

## HOSA: KADERS VOOR SECTOR- VOORZIENINGEN VAN DE TOEKOMST

Domeinarchitectuur  
Research Data  
Management



Auteurs:

Tom van Veen &  
Domeinarchitect RDM

Menno Scheers  
Lead Architect HOSA

Versie:

1.0

Datum:

December 2021

Aangejaagd door:

CIO's van de HO-instellingen, Architectenberaad HO, SURF, ....

## 1 Voorwoord/leeswijzer

Deze domeinarchitectuur schetst hoe de digitale ondersteuning rondom Research Data Management er over drie tot tien jaar uit zou moeten zien om te passen bij de gedragen ambities van de HO-sector, de veranderingen van onderzoek en internationale (technologische) ontwikkelingen. Deze domeinarchitectuur is vergelijkbaar met een bestemmingsplan en omvat een visie op de toekomst van Research Data Management en de samenhang hieromheen, met de daarbij behorende systemen. Tevens vormt het de basis voor een roadmap die het mogelijk maakt dit bestemmingsplan te realiseren en te prioriteren.

Op basis van een bestemmingsplan kan een gebouw echter nog niet gebouwd worden, daarvoor zijn vervolgarchitecturen en -ontwerpen vereist. Dit document biedt de kaders voor het toetsen van dergelijke vervolgarchitecturen en -ontwerpen. Om te beginnen, beschrijven we een eerste indicatie van de huidige situatie. Deze beschrijving is tot stand gekomen op basis van diverse gesprekken en sessies met de werkgroep en nodig om de domeinarchitectuur vorm te kunnen geven. Voor het daadwerkelijk realiseren van een sectorvoorziening is een nadere analyse van de huidige situatie vereist.

Voor het opstellen van de applicatiearchitectuur maken we in deze domeinarchitectuur gebruik van het concept van businessplatforms. Dit concept past bij komend decennium, ook wel de 'digital age' genoemd. Een businessplatform gaat uit van een business gebaseerd op waarde creërende interacties tussen externe producenten en consumenten. Het overkoepelende doel van het platform is matches te realiseren tussen gebruikers en het uitwisselen van producten, services, content, data of social currency te faciliteren, waarbij waarde wordt gecreëerd voor alle deelnemers (G. Parker, 2016).

We gebruiken dit concept omdat het inzicht geeft in zogenaamde online marktplaatsen, die door toenemende samenwerking in de sector op diverse plekken ontstaan. Hierbij leggen we vooral de nadruk op de term 'business'. Door dezelfde basisplaat als onderlegger te gebruiken, geven we overzicht en ontstaat een gezamenlijk vertrekpunt voor de discussies tussen de vele stakeholders. Daarnaast fungeert deze plaat als fundament om diverse initiatieven op een samenhangende wijze weer te geven.



## 2 Inhoudsopgave

|    |   |    |
|----|---|----|
| 1  | Voorwoord/leeswijzer.....               | 2  |
| 2  | Inhoudsopgave.....                      | 3  |
| 3  | Inleiding .....                         | 4  |
| 4  | Huidige situatie .....                  | 6  |
| 5  | Ambities van de sector .....            | 11 |
| 6  | Technologische ontwikkelingen.....      | 14 |
| 7  | Architectuurvisie .....                 | 16 |
| 8  | Platformstrategie .....                 | 21 |
| 9  | Stakeholders .....                      | 26 |
| 10 | Proces- en applicatiearchitectuur ..... | 28 |
| 11 | Principes.....                          | 41 |
| 12 | Bijlage .....                           | 43 |

## 3 Inleiding

### 3.1 Aanleiding

De HO-instellingen hebben geconstateerd dat instellingoverstijgende initiatieven zich uitbreiden en een grotere impact gaan hebben op de eigen voorzieningen. De VSNU onderzoekt in welke onderwerpen de universiteiten gezamenlijk willen gaan investeren, zoals onderzoeksfaciliteiten, ICT voor onderwijsvernieuwing of duurzame bedrijfsvoering. Daarnaast wordt in het hbo instellingoverstijgend samengewerkt, bijvoorbeeld rondom het thema onderzoeksondersteuning. De CIO's van instellingen, SURF en een aantal sectorpartners hebben daarom het initiatief genomen voor een architectuur voor digitale sectorvoorzieningen van de toekomst.

De ontwikkelingen op het gebied van onderzoeksdata en datamanagement leiden tot meer instellingoverstijgende samenwerking en organisatie op het gebied van gemeenschappelijke informatie- en ICT-voorzieningen (sectorvoorzieningen). Hierbij ontstaan bijvoorbeeld vragen als: 'Hoe zorgen we ervoor dat sectorvoorzieningen toekomstvast zijn?', 'hoe creëren we samenhang in de sectorvoorzieningen?', 'hoe maken we hergebruik mogelijk?' en, meer recentelijk: 'hoe borgen, beschermen en bevorderen we publieke waarden bij de digitalisering van onderwijs en onderzoek?'.

Sectorpartners als SURF, Studielink, DUO, DANS en NWO proberen alle instellingen hierbij zo veel mogelijk te faciliteren en te ondersteunen. Dit is echter een complex proces en doet de behoefte ontstaan aan een gezamenlijke architectuur: voor sectorbrede definitie, ontwikkeling en inzet van informatie- en ICT-voorzieningen is het noodzakelijk duidelijkheid te verschaffen over de vraag naar die voorzieningen, de eisen die er in samenhang aan gesteld worden, de vormgeving en inrichting ervan en de dienstverlening die ICT-dienstverleners aan instellingen bieden. Dit gemeenschappelijke kader willen we creëren vanuit een architectuurbenadering: de HO Sector Architectuur (HOSA).

### 3.2 Doelstelling

Het project HO Sector Architectuur heeft als doel een architectuur te definiëren voor sectorvoorzieningen die van belang zijn voor de strategische samenwerkingen tussen HO-instellingen, sectorpartners en marktpartijen. De basis van de HOSA is dan ook een optimale articulatie van de vraag van de sector met betrekking tot sectorvoorzieningen. Tevens biedt de HOSA een faciliterend kader voor de interoperabiliteit tussen instellingen en aanbieders van de gemeenschappelijke ICT-voorzieningen. Door de HOSA zullen lopende en nieuwe initiatieven op het gebied van sectorvoorzieningen sneller, toekomstgerichter en toekomstbestendiger tot stand komen. Sectorpartners en marktpartijen van ICT-voorzieningen kunnen met hun dienstenportfolio hier goed op inspelen.

### 3.3 Scope van de domeinarchitectuur

De scope van de domeinarchitectuur is in beginsel de HO-sector met wo en hbo. Gezien het feit dat onderzoek zich niet beperkt tot de HO-sector, is ook naar de internationale context van Research Data Management gekeken. De tijdshorizon betreft de middellange en lange termijn vooruitkijkend, met een inschatting van de functionele behoeften aan sectorvoorzieningen.

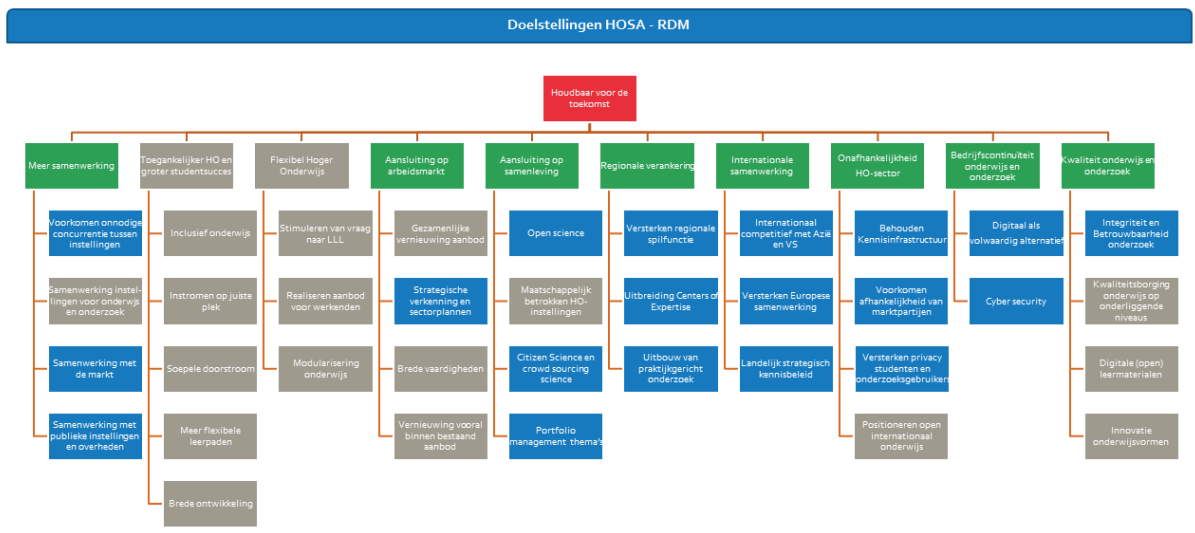
Onder Research Data Management verstaan we een expliciet proces met betrekking tot de creatie, het beheer en de verwerking van onderzoeksmaterialen om de toegankelijkheid en het gebruik ervan mogelijk te maken zolang ze waarde behouden. Met onderzoeksdata doelen we zowel op primaire onderzoeksdata (de onbewerkte, ruwe metingen of observaties) als op secundaire onderzoeksdata (data die ontstaan na bewerking door de onderzoeker (hercoderen, combineren, categoriseren, visualiseren en dergelijke). Hierbij hanteren we een brede blik en besteden we ook aandacht aan diensten, dienstverlening, processen, functionaliteit, data en technologie, governance, eigenaarschap, beheer en ondersteuning.

Deze domeinarchitectuur is een conceptuele beschrijving van de inrichting van sectorvoorzieningen voor onderzoeksdata in de brede zin van het woord. Zij beschrijft richtinggevend kaders voor positionering en functionaliteiten van de nieuwe sectorvoorzieningen rond onderzoeksdata. Bij het uitwerken van deze domeinarchitectuur hebben we de onderzoekscyclus en de onderzoeksgegevenscyclus zoals beschreven in HORA 2.1 als uitgangspunt genomen (Bedrijfsprocessen Onderzoeksgegevenscyclus, 2019).

De domeinarchitectuur is vergelijkbaar met een bestemmingsplan.

### 3.4 Doelenstructuur

Voor de HOSA is een doelenstructuur opgesteld, die inzicht geeft in de doelstellingen en ambities van de sector. De basis voor deze doelenstructuur is gelegd vanuit de strategische agenda van het Ministerie van OCW<sup>1</sup>. Deze is aangevuld met doelstellingen uit andere beleidsdocumenten (zie figuur 1). De domeinarchitectuur voor Research Data Management ondersteunt de realisatie van de aan onderzoeksdata gerelateerde HOSA-doelstellingen en ambities van de sector (blauw gemarkeerd).



Figuur 1: Doelenstructuur HOSA-RDM (zie bijlage 13.3 voor leesbare versie)

<sup>1</sup> Strategische agenda hoger onderwijs en onderzoek – Houdbaar voor de toekomst, Ministerie van OCW, Den Haag, december 2019

## 4 Huidige situatie

Voor het beschrijven van de huidige situatie op het vlak van Research Data Management (RDM) gaat deze domeinarchitectuur uit van de bestaande voorzieningen en veranderinitiatieven die vanuit verschillende organisaties en gremia zijn of worden ondernomen. Daarnaast hebben we door middel van interviews, rapporten en sessies diverse knelpunten rondom RDM geïnventariseerd.

Paragraaf 4.1 beschrijft kort een aantal initiatieven op het gebied van RDM. Dit overzicht is niet uitputtend, het geeft slechts inzicht in enkele van de vele initiatieven. Paragraaf 4.2 beschrijft de belangrijkste knelpunten die naar voren zijn gekomen.

### 4.1 Initiatieven

In deze paragraaf beschrijven we kort een aantal initiatieven (bestaand of in wording) die gerelateerd zijn aan het Research Data Management-vraagstuk. Tijdens de inventarisatie is gebleken dat het landschap aan initiatieven op het gebied van RDM zeer weids is. Deze initiatieven acteren op verschillende niveaus – van internationaal tot Europees tot regionaal – en geven stuk voor stuk invulling aan een behoefte die op dat niveau of voor een specifiek thema/vraagstuk aanwezig is. Dergelijke initiatieven staan dan ook vaak op zichzelf en kennen weinig tot geen onderlinge samenhang. Ze richten zich veelal op het vinden, verkrijgen en beschikbaar stellen van onderzoeksdata en veel minder op bijvoorbeeld het verwerken, creëren, analyseren of visualiseren ervan. De inventarisatie is dan ook verre van volledig, eenvoudigweg omdat het aantal initiatieven te talrijk en te divers is. Het geeft hoogstens inzicht in de diversiteit en fragmentatie.

#### DANS

DANS (Data Archiving and Networked Services), het Nederlandse instituut voor permanente toegang tot digitale onderzoeksdata, stimuleert onderzoekers om hun digitale onderzoeksdata vindbaar, toegankelijk, interoperabel en herbruikbaar te maken. Diensten van DANS zijn onder andere: DataverseNL voor databeheer op de korte termijn, EASY voor archivering op de lange termijn en NARCIS, het nationale portaal voor wetenschappelijke informatie. DANS participeert in tal van projecten die zich richten op het domein van RDM, waaronder de EOSC-hub en FAIR data self-assessment tool en CoreTrustSeal. DANS is een instituut van KNAW en NWO.

#### LCRDM

Het Landelijk Coördinatiepunt Research Data Management (LCRDM), een landelijk netwerk van experts op het gebied van RDM, faciliteert de koppeling tussen beleid en oplossing in nauwe samenspraak met onderwijs- en onderzoeksinstituten. Binnen het LCRDM werken experts samen om RDM-onderwerpen te agenderen die te groot zijn voor één instelling en die vragen om een gezamenlijke landelijke aanpak.

#### NPOS

In het Nationaal Programma Open Science (NPOS) komen de betrokken partijen uit het Nationaal Plan Open Science bijeen. Het programma zorgt ervoor dat Nederland vordert met de realisatie van de ambities op het gebied van Open Science en zicht houdt op de ontwikkelingen daarvan. De focus binnen het programma ligt op versnelling creëren op de drie speerpunten van het Nationaal Plan Open Science. De bij het nationaal plan betrokken partijen nemen deel aan regelmatig overleg in het programma. Het NPOS kent een stuurgroep, gevormd door de bestuurders van betrokken partijen, een adviesraad en verschillende projecten.

#### NPPO

Hogescholen doen veel praktijkgericht onderzoek, wat bijdraagt aan de vernieuwing van onderwijs, innovaties in het bedrijfsleven en maatschappelijke organisaties. Praktijkgericht onderzoek levert waardevolle kennis en producten op voor de beroepspraktijk en het onderwijs, die echter nog meer mensen kunnen bereiken.

Zichtbaarheid en vindbaarheid van informatie speelt hier een rol bij en dat is precies de reden voor oprichting van het Nationaal Platform Praktijkgericht Onderzoek (NPPO).

### **eScience Center**

Het Netherlands eScience Center, het nationale centrum voor de ontwikkeling en toepassing van onderzoekssoftware, brengt kennis, mensen en instellingen samen om software te bouwen die het gebruik van digitale methoden in onderzoek verbetert. Hierin wordt rechtstreeks samengewerkt met publieke onderzoeksorganisaties en onderzoekers in alle disciplines.

### **Digital Competence Centers**

Hogescholen gaan samen aan de slag in Digital Competence Centers (DCC's) voor Praktijkgericht onderzoek om RDM, FAIR-data en data-intensief onderzoek op hogescholen verder te faciliteren en zo hun open-science-ambities te realiseren. SURF ondersteunt het landelijk DCC voor Praktijkgericht onderzoek in de vorm van pilots, training en expertise. In zes werkpakketten gaan hogescholen urgente vraagstukken oppakken op het gebied van RDM, FAIR-data en data-intensief onderzoek.

### **4TU.ResearchData**

4TU.ResearchData is een initiatief van de 4TU.Federation en heeft tot doel een archief te bieden voor langetermijntoegang en het gestructureerd verzamelen, filteren, beoordelen, duiden, organiseren en delen van onderzoeksdata, met een focus op de technische wetenschappen. Dit archief ging in 2010 live en wordt sindsdien beheerd als een service voor onderzoekers (van universiteiten over de hele wereld) om hun onderzoeksdata te uploaden en te delen, en voor andere onderzoekers om onderzoeksdata te downloaden en te gebruiken in hun onderzoek.

### **NARCIS**

NARCIS is dé nationale portal voor wie informatie zoekt over wetenschappers en hun werk. Naast wetenschappers maken ook studenten, journalisten en medewerkers binnen onderwijs, overheid en bedrijfsleven gebruik van NARCIS. Deze portal biedt toegang tot wetenschappelijke informatie, waaronder (open access) publicaties afkomstig uit de repositories van alle Nederlandse universiteiten, KNAW, NWO en diverse wetenschappelijke instellingen, datasets van een aantal data-archieven, en ook beschrijvingen van onderzoeksprojecten, onderzoekers en onderzoeksinstituten.

### **Amsterdam Data Exchange**

De Amsterdam Data Exchange heeft als missie de academische wereld, het bedrijfsleven en de samenleving te verbinden in een divers Data Science- en AI-ecosysteem in de regio Amsterdam. In dit netwerk staan samenwerking, innovatie en betrokkenheid centraal.

