeld aan de wetenschapper. De actuele onderzoeksinformatie wordt hiermee in potentie zo compleet mogelijk. Om dit model te laten werken zullen de onderzoeksinformatiesystemen dus de mogelijkheid moeten hebben om meerdere identifiers per wetenschapper te registreren.

Het voordeel van het hybride model is dat voor het onderzoeksregistratieproces wordt aangesloten bij de beweging van de markt i.c. wetenschappelijke uitgeverij en leveranciers van onderzoeksinformatiesystemen die zich hebben georganiseerd rondom ORCID. Tegelijkertijd wordt de duurzaamheid en volledigheid worden geborgd door gebruik van ISNI in het catalogiseerproces. Door gebruik te maken van geautomatiseerde matching wordt de administratieve werklast rondom de onderzoeksinformatiesystemen tot het minimum gereduceerd.

Het nadeel is dat er ten opzicht van het eerste of tweede scenario extra lidmaatschapskosten gemaakt worden. Door in consortia te werken waarbij kwantumkortingen gelden, kunnen deze kosten beperkt gehouden worden.

8 AANBEVELINGEN

Overwegende het bovenstaande, adviseert de DAI werkgroep te kiezen voor het hybride model zoals beschreven in paragraaf 7.3.3.

Als enige voldoet dit model in de behoefte om aan te sluiten bij de marktontwikkelingen in de actuele wetenschappelijke informatievoorziening, maar tegelijkertijd een duurzaam en volledig overzicht biedt voor de toekomstige informatievoorziening.

Om dit model te realiseren worden daarbij de volgende aanbevelingen gedaan.

1. *Kies ISNI als duurzame identifier voor Auteursnamen.*

Als enige identifier borgt ISNI duurzame toegang. Anders dan het ORCID alternatief, vereist ISNI geen actieve betrokkenheid van de wetenschapper bij het toekennen van een identifier. Hierdoor kunnen ook retrospectief identifiers worden toegewezen aan bijvoorbeeld uit dienst getreden of overleden onderzoekers. Door kwaliteitsprocessen bij nationale bibliotheken worden fouten proactief ondervangen. En is de ISNI een logische keuze in de catalogusomgeving van de bibliotheek.

2. *Gebruik ORCID als identifier binnen de wetenschappelijke communicatie.*

Door de steun van uitgevers en standaard integratie mogelijkheden in de onderzoeksinformatiesystemen, is ORCID uitgegroeid tot de *de facto* standaard voor de actuele wetenschappelijke onderzoeksinformatievoorziening. Uitgevers en andere dienstverleners binnen de wetenschappelijke informatievoorziening zullen wetenschappers om hun ORCID vragen als identiteit. Bovendien heeft het initiatief zich door haar opzet verzekerd van betrokkenheid van de wetenschappelijke community, vergelijkbaar met de ontwikkeling van de DOI. Het ORCID model sluit daarom nauwer aan bij wereld waarbinnen de onderzoeker zijn of haar resultaten verspreidt en registreert. Daarnaast zullen ook bibliografische bronnen als Web of Science gebruik van ORCID als preferente auteursidentifier, waardoor het gebruik van ORCID belangrijke voordelen biedt voor het geautomatiseerd koppelen van gegevens in het onderzoeksinformatiesysteem.

3. *Motiveer onderzoekers hun ORCID te claimen.*

Vanuit de gedachte dat ORCID de *de facto* identifier voor auteurs zal vormen, zoals DOI dat is voor identifiers voor publicaties, is het van belang wetenschapper bewustwording te creëren bij wetenschappers van het bestaan van dit systeem. Het '*author-claim*' principe van ORCID vereist dat de wetenschapper zelf eigenaarschap neemt van zijn of haar identifier. Dat betekent dat de wetenschapper actief zijn of haar ORCID moet registreren en de werkgever moet machtigen dit voor

hem/haar te onderhouden.⁶ Tijdens het registratieproces is het van belang dat de onderzoeker zoveel mogelijk andere bekende identifiers, zoals een ResearcherID of ISNI, koppelt aan het ORCID account. Dit maakt de koppeling met infrastructures die (nog) gebruik maken van andere identifiersystemen zoals DAI en ISNI eenvoudiger.

