

Geïntegreerd Ontwerp Architectuur en Onderzoek

v1.1, Definitief, 14 maart 2018

Inhoud

Inleiding.....	2
Overeenkomsten en verschillen tussen de drie initiatieven	3
Metamodel en begrippenkader	3
Architectuurprincipes	4
Bedrijfsfuncties en bedrijfsprocessen.....	4
Informatieobjecten	4
Applicaties en applicatieplatform	4
Architectuurprincipes	5
Samenwerking.....	5
Beleidsvergangen.....	6
Gegevenskwaliteit.....	7
Archivering.....	8
Informatiebeveiliging en privacybescherming.....	9
Metamodel en begrippenkader	10
Bedrijfs-elementen.....	11
Applicatie-elementen.....	11
Technologie-elementen	11
Motivatie-elementen	11
Stakeholders en beleid.....	11
Bedrijfsfuncties en Bedrijfsprocessen.....	12
Onderzoekscyclus	13
Onderzoeksgegevenscyclus	14
Bedrijfsfuncties	14
Vorbereiden onderzoek	15
Uitvoeren onderzoek	17
Publiceren onderzoeksresultaten	18
Uitbaten onderzoeksresultaten	19
Informatieobjecten	20
Applicaties en applicatieplatform	22
Onderzoeksapplicatielandschap	22

De VRE.....	22
Applicatieplatform	23
Conclusies en aanbevelingen	24
Verantwoording	25
Bronmateriaal	26
LCRDM.....	26
D4LS	26
HORA.....	26
Begrippenlijst	26

Inleiding

Voor u ligt het *Geïntegreerd Ontwerp Architectuur en Onderzoek*, het eindrapport van de werkgroep *Architectuur en Onderzoek*. Het doel van het project *Architectuur en Onderzoek* is synergie te realiseren tussen de drie thans lopende architectuurinitiatieven met een focus op het domein Onderzoek:

- Werkgroep Architectuur en Onderzoek van het Architectenberaad HO (HORA)
- Subwerkgroep Architectuur van de werkgroep Faciliteiten en Datainfrastructuur Landelijk Coördinatiepunt Research Data Management (LCRDM)
- Architectuurwerkgroep van het programma Data4Lifesciences (D4LS)

Dit werk is uitgevoerd in opdracht van SURF, als onderdeel van het SURF Innovatieprogramma Efficiënte Service Delivery, project Architectuur en Standaarden. De opdracht aan de werkgroep luidde het bestaande architectuurmateriaal van de drie initiatieven te verzamelen en te analyseren op overeenkomsten en verschillen. De drie initiatieven blijken niet ver van elkaar af te liggen. Daardoor is de werkgroep erin geslaagd verder te gaan dan de oorspronkelijke doelstelling, en heeft zij een eerste aanzet kunnen doen tot een model waarin beelden uit de drie initiatieven samenkomen. Het resultaat is in dit *Geïntegreerd ontwerp* vastgelegd. Hoewel het model zeker nog verfijnd, verdiept en aangevuld kan worden – en in die zin als tussenproduct moet worden beschouwd – is de werkgroep verheugd dat het mogelijk bleek tot dit resultaat te komen.

De werkgroep heeft haar analyse uitgevoerd aan de hand van onderstaand raamwerk. Het beschikbare materiaal is vergeleken op de aspecten:

- Metamodel en voor architectuur relevante begrippen
- Architectuurprincipes
- Bedrijfsfuncties
- Informatieobjecten
- Bedrijfsprocessen
- Applicaties en applicatieplatform

Geïntegreerd ontwerp (GO)	LCRDM	D4LS	HORA
Metamodel en begrippenkader	<ul style="list-style-type: none"> - Architecture Framework LCRDM - Working model Catalogue Group - Glossary 	<ul style="list-style-type: none"> - Geen (expliciet) metamodel - De referentiearchitectuur wordt verder uitgesplitst in lagen (werk in uitvoering) 	HORA Kennismodel (NB. In ontwikkeling naar HORA 2.0)
Architectuurprincipes	<ul style="list-style-type: none"> - Toelichting bij poster van de Werkgroep Architectuur - Rubricering van policies 	Architectuurprincipes NFU D4LS WP2 (vastgesteld)	Principes voor informatievoorziening
Bedrijfsfuncties	<ul style="list-style-type: none"> - geen onderscheid tussen bedrijfsfuncties en –processen - Uitgangspunt is de data life cycle 	niet ingevuld in de D4LS referentiearchitectuur	Bedrijfsfunctiemodel
Informatieobjecten	Niet aanwezig	Niet aanwezig	Informatiemodel
Bedrijfsprocessen	<ul style="list-style-type: none"> - geen onderscheid tussen bedrijfsfuncties en –processen - Uitgangspunt is de data life cycle 	Uitgangspunt is de data life cycle	Bedrijfsprocesmodel
Applicaties en applicatieplatform	Technical / application level building blocks (as functions) Services catalogue?	High-level reference architecture (v1.1) + Architecture Workspace (v0.6)	Applicatiemodel + Applicatieplatformmodel

Figuur 1: Analyseraamwerk en bevindingen op hoofdlijnen

De werkgroep heeft vastgesteld dat er veel overeenkomsten bestaan tussen het materiaal uit de drie initiatieven, maar dat er tegelijkertijd ook duidelijke accentverschillen zijn. Hierover leest u in het hoofdstuk "Overeenkomsten en verschillen tussen de drie initiatieven". In de daaropvolgende hoofdstukken is op basis van het vergeleken materiaal een aanzet gedaan tot een 'geïntegreerd ontwerp' waarin alle deelnemende partijen zich herkennen. Deze hoofdstukken zijn gestructureerd conform het analyseraamwerk uit Figuur 1.

Overeenkomsten en verschillen tussen de drie initiatieven

De drie in de werkgroep vertegenwoordigde initiatieven – LCRDM, D4LS en HORA – houden zich alle bezig met architectuur in het onderzoeksdomein. Aan de hand van het analyseraamwerk is in kaart gebracht op welke manier de drie initiatieven invulling geven aan de verschillende architectuuraspecten.

Metamodel en begrippenkader

LCRDM heeft het gehanteerde metamodel vastgelegd in het *Architecture Framework LCRDM*. Dit raamwerk definieert een aantal views en een reeks enumeraties van onderdelen van de LCRDM referentiearchitectuur. De nadruk van in het raamwerk ligt op het expliciet modelleren van (de ondersteuning van) de collaboratie en van de verandering in scope van en overgangen in beleid. Daarnaast bestaat er een 'werkmodel' dat vooral gericht is op de totstandkoming van een catalogus. Dit *Working Model* van de *Catalogue Group* beschrijft architectuurelementen en hun onderlinge samenhang. Ten slotte is er een *Glossary* die een lijst van termen definieert zoals die binnen LCRDM-werkgroepen worden gehanteerd¹.

D4LS hanteert geen geëxpliciteerd metamodel. Parallel aan de werkzaamheden van de werkgroep Architectuur en Onderzoek is binnen D4LS gewerkt aan het verder uitsplitsen van de D4LS referentiearchitectuur in lagen, en het in algemene zin in lijn brengen van de D4LS architectuur met ArchiMate.

¹ https://www.edugroepen.nl/sites/RDM_platform/RDM%20Glossary

HORA is gestructureerd conform het HORA kennismodel, dat is opgesteld conform ArchiMate. Parallel aan de werkzaamheden van de werkgroep Architectuur en Onderzoek is binnen HORA gewerkt aan een nieuwe versie van het kennismodel, waarin ten opzichte van de eerdere versie met name aandacht wordt besteed aan de functionele beschrijving van systemen in termen van applicatiefuncties.

Een duidelijke overeenkomst tussen alle drie de initiatieven is het gebruik van ArchiMate als architectuur(modelleer)taal. De (expliciet of impliciet) gehanteerde metamodellen zijn in ieder geval alle te passen op het ArchiMate metamodel. In de invulling daarvan zijn accentverschillen waarneembaar:

- LCRDM benadrukt het bestaan en de rol van ‘polities’ en bestrijkt (in de vorm van de catalogus) ook de uiteindelijke ‘solutions’ (concrete software- en hardwareproducten).
- D4LS richt zich met name op de interne structuur van de virtual research environment.
- HORA kent, zeker in de nieuwe versie van het kennismodel, een meer geneste opzet van de bedrijfslaag.

Architectuurprincipes

LCRDM heeft een lijst van principes online gepubliceerd. Daarnaast heeft, na een inventarisatie van polities uit use cases, een rubricering van polities plaatsgevonden op basis waarvan voor architectuurprincipes relevante aspecten zijn benoemd.

D4LS heeft een vastgestelde lijst van architectuurprincipes.

HORA omvat een lijst van principes voor informatievoorziening.

Waar HORA primair kijkt vanuit het gezichtspunt van een individuele HO-instelling, vertrekt LCRDM vanuit organisatie-overstijgende samenwerking (de ‘onderzoekscollaboratie’). Inhoudelijk liggen de gehanteerde principes niet zelden in elkaars verlengde, maar de toepassing en voorbeelden uit HORA richten zich duidelijk vaker op interne bedrijfsvoering en studenten dan op onderzoekers en samenwerkingsverbanden. De LCRDM-principes leggen een duidelijk accent op instellingsoverstijgende en federatieve context en ondersteuning van veranderingen in (scope/toepassingsgebied van) beleid. De D4LS-principes leggen een duidelijk accent op informatiebeveiliging en het gebruik van persoonsgegevens, en bevatten veel gedetailleerde richtlijnen over hoe om te gaan met onderzoeksgegevens in een gedeelde, virtuele werkomgeving.

Bedrijfsfuncties en bedrijfsprocessen

LCRDM en **D4LS** maken beide geen onderscheid tussen bedrijfsprocessen en bedrijfsfuncties. **HORA** doet dat wel. Zowel **LCRDM** als **D4LS** richten zich op de onderzoeksgegevenscyclus. **HORA** heeft een veel uitgebreider bedrijfsfunctiemodel dat ook andere aspecten van onderzoek bestrijkt, en daarnaast ook sturing en bedrijfsvoering omvat.

Informatieobjecten

Alleen **HORA** heeft een uitgewerkt informatiemodel. De behandeling van informatieobjecten in **LCRDM** en **D4LS** heeft nog niet geleid tot informatiemodellen waarin die objecten gestructureerd zijn samengebracht.

Applicaties en applicatieplatform

In **LCRDM** is aan de hand van een reeks use cases een (eerste) inventarisatie gedaan van technische en applicatieve bouwstenen en bijbehorende functies.

In **D4LS** zijn functionaliteiten rondom de *Virtual Research Environment* uitvoerig en gedetailleerd in kaart gebracht, evenals de conceptuele structuur van de workspace.