### **IRODS**

IRODS staat voor Integrated Rule-Oriented Data System en betreft datamanagementsoftware die ongestructureerde data aan metadata koppelt. Voor onderzoekers die met grote hoeveelheden onderzoeksdata over langere tijd werken, is dit een goede manier om hun onderzoeksdata veilig te documenteren en op te slaan. IRODS kan onderzoekers helpen de integriteit van hun onderzoeksdata aan te tonen. Een belangrijke ontwikkeling voor het vergroten van de toepassing van IRODS in Nederland is dat SURF heeft aangekondigd een expertisecentrum in te richten rond IRODS.

### **European Open Science Cloud**

De European Open Science Cloud (EOSC) is een initiatief van de Europese Commissie dat tot doel heeft de virtuele omgeving van Europa te realiseren voor alle onderzoekers. Dit om onderzoeksdata op te slaan, te beheren, te analyseren en te hergebruiken voor onderzoeks-, innovatie- en onderwijsdoeleinden. Binnen de architectuur van EOSC bevindt zich de OpenAIRE Research Graph, een verzameling metadata en links die

onderzoeksentiteiten verbindt. Te denken valt aan artikelen, onderzoeksdata, software en andere onderzoeksproducten en entiteiten zoals organisaties, financiers, financieringsstromen, projecten, onderzoeksgemeenschappen en databronnen.

### **GO-FAIR**

GO FAIR is een bottom-up, stakeholder-driven en zelfsturend initiatief dat tot doel heeft de FAIR-dataprincipes te implementeren, zodat onderzoeksdata Findable, Accessible, Interoperable en Reusable (FAIR) worden. Het biedt een open en inclusief ecosysteem voor individuen, instellingen en organisaties die samenwerken via Implementation Networks (INs). De INs zijn actief in drie pijlers: GO CHANGE, GO TRAIN en GO BUILD. Via deze drie activiteitenpijlers pleit GO FAIR voor de FAIR-dataprincipes.

### **FAIRsFAIR**

FAIRsFAIR heeft als doel praktische oplossingen te bieden voor het gebruik van de FAIR-dataprincipes gedurende de hele levenscyclus van onderzoeksgegevens. De nadruk ligt op het bevorderen van een eerlijke datacultuur en het gebruik van 'good practices' bij het FAIR maken van gegevens. FAIRsFAIR wil een sleutelrol spelen bij de ontwikkeling van wereldwijde standaarden voor FAIR-certificering van repositories en de gegevens daarin. Tevens beoogt FAIRsFAIR een platform te bieden voor gebruik en implementatie van de FAIR-principes in het dagelijks werk van Europese aanbieders van onderzoeksgegevens en repositories.

### **Zenodo**

Zenodo is een multidisciplinaire open repository die door CERN wordt onderhouden. Datasets, documenten en ander onderzoeksmateriaal kunnen worden gelokaliseerd via de Zenodo-zoekmachine. Wetenschappers uit elke onderzoekdiscipline kunnen gegevens uploaden in elke bestandsindeling. Aan alle Zenodo-bestanden wordt automatisch een DOI (Digital Object Identifier) toegewezen. Details over het toewijzen van metagegevens aan onderzoeksgegevenssets staan in sectie 5 (onder c) van de EUI Library Research Data Guide.

### **Figshare**

Figshare is een online open access repository, waar onderzoekers hun onderzoeksresultaten kunnen bewaren en delen, waaronder cijfers, datasets, afbeeldingen en video's. Inhoud uploaden is gratis en de toegang is vrij, in overeenstemming met het principe van open data.

### **Mendeley**

Het Mendeley-platform, onderdeel van Elsevier, heeft tot doel de onderzoeker te helpen bij het organiseren, online samenwerken en nieuw onderzoek ontdekken. Elsevier stelt het platform gratis ter beschikking.

### **Nederlandse Gedragscode Wetenschappelijke Integriteit**

De Nederlandse Gedragscode Wetenschappelijke Integriteit definieert vijf principes van wetenschappelijke integriteit, 61 normen voor goede onderzoekspraktijken en zorgplichten voor instellingen. Met deze code sluit Nederland aan bij internationale ontwikkelingen op het gebied van wetenschappelijke integriteit. De code wordt onderschreven door de KNAW, NFU, NWO, de TO2-federatie, de VH en de VSNU. Deze organisaties dragen er zorg voor dat de instituten, universitair medische centra, hogescholen en universiteiten die zij vertegenwoordigen of die onder hen ressorteren, deze code onderschrijven. Ook andere instellingen, inclusief private ondernemingen, kunnen deze code confirmeren.

## **4.2 Aandachtspunten**

Hoewel de inventarisatie van de huidige situatie het Research Data Management-landschap niet volledig weergeeft, toont deze wel duidelijk de huidige grote fragmentatie. Daarnaast hebben we nog andere knelpunten geïdentificeerd, waarvan we de belangrijkste hierna kort beschrijven. De knelpunten en ook het gefragmenteerde landschap vormen input voor de verdere uitwerking van deze domeinarchitectuur.



### **Herkenbaarheid en afname vertrouwen**

Bij het uitvoeren van onderzoek wordt gebruik gemaakt van verschillende soorten onderzoeksdata. De betrouwbaarheid en kwaliteit van deze onderzoeksdata is niet altijd helder of gegarandeerd. Dit gebrek aan inzicht neemt steeds verder toe nu de samenleving steeds verder digitaliseert, waarbij grote hoeveelheden data geproduceerd worden. De herkomst of betrouwbaarheid van deze data is daarbij niet altijd duidelijk. Dit gebrek aan inzicht zet de betrouwbaarheid en kwaliteit van onderzoek onder druk en daarmee het vertrouwen in de uitkomsten van onderzoek, ondanks het feit dat de wetenschap beschikt over mechanismes die dergelijke onderzoeken eruit filteren. Voor de acceptatie van onderzoeksresultaten is inzicht in de betrouwbaarheid en kwaliteit van de onderzoeksdata essentieel. Een dergelijk inzicht zal bijdragen aan de herkenbaarheid en betrouwbaarheid van het onderzoeksresultaat.

### **Big data veranderen onderzoeksmethoden**

De digitalisering van onze samenleving leidt tot een grote groei aan beschikbare onderzoeksdata. Deze groei leidt, in combinatie met nieuwe technologieën, tot nieuwe onderzoeksvormen en methodieken. De mogelijkheden voor data-intensief onderzoek nemen zo verder toe. Artificial intelligence biedt inmiddels mogelijkheden om naar patronen in onderzoeksdata te zoeken, waardoor nieuwe onderzoeksvragen ontdekt worden die anders onopgemerkt waren gebleven. Het is belangrijk om deze nieuwe onderzoeksvormen te ondersteunen en zo onderzoek in brede zin te verrijken en te versnellen.

### **Datamanagement vraagt aandacht**

Voor goed onderzoek zijn juiste onderzoeksdata essentieel. Deze juiste onderzoeksdata moeten op kwaliteit worden gebracht en gehouden. Dit vraagt om management van onderzoeksdata (in de brede zin van het woord), iets wat steeds meer een specialisme is en continue aandacht vereist. Zo moet meta-datering op orde worden gebracht en gehouden, moeten onderzoeksdata op de juiste manier opgeslagen worden, enzovoort. Vaak doet de onderzoeker dit managen van onderzoeksdata 'erbij'. Wanneer het onderzoek is afgerond, is management van de onderzoeksdata echter niet meer vanzelfsprekend, terwijl deze onderzoeksdata nog steeds relevant kunnen zijn voor ander en toekomstig onderzoek. Daarom is het belangrijk voldoende en duurzaam tijd en aandacht te besteden aan het managen van onderzoeksdata met oog voor governance met de juiste rollen en functies.

### **Huidige voorzieningen ingericht op oude uitgangspunten**

De huidige IT-voorzieningen die de onderzoeker ondersteunen bij het verkrijgen, verwerken, toepassen, publiceren, etc. van onderzoeksdata, zijn in hun oorsprong niet ingericht op moderne uitgangspunten rondom onderzoeksdata. Ontwikkelingen als FAIR en Open Science stellen eisen aan onderzoeksdata die de huidige IT-voorzieningen niet altijd en/of niet eenvoudig ondersteunen. Dit kan ertoe leiden dat een onderzoeker essentiële onderzoeksdata mist of extra inspanning moet verrichten om de juiste onderzoeksdata te verkrijgen. Doorontwikkeling van bestaande IT-voorzieningen of het ontwikkelen van nieuwe voorzieningen op basis van de nieuwe uitgangspunten, bijvoorbeeld FAIR en Open Science, is essentieel om onderzoek toekomstbestendig te maken.

### **Weinig hergebruik van data**

Onderzoek is gebaat bij het eenvoudig kunnen vinden en gebruiken van juiste onderzoeksdata. Toch is het niet altijd eenvoudig onderzoeksdata te vinden, te verkrijgen of uit te wisselen. Dit leidt tot inefficiëntie in het onderzoeksproces. Onder andere het toepassen van standaardisatie en hanteren van FAIR-uitgangspunten kan de uitwisselbaarheid en vindbaarheid van onderzoeksdata vergroten. De efficiëntie van onderzoek zal daardoor toenemen.

### **Tegenstrijdige en veelheid aan eisen**

Onderzoeksdata bestaan er in vele soorten en maten. Hierbij worden verschillende eisen gesteld aan bijvoorbeeld opslag en verwerking van en toegang tot deze onderzoeksdata. Het is niet altijd duidelijk wat deze

eisen inhouden, bijvoorbeeld op het gebied van de AVG of ten aanzien van embargoperiodes. De eisen, maar ook de interpretatie van deze eisen, kunnen per aanbieder van onderzoeksdata verschillen. Eenduidigheid in de eisen rondom onderzoeksdata en helderheid in de interpretatie ervan is belangrijk om het managen van onderzoeksdata uitvoerbaar te houden.

### **Eigenaarschap, governance en financiering**

Onderzoeksdatamanagement is essentieel voor toekomstig onderzoek. Zo moeten onderzoeksdata bijvoorbeeld op de juiste manier gemetadateerd zijn, op de juiste manier opgeslagen zijn en op de juiste manier beschikbaar worden gesteld (zie ook 'Datamanagement vraagt aandacht'). Dit managen vereist niet alleen aandacht en het inrichten van eigenaarschap, maar ook de juiste financiering, financieringsstromen en governance rondom onderzoeksdata.

### **Kwaliteitsborging van data**

De toename van hoeveelheden onderzoeksdata maakt het voor een onderzoeker steeds moeilijker om vast te stellen welke onderzoeksdata voor het desbetreffende onderzoek voldoende kwalitatief zijn. Digitale ondersteuning en een procesmatige aanpak kunnen hierbij helpen.



## 5 Ambities van de sector

Het belang van goed en betrouwbaar onderzoek neemt steeds verder toe. Onderzoeksresultaten worden in toenemende mate toegepast bij beantwoording van maatschappelijke vraagstukken en bij het nemen van beslissingen op sociaal en economisch gebied. Ook in de politiek wordt steeds meer gebruik gemaakt van resultaten uit onderzoek. Tegelijkertijd nemen de complexiteit en de kosten van onderzoek sterk toe.

OCW heeft met de sector daarom een aantal doelstellingen geformuleerd in de strategische agenda<sup>2</sup>. De hoofddoelstelling daarbij is om onderzoek toekomstbestendig te maken, waarbij de nadruk ligt op (regionale, nationale, internationale) samenwerking, aansluiting op de samenleving en regionale verankering. Daarnaast hebben, onafhankelijk van de HO-sector, doelstellingen als bedrijfscontinuïteit van onderzoek en kwaliteit van onderzoek veel aandacht in de sector.

Vanuit dit perspectief heeft de HO-sector ambities geformuleerd om onderzoek toekomstbestendig te houden. In navolgende paragrafen beschrijven we deze ambities, die een sterke invloed hebben op de wijze waarop met onderzoeksdata wordt omgegaan.

### 5.1 Meer internationale samenwerking

Vernieuwing en innovatie zijn belangrijke drijfveren voor economische ontwikkeling. Aan deze vernieuwing en innovatie ligt veelal onderzoek ten grondslag. Onderzoek vormt daarmee een onmisbaar element voor economische ontwikkeling. Tegelijkertijd bevinden onderzoeksvraagstukken zich meer en meer op Europees of internationaal niveau. Om aan te haken bij deze ontwikkelingen, wil Nederland meer samenwerking op Europees niveau realiseren op het gebied van onderzoek en onderzoeksdata. Zo kan ons land, zeker gelet op de grote investeringen en inspanningen die de VS en Azië op dit gebied doen, mee blijven doen in de internationale top van onderzoek en de continuïteit van onderzoek waarborgen.

### 5.2 Meer regionale en nationale samenwerking

De HO-instellingen willen meer regionale en nationale samenwerking. Instellingen spelen in de toekomst op regionaal niveau een belangrijkere rol daar als het gaat om het aanjagen van kennisontwikkeling en innovatie. Steden en regio's krijgen steeds meer een eigen profiel, onderzoek en maatschappelijke vraagstukken sluiten steeds meer op elkaar aan. Dit vraagt om samenwerking tussen instellingen, bedrijven en maatschappelijke organisaties op regionaal niveau en een goede verankering van de instelling in de regio. De regio is hierbij niet beperkt tot een gebied binnen Nederland, maar kan ook een gebied binnen Europa of een werelddeel betreffen. Voorbeelden van genoemde ontwikkelingen zijn de initiatieven rondom het uitbouwen van praktijkgericht onderzoek en de ontwikkeling van Digital Data Competence Centers en Centers of Expertise. De instellingen vormen daarbij de spil van een slimme regio.

Onderzoek in Nederland is effectief door slimme samenwerkingscombinaties tussen diverse partijen. Om bijvoorbeeld gebruik te kunnen maken van benodigde geavanceerde apparatuur, is samenwerking met hoogtechnologische bedrijven onmisbaar. Om overlap tegen te gaan, meer synergie te creëren en onderlinge competitie te verminderen, is meer regionale en nationale samenwerking op het gebied van Research Data Management nodig.

---

<sup>2</sup> *Strategische agenda hoger onderwijs en onderzoek – Houdbaar voor de toekomst*, Ministerie van OCW, Den Haag, december 2019

### 5.3 Aansluiting op samenleving en onderwijs

Als hoeksteen van de samenleving dient onderzoek zich continu aan te sluiten bij deze samenleving. Dit om resultaten van onderzoek te kunnen omzetten naar bruikbare en toepasbare oplossingen voor maatschappelijke en economische vernieuwing. Aansluiting draagt ook bij aan het ontdekken van nieuwe onderzoeksvraagstukken.

De aansluiting tussen onderwijs en onderzoek in Nederland is van oudsher goed ontwikkeld. Studenten worden in Nederland in staat gesteld een onderzoekende houding en creatief denken te ontwikkelen. De sterke groei van het aantal studenten zet deze aansluiting echter onder druk, bijvoorbeeld als het gaat om de bekostiging van zowel onderwijs als onderzoek.

De ambitie is blijvende aansluiting te creëren bij de samenleving en het onderwijs. Dit onder andere door inzicht te bieden in maatschappelijke en economische ontwikkeling en deze te verbinden aan onderzoeksthema's. Daarmee wordt aanhaking van onderzoek op het landelijk strategisch kennisbeleid en portfoliomanagement op onderzoeksthema's ondersteund.

### 5.4 Open Science

Nederland heeft de ambitie om het onderzoeksproces open te maken. Onder de noemer Open Science wordt ernaar gestreefd de beschikbaarheid voor (her)gebruik uit te breiden, de deelname aan onderzoek te verbreden en de toepasbaarheid van onderzoek te vergroten. Wat begon met het actief openbaar maken van de eindpublicaties, worden in de toekomst het gehele proces inclusief de onderzoeksdata transparant en dragen burgers en professionals bij aan onderzoek via Citizen Science.

De ontwikkelingen richting Open Science veranderen het proces van onderzoek en maken dit transparanter. Anderen krijgen de gelegenheid mee te werken, bij te dragen aan en gebruik te maken van het wetenschappelijk proces. Zo kunnen gebruikers van buiten de wetenschap de onderzoekswereld beïnvloeden met vragen en ideeën en onderzoeksdata helpen verzamelen.

### 5.5 Borgen van kwaliteit van onderzoek

Onderzoek moet internationaal competitief en toonaangevend blijven. Daarvoor is het van belang dat de kwaliteit van onderzoek gewaarborgd blijft. Data vormen een belangrijke hoeksteen voor kwalitatief onderzoek.

Hergebruik van onderzoeksdata zal helpen de kwaliteit van onderzoek te waarborgen. Toch is hergebruik van goede onderzoeksdata niet vanzelfsprekend, onder andere omdat niet alle onderzoeksdata vindbaar, toegankelijk, uitwisselbaar en herbruikbaar zijn gemaakt. Door het expliciteren van het gebruik van FAIR (Findable, Accessible, Interoperable & Reusable)-principes en specifiek het aansluiten op internationale Linked Open Data-initiatieven, maar ook het inrichten van goed datamanagement (data stewardship) wordt hergebruik eenvoudiger gemaakt.

De ambitie is om de kwaliteit van onderzoek te waarborgen, onder andere door het expliciet toepassen van FAIR-principes en het inrichten van goed data stewardship voor onderzoek of daaruit voortkomende data.

### 5.6 Borgen onafhankelijkheid HO

In Nederland moet de onafhankelijkheid van onderzoek gewaarborgd blijven. Instellingen zijn in toenemende mate afhankelijk van commerciële platforms, die voor de individuele wetenschapper aantrekkelijk zijn omdat ze makkelijk toegankelijk zijn en weinig kosten. Hierdoor wordt onderzoek steeds afhankelijker van commerciële bedrijven, die veelal buiten Europa opereren (zie ook WTR-rapport *Samenwerking op service-*

*infrastructuur en competenties voor data*). De rectoren van de Nederlandse universiteiten hebben aangegeven dat het onwenselijk is wanneer onderwijs en onderzoek volledig in handen zijn van grote commerciële partijen. Het onderzoek in Nederland mag in hun ogen niet te zeer afhankelijk worden van deze platforms.