4. Registreer ORCID (en ISNI zie punt6) in het onderzoeksinformatiesysteem.

Om de geclaimde ORCID ook te gebruiken bij het beheer van de actuele onderzoeksinformatie en de verdere informatieketen te kunnen voeden, is het noodzakelijk de ORCID te registreren in het onderzoeksinformatiesysteem. Dit gebeurt handmatig nadat de onderzoeker de ORCID heeft aangemaakt of geautomatiseerd wanneer de universiteit de gegevens uit het onderzoeksinformatiesysteem met ORCID synchroniseert. Dit maakt het mogelijk bibliografische databases die gebruik maken van ORCID te koppelen voor geautomatiseerde invoer van bijvoorbeeld publicatiegegevens. Daarnaast kunnen institutionele repositories en aanpalende infrastructures zoals NARCIS en OpenAIRE worden gevoed met ORCID. Bij voorkeur wordt ook de ISNI in het onderzoeksregistratie systeem vastgelegd (zie 8.) zodat via het CRIS de benodigde matching tussen ISNI en ORCID zou kunnen plaatsvinden.

5. Bij voldoende animo een nationale licentie laten onderhandelen voor ORCID.

Een licentie biedt aanvullende mogelijkheden voor gedelegeerd beheer en automatische synchronisatie van gegevens tussen ORCID en onderzoeksinformatiesystemen. Een nationale licentie biedt de mogelijkheid voor kwantumkorting en is van strategische waarde bij het bepalen van de ontwikkeling van de dienst. De UKB kan SURFmarket vragen namens de universiteiten een licentie te onderhandelen. Hierbij zou ook aandacht moeten zijn voor de onderzoeksinstituten van bijvoorbeeld KNAW, NWO en de hogescholen.

6. Sluit onderzoeksinformatiesystemen aan op de ISNI infrastructuur.

Zowel Converis als PURE zijn actief in contact met ORCID. Als Nederlandse community kan druk worden gezet op de leveranciers om ook actief in contact te treden over koppeling met de ISNI infrastructuur. Zowel bij PURE als CONVERIS lijken hier zeker mogelijkheden voor te zijn.

Op korte termijn zal een pilot voor uitwisselen van onderzoeksinformatie (middels CERIF) starten die moet leiden tot een set van Nederlandse wensen gericht aan de leveranciers van de onderzoeksinformatiesystemen. Hierin kunnen aanbevelingen rondom het (automatisch) kunnen registreren van auteursidentifiers worden meegenomen.

7. Richt een Nederlands Registration Agency voor ISNI op.

Is door de KB al uitgevoerd. Daarmee kan de KB voor Nederlandse auteurs ISNI's (laten) registreren. Universiteitsbibliotheken en andere instellingen kunnen voor €800 per jaar gebruik maken van de

⁶ Registratie door de werkgever is alleen mogelijk met een licentie.

ISNI infrastructuur (geëxploiteerd door OCLC) en via de KB auteurs aandragen voor registratie (max. €0,50 per identifier).

8. *Universiteiten/UKB bibliotheken en KNAW/DANS-Narcis zouden ondersteund door OCLC een overgang van DAI naar ISNI moeten realiseren in 2015/2016.*

Dit kan het beste projectmatig gebeuren, bijvoorbeeld onder regie van de werkgroep innovatie ondersteund door OCLC. Een dergelijk project realiseert de volgende resultaten:

- Voor alle nieuwe wetenschappelijke auteurs wordt via de KB een ISNI registratie opgezet.
- In universitaire CRIS systemen en/of repositories (en bij andere instellingen) wordt de ISNI identifier geregistreerd naast eventueel ORCID. OCLC levert instellingen een ISNI/DAI concordantie tabel, waarmee de CRIS systemen en/of repositories initieel worden gevuld met ISNI's.
- In de uitwisseling van gegevens tussen repositories en NARCIS vervangt de ISNI de functie van de DAI. Daarnaast wordt het ORCID van auteurs opgenomen in NARCIS en aangeleverd via de geëigende harvesting infrastructuur door de CRIS systemen of repositories aan Narcis en KB.

9. *Een aantal universiteiten gaat samen met SURF in 2015/2016 in een pilotproject de mogelijkheden van ORCID voor versterking van de Nederlandse research infrastructuur onderzoeken.*

De volgende *use cases* zijn voorgesteld om te onderzoeken in een pilotproject:

- Is het mogelijk om de ORCID infrastructuur te koppelen aan de SURFconext infrastructuur? Minimaal kan ORCID worden aangesloten op SURFconext zodat onderzoekers met hun instellingsaccount kunnen inloggen bij ORCID. Hiertoe moet in ieder geval worden gekeken naar de juridische randvoorwaarden (zie juridisch normenkader cloudservices hoger onderwijs⁷). Mogelijk kan SURFconext ook een rol vervullen in het bevragen van de ORCID database en op basis hiervan de ORCID doorgeven aan andere op SURFconext aangesloten systemen waar de ORCID dient te worden geregistreerd.
- De deelnemende instellingen verkennen de mogelijke routes om hun auteurs centraal een ORCID te laten registreren in een CRIS. Tevens zal bekeken worden hoe een ISNI bij een ORCID kan worden opgeslagen.
- Onderzocht wordt of het mogelijk is om derde partijen van de mogelijkheid gebruik te laten maken om aan de hand van de ORCID de identiteit en affiliatie van een onderzoeker te laten valideren. Een voorbeeld van een dergelijke partij kan een uitgever die wil valideren dat een auteur aan een instelling verbonden is die een bepaalde licentie heeft afgesloten voor het publiceren van een (gold) open access artikel. Mogelijk kan SURFconext een brugfunctie vervullen indien er een koppeling nodig is met de instellingscredentials van de onderzoeker.

⁷ <https://www.surf.nl/kennis-en-innovatie/kennisbank/2013/juridisch-normenkader-cloud-services-hoger-onderwijs.html>

Voorts zal in de pilot moet worden vastgesteld of de mogelijkheden voor ORCID gebruik zonder licentie (author self claim) kunnen worden benut of dat een basic licentie (\$5000 p.j.) of een Premium licentie (\$6000 p.j., vanaf vijf instellingen) is gewenst of vereist.

De pilot zal worden aangekondigd via verschillende kanalen (SIG research information, UKB werkgroep Innovatie, SURF website). SURF voorziet daarbij in de coördinatie, kennisdeling en contacten met ORCID.

Deze pilot moet leiden tot een beter inzicht in de voordelen van ORCID implementatie en een eventuele (Nationale) licentie (en zo ja ook welk type). Ook zal worden meegenomen of een ORCID licentie een langdurig voordeel zou kunnen bieden of dat een licentie mogelijk slechts voor een aantal jaren essentieel is. Nadat een goede dekking van ORCID's is bereikt en een workflow voor ORCID registratie voor nieuwe onderzoekers via het CRIS is ingericht kan worden gebruik gemaakt van de gratis voorziening die ORCID al biedt aan via de author claim.

BIJLAGE : VERSIE BEHEER NOTA TOEKOMST NEDERLANDSE INFRASTRUCTUUR VOOR AUTEURSIDENTIFIERS

HISTORIE

Versie	Datum	Veranderingen	Auteur(s)
0.1	12-01-2014	Eerste concept	Magchiel Bijsterbosch
0.2	24-01-2014	Aanvullingen n.a.v. presentatie Janifer Gatenby (OCLC), commentaar Frits van Latum (TUD) en vergadering DAI werkgroep d.d. 13-01-2014.	Magchiel Bijsterbosch
0.3	19-05-2014	Uitwerking hybride model. Verwerking ontwikkelingen transitie UKB. Uitwerking aanbevelingen.	Magchiel Bijsterbosch
0.4	03-12-2014	Update huidige situatie	John Doove
0.5	09-04-2015	Laatste concept	Maarten Steenhuis
0.9	06-05-2015	Aanpassingen en Executive summary	John Doove
0.91	12-08-2015	Aanpassingen en correcties n.a.v. gesprek Marc van den Berg (UT) dd. 30-07-2015 en Niels van Dijk (SURFnet) dd 11-08-2015	John Doove
Pv0.1	14-08-2015	Publieksversie concept	John Doove
1.0	02-09-2015	Documentbeheer bijgewerkt	John Doove
Pv1.0	02-09-2015	Publieksversie definitief (documentbeheer in bijlage)	John Doove

VERSPREIDING

Naam	Rol	Datum verspreiding	Versie
DAI Werkgroep	Reviewer	12-01-2014	0.1
DAI Werkgroep	Reviewer	17-03-2014	0.2
DAI Werkgroep	Reviewer	30-09-2014	0.3
DAI Werkgroep	Reviewer	06-03-2015	0.4
DAI Werkgroep	Reviewer	20-04-2015	0.5
DAI Werkgroep	Reviewer	01-06-2015	0.9
UKD bestuur	Opdrachtgever	16-06-2015	0.9
DAI Werkgroep	Reviewer	13-08-2015	0.91
DAI Werkgroep	Reviewer	02-09-2015	1.0
DAI Werkgroep (pdf)	Reviewer	13-08-2015	Pv0.1
DAI Werkgroep (pdf)	Reviewer voor publicatie	13-08-2015	Pv1.0