Het **HORA** applicatiemodel beschrijft op een logisch niveau (onafhankelijk van specifieke productkeuzen) de applicaties die een instelling nodig heeft om haar processen te ondersteunen. Het HORA applicatieplatformmodel beschrijft de generieke software die gebruikt wordt voor applicaties en is daarmee onderdeel van de infrastructuur.

Architectuurprincipes

Het geïntegreerd ontwerp vertrekt vanuit een verkenning van gemeenschappelijke architectuurprincipes vanuit de drie deelnemende initiatieven. Architectuurprincipes zijn algemene regels en richtlijnen, bedoeld om duurzaam te zijn en zelden aangepast te worden, die de wijze waarop invulling wordt gegeven aan informatievoorziening binnen het onderzoeksdomein ondersteunen². Architectuurprincipes in het onderzoeksdomein richten zich met name op de volgende aspecten:

- Samenwerking
- Beleidsovergangen
- Gegevenskwaliteit
- Archivering
- Informatiebeveiliging en privacybescherming

Samenwerking

Principe: Ondersteun samenwerking op alle niveaus in de onderzoekscyclus.

Rationale: Onder invloed van globalisering en digitalisering zijn mensen steeds minder gebonden aan organisaties. Een toenemend aantal mensen wordt zelfstandig professional en werkt voor verschillende organisaties. Als zelfstandige professionals werken zij vaak samen aan initiatieven, waarbij de rol van de organisatie kleiner wordt. Bovendien vragen externe factoren ook om intensievere samenwerking. In veel wetenschappelijke vakgebieden wordt het aantal beschikbare financieringen minder, en het financieringsbedrag per financiering hoger. Het verwerven ervan vereist samenwerking. Onder invloed van beschikbare/open data ontstaat ook steeds meer multidisciplinair onderzoek waarbij collaboratie over instellingsgrenzen heen eerder regel dan uitzondering is.

Wetenschappelijk onderzoek is derhalve vaak organisatieoverstijgend van aard en onderzoeken worden veelal uitgevoerd door onderzoekers van verschillende onderzoeksinstellingen tezamen met medewerkers van bedrijven. Onderzoek vindt steeds meer plaats in een internationale setting van virtueel georganiseerde onderzoekers, waarbij verschillende instellingen betrokken zijn. Hiervoor is het nodig dat fysieke en virtuele omgevingen en organisaties kunnen worden verbonden.

Organisaties behouden hun eigen autonomie, en maken afspraken over de onderzoekssamenwerking. Waar samengewerkt wordt, is het daarom van belang dat die samenwerking op basis van een vertrouwensrelatie ('*trust*') kan worden ingevuld. Die trustrelatie komt tot uitdrukking in federatieve samenwerking op alle niveaus: business, IAA, datamanagement en netwerkmanagement. Zowel het primaire proces (tijdens het onderzoek) als het ondersteunende proces (diensten, infrastructuur) kan federatief worden uitgevoerd. In voorkomende gevallen

² Naar TOGAF: "Principles are general rules and guidelines, intended to be enduring and seldom amended, that inform and support the way in which an organization sets about fulfilling its mission."

kunnen ook andere vormen van samenwerking (bijvoorbeeld ketensamenwerking) mogelijk of nodig zijn.

Implicaties

Algemeen:

- De informatievoorziening regelt toegang op individuele basis, en niet op basis van organisatie.
- De informatievoorziening is organisatieoverstijgend en federatief ingericht.
- Naast processen die organisatieoverstijgend (in procesketens) worden uitgevoerd kunnen bepaalde processen ook federatief (in gezamenlijkheid) worden uitgevoerd.
- Datauitwisseling vindt plaats op basis van een onafhankelijk logisch model.
- Federatief en internationaal georiënteerd identity management zijn noodzakelijk

Onderzoeksdomein:

- Zowel het primaire proces (tijdens het onderzoek) als de ondersteuning daarvan (aanbodkant, registratie, infrastructuur) is federatief opgezet.
 - Gedistribueerd en federatief research data management
 - Federatieve diensten
 - Gemeenschappelijke afspraken over koppelvlakken op bedrijfs-, data-, applicatie- en infrastructuurniveau.
- Het gebruik van de gefedereerde diensten en afgesproken koppelvlakken wordt ondersteund vanuit een virtuele werkplek die als portaal fungeert en orkestratie biedt.

Beleidsvergangen

Principe: Ondersteun beleidsvergangen gedurende de onderzoekscyclus

Rationale: Onderzoek vindt steeds meer plaats in een internationale setting van virtueel georganiseerde onderzoekers, waarbij verschillende instellingen betrokken zijn. Hiervoor is het nodig dat fysieke en virtuele omgevingen en organisaties kunnen worden verbonden. Organisaties behouden hun eigen autonomie, en maken afspraken over de onderzoekssamenwerking. Onderdeel van die afspraken is het RDM-beleid dat binnen het onderzoekssamenwerkingsverband (de ‘collaboratie’) van toepassing is. Het gemeenschappelijke beleid van een consortium samenwerkende organisaties zal echter veelal verschillen van het RDM-beleid van individuele deelnemers aan het consortium, dat wordt gestuurd door andere stakeholders. Het databeleid in een pre-onderzoeksfase verschilt veelal van het RDM-beleid gedurende de uitvoering van het onderzoek, denk aan patiëntgegevens die gepseudonimiseerd moeten worden. Bij internationale samenwerking heeft men ook te maken met overschrijding van landsgrenzen waardoor andere wet- en regelgeving van toepassing is. Er kan ook verschil zijn tussen het RDM-beleid voor gepubliceerde resultaten (open access) en (nog) niet gepubliceerde data. En ook het beleid van bijvoorbeeld archiefinstanties kan verschillen van het beleid van het onderzoeksconsortium. Onderzoeksgegevens en -resultaten zullen daardoor op verschillende momenten in tijd en/of op verschillende locaties onder ander beleid vallen. Wanneer er verschillen in beleid zijn, moet men in staat zijn de nodige beleidsvergangen te ondersteunen. Door federatieve samenwerking en gemeenschappelijk RDM-beleid wordt het aantal beleidsvergangen geminimaliseerd.

Implicaties:

- Beleid en beleidsvergangen worden expliciet opgenomen in de architectuur.

- Beleidsovergangen kunnen leiden tot fouten. Onderzoekers dienen ontzorgd te worden bij het voldoen aan het vigerend beleid.
- Faciliteer datatransformaties (zoals pseudonimisering).

Gegevenskwaliteit

Principe: De kwaliteit van gegevens wordt expliciet geborgd.

Rationale: De kwaliteit van gegevens is de mate waarin deze geschikt zijn voor het bedoelde gebruik. Elk soort gebruik stelt eigen eisen op de verschillende dimensies zoals accuraatheid, compleetheid, actualiteit, beschikbaarheid, integriteit en betrouwbaarheid. Om (her)bruikbaarheid van gegevens vast te stellen en te ondersteunen moeten de kwaliteit expliciet worden geborgd.

Gegevens bepalen in sterke mate de productiviteit van organisaties; zonder gegevens kunnen processen niet worden uitgevoerd. Het belang van weloverwogen beheer van gegevens wordt in (wetenschappelijk) onderzoek ook steeds duidelijker. Gegevens zijn de basis voor onderzoeken en zorgen ervoor dat informatie en kennis ontstaat en kan worden overgedragen op studenten, bedrijven en de maatschappij. Goed databeheer versterkt de kwaliteit van data en de herbruikbaarheid ervan.

Het wordt steeds belangrijker om onderzoeksgegevens te delen met anderen. Dit maakt het mogelijk voor onderzoekers om daarop voort te bouwen en nog meer toonaangevend onderzoek uit te voeren. Daarnaast wordt het hierdoor mogelijk om de validiteit van onderzoeksresultaten vast te stellen. Deze controleerbaarheid verbindt gegevenskwaliteit aan de principes van wetenschappelijke integriteit [1].

Implicaties:

Algemeen:

- Het bedoelde gebruik en de gewenste kwaliteit van gegevens, waaronder metadata, wordt expliciet gemaakt bij inrichting en veranderingen in de informatievoorziening door hier specifieke eisen voor te definiëren. Bijvoorbeeld in ketenmanagement.
- Voor elke gegevensverzameling is een eigenaar aangewezen die verantwoordelijk is voor de kwaliteit en beschikbaarheid van de gegevens.
- Er is een controlemechanisme waarmee kan worden bepaald of gegevens de gewenste kwaliteit hebben.
- Voor elk gegeven wordt er, binnen de bedoelde gebruikscontext, een eenduidige gegevensdefinitie gebruikt die door alle betrokken actoren (menselijk dan wel machine) gehanteerd kan worden.
- Gegevens hebben één authentieke bron.
- Applicaties halen gegevens altijd uit de daarvoor aangewezen bronapplicatie.
- Bij hergebruik van gegevens moet worden vastgesteld in welke mate de kwaliteit en het bedoelde gebruik van de gegevens past bij het beoogde hergebruik.

Onderzoeksdomein:

- Onderzoeksgegevens voldoen aan de FAIR-principes [3].
- Gepresenteerde informatie is controleerbaar. Als onderzoeksresultaten openbaar worden gemaakt, blijkt duidelijk waar de gegevens en de conclusies op zijn gebaseerd, waaraan ze zijn ontleend en waar ze te controleren zijn. [1]
- Er zijn voorzieningen beschikbaar voor het beheren van onderzoeksgegevens.

- Onderzoekers worden ondersteund en ontzorgd door ‘datapprofessionals’ [2] die hen helpen bij alle gebruik van gegevens in de onderzoekscyclus, zoals het opstellen van datamanagementplannen, het beheren van onderzoeksgegevens, en de verbetering van de kwaliteit van gegevens.
- De context van de totstandkoming en/of het beoogde gebruik van gegevens wordt bij de gegevens opgeslagen als metadata.

[1] VSNU, De Nederlandse Gedragscode Wetenschapsbeoefening – Principes van goed wetenschappelijk onderwijs en onderzoek, herziening 2014

[2] Rob Grim, Marianne van der Heijden, Madeleine de Smaele, Ellen Verbakel: Witboek Dataprofessionals in Nederland, SURF, augustus 2011.

[3] Wilkinson et al., The FAIR Guiding Principles for scientific data management and stewardship, Scientific Data 3:160018 (2016), doi:10.1038/sdata.2016.18

Archivering

Principe: Archiefwaardige informatie wordt in een geschikt archief bewaard

Rationale: Organisaties kunnen archiefplichtig zijn op basis van (lokale) wet- en regelgeving. Zo staat er in de selectielijst voor hogescholen bijvoorbeeld dat er een verplichting is voor onderwijsmateriaal, onderzoeksgegevens en publicaties. Daarnaast kunnen belanghebbenden (zoals onderzoeksinstituten of financiers) het waardevol vinden dat zaken worden gearchiveerd. Archivering is bovendien een belangrijke randvoorwaarde voor het kunnen realiseren van diverse gebruiksdoelen zoals FAIR, reproduceerbaarheid of controleerbaarheid. Welke informatie archiefwaardig (verplicht of anderszins waardevol) is, wordt bepaald door de combinatie van policies waaronder het onderzoek is uitgevoerd. Afhankelijk van het soort gegevens en het beoogde doel van archivering, worden eisen gesteld aan de manier waarop de informatie gearchiveerd wordt – bijvoorbeeld in een gecertificeerde omgeving.