### **5.7 Continuïteit van onderzoek**

De toegenomen digitale ondersteuning van onderzoek brengt ook continuïteitsrisico's voor onderzoeksdata met zich mee. Instellingen zijn bijvoorbeeld kwetsbaar voor cybercriminaliteit, waarbij onderzoeksdata kunnen worden gecompromitteerd of onderzoeksinformatie kan worden gestolen. Het is daarom zaak dergelijke continuïteitsrisico's op het gebied van onderzoeksdata te verminderen door onderzoeksdata veilig te stellen en de beschikbaarheid ervan te garanderen.

## 6 Technologische ontwikkelingen

De digitalisering van onze samenleving zorgt voor een exponentiële toename in hoeveelheden onderzoeksdata. Deze toename in onderzoeksdata in combinatie met door technologie gedreven ontwikkelingen heeft invloed op onderzoek en specifiek het Research Data Management-vraagstuk. In navolgende paragrafen beschrijven we een aantal geïdentificeerde ontwikkelingen en hun impact op het RDM-vraagstuk.

### 6.1 Internet of Things

De ontwikkeling van IoT zal bijdragen aan een sterke toename van de hoeveelheid onderzoeksdata. Veel apparaten en voorzieningen worden uitgerust met sensoren en vormen gezamenlijk grote netwerken die allerlei verschillende omgevingsdata genereren. IoT-devices kunnen worden toegepast om data te genereren waarop onderzoek kan plaatsvinden. Netwerken van bijvoorbeeld slimme steden zijn in de toekomst te gebruiken voor sociaal onderzoek naar hoe mensen leven en werken. Ook zijn de ontwikkelingen rondom IoT te combineren met artificial intelligence om zo trends of patronen te herkennen. Daarnaast kan IoT bijdragen aan validatie van data door te fungeren als referentie en door signalering van trends.

### 6.2 Artificial intelligence

Uiteenlopende vormen van artificial intelligence (AI) bieden zeer krachtige hulpmiddelen voor alle facetten van het proces van onderzoek. Denk aan het verzamelen van data rond een specifieke onderzoeksvraag, het ontdekken van patronen in onderzoeksdata of het geautomatiseerd schrijven van publicaties. Hierbij is het van belang steeds te kunnen vaststellen op welke wijze AI met de data omgaat.

Ook diverse andere landen zien AI als een cruciale voorziening in de onderlinge concurrentie. Verschillende initiatieven in Nederland en Europa spelen hierop in, zoals het AI-ROBOTICS-initiatief van de Europese Commissie of de ontwikkeling van een snuffelrobot door Shell in Pernis.

### 6.3 Blockchain

Hoewel blockchain nog volop in ontwikkeling is, biedt deze technologie mogelijkheden om vertrouwen toe te voegen aan de informatievoorziening van veel facetten in de maatschappij. Waar vertrouwen voorheen gebaseerd was op institutionele instanties of rechtspersonen, zorgt blockchain dat deze instanties en rechtspersonen niet meer of minder nodig zijn. In diverse sectoren wordt breed geëxperimenteerd met het vinden van toepassingen. Zo heeft de financiële sector behoorlijk geïnvesteerd in blockchain, maar heeft dit nog geen breed gebruikte toepassing opgeleverd. Blockchain biedt veel mogelijkheden voor diverse onderdelen van het onderzoeksproces en mogelijkheden om betrouwbaarheid van onderzoeksdata te waarborgen. Daarnaast kan deze technologie ingezet worden voor het borgen van de betrouwbaarheid van het onderzoeksproces en de onderzoeksdata.

### 6.4 Data Science

Data Science biedt met technologie als Data Lakes, Scalable Data Analytics nieuwe mogelijkheden voor onderzoek, waardoor dit meer datagestuurd kan worden. De groei van het aantal data heeft meer waarde naarmate data te combineren zijn. Technologische ontwikkelingen als data analytics kunnen hierbij helpen en bijdragen aan het genereren van nieuwe onderzoeksdata uit bestaande data.

### 6.5 Computing

Rekenkracht blijft zich doorontwikkelen. De ontwikkelingen op het gebied van quantum computing, parallel computing en distributed computing gaan eraan bijdragen dat meerdere scenario's verwerkt kunnen worden.

Ook zullen hierdoor andere vormen van onderzoek gaan ontstaan, wat zijn weerslag zal hebben op data ten behoeve van onderzoek.

## 6.6 Trusted (third) parties

TTP's kunnen een rol spelen bij het beheer, de verwerking (bijvoorbeeld pseudonimisatie, versleuteling) en de uitwisseling van onderzoeksdata tussen partijen. De TTP zorgt voor naleving van wet- en regelgeving rond privacy van onderzoeksdata en databescherming. Verder kunnen TIP's bijdragen aan vertrouwen in de wetenschap door op internationale normeringen te certificeren.

## 6.7 Platforms en infrastructuur

Onderzoek wordt steeds vaker uitgevoerd door meerdere onderzoekers. Daarbij is het van belang dat internationale onderzoeksresultaten goed uitgewisseld kunnen worden. Virtual Research Environments (VRE's) en datamanagementplatforms bieden (online) werkomgevingen die specifiek zijn gericht op onderzoekers, vaak inclusief opslag en distributie. Het kunnen gebruiken van VRE's met internationale standaarden vereenvoudigt de benodigde samenwerking.

Daarnaast bieden concepten als Hybrid Cloud Infrastructuur flexibiliteit en schaalbaarheid voor bijvoorbeeld dataopslag, rekenkracht en bandbreedte. De hybride inrichting kan onderzoek versnellen, daar waar nodig in combinatie met de vereiste vertrouwelijkheid en integriteit.

## 7 Architectuurvisie

In diverse rapporten wordt de noodzaak tot meer samenhang in de ICT-voorzieningen en infrastructuur voor research datamanagement benoemd. NWO geeft in een publicatie aan dat er behoefte is aan meer samenhang tussen de diverse sectorvoorzieningen voor RDM en dat CIO's van de instellingen en SURF het voortouw zouden moeten nemen. Ook het Nationaal Programma Open Science stelt in haar eindrapport dat een gezamenlijke architectuur ontwikkeld zou moeten worden. Deze behoefte aan architectuur speelt niet alleen nationaal. Volgens bijvoorbeeld EU-commissaris Carlos Moedas is er niet een tekort aan initiatieven, maar ontbreekt er wel een soort one-stop-shop voor onderzoekers en is er geen architectuur die onderzoekers in staat stelt om aan te haken. Maar hoe zou een dergelijke architectuur eruit moeten zien?

### 7.1 Architectuurvisie

HOSA biedt een architectuurvisie voor de HO-sector met een horizon van drie tot tien jaar vooruit. Kwalitatief en betrouwbaar onderzoek is cruciaal voor de ontwikkeling van de samenleving. Door digitalisering verandert onderzoek continu. Die veranderingen hebben impact op de wijze waarop met onderzoeksdata wordt omgegaan en wat voor ondersteuning vanuit sectorvoorzieningen nodig is. Deze architectuurvisie is gebaseerd op initiatieven en ambities van de sector, de technologische ontwikkelingen en de aandachtspunten in de huidige situatie.

Nederland heeft een aantal ambities die worden gefaciliteerd vanuit onderzoek van universiteiten en hogescholen. Bijvoorbeeld om van ons land een meer data gedreven kenniseconomie te maken, zoals benoemd in het Kennis- en Innovatieconvenant. Of innovatie in regio's, waarbij deze regio's graag samen met HO-instellingen de regio smart maken om zo beter aan te sluiten op nieuwe ontwikkelingen. OCW heeft met de HO-sector een aantal doelstellingen geformuleerd. Ook organisaties als VSNU, VH, NWO, KNAW, DANS en SURF hebben doelstellingen geformuleerd. In de HOSA zijn deze doelstellingen verbonden en weergegeven in de doelenstructuur. Sectorvoorzieningen die zijn gebaseerd op principes en uitgangspunten die passen bij deze doelen zijn vereist om deze doelstellingen te halen. De HOSA heeft deze uitgewerkt tot een kader.

#### Samenwerking

De economische en maatschappelijke waarde van onderzoek neemt toe. Onderzoek is een concurrerende factor. Tegelijkertijd nemen de kosten van onderzoek sterk toe. Voor een relatief klein land als Nederland wordt het zowel financieel als organisatorisch moeilijker om kwalitatief hoogwaardig onderzoek te blijven uitvoeren. Daarom is samenwerking op regionaal, nationaal, Europees en mondiaal niveau noodzakelijk. Krachtenbundeling op regionaal, nationaal en internationaal niveau vergroot de efficiëntie en effectiviteit, doordat onderzoeksdata en -voorzieningen met elkaar worden gedeeld. Samenwerking op regionaal, nationaal en Europees niveau gaat een belangrijke rol spelen in het vergroten van de concurrerende positie van onderzoek in Nederland, en ook in het behoud van hoogwaardig onderzoek in ons land. Daarnaast neemt de complexiteit van maatschappelijke vraagstukken toe. Dit vraagt om bijvoorbeeld meer interdisciplinaire samenwerking tussen diverse vakgebieden. Hierdoor ontstaat de noodzaak om op termijn veel meer verbinding te leggen tussen infrastructuur die nu veelal per discipline zijn georganiseerd.

#### Open onderzoek

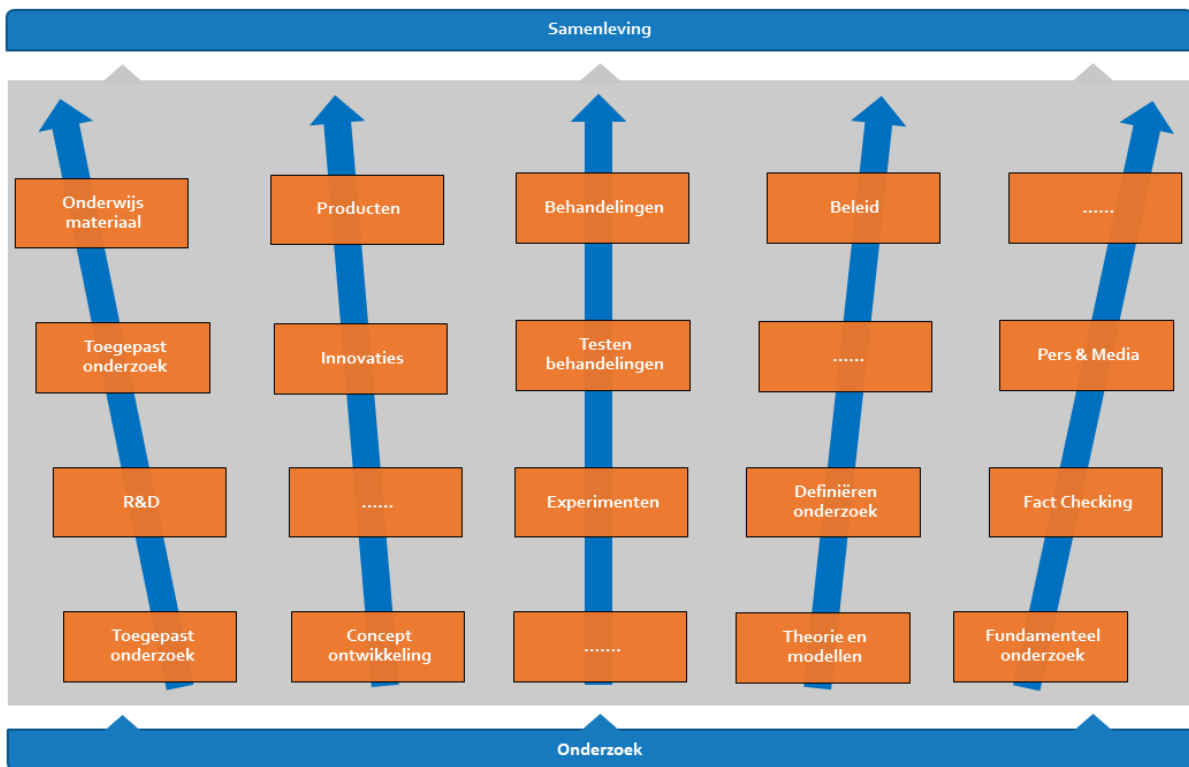
De eisen die gesteld worden aan de openheid en transparantie van onderzoek nemen toe. Niet alleen de beweging naar Open Science speelt hier een aanjagende rol in, maar ook de wens om onderzoek sneller tot resultaat te laten komen. Dit zal ertoe leiden dat onderzoek niet meer alleen binnen de muren van een onderzoeksinstituut zal plaatsvinden. Was voorheen het onderzoeksproces nog minder open, met als eindresultaat een artikel dat beschikbaar werd gesteld, straks worden tussenstappen en de onderliggende data in toenemende mate open beschikbaar gesteld. Met behulp van sectorvoorzieningen wordt Open Science aangejaagd en versterkt.



Het effect hiervan zal zijn dat gespecialiseerde partijen kunnen participeren in het onderzoek. Zo ontstaan verschillende samenwerkingsverbanden met wisselende samenstelling van zowel publieke als private partijen. Bedrijven, ziekenhuizen, media, overheden nemen hieraan deel, door data aan elkaar beschikbaar te stellen of actief bij te dragen aan onderzoek. Een voorbeeld van deze samenwerking en een beweging naar open data op dit moment is de doelstelling van de Nederlandse overheid om alle publieke databestanden actief beschikbaar te stellen voor gebruik in onderzoek. Hierdoor ontstaat een situatie waarin onderzoekers hun databestanden steeds vaker wederzijds beschikbaar stellen (met inachtneming van de privacyregels). Deze onderzoeksdata worden vervolgens weer toegepast bij de vorming van beleid of gebruikt bij ander onderzoek.

Om deze ontwikkelingen te faciliteren en tegelijkertijd aspecten als doelmatigheid, veiligheid en betrouwbaarheid te waarborgen, zal sectoraal moeten worden samengewerkt en zullen sectorvoorzieningen ontstaan die de beweging naar open onderzoek maximaal ondersteunen over de grenzen van organisaties heen. Waar voorheen de focus lag op de universiteiten, hogescholen en onderzoeksinstituten komen nu bijvoorbeeld ook overheden, bedrijven en ziekenhuizen in beeld. Hierbij gaat het niet alleen om het managen van de onderzoeksdata, maar ook om heldere en transparante dienstverlening richting betrokkenen in het onderzoeksproces.

**Verschuiving van focus**



Figuur 2: Verandering van focus

De snelheid waarmee de samenleving verandert door de toegenomen digitalisering en globalisering, leidt tot toenemende druk op onderzoek om sneller tot bruikbare resultaten te komen, bijvoorbeeld in de vorm van toepasbare oplossingen, bruikbare kennis en toepasbare informatie. De behoefte ontstaat om de time-to-

markt van onderzoek te verkorten. Daarmee verschuift ook de focus van onderzoek. Niet alleen het onderzoek staat centraal, maar ook de ontvanger van het onderzoeksresultaat. Dit onderzoeksresultaat of tussenresultaat zal worden afgestemd op de ontvanger en de toepasbaarheid van onderzoek versnellen. De ontvanger of gebruiker van het onderzoeksresultaat vervult uiteenlopende rollen, variërend van professor tot journalist tot ondernemer en burger.

De COVID-19-pandemie heeft deze verandering al sterk geïllustreerd. De behoefte aan een snel en toch veilig en betrouwbaar vaccin vraagt om een andere aanpak van het onderzoeksproces, een veel opener proces waarin vaker kennis of tussenresultaten worden gedeeld en beschikbaar worden gesteld. Dit beschikbaar stellen van kennis of tussenresultaten vraagt om een onderliggende infrastructuur die het mogelijk maakt kennis over tussenresultaten eenvoudig beschikbaar te stellen aan vele doelgroepen, zodat deze regionaal, nationaal en internationaal kan worden toegepast.

Deze focusverschuiving zal andere eisen stellen aan de wijze waarop met onderzoeksdata wordt omgegaan. Onderzoeksdata, waaronder dus ook resultaten en tussenresultaten, moeten worden afgestemd op de ontvanger. Classificatie van onderzoeksdata wordt daarmee belangrijker. Ethiek en privacy moeten worden bewaakt. Sectorale samenwerking ondersteunt en faciliteert deze ontwikkelingen.

### **Maatschappelijke participatie**

Onder de noemer van Citizen Science dragen burgers in verschillende rollen en op verschillende niveaus actief bij aan onderzoek. Voorbeelden van Citizen Science zijn onder andere Hanny's Voorwerp, een reflectieniveau in 2007 ontdekt door de Nederlandse onderwijzeres Hanny van Arkel uit Heerlen, en de tekenradar ([www.tekenradar.nl](http://www.tekenradar.nl)) waar informatie over de ziekte van Lyme wordt gedeeld. Een recenter voorbeeld betreft de wijze waarop het RIVM informatie en onderzoeksdata deelt over de ontwikkelingen rondom COVID-19. Op basis van feedback op deze, door het RIVM gedeelde data, worden steeds meer aanvullende data beschikbaar gesteld waarmee onderzoekers, in de brede zin van het woord, zelf aan de slag kunnen en nieuwe inzichten kunnen ontwikkelen en toevoegen. Dit soort ontwikkelingen zal steeds meer voorzien in de behoefte van burgers om beschikbare data en informatie op juistheid te kunnen controleren of actief bij te dragen aan onderzoek. Sectorvoorzieningen van de toekomst zouden in staat moeten zijn deze verscheidenheid aan rollen en niveaus van participatie en informatiebehoefte te ondersteunen. De participatie van burgers in onderzoek roept wel nieuwe vragen op rondom sectorvoorzieningen. Hoe toegankelijk moeten ze bijvoorbeeld zijn? Welk kennisniveau is vereist? En hoe regelen we toegang tot gegevens die enige afscherming vereisen?

### **Specialisatie in deelprocessen van onderzoek**

De beweging naar open onderzoek leidt mogelijk tot een meer modulair onderzoekproces. Waar voorheen vooral gekeken werd naar het aantal publicaties en de impact hiervan op basis van citatiescore van het journal waarin publicaties verschijnen, wordt het ontwikkelen van hoogwaardige onderzoeksdata of valorisatie van onderzoek in toenemende mate gewaardeerd. Partijen kunnen bijvoorbeeld gespecialiseerd zijn in het creëren van hoogwaardige onderzoeksdata of het verwerken daarvan. Toekomstige RDM-platforms dragen hieraan bij door deze processen onafhankelijk van elkaar te kunnen laten uitvoeren. Modularisatie en de daaruit voortvloeiende specialisatie van onderzoekers zijn uitgangspunten van de sectorarchitectuur.

Voor iedere fase van het onderzoek kunnen deze gespecialiseerde partijen een bijdrage leveren. Elke fase van het onderzoeksproces vraagt om andere samenwerkingsverbanden. Resultaten worden sneller, open en transparant beschikbaar gesteld en vormen vervolgens weer input voor een vervolgstap in het onderzoek of worden toegepast in een ander onderzoek. Sectorvoorzieningen zouden dit op eenzelfde wijze moeten kunnen ondersteunen.

## Digitalisering

Samenlevingen leunen meer en meer op informatie en informatietechnologie. Bedrijven, overheden en instellingen transformeren naar digitale organisaties, waarin digitale informatie en informatietechnologie de eindproducten zijn. Processen worden dan volledig digitaal uitgevoerd. Veranderingen en ontwikkelingen worden ingericht en gestuurd op basis van de data die daarbij ontstaan en worden verzameld. De verregaande digitalisering van onze samenleving levert grote hoeveelheden data op. Deze data zijn ook van waarde voor onderzoek. De grote hoeveelheden data maken dat factoren als vindbaarheid, toegankelijkheid, uitwisselbaarheid en herbruikbaarheid belangrijk zijn om onderzoek efficiënt en effectief te houden. Digitalisering zal leiden tot een grote vraag naar voorzieningen die het werken met deze data goed ondersteunen.

## 7.2 Uitgangspunten

Genoemde ontwikkelingen vinden plaats in een decennium dat ook wel wordt aangeduid als de digital age. Veel bedrijven maken hierin de transitie naar digital companies. Het eindproduct en de bijbehorende services zijn primair digitaal en bestaan in belangrijke mate uit bijvoorbeeld data of software. Ook de ondersteunende processen zijn volledig gedigitaliseerd. HO-instellingen en de sector zullen deze ontwikkeling eveneens doormaken. Bepaalde handmatige handelingen verdwijnen en ook advisering zal niet altijd meer door mensen gebeuren. Ook zullen digitale concepten en technologieën hun intrede doen in de sector, zoals businessplatforms, AI en blockchain.

### Internationale samenwerking

Wetenschap is veelal internationaal georiënteerd en georganiseerd. De EU investeert bijvoorbeeld breed in Europese samenwerking om zo concurrentie uit de VS en Azië het hoofd te bieden. Is het nog wel nodig om Nederlandse sectorvoorzieningen te organiseren? Vanuit de belangen van Nederlandse burgers, bedrijven, regio's en andere actoren is deze behoefte wel zichtbaar. Dit betekent allereerst dat HOSA moet kunnen aansluiten op Europese initiatieven en standaarden. Daarnaast moeten initiatieven vanuit Nederland hergebruikt of opgeschaald kunnen worden in Europese landen. Dit is benoemd in de overkoepelende principes van de HOSA. Voor gegevensuitwisseling met het buitenland moeten integratiepatronen worden afgesproken.

### Innovatie en autonomie

Eén van de vraagstukken die in architectuur vaak naar voren komt is in welke mate onderliggende technologie voor iedereen gestandaardiseerd moet worden. Het idee erachter is dat hiermee kosten bespaard kunnen worden en interoperabiliteit verhoogd kan worden. In het veld van onderzoek is dit onrealistisch. HOSA doet daarom geen uitspraken over de technologie die onderzoekers, onderzoeksgroepen of vakgebieden zelf gebruiken. Innovatie in onderzoek zou verminderen en autonomie van de wetenschap zou in het geding komen. De HOSA geeft wel richting aan het verbinden van de vele partijen en vakgebieden die bij onderzoek betrokken zijn, de afnemers van onderzoeksresultaten en de intermediaire technologie die nodig is om de partijen te verbinden.

### Platformstrategie

De ontwikkelingen in onderwijs en onderzoek vragen in toenemende mate om uitwisseling tussen verschillende partijen, zoals studenten, docenten, onderzoekers, bedrijven, overheden en burgers. Het gaat dan om uitwisseling van bijvoorbeeld content, data en voorzieningen. De HOSA maakt daarom bij het opstellen van de applicatiearchitectuur voor de sectorvoorzieningen gebruik van het concept van businessplatforms. Dit concept geeft inzicht in zogenaamde online marktplaatsen, die mede door toenemende samenwerking in de sector op diverse plekken ontstaan.

Businessplatforms vinden hun oorsprong bij de techreuzen die dienstverlening op een andere wijze vormgeven. Voor de HO-sector benoemt de HOSA vergelijkbare platforms, maar dan door de bril van de publieke sector en publieke waarden. Deze 'digital public sector platforms' faciliteren interactie en samenwerking, brengen vraag en aanbod bij elkaar en sturen op gezamenlijke waarden en normen. Hiermee versterken we onafhankelijkheid van onderzoek.

### **Sectorvoorzieningen ondersteunen verschillende typen samenwerking**

Binnen de doelen op het gebied van onderzoek komt samenwerking op vele niveaus terug: tussen instellingen lokaal in de regio, internationaal, binnen vakgebieden of met het bedrijfsleven. Voor instellingen is het ook steeds de vraag hoe ze hun samenwerkingsverbanden op de verschillende niveaus goed ondersteund krijgen. Op dit moment worden deze samenwerkingsverbanden veelal per geval ingericht, met als nadeel dat ze moeilijk schaalbaar en interoperabel zijn. Met het oog op de ambities van de HO-sector op het gebied van samenwerking is deze schaalbaarheid wel vereist. HOSA maakt de sectorvoorzieningen geschikt voor samenwerking op genoemde niveaus door ervoor te zorgen dat dezelfde services tegelijkertijd aangeboden kunnen worden in portalen op het niveau van een instelling, een samenwerkingsverband, een vakgebied en een nationaal of internationaal initiatief.

### **Diversiteit in interfaces en kanalen**

Door het uitgangspunt dat sectorvoorzieningen samenwerkingsverbanden op diverse niveaus ondersteunen, ontstaan op termijn meerdere portalen die gebruik maken van onderliggende services vanuit sectorvoorzieningen. Denk aan een landelijk portaal of een portaal van een instelling. Door technologische ontwikkelingen zoals omnichannel en multiple user interfaces zal de gebruikelijke interactie via tekst in een portaal over enkele jaren onvoldoende zijn. Daarom worden sectorvoorzieningen zo opgezet dat ze interactie via een grote diversiteit aan interfaces en kanalen ondersteunen, waaronder chatbots, mobile, gebaren en gaming.

### **Gegevensuitwisseling**

Om gegevensuitwisseling ook echt mogelijk te maken, is het belangrijk afspraken te maken over gegevens van onderzoek over de instellingen heen, en gegevens over bijvoorbeeld datasets, publicaties en onderzoeksprojecten te standaardiseren. Gegevensuitwisseling zal plaatsvinden tussen vele partijen waaronder publieke, private en buitenlandse instellingen. De sector zal actief participeren in het ontwikkelen van internationale standaarden en kennis hierover opbouwen. Hiermee borgt de sector niet alleen internationale, maar ook nationale samenwerking op de verschillende niveaus.

### **Financiële verrekening**

In het kader van de flexibilisering houdt de HOSA rekening met de verwachte behoefte aan verrekening. Daarbij wil de HOSA alle verrekening binnen het onderzoeksdomein op een vergelijkbare wijze afhandelen. Eventuele onderlinge verrekening tussen HO-instellingen, passend bij het gekozen modulaire niveau, zou binnen HOSA gefaciliteerd moeten kunnen worden. Dit rekening houdend met de specialisatie en differentiatie van onderzoek binnen de sector. Ook kan een sectorvoorziening naar verwachting ondersteuning bieden bij het onderling toerekenen van kosten tussen instellingen en een flexibel onderzoeksbudget.

## 8 Platformstrategie

### 8.1 Wat is een platform?

Wereldwijd is er een ontwikkeling gaande waarin het concept van businessplatforms centraal staat. Bekende voorbeelden zijn Uber, Bol en AirBnB. Steeds meer traditionele bedrijven zoals General Electric, Philips en Nike laten zich inspireren door dit model en hebben eigenschappen van businessplatforms overgenomen. Ook in de publieke sector zijn platforms in opkomst, zoals EDx, Coursera, Amazon Ignite en Woolf in het onderwijsdomein. Deze platforms bieden een omgeving die waardecreërende interacties tussen producenten en consumenten mogelijk maakt. Een platform biedt open, toegankelijke faciliteiten voor deze interacties en regelt de onderliggende governance.

Het overkoepelende doel van de platforms is het koppelen van vraag en aanbod bij afnemers en aanbieders van diensten, content, kennis en data, en het vergemakkelijken van de uitwisseling hiervan. Businessplatforms kunnen hun dienstverlening makkelijk opschalen doordat ze de diensten, data, kennis en content veelal niet zelf produceren. Uber heeft bijvoorbeeld zelf geen taxi's en taxichauffeurs en hoeft ze dus niet zelf aan te schaffen of aan te nemen als er meer vraag is. Ook hoeft Uber geen onderhoud van een wagenpark te organiseren. En Coursera produceert zelf geen onderwijs, maar brengt vraag van lerenden en aanbod van andere organisaties bij elkaar.

Een platform integreert functies en competenties die nodig zijn om diensten te leveren aan gebruikers. Het platform kan helpen bij het vinden van het juiste aanbod, vereenvoudigt het aanbieden van nieuwe diensten, ondersteunt de transactie tussen aanbieder en afnemer, monitort gebruik en bewaakt de kwaliteit. Het platform vergemakkelijkt de interactie tussen aanbieder en afnemer. Een businessplatform biedt een open, samenwerkingsgerichte infrastructuur voor interacties tussen de op het platform opererende partijen met een governance die een gezond ecosysteem waarborgt.

### 8.2 Waarom RDM-platforms?

Vanuit de geïnventariseerde strategische ambities in de sector ontstaat er samenwerking op allerlei niveaus, wat leidt tot behoefte aan veel meer interactie. Denk aan meer samenwerking tussen instellingen, internationale samenwerking, samenwerking met bedrijven, samenwerking in de regio en samenwerking met burgers. Dit vereist een onderliggend mechanisme dat invulling kan geven aan de vele interacties en samenwerking op verschillende niveaus. Het concept van businessplatforms biedt hier aanknopingspunten voor.

De ontwikkelingen in onderwijs en onderzoek vragen om uitwisseling tussen allerlei actoren. Het concept van businessplatforms biedt 'marktplaatsen' waar deze uitwisseling centraal staat.

In de huidige praktijk van Research Data Management (RDM) ontstaat reeds een aantal initiatieven die al een beetje richting businessplatforms bewegen, zoals Amsterdam Data Exchange en European Open Science Cloud. Amsterdam Data Exchange heeft als ambitie een infrastructuur aan te bieden, waarbinnen aan de hand van gemeenschappelijke regels een vertrouwde en veilige omgeving voor realtime datagestuurde samenwerking wordt gewaarborgd en waar geïnteresseerde partners zich kunnen aansluiten. European Open Science Cloud wil onderzoeksinfrastructuren koppelen en zo onderzoekers en wetenschappelijke professionals een omgeving bieden om data op te slaan, te delen en te gebruiken.

Paragraaf 4.1 liet al zien dat er op het gebied van RDM veel initiatieven zijn. We constateerden dat hiertussen slechts beperkt samenhang en daarnaast ook overlap bestaat. Dit leidt tot een onoverzichtelijk landschap van RDM-diensten, waarbij de afnemer zelf zijn eigen weg moet vinden, er onnodige concurrentie ontstaat en weinig tot geen zicht en sturing is op het gebied van kwaliteit en betrouwbaarheid. Daarnaast nemen veel initiatieven slechts één of meer delen van het onderzoeksproces als uitgangspunt, terwijl de toekomstige

ontwikkelingen, zoals geschetst in hoofdstuk 7, de waardeketen van onderzoek als uitgangspunt nemen. Deze ontwikkelingen vragen om een onderliggende infrastructuur die de gehele waardeketen van onderzoek geïntegreerd kan ondersteunen.

Het model van businessplatforms binnen de HOSA ondersteunt het creëren van samenhang in de verschillende initiatieven in het domein van onderzoek en ook onderwijs. Het gaat uit van dienstverlening en governance die passen bij de sector en publieke waarden. Het brengt structuur, een gemeenschappelijke taal, overzicht en een checklist bij nieuwe initiatieven. Hierdoor kan het gesprek tussen diverse stakeholders gestructureerder gevoerd worden en kan besluitvorming over afwijkingen van de architectuur expliciet plaatsvinden.

Publieke businessplatforms gaan uit van dienstverlening en governance die passen bij de sector en publieke waarden.

Hierna worden de verschillende platforms van HOSA op hoofdlijnen uitgewerkt. De datamanagementplatforms brengen vraag en aanbod van data- en datamanagementdiensten bijeen, bundelen krachten en jagen samenwerking en innovatie op het gebied van onderzoeksdata aan. Aanbieders wordt een omgeving geboden waarin ze diensten eenvoudig kunnen aansluiten en beschikbaar kunnen stellen aan afnemers. Afnemers wordt een omgeving geboden waarin eenvoudig diensten vindbaar en af te nemen zijn. Er ontstaan mogelijkheden tot het realiseren van efficiency rond zaken als verrekening, licentiëring en ondersteuning.

De platformbenadering sluit aan bij de verwachtingen in de samenleving rond moderne dienstverlening, waaronder gebruiksgemak, transparantie en beleving. Deze benadering draagt via haar governance-inrichting bij aan het waarborgen van onafhankelijkheid van onderzoek. Tot slot biedt de gekozen opzet de mogelijkheid om te schakelen tussen de verschillende niveaus waarop een wetenschapper actief is, zoals internationaal, nationaal, regionaal en binnen onderzoeksdomeinen.

### 8.3 RDM-platforms

Zoals aangegeven in paragraaf 3.3, hebben we de in de HORA 2.1 beschreven onderzoekscyclus en onderzoeksgegevenscyclus als uitgangspunt genomen voor de beschrijving van deze domeinarchitectuur. Vanuit dit perspectief is het onderzoeksproces grotendeels in vier hoofdprocessen in te delen, te weten:

- *Definiëren onderzoek*. Dit hoofdproces richt zich op het in kaart brengen en definiëren van onderzoeksthema's.
- *Vorbereiden onderzoek*. Dit hoofdproces is gericht op het voorbereiden van het beoogde onderzoek.
- *Uitvoeren onderzoek*. Dit hoofdproces richt zich op het deel waarin het daadwerkelijke onderzoek uitgevoerd wordt.
- *Uitnutten kennis*. Bij dit hoofdproces ligt de focus op het aan een breed publiek beschikbaar stellen van onderzoeksresultaten (publicaties, onderzoeksdata, enzovoort).

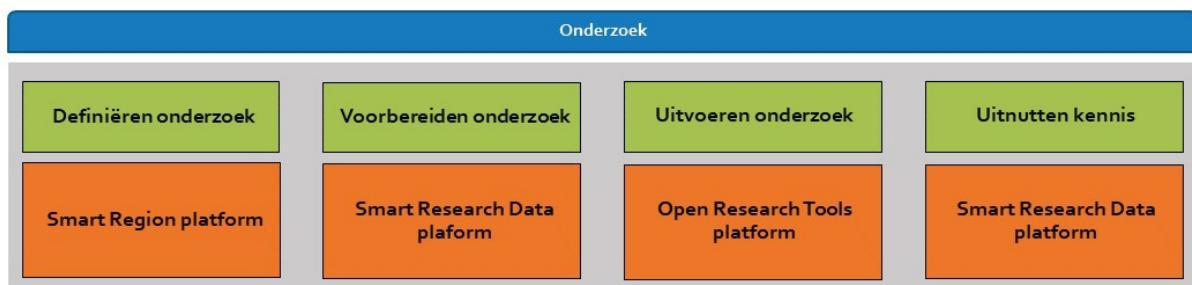
In deze hoofdprocessen worden (een vorm van) onderzoeksdata gecreëerd, uitgewisseld, verwerkt, enzovoort. Drie van de vier zijn duidelijk herkenbaar terug te vinden in de HORA onder het thema 'Onderzoekscyclus'. Het proces 'Vorbereiden onderzoek' heeft een relatie met de bedrijfsfunctie 'Onderzoeksontwikkeling', het proces 'Uitvoeren onderzoek' is gerelateerd aan de bedrijfsfunctie 'Onderzoeksuitvoering' en het proces 'Uitnutten van kennis' heeft een relatie met de bedrijfsfuncties 'Onderzoeksdisseminatie' en 'Kennisuitnutting'. Het proces 'Definiëren onderzoek' is niet direct een onderdeel van de onderzoekscyclus zelf, maar een essentiële, meer strategische processtap om richting te geven aan toekomstig onderzoek.