Implicaties:

Algemeen

- Er is een vastgesteld selectiemechanisme om te bepalen wat archiefwaardig is
 - Organisaties werken met een Basis Selectie Document (BSD) waarin alle formele soorten documenten, hun bewaartermijn en/of vernietigingstermijn zijn beschreven.
- Het archiveringsdoel en de bijbehorende bewaartermijnen en beoogde gebruikshorizon bepalen hoe en waar gegevens worden gearchiveerd zodat zij op een later moment kunnen worden ontsloten of gereproduceerd.
- Systemen zorgen ervoor dat de gegevens die ze beheren op de juiste momenten worden bewaard of vernietigd.
- Gegevens die lang bewaard moeten worden, blijven leesbaar doordat de daarvoor noodzakelijke apparatuur en programmatuur wordt bewaard of doordat ze wanneer nodig worden omgezet in een ander formaat.

Onderzoeksdomein

- De kwaliteit van dataverzameling, data-invoer, dataopslag en dataverwerking wordt goed bewaakt. Goede verslaglegging van alle stappen en controle op de uitvoering is noodzakelijk

(labjournaals, voortgangsrapportages, documentatie van afspraken en beslissingen enz.). Goede archivering stelt eisen aan de hele keten, beginnend met het datamanagementplan.

- Voor elk onderzoek geeft een datamanagementplan inzicht in de kaders en wensen met betrekking tot archivering.
- Onderzoeksresultaten en de daarbij behorende onderzoeksgegevens worden na publicatie opgeslagen in repositories of archieven. Indien nodig wordt gebruik gemaakt van een gecertificeerde omgeving
- De bewaartermijn van onderzoeksgegevens is vastgesteld conform vigerend beleid en/of geldende wet- en regelgeving (cf. [1], minimaal 10 jaar). Deze gegevens worden op aanvraag ter beschikking gesteld aan andere wetenschapsbeoefenaren, tenzij wettelijke bepalingen of afgesproken beperkingen zich daartegen verzetten. (cf. [1])
- Onderzoeksgegevens worden zodanig gearchiveerd dat deze binnen de gestelde bewaartermijn met een minimum aan tijd en handelen (passend bij de grootte en complexiteit van de gegevensverzameling en het doel van de archivering) kunnen worden geraadpleegd. (cf. [1])

[1] VSNU, De Nederlandse Gedragscode Wetenschapsbeoefening – Principes van goed wetenschappelijk onderwijs en onderzoek, herziening 2014

Informatiebeveiliging en privacybescherming

Principe: Gegevens zijn beveiligd op basis van hun risicoclassificatie

Rationale: Onderdeel van de kwaliteit van gegevens zijn de aspecten Beschikbaarheid, Integriteit en Vertrouwelijkheid (BIV) die als kernonderdeel worden gezien van informatiebeveiliging. Ontwikkelingen als consumerization, tijd- en plaatsafhankelijk werken en cloud computing maken informatiebeveiliging een actueel onderwerp. Grenzen van organisaties vervagen en traditionele beveiligingsmaatregelen passen niet meer. Incidenten waardoor de beschikbaarheid, integriteit of vertrouwelijkheid van gegevens in gevaar wordt gebracht zorgen voor ernstige ontregeling van organisaties. Het is daarom belangrijk de risico's expliciet te maken. Hierdoor kunnen de meest passende maatregelen worden genomen en worden overmatige maatregelen vermeden.

Bovenop deze algemene tendens kan wetenschappelijk onderzoek betrekking hebben op personen (bijvoorbeeld patiënten), en daardoor op (bijzondere) persoonsgegevens. Dergelijke gegevens vragen om aanvullende maatregelen, in lijn met de AVG.

Implicaties:

Algemeen:

- Gegevens zijn door de gegevensverantwoordelijke voorzien van een BIV-classificatie die aangeeft wat het gewenste niveau van Beschikbaarheid, Integriteit en Vertrouwelijkheid is.
- Informatiebeveiligingsmaatregelen zijn gebaseerd op het informatiebeveiligingsbeleid en de BIV-classificatie van de betrokken gegevens. Maatregelen worden ook gebaseerd op een risicoanalyse vanuit procesperspectief.
- Informatiebeveiliging wordt integraal meegenomen bij het ontwerp en de inrichting van applicaties en infrastructuur.
- Naleving van informatiebeveiligingsmaatregelen is een verantwoordelijkheid van alle betrokkenen en wordt onder meer geborgd door periodieke interne en externe audits, passend bij de risicoclassificatie.

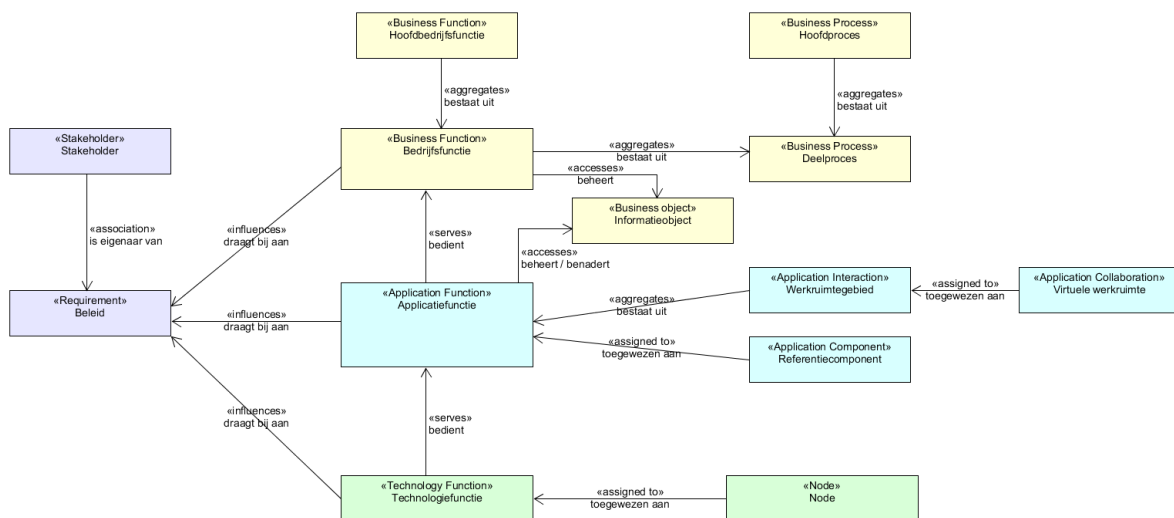
- Wanneer de verantwoordelijkheid voor de data overgaat naar een andere partij, wordt eventueel een gebruiksovereenkomst afgesloten.
- Wanneer de data verwerkt worden door een andere partij, maar de verantwoordelijkheid niet meegaat, wordt een be-/verwerkersovereenkomst afgesloten.

Onderzoek:

- Er is vastgelegd welke soorten data bij een onderzoek betrokken zijn en/of daaruit voortvloeien en hoe gebruikers van die data daarmee om moeten gaan.
- De (technische) integriteit van onderzoeksgegevens is op zo'n manier gewaarborgd dat de wetenschappelijke integriteit van het onderzoek niet in gevaar kan komen door, bijvoorbeeld, ongeautoriseerde aanpassingen of gegevensdegradatie.
- De privacy van onderzoeksobjecten wordt gewaarborgd, doordat onderzoeksgegevens die op personen betrekking hebben alleen toegankelijk zijn voor geautoriseerde personen en waar mogelijk worden gepseudonimiseerd of geanonimiseerd.
- Onderzoeksomgevingen zijn van elkaar gescheiden op basis van vertrouwelijkheid.

Metamodel en begrippenkader

Onderstaande figuur geeft het metamodel weer dat in dit geïntegreerd ontwerp is gehanteerd. Het metamodel is gebaseerd op de architectuurprincipes en de (gereconstrueerde) metamodellen die door de verschillende initiatieven gehanteerd worden. Het metamodel biedt een overkoepelend begrippenkader voor de elementen en relaties die in het ontwerp worden gehanteerd.



Het metamodel is gemapt op de overeenkomstige concepten uit de ArchiMate architectuurmodelleertaal³. De kleuren in het diagram komen overeen met de kleuren die in de ArchiMate-specificatie gebruikt worden om concepten uit verschillende lagen aan te duiden:

- Geel: bedrijfselementen
- Turquoise: applicatie-elementen
- Groen: technologie-elementen
- Lavendel: motivatie-elementen

³ The Open Group, ArchiMate 3.0.1 Specification, augustus 2017

Bedrijfselementen

Op de bedrijfslaag onderscheiden we Hoofdbedrijfsfuncties die bestaan uit Bedrijfsfuncties, en Hoofd(bedrijfs)processen die bestaan uit (Deel)processen (zie het hoofdstuk Bedrijfsfuncties en Bedrijfsprocessen). Deze elementen representeren het gedrag van een organisatie in het onderzoeksdomein. De gegevens die in het kader van onderzoek worden beheerd en gebruikt, worden als informatieobjecten gerepresenteerd (zie het hoofdstuk Informatieobjecten).

Applicatie-elementen

Applicaties worden functioneel beschreven aan de hand van Applicatiefuncties.

Referentiecomponenten zijn toegewezen aan een of meer applicatiefuncties. Deze componenten worden soms ook wel aangeduid als 'logische applicatiecomponenten' en bevinden zich op het niveau van 'architecture building blocks'⁴. Referentiecomponenten kunnen bijvoorbeeld dienen als een facet op een softwarecatalogus waarin concrete softwareproducten (c.q. solution building blocks) zijn ondergebracht. Solution building blocks maken geen onderdeel uit van dit ontwerp.

(Federatieve) samenwerking en bijbehorende virtuele werkruimtes hebben een nadrukkelijke positie in het onderzoeksdomein. Dergelijke werkruimten zijn toegewezen aan 'werkruimtegebieden' (areas) die bestaan uit applicatiefuncties die op hun beurt weer aan verschillende referentiecomponenten kunnen zijn toegewezen. Zo werken meerdere referentiecomponenten dus samen om de benodigde functionaliteit voor een virtuele werkruimte te leveren.

Technologie-elementen

Analoog aan Referentiecomponenten en Applicatiefuncties zijn op de technologielaag Nodes toegewezen aan Technologiefuncties. Ook hier bevinden Nodes zich op het niveau van architecture building blocks waaraan in verdere uitwerking concrete producten als solution building blocks kunnen worden gerelateerd.