Figuur 3: Relatie HOSA RDM en HORA

Op basis van de vier hoofdprocessen en de platformstrategie is een inrichting op basis van een drietal datamanagementplatforms te herkennen. Deze platforms richten zich op het ondersteunen van het RDM-proces, dit zijn het:

- Smart Region-platform. Voor de ontwikkeling van slimme regio's, variërend van lokale, nationale tot internationale regio's, is het noodzakelijk gezamenlijk inzicht te ontwikkelen in de kennisbehoefte van de regio's voor nu en in de toekomst. Het Smart Region -platform biedt mogelijkheden om dit inzicht met elkaar te creëren en op elkaar af te stemmen. Op basis van data en data-analyse worden onderzoekthema's in kaart gebracht en gedefinieerd. Het platform biedt mogelijkheden voor onderzoekers om zich te informeren over lopende onderzoekstrajecten.
- Smart Research Data-platform. Instellingen vormen met hun campus een belangrijk onderdeel in de ontwikkeling van de samenleving. Nu al huisvesten allerlei partners zich op of in de buurt van de campus. In de digitale wereld kan de instelling een vergelijkbare rol krijgen. Goede onderzoeksdata zijn essentieel voor het doen van goed onderzoek. Instellingen kunnen met behulp van het Smart Research Data-platform een basis leggen voor de uitwisseling van onderzoeksdata en daarmee voor de ontwikkeling van een digitale campus. Partners zoals bedrijven, overheden en ziekenhuizen kunnen zich vestigen op dit platform. Het Smart Research Data-platform ondersteunt de vraag naar en aanbod van onderzoeksdata en biedt dienstverlening hierop.
- Open Research Tools-platform. Voor het bewerken, analyseren, visualiseren, enzovoort van onderzoeksdata zijn specialistische en veelal kostbare diensten nodig. Het is onwenselijk dat iedere instelling die voor zichzelf inricht, enerzijds vanuit het oogpunt van het betaalbaar houden van onderzoek, anderzijds vanuit het perspectief van samenwerking. Het Open Research Tools-platform brengt diverse RDM-diensten slim bijeen. Instellingen kunnen daarmee niet alleen hun eigen diensten beschikbaar stellen of die van andere instellingen gebruiken, maar ook die van partners. Zo versterken ze de ontwikkeling van de digitale campus en van de regio.



Figuur 4: Platforms HOSA RDM

De platforms ondersteunen het gehele RDM-proces van onderzoeksdata en sluiten naadloos op elkaar aan. In hoofdstuk 10 worden de platforms nader toegelicht en conceptueel beschreven.

## 8.4 Inrichting van een platform

De platformstrategie vraagt om een inrichting die recht doet aan de architectuurvisie.

### *Integer, transparant en betrouwbaar*

Algemeen worden de platforms zodanig ingericht dat ze zich gedragen als integere, transparante en betrouwbare dienstverleners richting de diverse stakeholders zoals burgers, onderzoekers, regio's, publieke instellingen en private partijen.

### *Open en modulaire inrichting*

De RDM-platforms kennen een open en modulaire inrichting, waar verschillende functies op het gebied van RDM bijeenkomen en beschikbaar worden gesteld. De componenten zijn 'loosely coupled', waardoor het platform eenvoudig aan te passen is aan veranderende wensen en eisen van participerende partijen. Het gebruik van open standaarden vormt een belangrijk element: hiermee kunnen bijvoorbeeld datamanagementdiensten eenvoudig beschikbaar worden gesteld. De modulaire opzet en open architectuur stimuleren hergebruik en aansluiting op de platforms, het (ont)koppelen van data en diensten kan eenvoudig plaatsvinden, en ook de interoperabiliteit tussen de platforms.

### *Ontsluiting van diensten voor en door instellingen*

De open en modulaire architectuur in combinatie met het gebruik van standaarden stelt instellingen in staat diensten van het platform via de eigen portalen te ontsluiten en beschikbaar te stellen. Daarnaast kunnen instellingen besluiten via de platforms eigen diensten met specifieke kenmerken (landelijk) als dienst aan te sluiten en beschikbaar te stellen. Zo ondersteunen de platforms de instellingen om de uitstraling van de fysieke samenwerking op de fysieke campus mee te nemen in de digitale wereld. Ook stimuleert dit de ontwikkeling van de digitale campus, wat instellingen mogelijkheden biedt om de spil te gaan vormen in de slimme steden van de toekomst.

### *Internationale aansluiting*

De platforms kunnen ook worden aangesloten op internationale initiatieven of internationaal worden ingezet. Door het hanteren van een gemeenschappelijke basisplaat als onderlegger wordt het eenvoudiger om lokale, nationale en internationale initiatieven op elkaar aan te sluiten en zo een gezamenlijk ecosysteem te ontwikkelen rondom RDM.

### *Intermediaire rol*

De platforms fungeren als een marktplaats voor onderzoeksdata en onderzoeksdatadiensten. Ze brengen vraag en aanbod rond deze data en diensten bij elkaar. De platforms vervullen een intermediaire rol tussen deze verschillende partijen en bieden overkoepelende governance voor het bereiken en behouden van de juiste kwaliteit, veiligheid, betrouwbaarheid en integriteit van de onderzoeksdata en diensten.

De platforms bieden de kern van wat instellingen niet zelf kunnen regelen.

### *Governance*

De inrichting van de governance van deze platforms is zodanig dat aspecten als onafhankelijkheid, transparantie, integriteit en betrouwbaarheid zijn gewaarborgd. Deze governance vertegenwoordigt de belangrijkste stakeholders.

### *Regievoering*

Platformmanagement voert regie over de aangeboden data en diensten en waarborgt kwaliteitskaders en serviceniveaus. Het is gemandateerd en verantwoordelijk voor de ontwikkeling, implementatie en monitoring van kwaliteitsborging, integriteit en betrouwbaarheid van het platform. Het platformmanagement signaleert diensten en data die niet voldoen aan de vastgestelde normen en acteert hierop. Dit vraagt om overkoepelende afspraken en normen, maar ook om mogelijkheden om deze afspraken en normen te



monitoren, en verder het signaleren van benodigde veranderingen aan deze afspraken en normen, bijvoorbeeld als gevolg van veranderingen in landelijke afspraken. Het platformmanagement draagt zorg voor inrichting, (door)ontwikkeling en implementatie van het platform in samenwerking met de belangrijkste stakeholders.

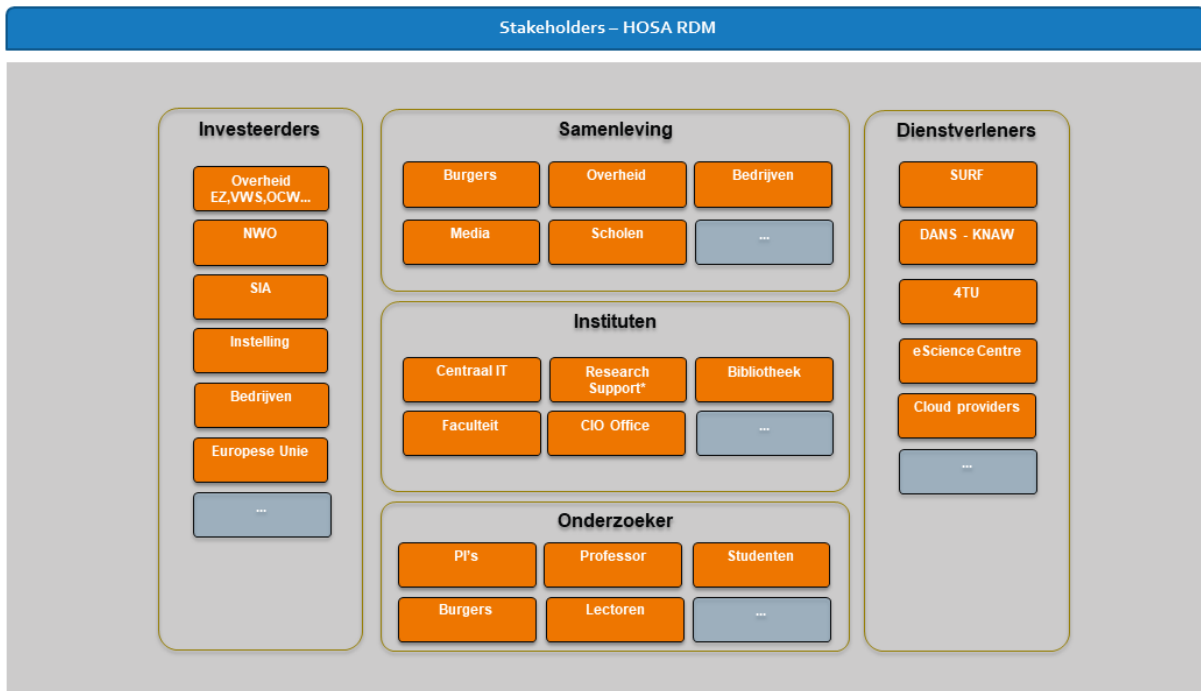
De platformstrategie ondersteunt het ontstaan van een landelijk ecosysteem op het gebied van RDM dat ook internationaal kan worden ingezet of aansluit op internationale initiatieven en ontwikkelingen. De platforms brengen ordening in het weidse landschap van RDM-initiatieven en ondersteunen de ontwikkeling van coherent RDM.

## **8.5 Complementair**

In de internationale context zijn initiatieven te herkennen die analoog zijn aan de hiervoor beschreven platformstrategie of delen daarvan invullen. De platforms in deze domeinarchitectuur zijn hier complementair aan; ze vervangen deze niet. De platformstrategie maakt het daarnaast mogelijk om voorzieningen op zowel regionaal als nationaal of internationaal niveau aan te bieden, waarbij via het platform ondersteuning richting de eindgebruiker is georganiseerd (zie ook beschrijving van de platforms).

## 9 Stakeholders

De stakeholders binnen het HOSA-domein Research Data Management (RDM) zijn op te delen in vijf groepen, te weten: Investeerders, Samenleving, Dienstverleners, Instituten en Onderzoekers. Binnen deze groepen opereren partijen met een, in hoofdlijnen, zelfde soort belang op het gebied van onderzoeksdata. Ze kennen gedeeltelijk een onderlinge afhankelijkheid op het RDM-domein, zonder dat ze daar individueel een doorslaggevende rol spelen. De platforms worden hierbij vormgegeven vanuit samenwerkingsverbanden tussen twee of meer stakeholders binnen of tussen de groepen.



Figuur 5: Stakeholders

N.B. partijen in dit overzicht kunnen meerdere rollen vervullen. Ze fungeren als voorbeeld in de context van stakeholdergroepen.

### 9.1 Samenleving

De eerste groep stakeholders betreft de samenleving. De samenleving profiteert van onderzoek onder andere door nieuwe producten, verrijking van onderwijs of kennis en informatie ten behoeve van besluitvorming. Het gaat hier bijvoorbeeld om burgers, media, overheden, enzovoort.

Deze groep heeft baat bij en wil snel inzicht in transparant, herkenbaar, betrouwbaar en kwalitatief onderzoek. Ook profiteert zij ervan als de betrouwbaarheid van onderzoek eenvoudig is vast te stellen, maar ook als onderzoeksresultaten snel en eenvoudig gevonden en toegepast kunnen worden. De platforms bieden deze betrouwbaarheid, toegankelijkheid en transparantie.

De meerwaarde voor deze stakeholdersgroep ligt in de bijdrage die de platforms gaan leveren aan het stimuleren en ondersteunen van open onderzoek, waarin eenvoudig kan worden geparticipeerd en waar betrouwbare resultaten uit het onderzoek te vinden zijn.

### 9.2 Investeerders

De tweede groep betreft investeerders in onderzoek. Overheden (of daaraan gelieerde organisaties zoals NWO) en bedrijfsleven investeren jaarlijks veel geld in onderzoek. Ze worden echter geconfronteerd met een

sterke toename van de kosten en complexiteit van onderzoek. Tegelijkertijd stijgt de behoefte om de waarde van onderzoek veel sneller te kunnen vermarkten, waardoor de behoefte aan ondersteuning van onderzoekers groeit. Samenwerking is noodzakelijk om onderzoek ook in de toekomst betaalbaar te houden en de complexiteit het hoofd te kunnen bieden. Daarnaast is samenwerking vereist om de wens voor een snelle time-to-market van onderzoek te kunnen vervullen.

Voor investeerders ligt de meerwaarde in het feit dat de platformopzet de samenwerking en bundeling van krachten faciliteert door een open en transparante inrichting, met inachtneming van aspecten als onafhankelijkheid, betrouwbaarheid, privacy, security en integriteit.

### 9.3 Instituten

De derde groep betreffen de instituten, zijnde de instellingen, onderzoeksinstituten en researchcentra. Onderzoek, van fundamenteel tot toegepast onderzoek, wordt voor een groot deel uitgevoerd bij en aangejaagd door deze instituten. Zij vormen daarmee een belangrijke factor voor innovatie in de regio waar ze zijn gesitueerd. Tegelijkertijd wordt het voor deze instituten steeds complexer het gehele onderzoeksproces volledig te blijven ondersteunen. Slimme samenwerkingsverbanden tussen instituten onderling en met private of publieke partijen kunnen die complexiteit beheersbaar maken en tegelijkertijd innovaties aanjagen.

De meerwaarde voor deze instituten ligt erin dat zij via de platforms zowel eigen faciliteiten slimmer beschikbaar kunnen stellen alsook faciliteiten van derden kunnen aantrekken voor eigen onderzoek. Dit voorkomt onder andere dat instituten diensten gaan inrichten die al beschikbaar zijn bij andere instituten. Door het voorkomen van dergelijke dubbelingen ontstaat meer ruimte voor diversiteit en innovatie.

### 9.4 Onderzoekers

Een vierde groep van stakeholders betreft de onderzoekers. Onderzoekers hebben baat bij ondersteuning die hen in staat stelt eenvoudig de faciliteiten, tooling en onderzoeksdata te verkrijgen die nodig zijn voor hun onderzoek. Ook hebben onderzoekers profijt van het uitvoeren van zo min mogelijk administratieve werkzaamheden, zodat ze zich maximaal kunnen richten op het onderzoek zelf.

Onderzoekers zullen meerwaarde ervaren doordat de platforms hen helpen bij het vinden en verkrijgen van de benodigde diensten en onderzoeksdata, maar ook bij het verspreiden van kennis en resultaten die voortkomen uit het onderzoek.

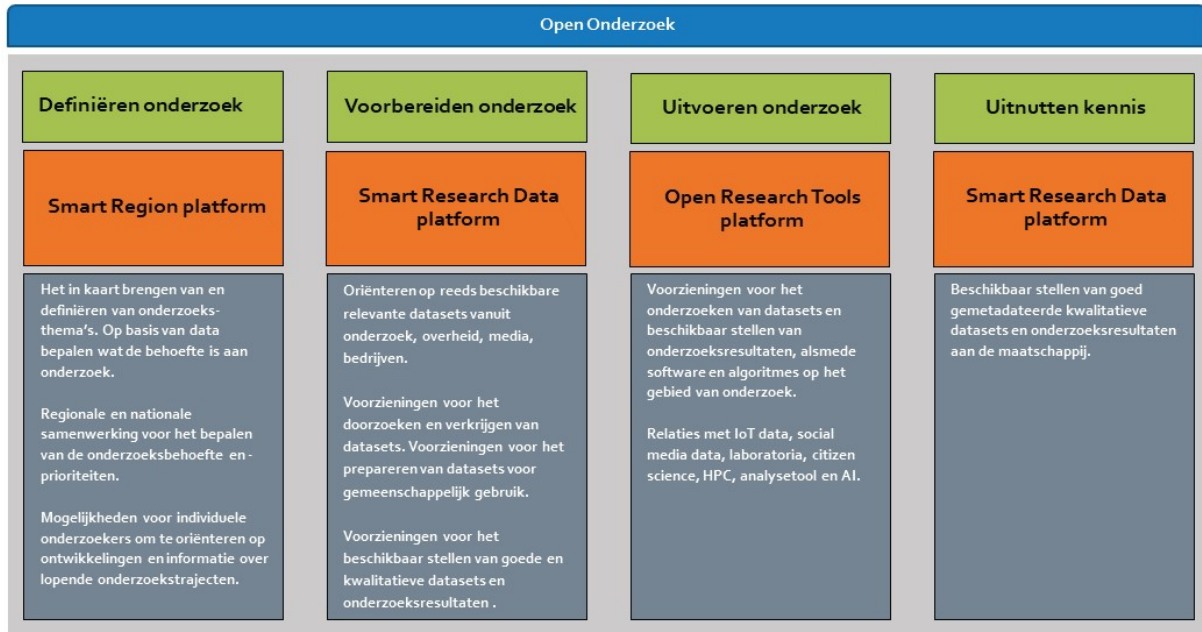
### 9.5 Dienstverleners

De laatste groep stakeholders omvat de dienstverleners rondom onderzoek. Leveranciers van onderzoeksdiensten worden geconfronteerd met een ondoorzichtig landschap aan onderzoeken. Soms is onduidelijk waar behoefte aan is en vormt het gefragmenteerde landschap een barrière in het aanbieden van diensten, soms is de behoefte duidelijk maar is veel maatwerk nodig, wat complex en duur is.