Motivatie-elementen

Bij onderzoek zijn verschillende belanghebbenden betrokken die middels beleid invloed uitoefenen op het onderzoek. Bedrijfs-, applicatie- en technologiefuncties dragen bij aan de realisatie van het gevoerde beleid. Zie voor meer informatie het hoofdstuk Stakeholders en beleid.

Stakeholders en beleid

In onderzoekscollaboraties komen verschillende vormen van beleid samen. Dat beleid is niet altijd met elkaar verenigbaar. Wanneer onderzoeksgegevens en onderzoeksresultaten onder ander beleid gaan vallen (anders gezegd: wanneer sprake is van een beleidsovergang) dienen daarom technische en organisatorische maatregelen te worden genomen die de gegevens en resultaten geschikt maken voor het vigerend beleid.

Hoewel stakeholders, hun beleid, en de wijze waarop bedrijfs-, applicatie- en technologiefuncties daaraan invulling geven pas concreet worden in een specifieke onderzoekscontext, loont het om als referentie een opsomming te geven van veelvoorkomende soorten stakeholders en beleid.

Stakeholder	Beleid
Individuele onderzoekinstelling	Beleidskader RDM
Individuele onderzoekinstelling	Beleidskader onderzoek

⁴ cf. The Open Group, TOGAF Version 9.1, 2011

Individuele onderzoekinstelling	Enterprise en Domeinarchitectuur
Individuele onderzoekinstelling	Maatregelen security / kwaliteit
Onderzoekscollaboratie	Samenwerkingsafspraken; alle soorten beleidselementen van Individuele onderzoekinstelling kunnen voorkomen
Europese unie	General data protection regulation (GDPR, aka AVG)
Andere overheden	Lokale wet- en regelgeving
Nederlandse universiteiten	Gedragcode VSNU
Zorgverlener	Wet op de geneeskundige behandelovereenkomst (WGBO)
Onderzoekssubject	Informed consent
Wetenschapsdiscipline	Vakgebiedspecifieke gedragscodes
Financier	Voorwaarden financier
Leverancier onderzoeksgegevens of -diensten	Voorwaarden leverancier

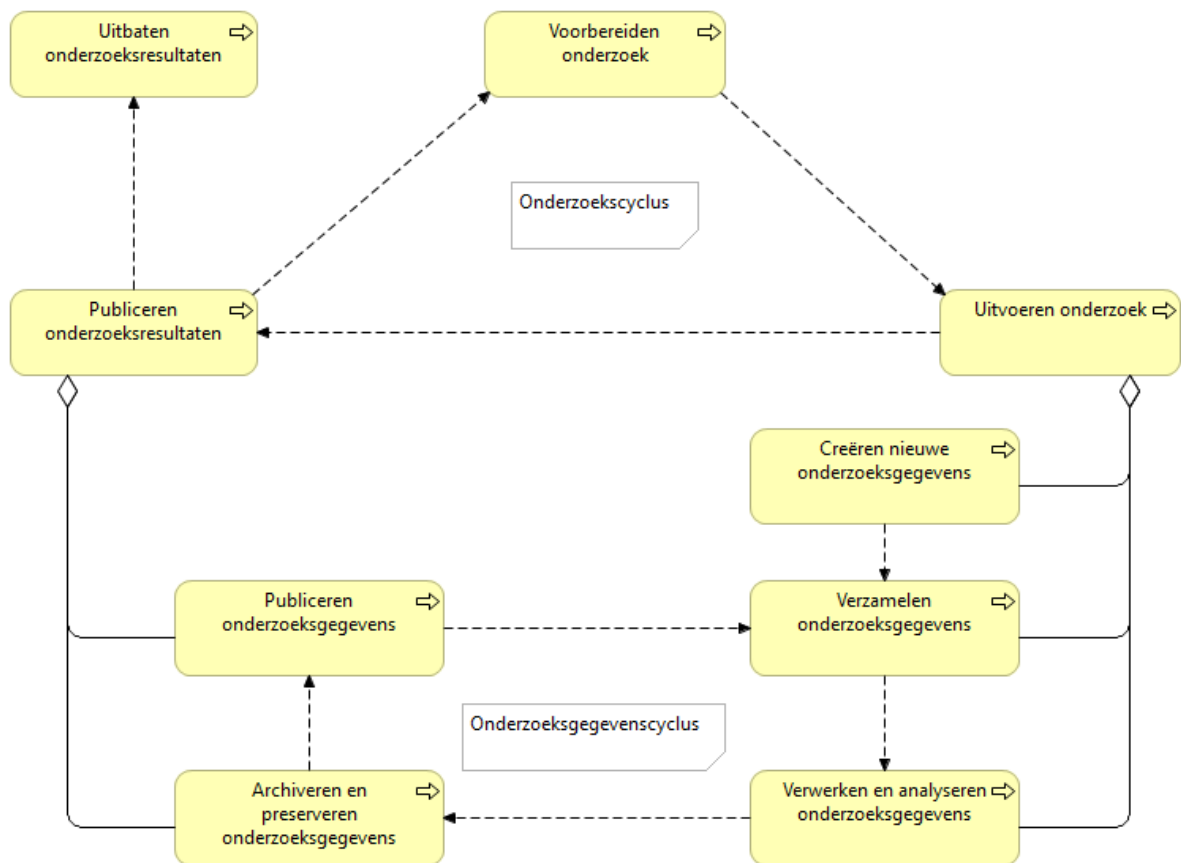
Bedrijfsfuncties en Bedrijfsprocessen

HORA definieert bedrijfsfuncties als logische eenheden van activiteiten die soortgelijke kennis en competenties vragen⁵. Bedrijfsfuncties beschrijven wat een organisatie doet, onafhankelijk van hoe het wordt uitgevoerd. Bedrijfsprocessen zijn sterk gerelateerd aan bedrijfsfuncties. Waar het bedrijfsfunctiemodel benadrukt wat er wordt gedaan, legt het bedrijfsprocesmodel de nadruk op hoe dat gebeurt. In een bedrijfsprocesmodel komt volgorde van activiteiten tot uitdrukking⁶.

In het onderzoeksdomein kunnen we twee aan elkaar gerelateerde cycli onderscheiden: de *onderzoekscyclus* en de *onderzoeksgegevenscyclus*. Deze zijn zeer sterk gerelateerd maar niet één-op-één af te beelden, vandaar dat het van belang is te bepalen vanuit welke cyclus we kijken bij het nadenken over onderzoek – data of het onderzoek zelf. Het denken in cycli helpt bij het vormen van gedachten rondom hergebruik. Anderzijds helpen ze bij de ketengedachte, dat er meerdere partijen zijn betrokken in de verschillende stadia van onderzoek.

⁵ ArchiMate 3.0.1: "A business function is a collection of business behavior based on a chosen set of criteria (typically required business resources and/or competencies), closely aligned to an organization, but not necessarily explicitly governed by the organization."

⁶ ArchiMate 3.0.1: "A business process represents a sequence of business behaviors that achieves a specific outcome such as a defined set of products or business services."



Figuur 2: De onderzoekscyclus en de onderzoeksgegevenscyclus

Onderzoekscyclus

De *onderzoekscyclus* plaatst het uitvoeren van onderzoek in een breder perspectief waarbij ook het voorbereiden van onderzoek, het beheer en ontsluiten van onderzoeksresultaten en het uitbaten van onderzoeksresultaten zijn betrokken. De onderzoekscyclus bestaat uit de volgende fasen:

- Voorbereiden onderzoek: Het uitvoeren van voorbereidende handelingen die erop gericht zijn een onderzoek succesvol te kunnen uitvoeren;
- Uitvoeren onderzoek: Het uitvoeren van een zorgvuldige, verifieerbare en systematische studie van een bepaalde kwestie ter beantwoording van onderzoeksvragen;
- Publiceren onderzoeksresultaten: Het aan een breder publiek beschikbaar stellen van resultaten uit een onderzoek;
- Uitbaten onderzoeksresultaten: Het gebruiken van onderzoeksresultaten om daar voordeel mee te behalen.

Onderzoeksresultaten omvatten de antwoorden op onderzoeksvragen en alle bij de beantwoording van de onderzoeksvragen opgedane kennis en expertise, vastgelegd in welke vorm dan ook. Voorbeelden van onderzoeksresultaten zijn wetenschappelijke artikelen, proefschriften, ontwikkelde of verbeterde onderzoeksinstrumenten, en onderzoeksgegevenssets.

De cyclus 'sluit' wanneer het onderzoek leidt tot vervolgonderzoek.

Onderzoeksgegevenscyclus

De *onderzoeksgegevenscyclus* vindt plaats in de context van het uitvoeren van onderzoek en het beheren en ontsluiten van onderzoeksresultaten. De cyclus richt zich op de omgang met onderzoeksgegevens en kent de volgende fasen:

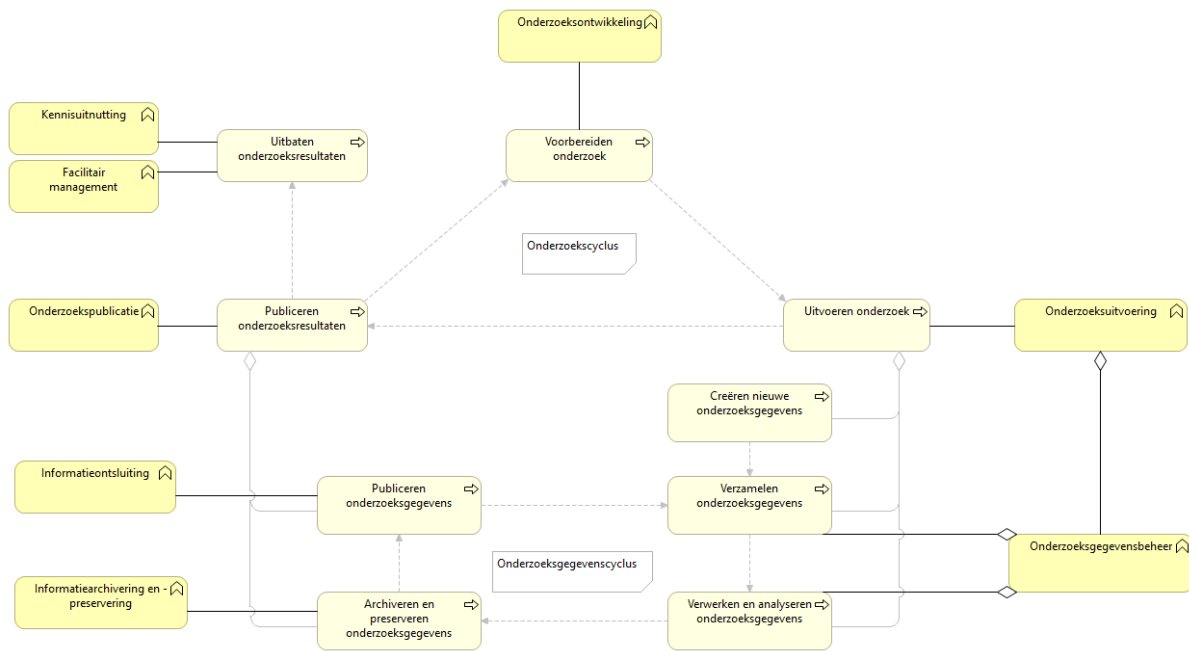
- Creëren nieuw onderzoeksgegevens: Het genereren van nog niet eerder vastgelegde onderzoeksgegevens, ook wel primaire onderzoeksgegevens genoemd;
- Verzamelen onderzoeksgegevens: Het verzamelen van bestaande en nieuw gecreëerde onderzoeksgegevens;
- Verwerken en analyseren onderzoeksgegevens: Het transformeren, combineren of anderszins manipuleren van onderzoeksgegevens;
- Archiveren en conserveren onderzoeksgegevens: Het gedurende langere tijd bewaren en beschikbaar houden van onderzoeksgegevens;
- Publiceren onderzoeksgegevens: Het beschikbaar stellen en ontsluiten van onderzoeksgegevens.

Onderzoeksgegevens zijn gegevens die in het kader van het uitvoeren van onderzoek ontstaan en/of worden gebruikt. Het begrip ‘gegevens’ dient hier zo breed mogelijk te worden opgevat, en omvat bijvoorbeeld ruwe meetpunten en analyseresultaten maar ook scripts en software, beschrijvingen van apparatuur en onderzoekopstellingen.

De cyclus ‘sluit’ bij het hergebruik van onderzoeksgegevens, i.e. wanneer eerder gepubliceerde gegevens verzameld worden in het kader van een nieuw onderzoek.

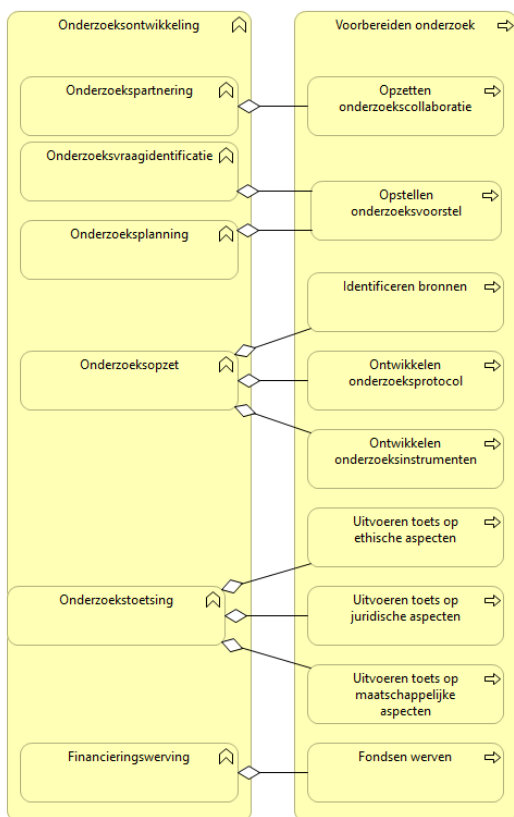
Bedrijfsfuncties

De fasen in de onderzoekscyclus en de onderzoeksgegevenscyclus omvatten wezenlijk andere activiteiten. Daar zijn in de tijd ook andere actoren bij betrokken. Bij de publicatie van onderzoeksgegevens zijn bijvoorbeeld vaak archiefinstanties betrokken, het schrijven van een artikel als onderzoekspublicatie ligt primair bij de onderzoeker zelf. Onderstaande figuur laat de verbinding zien tussen de fasen in de twee cycli en de daaraan gerelateerde bedrijfsfuncties die soortgelijke kennis en competenties vragen, en daardoor toegewezen kunnen zijn aan verschillende organisaties of organisatieonderdelen.



In de volgende paragrafen zijn voor elk van de hoofdprocessen uit de onderzoekscyclus de bijbehorende deelprocessen en de mapping op bedrijfsfuncties in detail uitgewerkt⁷.

Voorbereiden onderzoek



⁷ De uitgewerkte mapping zijn een eerste aanzet en voer voor discussie. De werkgroep is van mening dat ze voor verbetering vatbaar zijn. Zie ook de conclusies en aanbevelingen.

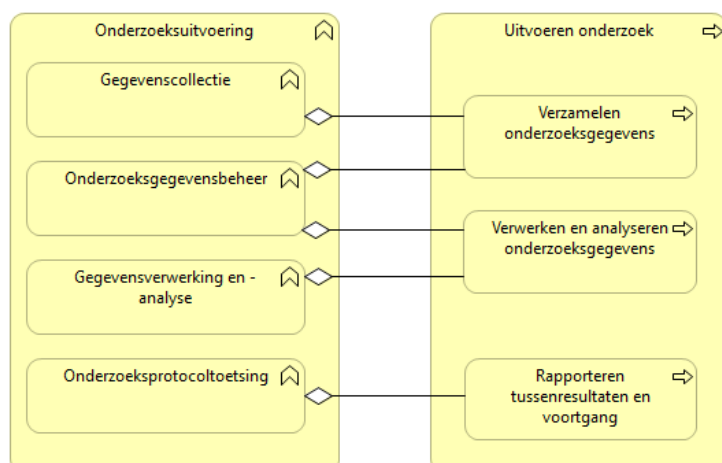
Het hoofdproces *Vorbereiden onderzoek* kent de volgende deelprocessen:

- Opzetten onderzoekscollaboratie
- Opstellen onderzoeksvoorstel
- Identificeren bronnen
- Ontwikkelen onderzoeksprotocol
- Ontwikkelen onderzoeksinstrumenten
- Uitvoeren toets op ethische aspecten
- Uitvoeren toets op juridische aspecten
- Uitvoeren toets op maatschappelijke aspecten
- Fondsen werven

Deze deelprocessen vallen binnen de volgende bedrijfsfuncties, die alle onderdeel uitmaken van de hoofdbedrijfsfunctie *Onderzoeksontwikkeling*:

Bedrijfsfunctie	Definitie	Deelprocessen
Onderzoekspartnering	Het zoeken van partijen die een bijdrage willen leveren aan onderzoek en het maken van afspraken met die partijen.	Opzetten onderzoekscollaboratie
Onderzoeksvraagidentificatie	Het verkennen van waardevolle onderzoeksonderwerpen.	Opstellen onderzoeksvoorstel
Onderzoeksplanning	Het opstellen van een voorstel voor een onderzoek en het indienen ervan voor goedkeuring.	Opstellen onderzoeksvoorstel
Onderzoeksofzet	Het zorgen dat de randvoorwaarden voor het onderzoek zijn geborgd.	Identificeren bronnen, Ontwikkelen onderzoeksprotocol, Ontwikkelen onderzoeksinstrumenten
Onderzoekstoetsing	Het toetsen of het voorstel voor een onderzoek voldoet aan externe en interne kaders en de plaats binnen of tussen lopende onderzoeksprojecten.	Uitvoeren toets op ethische aspecten Uitvoeren toets op juridische aspecten Uitvoeren toets op maatschappelijke aspecten
Financieringswerving	Het aanvragen van een financiering voor een onderzoek.	Fondsen werven

Uitvoeren onderzoek



Het hoofdproces *Uitvoeren onderzoek* kent de volgende deelprocessen⁸:

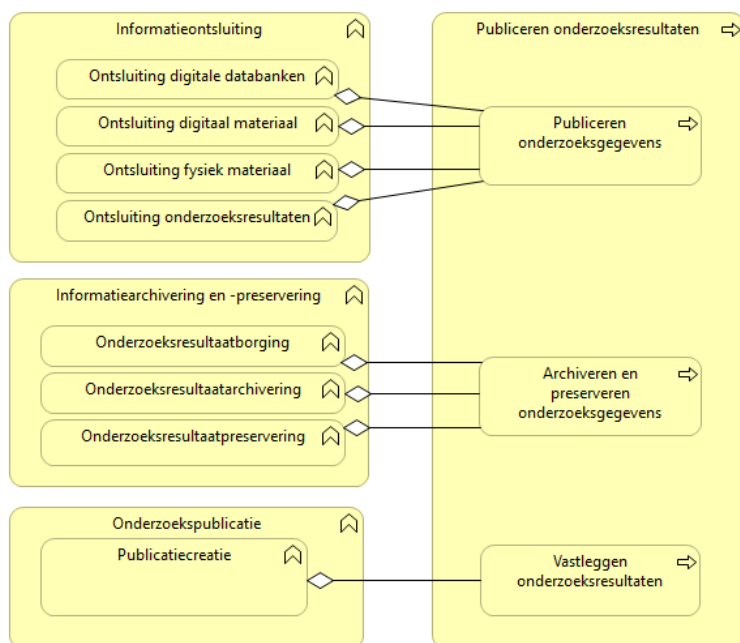
- Verzamelen onderzoeksgegevens^(*)
- Verwerken en analyseren onderzoeksgegevens^(*)
- Rapporteren tussenresultaten en voortgang

Deze deelprocessen vallen binnen de volgende bedrijfsfuncties, die deel uitmaken van de hoofdbedrijfsfuncties *Onderzoeksuitvoering* en *Onderzoekspublicatie*:

Bedrijfsfunctie	Definitie	Deelprocessen
Gegevenscollectie	Het verzamelen van nieuwe en relevante eerder gepubliceerde onderzoeksgegevens.	Verzamelen onderzoeksgegevens
Gegevensverwerking en -analyse	Het verwerken (extraheren, converteren, opschonen, ordenen e.d.), analyseren en interpreteren van de verzamelde onderzoeksgegevens, waaronder modellen en hypothesen.	Verwerken en analyseren onderzoeksgegevens
Onderzoeksgegevensbeheer	Het ervoor zorgen dat onderzoeksgegevens met voldoende kwaliteit beschikbaar zijn voor het onderzoek.	Verzamelen onderzoeksgegevens, Verwerken en analyseren onderzoeksgegevens
Onderzoeksprotocoltoetsing	Het periodiek uitvoeren van audits op de kwaliteit en juiste uitvoering van het onderzoek, al dan niet door de (externe) financiers of andere stakeholders.	Rapporteren tussenresultaten en voortgang

⁸ Processen gemarkeerd met ^(*) zijn onderdeel van de onderzoeksgegevenscyclus

Publiceren onderzoeksresultaten



Het hoofdproces *Beheren en ontsluiten onderzoeksresultaten* kent de volgende deelprocessen⁹:

- Publiceren onderzoeksgegevens^(*)
- Archiveren en preserveren onderzoeksgegevens^(*)
- Vastleggen onderzoeksresultaten

Deze deelprocessen vallen binnen de volgende bedrijfsfuncties, die deel uitmaken van de hoofdbedrijfsfuncties *Informatieontsluiting* en *Informatiearchivering en -preservering*:

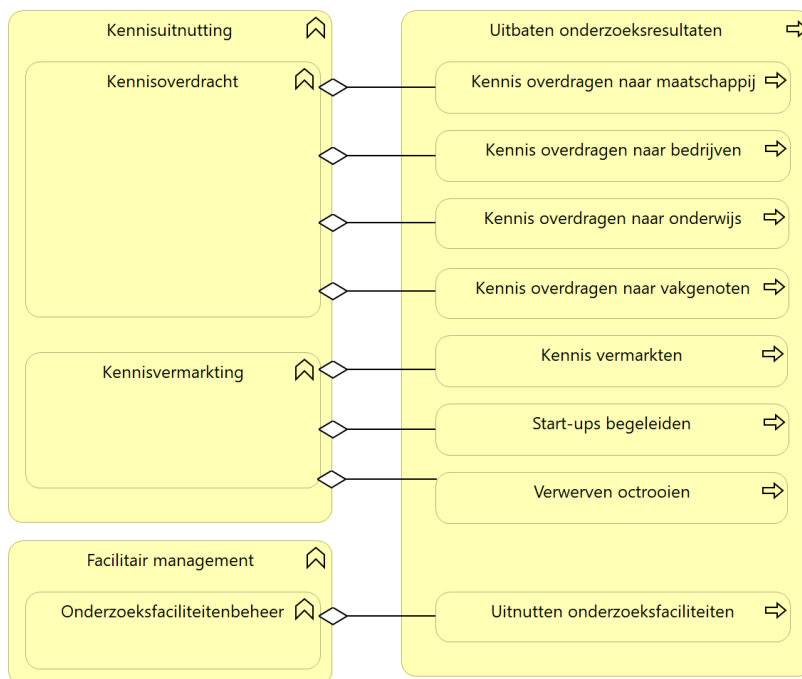
Bedrijfsfunctie	Definitie	Deelprocessen
Ontsluiting digitale databanken	Het toegang geven tot gegevensverzamelingen die in digitale vorm aanwezig zijn.	Publiceren onderzoeksgegevens
Ontsluiting digitaal materiaal	Het beschikbaar stellen van elektronische boeken en tijdschriften ¹⁰ .	Publiceren onderzoeksgegevens
Ontsluiting fysiek materiaal	Het beschikbaar stellen van fysieke middelen zoals boeken, tijdschriften en Dvd's.	Publiceren onderzoeksgegevens
Ontsluiting onderzoeksresultaten	Het beschikbaar stellen van de resultaten van onderzoeken.	Publiceren onderzoeksgegevens
Onderzoeksresultaatborging	Het collectioneren van onderzoeksgegevens die bewaard moeten blijven na	Archiveren en preserveren onderzoeksresultaten

⁹ Processen gemarkeerd met ^(*) zijn onderdeel van de onderzoeksgegevenscyclus

¹⁰ Het onderscheid tussen digitale databanken, digitaal materiaal en fysiek materiaal lijkt niet goed te passen op de huidige onderzoekswerkelijkheid, waarbij bijvoorbeeld gegevensverzamelingen vaak nog als publicaties in klassieke zin (boeken, tijdschriften) beschikbaar worden gesteld en waarin bijvoorbeeld software en scripts ook als digitale onderzoeksgegevens beschikbaar worden gemaakt.

	het onderzoek en het registreren van de onderzoeksresultaten in de relevante interne en externe administraties.	
Onderzoeksresultaatarchivering	Het op gecontroleerde wijze beheren van onderzoeksresultaten.	Archiveren en preservareren onderzoeksresultaten
Onderzoeksresultaatpreservering	Het gedurende lange tijd beschikbaar houden van onderzoeksresultaten en onderzoeksmaterialen.	Archiveren en preservareren onderzoeksresultaten
Publicatiecreatie	Het verwerken van de onderzoeksresultaten tot publicaties zoals benoemd in de definitieafspraken wetenschappelijk onderzoek van de VSNU, inclusief het verwerken van reviewcommentaar.	Vastleggen onderzoeksresultaten

Uitbaten onderzoeksresultaten



Het hoofdproces *Uitbaten onderzoeksresultaten* kent de volgende deelprocessen:

- Kennis overdragen naar maatschappij
- Kennis overdragen naar bedrijven
- Kennis overdragen naar onderwijs
- Kennis overdragen naar vakgenoten
- Kennis vermarkten
- Start-ups begeleiden

- Verwerven octrooien
- Uitmatten onderzoeksfaciliteiten

Deze deelprocessen vallen binnen de volgende bedrijfsfuncties, die deel uitmaken van de hoofdbedrijfsfuncties *Onderzoeksuitvoering* en *Onderzoekspublicatie*:

Bedrijfsfunctie	Definitie	Deelprocessen
Kennisoverdracht	Het overdragen van kennis over onderzoeksresultaten (publicaties, onderzoeksgegevens)	Kennis overdragen naar maatschappij, Kennis overdragen naar bedrijven, Kennis overdragen naar onderwijs, Kennis overdragen naar vakgenoten
Kennisvermarkting	Het daadwerkelijk vermarkten van de kennis die voortvloeit uit de onderzoeksresultaten (publicaties, onderzoeksgegevens).	Kennis vermarkten, Start-ups begeleiden, Verwerven octrooien
Onderzoeksfaciliteitenbeheer	Het beheren van onderzoeksfaciliteiten, -middelen en -diensten zoals laboratoria, onderzoeksapparatuur, High Performance Computing en IT-diensten	Uitmatten onderzoeksfaciliteiten ¹¹

Informatieobjecten

De informatieobjecten zijn in de drie deelnemende initiatieven vooralsnog betrekkelijk dun uitgewerkt. Alleen HORA biedt een overzicht van aan onderzoek gerelateerde objecten, waarvoor de projectgroep heeft geconstateerd dat deze verder, verbreed, verfijnd en op onderdelen scherper gedefinieerd zouden moeten worden om beter bruikbaar te zijn.

¹¹ Onderzoeksfaciliteiten worden hier als onderzoeksresultaat gepositioneerd. Voor 'kleinschalige' faciliteiten lijkt dat een valide uitgangspunt: uit een onderzoek komt nieuwe of verbeterde onderzoeksapparatuur, die wordt uitgenut als nieuwe faciliteit. Ditzelfde proces speelt ook op de langere termijn (aan de ontwikkeling van een stofvrij lab heeft ook onderzoek ten grondslag gelegen), maar gevoelsmatig past dat niet goed onder 'onderzoeksresultaat'. Ook het verband tussen Facilitair management en Onderzoeksfaciliteitenbeheer kan worden aangescherpt; er is waarschijnlijk meer nodig voor onderzoeksfaciliteitenbeheer dan alleen facilitair management, maar op basis van het beschikbare materiaal is de getoonde mapping op dit moment de best mogelijke.

Onderzoek

Naam	Beschrijving
Onderzoek	Een zorgvuldige, verifieerbare en systematische studie van een bepaalde kwestie.
Onderzoeker	Een individu die iets op een wetenschappelijke manier onderzoekt.
Onderzoeksgegevens	De tijdens het onderzoek verzamelde gegevens of een afgeleide daarvan.
Onderzoeksubject	Een object dat wordt onderzocht in een onderzoek (kan ook een subject/persoon zijn).
Proefschrift	Een wetenschappelijke verhandeling die een promovendus moet schrijven en openbaar verdedigen om de graad van doctor te krijgen.
Promotieonderzoek	Een onderzoek met promotiemogelijkheden
Promotietrajectovereenkomst	Een afspraak die de rechten en verplichtingen regelt in promotiegerelateerde zaken tussen de instelling en de promovendus
Promovendus	Een individu die door het zelfstandig beoefenen van wetenschappelijk onderzoek en, hierop aansluitend, het schrijven van een proefschrift wil promoveren tot de academische graad van doctor.
Publicatie	Een resultaat van een onderzoek zoals beschreven als publicatie in de Definitie afspraken Wetenschappelijk Onderzoek van de VSNU.
Samenwerkingsverband	Een verzameling organisaties die gezamenlijk een onderzoek uitvoert.
Subsidieovereenkomst	Een overeenkomst tussen een onderwijsinstelling en een subsidieverstrekker voor het uitvoeren van een onderzoek.
Subsidieprogramma	Een door subsidieverstrekker geïdentificeerd domein waarvoor subsidie kan worden verkregen.

Valorisatie

Naam	Beschrijving
Octrooi	Een door de overheid of bevoegd gezag verleend exclusief recht tot het maken of verkopen van een industrieel product dat is verleend in een verslagjaar.

Informatieontsluiting

Naam	Beschrijving
Expressie	De specifieke intellectuele of artistieke vorm waarin een werk wordt gerealiseerd (uitdrukkingsvorm).
Item	Een enkelvoudig exemplaar van een manifestatie.
Manifestatie	De fysiek verschijningsvorm van de expressie van een werk.
Onderzoeksgegevens	De tijdens het onderzoek verzamelde gegevens of een afgeleide daarvan.
Publicatie	Een resultaat van een onderzoek zoals beschreven als publicatie in de Definitie afspraken Wetenschappelijk Onderzoek van de VSNU.
Uitleen	Het tijdelijk gebruik maken van een materiaal.
Werk	Een te onderscheiden intellectueel of artistiek werk.

Figuur 3: Onderzoeksgerelateerde informatieobjecten uit HORA

Een mogelijke inspiratiebron voor het informatiemodel is CERIF, het Common European Research Information Format. Dit model specificeert de (onderzoeks)wereld in termen van onderzoeker, funder, instrument, et cetera.