De meerwaarde van de platforminrichting voor dienstverleners is gelegen in het feit dat hun diensten veel eenvoudiger kunnen worden gevonden, afgenomen en samengebracht. Ook ontstaat voor dienstverleners een duidelijke demand-supply situatie.

## 10 Proces- en applicatiearchitectuur

Uitgangspunt voor het beschrijven van de proces- en applicatiearchitectuur zijn de in hoofdstuk 8 beschreven platformbenadering en platformindeling. Ook de in hoofdstuk 7 beschreven veranderingen ten aanzien van het onderzoeksproces, en specifiek de ontwikkelingen rondom open onderzoek, liggen ten grondslag aan de uitwerking van de proces- en applicatiearchitectuur. Dit geeft het volgende beeld:



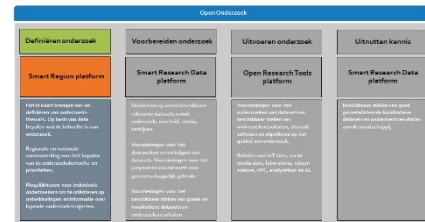
Figuur 6: Processen in relatie tot HOSA-platforms

In navolgende paragrafen beschrijven we kort de hoofdprocessen en geven we aan de hand daarvan inzicht in de conceptuele inrichting van de bijbehorende platforms.

## 10.1 Definiëren onderzoek

### Wat gebeurt er in dit proces?

In het proces ‘Definiëren onderzoek’ worden onderzoeksthema’s in kaart gebracht en gedefinieerd. In de huidige situatie lijkt dit een diffuus en grotendeels niet op data gebaseerd proces, waarbij vele stakeholders betrokken zijn die ‘handmatig’ gegevens uitwisselen. De strategische agenda voor de HO-sector streeft naar betere afbakening en bredere afstemming in en met bijvoorbeeld regio’s. In dit proces werken met name functionarissen van zowel instellingen, sectorpartners als overheden in de rol van beleidsmakers samen, maar ook onderzoekers die zich oriënteren op nieuwe onderzoeksthema’s. Daarnaast zijn ook andere publieke en private sectoren betrokken in het proces. Alle betrokken partijen geven input over de behoefte aan nieuwe kennis over een paar jaar, van waaruit samenwerking op regionaal, landelijk of breder (bijvoorbeeld Europees) niveau ontstaat bij het bepalen van de onderzoeksbehoefte.



Genoemde functionarissen en onderzoekers verzamelen, verwerken, analyseren en visualiseren onderzoeksdata voor het bepalen van de onderzoeksbehoefte. Hierbij wordt op basis van data-analyse en -visualisaties inzicht gecreëerd in de toekomstige kennisbehoefte. Dit inzicht wordt gecombineerd met het inzicht in reeds lopende onderzoeken op gedefinieerde onderzoeksthema’s. Die combinatie stelt functionarissen en onderzoekers in staat onderzoeksthema’s voor de toekomst te definiëren en hier richting aan te geven. Deze behoeftebepaling kan op verschillende niveaus plaatsvinden, variërend van lokaal, regionaal, landelijk tot internationaal.

Waarschijnlijk hebben betrokkenen zowel een vraag naar als een aanbod van relevante onderzoeksdata. In de huidige situatie zoeken partijen voor het bepalen van de onderzoeksbehoefte voornamelijk zelf de benodigde onderzoeksdata bij elkaar en worden deze door henzelf verwerkt, geanalyseerd en gevisualiseerd. Zowel de verzamelde en verwerkte onderzoeksdata als de uitkomsten van de analyses en visualisaties worden vervolgens lokaal opgeslagen. Door hierin gezamenlijk op te trekken, ontstaan minder dubbelingen in onderzoek, worden toekomstige thema’s sterker zichtbaar en wordt samenwerking op onderzoeksthema’s gestimuleerd en zichtbaar.

### Wat is de rol van het platform in dit proces?

OCW benoemt in de strategische agenda de ambitie om meer te werken met sectorplannen. Dit moet onder andere samenwerking en integrale aanpak op onderzoeksthema’s stimuleren. De nadruk bij het Smart Region-platform ligt op het ondersteunen van de samenwerking, de interactie en integrale aanpak bij het definiëren van onderzoeksthema’s, en ook het geven van inzicht in het huidige onderzoeksportfolio. Met behulp van het platform kunnen ontwikkelingen op het gebied van onderzoek landelijk zichtbaar gemaakt worden. Het platform ondersteunt dit door het samenbrengen van faciliteiten die dit proces ondersteunen.

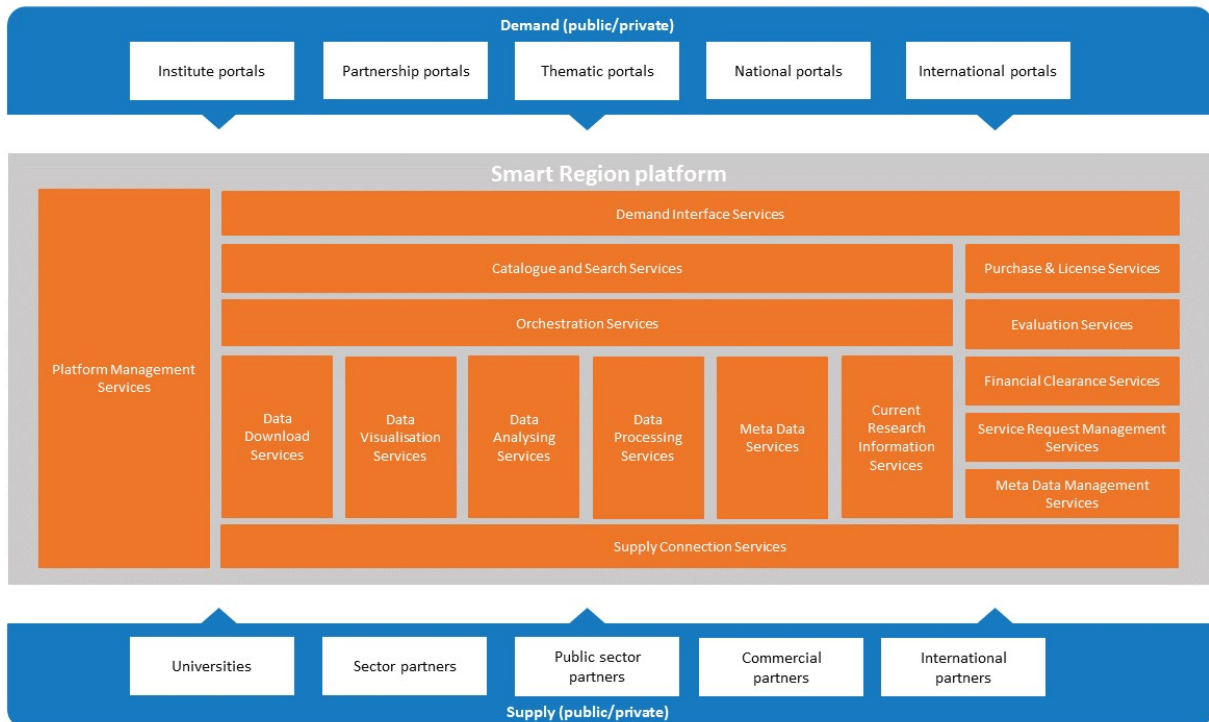
Ook wordt relevante en actuele beleidsinformatie op één openbaar platform beschikbaar gesteld. (Strategische agenda hoger onderwijs en onderzoek – Houdbaar voor de toekomst, Ministerie OCW, Den Haag, december 2019).

Naast het ondersteunen van beleidsontwikkeling rondom onderzoek in het onderzoeksdomein heeft dit platform eenzelfde rol in het definiëren van onderwijs in het domein onderwijs. Betrokken partijen kunnen in brede zin de behoefte aan kennis voor de toekomst in de regio in kaart brengen.

### Uit welke componenten bestaat het platform?

Het Smart Region-platform is ingericht om het in kaart brengen van en definiëren van onderzoeksthema’s te ondersteunen. Het ondersteunt medewerkers in rollen als beleidsmaker, onderzoeker of data scientist bij

instellingen, partners of overheden. Dit platform biedt een omgeving waar partijen data met elkaar kunnen uitwisselen en samenbrengen. Het betreft hier data over onderzoek zelf, maar ook macro-economische en trenddata. Hierbij staan niet alle data op het platform zelf, maar biedt het platform mogelijkheden om data van participerende partijen, beschikbaar in andere bronnen, af te nemen. Het platform kan omgaan met vraag en aanbod rond zowel gestructureerde als ongestructureerde data. Logische aanbieders van data zijn instellingen, NWO, KNAW, OCW en het CBS.



Figuur 7: Smart Region-platform

### Diensten die op het platform worden aangeboden

#### *Data Download Services*

Het platform biedt diensten om gegevens, beschikbaar in andere bronnen, op te halen of mee te koppelen. Bijvoorbeeld uit bronnen van op het platform opererende en participerende partijen.

#### *Processing, Analysing & Visualisation Services*

Op het platform worden zowel gestructureerde als ongestructureerde data aangeboden. Daarvoor biedt het platform voorzieningen aan die medewerkers ondersteunen bij het verwerken van deze data (*Data Processing Services*). Daarnaast verstrekt het platform voorzieningen voor het analyseren van data (*Data Analysing Services*) en het visualiseren van uitkomsten (*Data Visualisation Services*). Deze diensten zijn lichter van aard dan bijvoorbeeld op het Open Research Tools-platform van toepassing is, omdat de data die op het Smart Region-platform worden uitgewisseld minder intensieve bewerking vragen.

#### *Meta-Data Services*

Het platform omvat voorzieningen die aanbieders ondersteunen bij het meta-dateren van hun data (*Meta-Data Services*). Een voorbeeld hiervan zijn diensten voor het valideren van meta-data. De meta-datering maakt het mogelijk om in data te zoeken en verbanden te leggen tussen data. Voor het meta-dateren worden standaarden gehanteerd. Hiervoor wordt governance ingericht. Daarnaast wordt gebruik gemaakt van een community om toe te kunnen zien op de inzet van deze standaarden.

### *Current Research Information Services*

Het creëren van inzicht in het huidige onderzoeksportfolio ondersteunt het platform door koppelingen mogelijk te maken met de Current Research Information Systems van instellingen. Het platform biedt mogelijkheden tot het doorzoeken van deze informatie. Hierdoor ontstaat een samenhangend beeld van de thema's waarop al onderzoek wordt gedaan. Op basis hiervan zijn onafhankelijk landelijke ontwikkelingen zichtbaar te maken.

### Functies die op het platform aanwezig zijn

De gebruikersomgeving richt zich op de professionele medewerker in bijvoorbeeld de rol van beleidsmaker of data scientist bij instellingen, sectorpartners of overheden. Op het platform kan met behulp van een catalogus en zoekfunctie (*Catalogue and Search*) naar gegevens en voorzieningen worden gezocht. Goede filters en zoekalgoritmes ondersteunen de gebruiker bij het vinden van relevante gegevens. Als het platform succesvol wordt, zal de hoeveelheid beschikbare gegevens en aangesloten voorzieningen toenemen. Goede filtermechanismen en zoekalgoritmes zijn dan essentieel. Daarnaast biedt het platform een omgeving waar de gebruiker de aangeboden data en diensten kan aanschaffen (*Purchase & Licensing Services*). Verrekening van eventuele afdrachten over, van, voor het gebruik van data en diensten wordt eveneens ondersteund (*Financial Clearance Services*). De gebruikersomgeving ondersteunt verschillende devices en is eenvoudig, laagdrempelig en toegankelijk in gebruik.

### *Service en Evaluatie*

Alle diensten die op het platform worden aangeboden, zijn voorzien van transparante serviceniveaus. Gebruikers krijgen daarbij ondersteuning in het gebruik van data of diensten (*Service Request Management Services*), die hun vragen routeert naar de desbetreffende aanbieder. Daarnaast hebben afnemers van data en diensten de mogelijkheid om afgenomen data en diensten te beoordelen (*Evaluation Services*). De evaluatie wordt getoond bij de desbetreffende dienst of dataset die de medewerker wil afnemen. Aanbieders van onderzoeksdata en diensten hebben de mogelijkheid te reageren op de feedback bij de evaluatie.

### *Meta Data Management Services*

Het platform ondersteunt aanbieders van data en diensten bij het vastleggen van relevante informatie over de aangeboden data of diensten. Deze meta-data worden gebruikt bij de catalogus en de zoekfunctie.

### *Supply Connection Services*

Voor het koppelen van diensten en het beschikbaar stellen van onderzoeksdata is een koppelcomponent aanwezig. Deze component helpt aanbieders van data en diensten eenvoudig en gestandaardiseerd hun diensten en data via het platform beschikbaar te stellen. Verder kunnen instellingen dankzij deze component eigen voorzieningen als dienst aanbieden op het platform. Deze functionaliteit bevordert onder andere de samenwerking, het hergebruik en de differentiatie van diensten.

### *Orchestration Services*

Een orkestratiefunctie draagt zorg voor de verbinding van individuele componenten van het platform. Door deze functie wordt het verkeer op het platform goed geregeld.

### *Platform Management Services*

Het Smart Region-platform biedt mogelijkheden voor het borgen en nakomen van afspraken, bijvoorbeeld over de kwaliteit van aangeboden voorzieningen en gegevenssets. Zaken als ongewenste content of niet-presterende partners moeten gesignaleerd kunnen worden. Hiervoor zijn overkoepelende afspraken nodig. Ook dienen afspraken gemaakt te worden over welke rol het platform en data van het platform spelen bij een instellingsaccreditatie (ook commerciële partijen). Het platformmanagement (*Platform Management Services*) draagt hiervoor zorg, evenals voor de inrichting, (door)ontwikkeling en implementatie van het platform in samenwerking met de belangrijkste stakeholders.

**Relatie met andere platforms**

Het Smart-Region-platform voor onderzoek is vergelijkbaar met het Smart Region-platform voor onderwijs. Vanuit dit platform is op basis van data in te schatten welke kennis moet worden ontwikkeld en verspreid binnen een regio. De platforms lijken op elkaar voor wat betreft de onderliggende technologieën. Hergebruik van building blocks ligt voor de hand. Bij het Smart Region-platform zal de nadruk liggen op het delen, vinden, verkrijgen, verwerken, analyseren en visualiseren van trenddata.

**Reeds zichtbare initiatieven**

Momenteel is al een aantal initiatieven zichtbaar die kunnen dienen als voorbeeld voor een onderdeel van het toekomstige platform. Dit zijn:

- euroCRIS
- Trends.google.nl
- Trenddata.com
- Data.overheid.nl



## 10.2 Voorbereiden onderzoek

### Wat gebeurt er in dit proces?

Het proces ‘Voorbereiden onderzoek’ richt zich op de onderzoeksprofessional die zich bezighoudt met het voorbereiden van een specifiek onderzoek. Op basis van de processtap ‘Definiëren onderzoek’ zijn onderzoeksthema’s op regionaal, landelijk of internationaal niveau gedefinieerd. Aan de hand van deze thema’s wordt in dit proces het onderzoek voorbereid, veelal resulterend in een goedgekeurd onderzoeksvoorstel. Omdat hierbij wordt samengewerkt op regionaal, landelijk en internationaal niveau, is het van belang dat dit proces kan aansluiten bij Europese samenwerking op het gebied van onderzoek en onderzoeksondersteuning.



De onderzoeker ontplooit activiteiten om alle beschikbare onderzoeksdata te verkrijgen voor het te voeren onderzoek, waarbij hij in dit proces op zoek gaat naar kwalitatieve data. Waar nodig wordt de onderzoeker ondersteund bij de zoektocht en aanschaf van benodigde onderzoeksdata, inclusief eventuele financiële afhandeling daarvan. In dit proces opereren partijen met een vraag naar onderzoeksdata, maar ook partijen die onderzoeksdata aanbieden.

Rondom onderzoek werken instellingen met veel partijen samen. Zij vinden samenwerking belangrijk en proberen actief partijen naar de campus te krijgen en te verbinden. Dergelijke partijen kiezen er soms bewust voor om zich op of rond de campus te vestigen. Een belangrijk onderdeel in de samenwerking is het delen van assets, zoals laboratoria, faciliteiten, maar vooral ook onderzoeksdata.