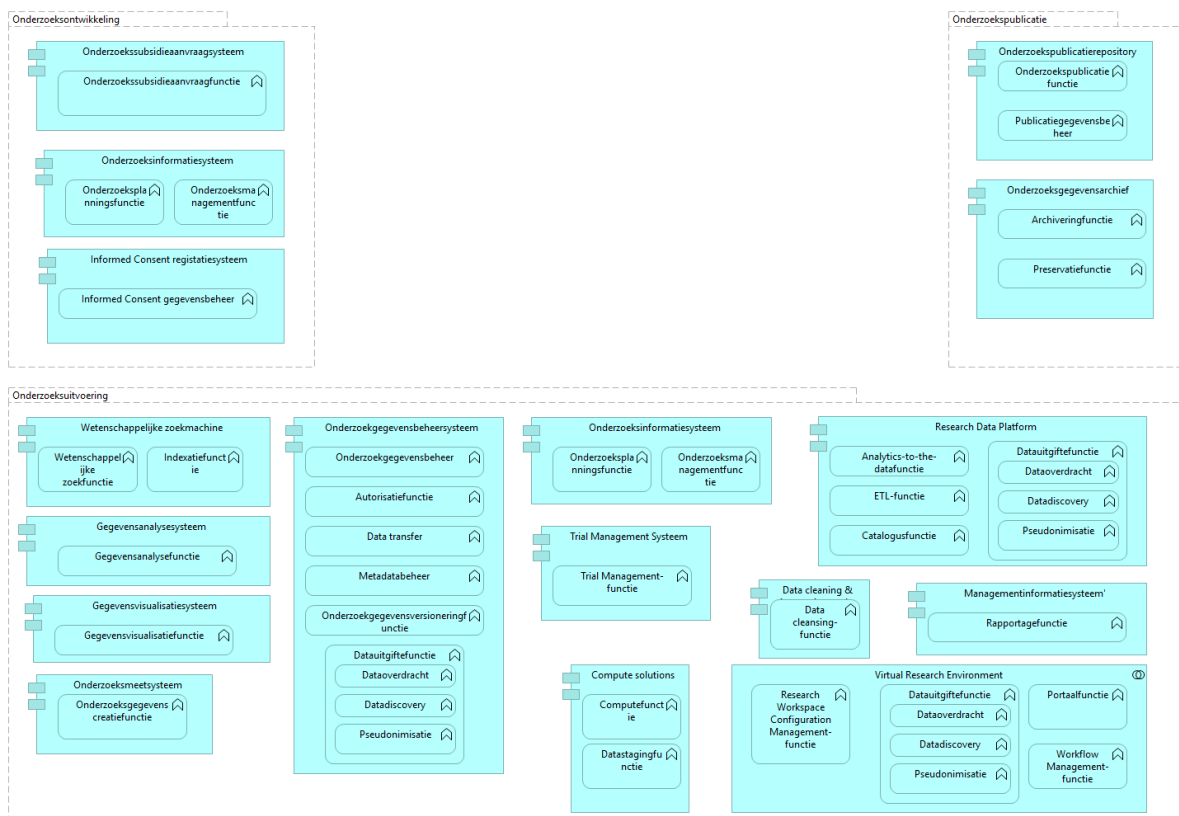
Vanwege het belang van informatiebeveiliging en privacybescherming bij onderzoek, zijn BIV-classificaties (beschikbaarheid, integriteit, vertrouwelijkheid) bij de informatieobjecten van belang. Enerzijds om een indicatie te geven van de mogelijke risicoclassificatie voor systemen en processen die met de desbetreffende gegevens werken, anderzijds om relevante informatieobjecten te kunnen differentiëren. Zo kunnen ruwe patiëntgegevens en resultaten uit astrofysisch onderzoek beide gekwalificeerd worden als 'onderzoeksgegevens', maar omdat de eerste tot op de persoon herleidbare, (bijzondere) persoonsgegevens betreft zal de BIV-classificatie voor dit type gegevens anders zijn.

Applicaties en applicatieplatform

Onderzoeksapplicatielandschap

Onderstaande figuur toont het onderzoeksapplicatielandschap, gegroepeerd naar de bedrijfsfuncties die door de verschillende applicaties worden bediend. De applicatiecomponenten zijn functioneel benoemd, op het niveau van 'Architecture Building Blocks'. De *Virtual Research Environment* (VRE) is weergegeven als applicatiecollaboratie c.q. virtuele werkruimte (cf. Metamodel en begrippenkader). Dit laat zien dat de VRE bestaat uit een samenwerking van applicatiecomponenten die elk een functionele bijdrage aan de VRE leveren. In Figuur 4 zijn voor de VRE alleen de functies opgenomen die de VRE zelf toevoegt aan de functionaliteiten die vanuit de samenwerkende componenten worden geleverd. Zie de paragraaf "De VRE" hieronder voor een nadere uitwerking waarin ook de relatie van de VRE naar de samenwerkende componenten uit het applicatielandschap wordt getoond.

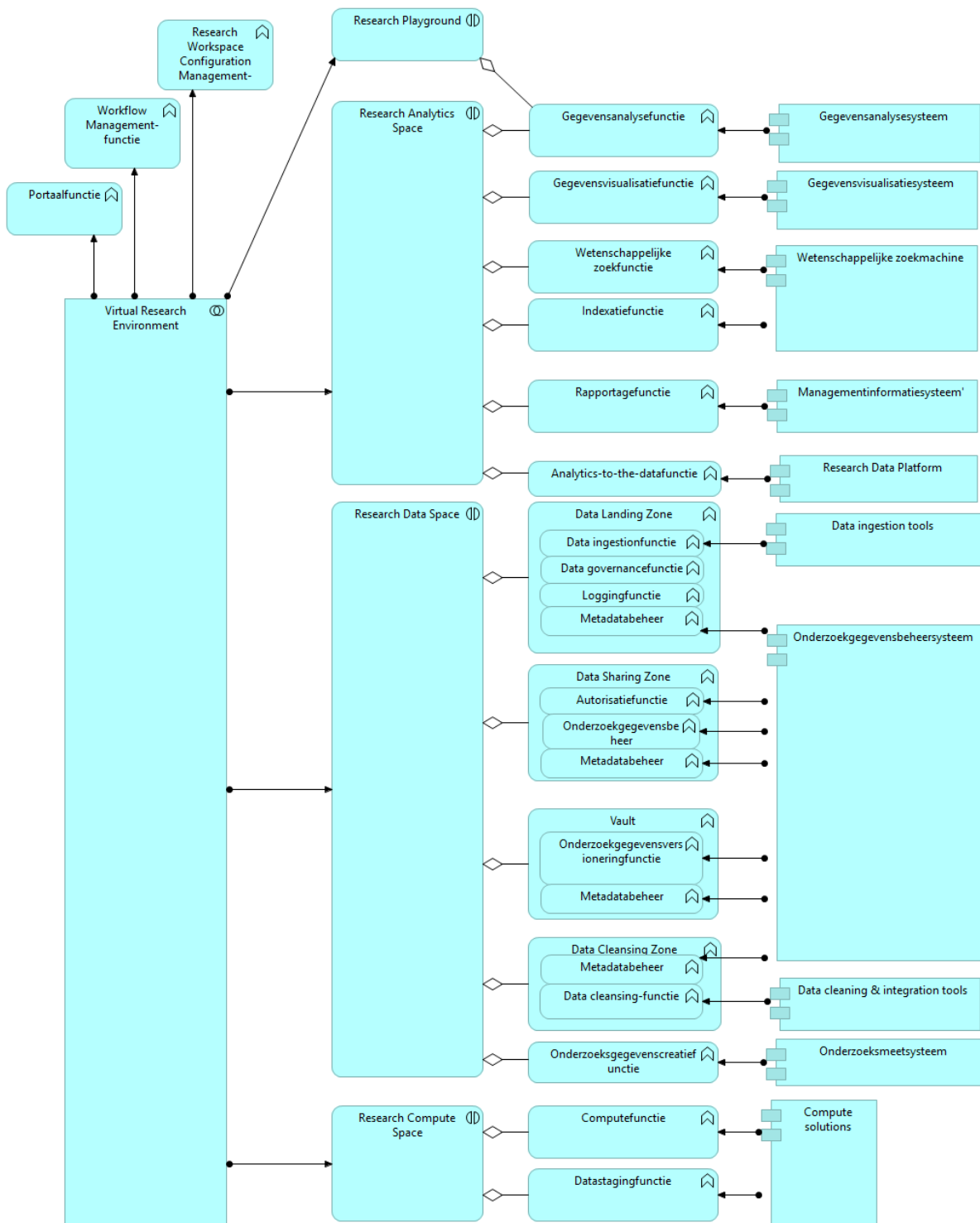
Sommige applicatiecomponenten kunnen als bron dienen voor onderzoeksgegevenssets die in andere applicatiecomponenten (her)gebruikt worden. Deze bronnen zijn toegewezen aan een 'Datauitgiftefunctie', die bestaat uit dataoverdracht, datadiscovery en pseudonimisatie. Het gaat in het bijzonder om Onderzoeksgegevensbeheersysteem, Research Data Platform en VRE.



Figuur 4: Onderzoeksapplicatielandschap

De VRE

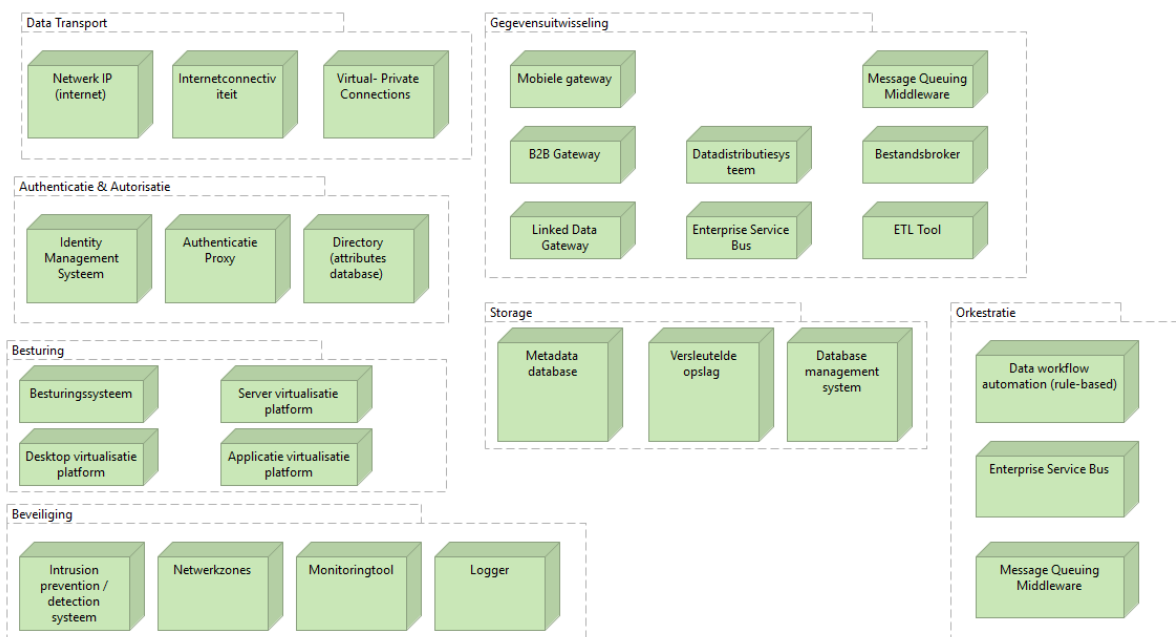
Onderstaande view toont de functionaliteit van de VRE, uitgedrukt in functionaliteiten uit het onderzoeksapplicatielandschap. De view toont hoe verschillende applicatiecomponenten gezamenlijk de complete functionaliteit voor de virtuele werkruimte realiseren.



Figuur 5: De Virtual Research Environment als virtuele werkruimte (applicatiecollaboratie)

Applicatieplatform

De functionaliteiten in het onderzoeksapplicatielandschap worden waar nodig ondersteund door technologiefuncties uit het applicatieplatform. Onderstaande figuur toont diverse nodes (op het niveau van Architecture Building Blocks) die de applicatiefuncties kunnen ondersteunen, gegroepeerd op technologiefunctie.