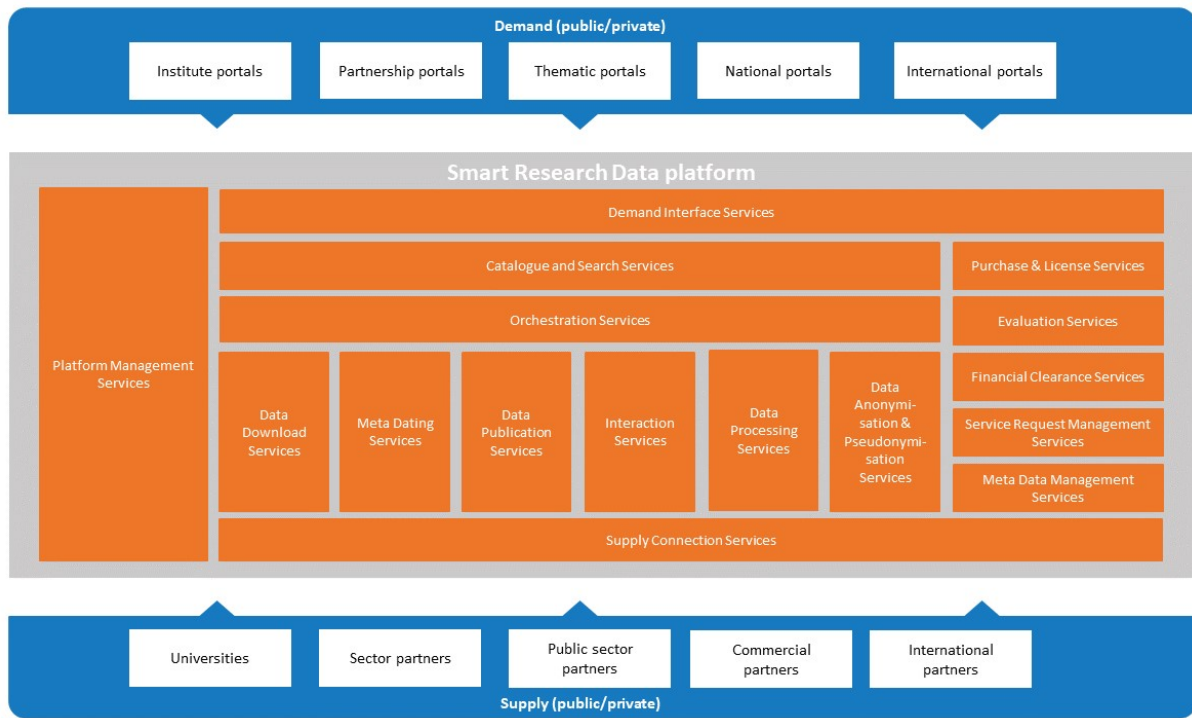
De HO-sector streeft naar een open toegang tot onderzoeksdata en publicaties. Bij partijen waarmee instellingen samenwerken, zal dit niet altijd het geval zijn. Dus zullen onderzoeksdata van deze partijen niet altijd open toegankelijk zijn. Een voorbeeld hiervan zijn onderzoeksdata van partijen die wel beschikbaar zijn voor het doen van wetenschappelijk onderzoek, maar niet breder. In dit proces worden hiervoor strikte afspraken gemaakt over zowel het gebruik als de afscherming van onderzoeksdata.

### Wat is de rol van het platform in dit proces?

Het Smart Research Data-platform ondersteunt het voorbereiden van het onderzoek als het gaat om de onderzoeksdata die nodig zijn om het onderzoek te kunnen uitvoeren. Op het platform opereren partijen zoals overheden, bedrijven, ziekenhuizen, media en de instellingen zelf, maar ook andere private of publieke partijen.

### Uit welke componenten bestaat het platform?

Het Smart Research Data-platform is ingericht om de uitwisseling van allerlei vormen van kwalitatief hoogwaardige onderzoeksdata te bevorderen. Het platform ondersteunt medewerkers in rollen als onderzoeker of onderzoeksondersteuner bij instellingen, sectorpartners of overheden. Het biedt een omgeving om onderzoeksdata van participerende partijen, beschikbaar in andere bronnen, te vinden en af te nemen, en ook om (tussen)resultaten van onderzoek te publiceren. Het platform kan omgaan met vraag en aanbod rond zowel gestructureerde als ongestructureerde data.



Figuur 8: Smart Research Data-platform

Diensten die op het platform worden aangeboden

*Data Download en Publicatie*

Het platform biedt voorzieningen om gegevens, beschikbaar in andere bronnen, op te halen of mee te koppelen. Bijvoorbeeld uit bronnen van op het platform opererende en participerende partijen. Voor het beschikbaar stellen van onderzoeksdata, bijvoorbeeld in de vorm van resultaten uit onderzoek, worden diensten aangeboden die ondersteuning bieden bij de publicatie van data (*Data Publication Service*).

*Data Processing, Anonymisation & Pseudonymisation Services*

Op het platform worden zowel gestructureerde als ongestructureerde data aangeboden. Daarvoor biedt het platform voorzieningen aan die de medewerker ondersteunen bij het verwerken van deze data (*Data Processing Services*) tot bruikbare data. Soms moeten data eerst geanonimiseerd of gepseudonimiseerd worden, alvorens ze via het platform worden aangeboden. Het platform biedt in beide gevallen diensten aan die de aanbieder daarbij ondersteunen (*Data Anonymisation & Pseudonymisation Services*).

*Meta-Data Services*

Net als bij het Smart Region-platform biedt het Smart Research Data-platform diensten aan die aanbieders ondersteunen bij het meta-dateren van hun data. Ook hier is standaardisatie een belangrijk aspect bij meta-datering. Hiervoor is een community nodig met bijbehorende governance-inrichting.

*Interaction Services*

Tot slot biedt dit platform diensten aan rond interactie tussen aanbieders en afnemers van data. Deze interactie richt zich op het gebruik en het bereiken van de juiste kwaliteit van de data. Voor de werking van het platform is het essentieel dat deze interactie goed is ontworpen.

Functies die op het platform aanwezig zijn

Net als op het Smart Region-platform bevat het Smart Research Data-platform functies voor het creëren van de gebruikersomgeving (*Demand Interface Services*) en het kunnen zoeken en vinden van diensten en data (*Catalogue and Search Manager*). Voor het ondersteunen van gebruikers bij het gebruik van diensten en data is

een serviceomgeving beschikbaar (*Service Request Management Services*). Gebruikers kunnen diensten en data beoordelen met behulp van evaluatiefuncties (*Evaluation Services*).

Aanbieders van onderzoeksdata en -diensten worden ondersteund bij het vastleggen van relevante informatie en kwaliteitskenmerken over de data of dienst. Enkele voorbeelden zijn mate van robuustheid en veiligheid, toepassingsgebied of gegarandeerde bestaansduur van de dienst. Deze meta-data worden toegepast en getoond in de *Catalogue and Search Services*. Aanbieders en afnemers worden ondersteund bij de aanschaf en de licentietechnische afhandeling (*Purchase & License Services*) en verrekening (*Financial Clearance Service*) die bij het afnemen en gebruik van de data en diensten van toepassing is.

#### *Connection en Orchestration*

Voor het koppelen van diensten en het beschikbaar stellen van onderzoeksdata is een koppelcomponent beschikbaar, waarmee koppelingen eenvoudig en gestandaardiseerd kunnen worden gelegd. Instellingen kunnen hiermee eigen voorzieningen als dienst aanbieden op het platform. Deze functionaliteit bevordert onder andere de samenwerking, het hergebruik en de differentiatie van diensten. Een orkestratiefunctie draagt zorg voor de verbinding van individuele componenten van het platform. Door deze functie wordt het verkeer op het platform goed geregeld.

#### *Platform Management Services*

Het Smart Research Data-platform biedt mogelijkheden voor het borgen en nakomen van afspraken, bijvoorbeeld over de kwaliteit van aangeboden voorzieningen en gegevenssets. Zaken als ongewenste content of niet-presterende partners moeten gesignaleerd kunnen worden. Hiervoor zijn overkoepelende afspraken nodig. Ook dienen afspraken gemaakt te worden over welke rol het platform en data van het platform spelen bij een instellingsaccreditatie (ook commerciële partijen). Het platformmanagement (*Platform Management Services*) draagt hiervoor zorg, evenals voor de inrichting, (door)ontwikkeling en implementatie van het platform in samenwerking met de belangrijkste stakeholders.

#### **Relatie met andere platforms**

Het Smart Research Data-platform heeft een relatie met het Smart Region-platform als het gaat om het zoeken, verkrijgen en delen van onderzoeksdata. De platforms hebben op bepaalde onderdelen gelijksoortige onderliggende technologieën. Hergebruik van building blocks ligt voor de hand. Bij het Smart Research Data-platform zal de nadruk liggen op het aanbieden en publiceren van onderzoeksdata.

#### **Reeds zichtbare initiatieven**

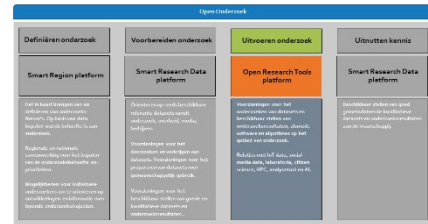
Momenteel is er al een aantal initiatieven zichtbaar die kunnen dienen als voorbeeld voor een onderdeel van het toekomstige platform. Dit zijn:

- Re3data.org
- DataCite Metadata Search
- Biological Magnetic Resonance Databank
- Datadash
- Research Drive
- Deltares BlueEarth Data

### 10.3 Uitvoeren onderzoek

#### Wat gebeurt er in dit proces?

Nadat de voorbereidingen voor het onderzoek in het vorige proces ‘Voorbereiden onderzoek’ zijn afgerond (veelal resulterend in een goedgekeurd onderzoeksvoorstel), kan het onderzoek worden gestart. In het proces ‘Uitvoeren onderzoek’ gaat de onderzoeker daadwerkelijk aan de slag met het onderzoek. Het proces richt zich op de onderzoeker die zich bezighoudt met het uitvoeren van het onderzoek zelf. Omdat onderzoek op verschillende niveaus kan plaatsvinden, is het van belang dat dit proces aansluit bij regionale, landelijke en ook Europese samenwerking op het gebied van onderzoek en onderzoeksondersteuning.



De onderzoeker ontplooit activiteiten om alle beschikbare onderzoeksdata (zie proces ‘Voorbereiden onderzoek’) te analyseren, hierbij ondersteund door verschillende onderzoeksfaciliteiten en algoritmes. Daarnaast kan de onderzoeker op zoek gaan naar kwalitatieve onderzoeksdata voor verdere ontwikkeling van het onderzoek, dit kunnen zowel statische als semi-realtime data zijn. Indien nodig kan de onderzoeker ondersteund worden in de zoektocht en aanschaf van benodigde faciliteiten, algoritmes en onderzoeksdata, inclusief de eventuele financiële afhandeling daarvan.

In dit proces zijn partijen betrokken die vanuit een onderzoeksvraag faciliteiten, algoritmes en onderzoeksdata op het gebied van onderzoek nodig hebben dan wel aanbieden. Logische partijen zijn onder andere onderzoekers (onderzoekconsortia), instellingen, SURF, maar ook private partijen.

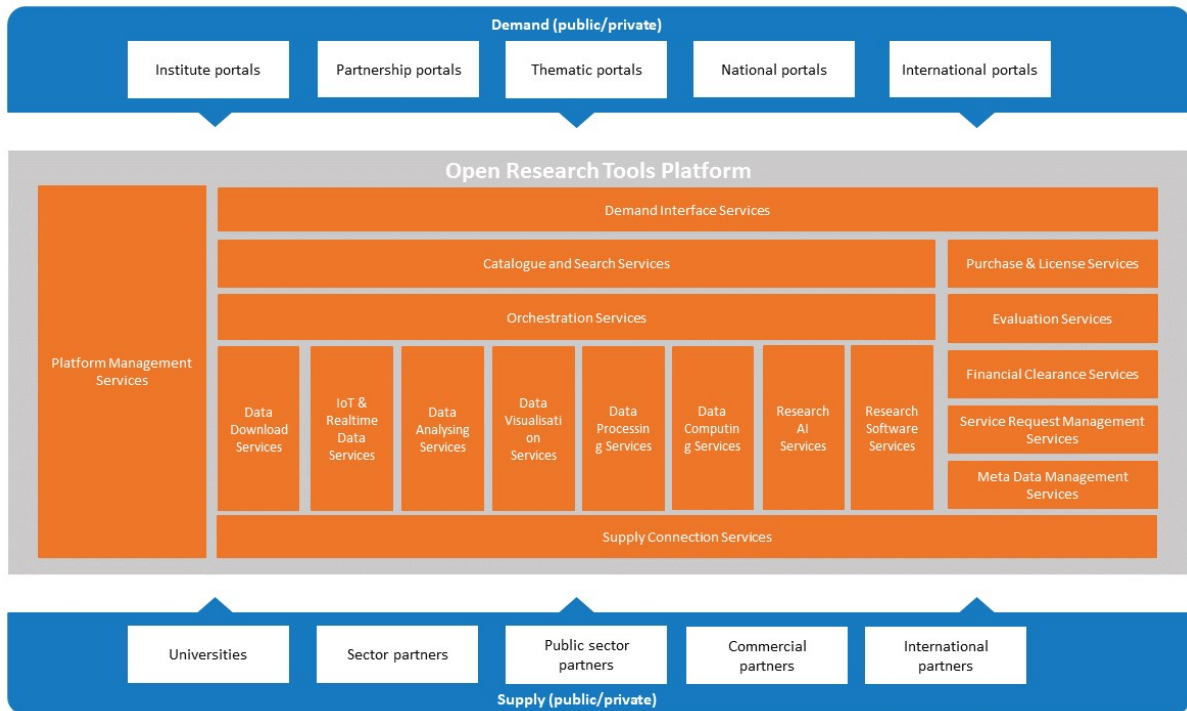
#### Wat is de rol van het platform in dit proces?

De rol van het Open Research Tools-platform is om de onderzoeker tijdens het daadwerkelijke onderzoek te ondersteunen bij het onderzoeken van data. Via dit platform kan de onderzoeker onder andere gebruik maken van realtime data uit IoT-netwerken, smart cities, social media en citizen science-netwerken, maar ook van rekencapaciteit voor het analyseren van data. Door het generiek beschikbaar stellen van deze data en diensten ontstaan nieuwe mogelijkheden voor onderzoek. Daarbij kan de sector een belangrijke bijdrage leveren aan de ontwikkeling van slimme steden. Onderzoeksdata staan niet op het platform zelf.

Instellingen ontwikkelen software, onderzoeksalgoritmes en andere instrumenten voor onderzoek die breder toepasbaar zijn. Het platform ondersteunt de uitwisseling van deze instrumenten.

#### Uit welke componenten bestaat het platform?

Het Open Research Tools-platform is ingericht ter ondersteuning van het voeren van onderzoek. Het ondersteunt de medewerker in de rol van onderzoeker. Het platform biedt een omgeving waar de onderzoeker onderzoeksdata met behulp van verschillende hulpmiddelen kan onderzoeken en van waaruit onderzoeksresultaten ontstaan. Veel van deze onderzoeksdata staan niet op het platform zelf, maar het platform biedt mogelijkheden om data van participerende partijen af te nemen. Verder levert het platform ook diensten die de onderzoeker ondersteunen bij het onderzoeken van onderzoeksdata.



Figuur 9: Open Research Tools-platform

Diensten die op het platform worden aangeboden

*Data Download, IoT & Realtime Services*

Het platform biedt voorzieningen om gegevens, beschikbaar in andere bronnen, op te halen of mee te koppelen (*Data Download Services*). Bijvoorbeeld uit bronnen van op het platform opererende en participerende partijen. Ook worden voorzieningen aangeboden waarmee koppelingen kunnen worden gerealiseerd naar IoT- en realtime datastromen (*IoT & Realtime Data Services*).

*Processing, Analysing & Visualisation Services*

Voor het verwerken van data biedt het platform diensten (*Data Processing Services*) aan die de medewerker hierbij ondersteunen. Ook voor het analyseren van data worden op het platform diensten (*Data Analysing Services*) aangeboden. Visualisatiediensten (*Data Visualisation Services*) ondersteunen de medewerker bij het presenteren van de analyse-uitkomsten. In tegenstelling tot de gelijknamige diensten op het Smart Region-platform zijn deze diensten geschikt voor zwaardere analyses, verwerking en visualisaties.

*Data Computing Services*

Voor hele complexe en zware analyses van onderzoeksdata biedt het platform diensten aan op het gebied van rekenkracht of brengt het deze bij elkaar. Op basis van criteria wordt (bij voorkeur geautomatiseerd) bepaald welke computing service het meest geschikt is. Voorbeelden zijn grid computing, high performance computing of generieke cloud computing. Deze computing-diensten worden beschikbaar gesteld door instellingen of door publiek/private partners. Het platform voert de regie over de computing services.

*Research AI Services*

Kunstmatige intelligentie zal in toenemende mate ingezet gaan worden in de maatschappij, waarbij onderzoek vooroploopt. In onderzoek worden algoritmes gecreëerd die fungeren als basis onder kunstmatige intelligentie. Dit zijn belangrijke onderdelen van onderzoeksresultaten en er zal behoefte ontstaan om deze te kunnen hergebruiken. Het Open Research Tools-platform biedt diensten die het mogelijk maken om succesvolle algoritmes beschikbaar te stellen aan collega-onderzoekers en partners.

### *Research Software Services*

Analoog aan kunstmatige intelligentie worden diensten aangeboden voor het delen van specifieke onderzoekssoftware. De onderzoeker heeft op het platform de mogelijkheid deze specifieke software te vinden en te verkrijgen. Ook ondersteunt het platform de onderzoeker bij het beschikbaar stellen van dergelijke software.

### Functies die op het platform aanwezig zijn

Ook op dit platform zijn de eerdergenoemde generieke functies beschikbaar ter ondersteuning van de interactie tussen aanbieders en afnemers. De *Catalogue and Search Services* ondersteunen het zoeken, de *Purchase & License Services* en de *Financial Clearance Services* ondersteunen de aanschaf en verrekening van aanschaf en gebruik. *Service Request Management Services* ondersteunen de vraagafhandeling bij het gebruik van de data en diensten op het platform.

Aanbieders van onderzoeksdata en onderzoeksdiensten worden bij het aanbieden ondersteund bij het vastleggen van relevante informatie en kwaliteitskenmerken over de data of dienst (*Meta Data Management Services*). Vanwege de aard van dit platform omvatten de kwaliteitskenmerken niet alleen zaken als robuustheid en veiligheid, toepassingsgebied of bestaansduur van data of dienst, maar ook issues als afscherming tegen spionage, snelheid van wegschrijven van data, bandbreedte naar de opslag, rekenkracht in nabijheid van de opslag en back-up regime. Deze meta-data worden toegepast en zichtbaar gemaakt via de *Catalogue and Search Manager*.

De koppelfunctie (*Supply Connection Services*) zorgt ervoor dat koppelingen eenvoudig en gestandaardiseerd kunnen worden gelegd. Met deze functie kunnen instellingen eigen voorzieningen als dienst aanbieden op het platform. Deze functionaliteit bevordert onder andere de samenwerking, het hergebruik en de differentiatie van diensten. Een orkestratiefunctie draagt zorg voor de verbinding van individuele componenten van het platform. Door deze functie wordt het verkeer op het platform goed geregeld.

Alle functionaliteiten die via het platform beschikbaar zijn, zijn via de (gebruikers)interface eenvoudig en laagdrempelig toegankelijk (*Demand Interface Services*). Deze interface ondersteunt verschillende devices.

Het Open Research Tools-platform biedt mogelijkheden voor het borgen en nakomen van afspraken, bijvoorbeeld over de kwaliteit van aangeboden voorzieningen en gegevenssets. Zaken als ongewenste content of niet-presterende partners moeten gesignaleerd kunnen worden. Hiervoor zijn overkoepelende afspraken nodig. Ook dienen afspraken gemaakt te worden over welke rol het platform en data van het platform spelen bij een instellingsaccreditatie (ook commerciële partijen). Het platformmanagement (*Platform Management Services*) draagt hiervoor zorg, evenals voor de inrichting, (door)ontwikkeling en implementatie van het platform in samenwerking met de belangrijkste stakeholders.

### **Relatie met andere platforms**

Het Open Research Tools-platform heeft een relatie met het Smart Research Data-platform en met het Smart Region-platform. De platforms hebben op bepaalde onderdelen gelijksoortige onderliggende technologieën. Hergebruik van building blocks ligt voor de hand.

Randvoorwaardelijk voor het Open Research Tools-platform is het managen van toegang tot onderzoek. De onderzoeker stelt vast in welke mate zijn onderzoek open moet zijn en geeft verschillende doelgroepen toegang. Dit wordt in belangrijke mate bepaald door afspraken die voorafgaand aan het onderzoek gemaakt zijn met bijvoorbeeld degene die het onderzoek financiert. De diensten moeten de onderzoeker hierin voldoende ondersteunen.

### **Reeds zichtbare initiatieven**

Momenteel is reeds een aantal initiatieven zichtbaar die als voorbeeld kunnen dienen voor een onderdeel van het toekomstige platform. Dit zijn:

- DataverseNL
- EASY
- Nationale Supercomputer Cartesius
- HPC Cloud
- Jupyter Notebook

## 10.4 Uitnutten kennis

### Wat gebeurt er in dit proces?

Het proces ‘Uitnutten kennis’ richt zich op het aan een breder publiek beschikbaar stellen van onderzoeksresultaten (publicaties, onderzoeksdata, tussenresultaten). De onderzoeker wil de resultaten uit het proces ‘Uitvoeren onderzoek’ delen met collega-onderzoekers voor collegiale toetsing of kennisoverdracht. Ook wil de onderzoeker in dit proces (tussen)resultaten beschikbaar stellen (publiceren) voor een breder publiek.



In dit proces kunnen onderzoekers zich oriënteren op nieuwe data relevant voor hun eigen onderzoek, maar ook op eventuele tussen- en eindresultaten uit eerder onderzoek. Hierbij wordt hetzelfde platform ingezet als bij het proces ‘Voorbereiden onderzoek’: het Smart Research Data-platform.

### Wat is de rol van het platform in dit proces?

Het Smart Research Data-platform ondersteunt het zoeken, afnemen en publiceren van onderzoeksdata in de vorm van publicaties of tussenresultaten van onderzoek. Partijen die op het platform opereren, zijn instellingen, overheden, bedrijven, ziekenhuizen en media, maar ook andere private of publieke partijen.

### Uit welke componenten bestaat het platform?

Zie voor een beschrijving van de componenten van dit platform paragraaf 10.2.

### Relatie met andere platforms

Het Smart Research Data-platform heeft een relatie met het Education Exchange-platform voor flexibilisering van onderwijs. Bij het ontwikkelen van onderwijs wordt onder andere gebruik gemaakt van resultaten uit onderzoek. Daarnaast hebben de datamanagementplatforms op bepaalde onderdelen gelijksoortige onderliggende technologieën. Hergebruik van building blocks ligt voor de hand.

### Reeds zichtbare initiatieven

In de huidige situatie is al een aantal initiatieven zichtbaar die als voorbeeld kunnen dienen voor een onderdeel van het toekomstige platform. Dit zijn:

- TU Delft Repository
- Wetenschapswinkel van de WUR
- HBO Kennisbank
- VU-research portal
- Google Scholar



## 11 Principes

Kenmerkend voor de domeinarchitectuur voor Research Data Management (RDM) is de sterke verbinding tussen diverse platforms. Daarbij wordt een groot aantal processen en applicaties gevoed met informatie vanuit verschillende componenten binnen en buiten de HO-sector. De onderstaande principes zijn specifiek voor het RDM-domein en deels een verbijzondering van principes die op HOSA algemeen niveau zijn gedefinieerd. Hierbij is onder andere aandacht besteed aan de beheersbaarheid van de platforms, omdat daar de grootste afbreukrisico's te vinden zijn als het gaat om de betrouwbaarheid van de platforms. De architectuurprincipes bieden richtlijnen voor het inrichten van de gewenste architectuur. Ze dienen als instrument voor en verantwoording bij inrichtingsbeslissingen. De principes ondersteunen een consistente inrichting.

| Principe  | Rationale  | Consequenties   |
|---|--|---|
| RDM – 001 – stimuleren van samenwerking                     | <ul style="list-style-type: none"> <li>Om onderzoek toekomstbestendig te maken, is nationale en internationale samenwerking essentieel.</li> <li>Voor het verkorten van de time-to-market van onderzoek is samenwerking noodzakelijk.</li> </ul>   | <ul style="list-style-type: none"> <li>De diensten die via de platforms worden aangeboden en afgenomen, zijn laagdrempelig en eenvoudig in gebruik.</li> <li>De diensten die via de platforms worden aangeboden en afgenomen, ondersteunen meertaligheid.</li> <li>De platforms ondersteunen cross-border samenwerking.</li> <li>De platforms geven inzicht in vraag en aanbod rond onderzoeksdata en onderzoeksdatadiensten.</li> </ul>  |
| RDM – 002 – ondersteunen van diversiteit en verscheidenheid | <ul style="list-style-type: none"> <li>Voor goed onderzoek is divers aanbod van kwalitatieve onderzoeksdata noodzakelijk.</li> <li>Modulariteit van onderzoek zal leiden tot diversiteit en verscheidenheid in onderzoek en in participatie aan onderzoek.</li> </ul>  | <ul style="list-style-type: none"> <li>Platforms ondersteunen de uitwisseling van verschillende vormen van data.</li> <li>Platforms zijn flexibel ingericht en eenvoudig aan te passen.</li> <li>De platforms en diensten die via de platforms worden aangeboden en afgenomen, ondersteunen moderne en bij voorkeur open (toegankelijkheids-) standaarden en uitwisselingsprotocollen.</li> <li>De platforms bieden ruimte aan diverse aanbieders van data en datamanagementdiensten. Denk aan SURF, instellingen, private aanbieders, maar ook aan innovatieve start-ups.</li> </ul> |
| RDM – 003 – open en transparant                             | <ul style="list-style-type: none"> <li>Open onderzoek verkort de time-to-market van onderzoek.</li> <li>Het open beschikbaar stellen van onderzoeksdata zal bijdragen aan een efficiënter en effectiever onderzoeksproces.</li> <li>Het open beschikbaar stellen van onderzoeksdata (en resultaten) zal kennisontwikkeling en toepassing van onderzoeksresultaten verhogen.</li> </ul> | <ul style="list-style-type: none"> <li>Onderzoeksdata aangeboden op de platforms kennen een unieke identifier.</li> <li>Onderzoeksdata aangeboden op de platforms zijn gemetadateerd volgens de FAIR-principes.</li> <li>HO-instellingen bieden hun onderzoeksdata en onderzoeksdatadiensten aan via de platforms.</li> <li>De platforms geven inzicht in de beschikbaarheid en status van de aangeboden diensten.</li> <li>Onderzoeksdata en diensten aangeboden op de platforms zijn voorzien van heldere licentiestructuren, service- en contractafspraken.</li> </ul>             |
| RDM – 004 – inzicht in kwaliteit                            | <ul style="list-style-type: none"> <li>Inzicht in de kwaliteit, betrouwbaarheid, integriteit en vertrouwelijkheid van onderzoeksdata en diensten zal de kwaliteit en herkenbaarheid van onderzoeksresultaten verder verhogen.</li> </ul>   | <ul style="list-style-type: none"> <li>Onderzoeksdata en diensten die op het platform worden aangeboden, zijn voorzien van actuele meta-data waaronder kwaliteitskenmerken als betrouwbaarheid, integriteit en vertrouwelijkheid.</li> </ul>  |

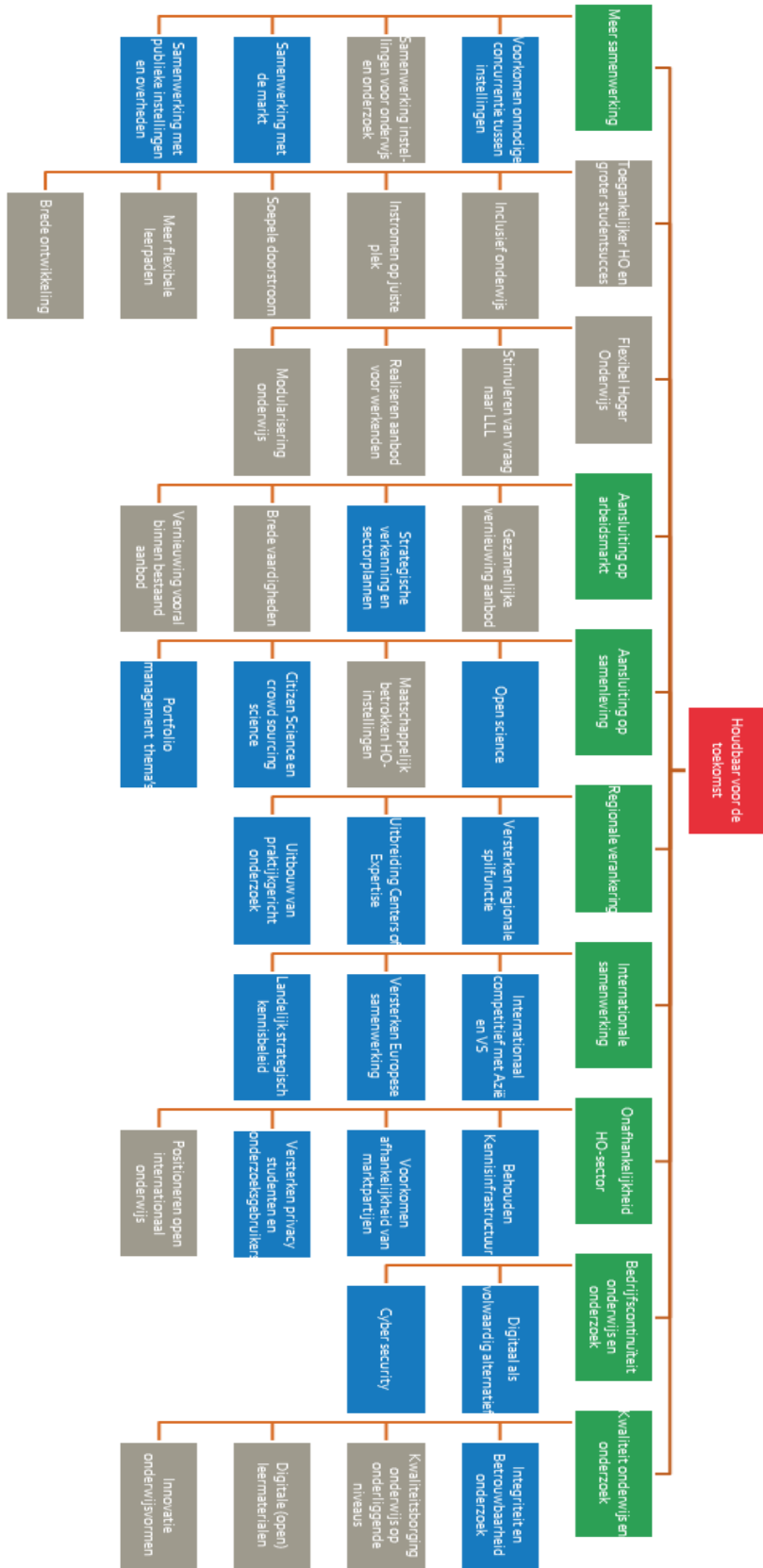
|                                |  |   |
|--------------------------------|--|---|
|                                |  | <ul style="list-style-type: none"><li>• Onderzoeksdata worden gemanaged aan de hand van een data life cycle.</li></ul>  |
| RDM – 005 – ontzorging gebruik | <ul style="list-style-type: none"><li>• Toegankelijk onderzoek begint met toegankelijkheid van onderzoeksdata en diensten. Door 'ontzorging' wordt onderzoek toegankelijk gemaakt.</li></ul> | <ul style="list-style-type: none"><li>• Aanbieders van onderzoeksdata en diensten zijn verantwoordelijk voor de actualiteit van de meta-data op het platform.</li><li>• De aanbieders zijn verantwoordelijk voor het leveren van gebruikersondersteuning in overeenstemming met overeengekomen serviceniveaus.</li><li>• Onderzoeksdata en diensten zijn eenvoudig af te nemen en te gebruiken.</li></ul> |

## 12 Bijlage

### 12.1 Geraadpleegde personen & gremia

|  |  |
|--|--|
| <p>Werkgroep RDM</p> <ul style="list-style-type: none"> <li>- Wilmar de Lange (Hogeschool Rotterdam)</li> <li>- Maarten Hoogerwerf (Universiteit Utrecht)</li> <li>- Louis Stevens (SURF)</li> <li>- Rudy Dokter (Hogeschool Saxion)</li> <li>- Mark Cole (SURF)</li> <li>- Els Velraeds (Fontys Hogescholen)</li> <li>- Taco Horsley (Universiteit van Amsterdam)</li> <li>- Joachim Rijsdam (Leiden Universiteit)</li> </ul>   | <p>Stuurgroep HOSA</p> <ul style="list-style-type: none"> <li>- Anton Opperman (EUR)</li> <li>- Rose van Iperenburg (HAN)</li> <li>- John Kropman (Fontys)</li> <li>- Jan-Willem Brock (Leiden Universiteit)</li> <li>- Hans Louwhoff (SURF)</li> <li>- Rene Schenk (Avans)</li> </ul>   |
| <p>Gesprekken</p> <ul style="list-style-type: none"> <li>- Wendy Sahier (SURF)</li> <li>- John Doove (SURF)</li> <li>- John Meurs (SURF)</li> <li>- Bram Gakeer / Juriaan van Kan (OCW)</li> <li>- Pim Jorg, Melanie Lemmen, Stephan van Duren, Danny ten Bosch (NWO)</li> <li>- Maurice Bouwhuis, Magchiel Bijsterbosch (SURF)</li> <li>- Jeroen Belien (VUMC)</li> <li>- Ronald Mens (TNO)</li> <li>- Willem Koeman (AmDex)</li> <li>- David Groep (Nikhef)</li> <li>- Hylke Koers (SURF)</li> <li>- Wim Hugo (DANS)</li> <li>- Menno Nonhebel (KNAW)</li> <li>- DCC HBO werkgroep ICT</li> <li>- DCC WO project “De verbinding tussen Research Cloud en iRods en Yoda”</li> <li>- DCC WO project “Koppelen DCC services aan de Research Cloud”</li> <li>- Mark van de Sanden (SURF)</li> <li>- TNC’21 (GEANT)</li> <li>- Programmteam VU</li> <li>- Webinar SURF “Platformen hebben de toekomst”</li> <li>- RDM Regiegroep stand van zaken HOSA</li> <li>- Programma RDS</li> <li>- Landelijk netwerk informatiemangers</li> <li>- NLR (Erik Baalbergen)</li> </ul> | <p>Review</p> <ul style="list-style-type: none"> <li>- Wilmar de Lange (Hogeschool Rotterdam)</li> <li>- Maarten Hoogerwerf (Universiteit Utrecht)</li> <li>- Louis Stevens (SURF)</li> <li>- Rudy Dokter (Hogeschool Saxion)</li> <li>- Mark Cole (SURF)</li> <li>- Joachim Rijsdam (Leiden Universiteit)</li> <li>- Els Velraeds (Fontys Hogescholen)</li> <li>- Taco Horsley (Universiteit van Amsterdam)</li> <li>- Tim van Neerbos (Universiteit Utrecht)</li> <li>- Frank Snels (Universiteit van Twente)</li> <li>- ArchitectenBeraad HO             <ul style="list-style-type: none"> <li>o Dick van der Linden (Hogeschool Leiden)</li> <li>o John van den Berge (Technische Universiteit Eindhoven)</li> <li>o Patrick van der Veer (Universiteit Utrecht)</li> </ul> </li> <li>- John Doove (SURF)</li> <li>- David Groep (Nikhef)</li> <li>- Hylke Koers (SURF)</li> <li>- Menno Nonhebel (KNAW)</li> <li>- Wim Hugo (DANS)</li> <li>- Erik Baalbergen (NLR)</li> </ul> |

## 12.2 Doelenstructuur HOSA-RDM



Doelstellingen HOSA - RDM