Figuur 6: Applicatieplatform¹²

Conclusies en aanbevelingen

De werkgroep is - op basis van het beschikbare materiaal uit LCRDM, D4LS en HORA - in staat geweest om een samenhangend beeld te schetsen waarin vertegenwoordigers van alle drie de initiatieven zich herkennen. Het is nu aan elk van de initiatieven om te bepalen wat de impact is van de resultaten van deze werkgroep op hun eigen architectuur.

Veel van de gedurende de analyse gevonden concepten vragen in ieder geval om verduidelijking. In dit geïntegreerd ontwerp is daar op onderdelen een aanzet toe gedaan, met name rondom de onderzoekscyclus en onderzoeksgegevenscyclus. Toch vereist dit nog de nodige aandacht. Met name de gevonden bedrijfs- applicatie- en technologiefuncties zijn vaak op betrekkelijk hoog niveau beschreven, net als de informatieobjecten uit HORA. Dit leidt snel tot interpretatieverschillen.

De verdieping die wordt geboden door de mapping van bedrijfsfuncties op bedrijfsprocessen lijkt voorsnog niet heel groot. De mapping ligt nog erg dicht tegen een 1-op-1-verband. Vanaf deze mapping de stap maken naar de specifieke invulling vanuit het applicatielandschap vraagt om een procesuitwerking die dieper gaat. Hierbij zou een referentiearchitectuur als HORA idealiter meer kaders kunnen geven die duidelijk maken hoe de procesmodellen zich moeten verhouden tot de bedrijfsfuncties, en wanneer iets als proces dan wel als functie opgenomen dient te worden.

Gedurende de analyse zijn we tot de conclusie gekomen dat er op het gebied van onderzoeksarchitectuur een gebrek is aan een eenduidig informatiemodel. Zo'n model is nodig, ook om de bedrijfsfuncties en –processen beter te kunnen duiden. Een verder uitgewerkt informatiemodel zou handvatten moeten bieden voor consistentie en definities rondom begrippen als ‘materiaal’, ‘gegevens’, ‘modellen’, et cetera.

¹² ESB en Message Queuing Middleware komen in het diagram beide dubbel voor, om zowel de gegevensuitwisseling- als orkestratieaspecten ervan te benadrukken. Een vervolgitwerking zou deze concepten functioneel verder uit kunnen splitsen.

In sommige bedrijfsfuncties klinken nu informatieobjecten door, waarbij het de vraag is of die objecten niet tot het informatiemodel beperkt zouden moeten blijven. De vraag rijst of elk informatieobject een eigen bedrijfsfunctie behoeft. Dit is een punt van aandacht voor een vervolg.

Hoewel er duidelijke accentverschillen bestaan tussen de drie initiatieven, is de mate van overlap zodanig dat er wat de werkgroep betreft veel potentie is voor verdere harmonisering. Elk initiatief zal zijn eigen accenten blijven leggen, maar er lijkt voldoende basis voor een verder uitgewerkte, gedeelde onderzoeksarchitectuur. Dit geïntegreerd ontwerp is daarin de eerste stap. De werkgroep ziet vooral het uitgewerkte onderscheid tussen de onderzoekscyclus en de onderzoeksgegevenscyclus als een waardevolle bijdrage, en als een belangrijke kapstok voor een eventueel vervolg.

De werkgroep roept op om HORA de ‘moeder’ te laten zijn die het resultaat uit deze analyse borgt en beheert. Wanneer doorontwikkeling plaatsvindt, kunnen ook die resultaten teruggelegd worden bij HORA. De partijen die aan de werkgroep hebben deelgenomen, kunnen elkaar voor die doorontwikkeling weer opzoeken.

De werkgroep roept SURF op om waar nodig zorg te dragen voor het vervolg. Dat betekent: houdt voeling met de drie programma’s, en organiseer en faciliteer een vervolg wanneer daar vanuit deze partijen behoefte aan is.

Samengevat

1. Laat elk van de deelnemende initiatieven de impact bepalen van de resultaten van deze werkgroep op hun eigen architectuur;
2. Werk de gevonden bedrijfs- applicatie- en technologiefuncties verder uit.;
3. Werk de processen verder uit; geef hiervoor vanuit HORA kaders mee.
4. Werk een eenduidig onderzoeksinformatiemodel uit, mogelijk geïnspireerd op CERIF en rekening houdend met BIV-classificaties van informatieobjecten. Gebruik het informatiemodel om de bedrijfsfuncties en processen beter te duiden. Besteed daarbij ook aandacht aan de vraag in hoeverre informatieobjecten 1-op-1 door zouden moeten klinken in (de benaming van) de bedrijfsfuncties;
5. Laat HORA de 'moeder' zijn die het resultaat uit deze analyse borgt en beheert;
6. Laat SURF waar nodig zorg dragen voor een vervolg

Verantwoording

Het project *Architectuur en Onderzoek* is uitgevoerd in opdracht van SURF, als onderdeel van het SURF Innovatieprogramma Efficiënte Service Delivery, project Architectuur en Standaarden. In de periode oktober 2017 – februari 2018 heeft de werkgroep in zes sprints gewerkt aan de analyse van materiaal uit LCRDM, D4LS en HORA op overeenkomsten en verschillen.

De werkgroep kende de volgende leden:

Lid	Organisatie	Namens
Caspar Terheggen	Radboud Universiteit	LCRDM, HORA
Joyce Nijkamp	Universiteit van Amsterdam	LCRDM
Maarten Hoogerwerf	Universiteit Utrecht	LCRDM, HORA
Alex van den Berg	ErasmusMC	D4LS
Jeroen Beliën	VUMC	D4LS
Hans van den Berg	AMC	D4LS
Sir Bakx (opdrachtgever)	SURF	

Bronmateriaal

LCRDM¹³

- Preso Architectuur LCRDM aan AB, 30 maart 2017
- RDM Architecture Framework, juni 2017
- Poster WG Architectuur, feb. 2017
- Poster WG Catalogus, 8 jun 2017, final
- Poster WG Glossary, 8 jun 2017, v2
- LCRDM RDM Architectuurprincipes, sep. 2017
- Template Preparation Architectural Views, v1.0
- Aggregated Policy and Business Layer (v. 24-11-2017)
- Inventarisatie Technology, 14-09-2017, v0.3
- Verwerking GDPR in RDM Framework, dec 2017
- Stand van Zaken Subwerkgroep Architectuur, 26 september 2017
- UvA: Research landscape vs2, okt. 2017

D4LS

- Slides Data4lifesciences – Een gezamenlijke infrastructuur voor biomedisch onderzoek, zonder datum
- Slides Federation for Research, Irene Nooren, 25 mei 2017
- Slides Next-generation design of the research environments for UMCs, Hans van den Berg, 11 sep 2017 (D4LS symposium)
- ErasmusMC: Referentie Architectuur Onderzoek, juli 2017
- Leaflet Een gedeelde data infrastructuur voor biomedisch onderzoek (zonder datum)
- Presentatie High level referentie architectuur, WP2 (Hans van de Berg, Jeroen Beliën), 1 november 2016
- Toelichting high level referentie architectuur onderzoekomgeving, versie 1.2, J. Belien, J.T. van den Berg, 1 november 2016
- Presentatie Workspaces, concept, 3 november 2016
- Workspaces Architecture, Use cases descriptions, 1 november 2016
- D4LS Onderzoeksarchitectuur (ArchiMate), v8, 17 december 2017

HORA

- Hoger Onderwijs Referentie Architectuur, <http://www.wikixl.nl/wiki/hora>
- Slides HORA voor Onderzoek @ Universiteit Utrecht, zonder datum
- Bijgesteld kennismodel HORA 2.0

Begrippenlijst

Onderstaande begrippenlijst toont een beknopt overzicht van in dit rapport gehanteerde begrippen. De lezer zij verwezen naar de LCRDM Begrippenlijst¹⁴ (waarin tevens begrippen uit D4LS zijn

¹³ Materiaal is beschikbaar op

https://www.edugroepen.nl/sites/RDM_platform/Faciliteiten/RDM%20architectuur.aspx

¹⁴ https://www.edugroepen.nl/sites/RDM_platform/RDM%20Glossary/

opgenomen) en de HORA¹⁵ voor aanvullende definities van begrippen en concepten afkomstig uit deze initiatieven.

Begrip	Verklaring
Analytics tot he data	De analyse wordt verzonden naar een of meer federatieve opgeslagen gegevens die zich buiten de werkruimte bevinden (kan een andere werkruimte zijn) en retourneert (geaggregeerde en / of pseudo / geanonimiseerde) uitkomsten naar de werkruimte. Er is geen (directe) menselijke toegang tot de beoogde gegevens. (Bron: LCRDM Begrippenlijst)
ArchiMate	Een enterprise architectuur modelleertaal
Archiveren en preserven onderzoeksgegevens	Het gedurende langere tijd bewaren en beschikbaar houden van onderzoeksgegevens.
AVG	Algemene Verordening Persoonsgegevens (ook bekend als GDPR)
BIV	Beschikbaarheid, Vertrouwelijkheid, Integriteit
CERIF	Common European Research Information Format
Creëren nieuwe onderzoeksgegevens	Het genereren van nog niet eerder vastgelegde onderzoeksgegevens, ook wel primaire onderzoeksgegevens genoemd.
Datauitgiftefunctie	De ondersteuning door een onderzoeksgegevensbron van dataoverdracht, datadiscovery en pseudonimisatie.
D4LS	Data4lifesciences
GDPR	General Data Protection Regulation (ook bekend als AVG)
Ethische aspecten van onderzoek	Samen met juridische en maatschappelijke aspecten onderdeel van de zgn. ELSI of ELSA (ethical, legal, societal implications/aspects) van onderzoek.
Facilitair management	Het zorgdragen voor de faciliteiten van de organisatie. (Bron: HORA)
FAIR	Acroniem voor Findable, Accessible, Interoperable, Reusable (vindbaar, toegankelijk, interoperabel, herbruikbaar).
Federatieve samenwerking	Samenwerking waarin partijen hun zelfstandigheid behouden, gebaseerd op afspraken en onderling vertrouwen.
HORA	Hoger Onderwijs Referentie Architectuur
IAA	Identificatie, authenticatie, autorisatie
Informatieontsluiting	Het beheeren en beschikbaar stellen van informatie.
Informatiearchivering en -preservering	Het gedurende langere tijd bewaren en beschikbaar houden van informatie.
Juridische aspecten van onderzoek	Samen met ethische en maatschappelijke aspecten onderdeel van de zgn. ELSI of ELSA (ethical, legal, societal implications/aspects) van onderzoek.
Kennisuitnutting	Het beschikbaar en geschikt maken van onderzoeksresultaten (publicaties, onderzoeksgegevens) voor de toepassing in producten, processen en diensten zoals interactie met de maatschappij en private organisaties. (Bron: HORA)
LCRDM	Landelijk Coördinatiepunt Research Data Management
Maatschappelijke aspecten van onderzoek	Samen met ethische en juridische aspecten onderdeel van de zgn. ELSI of ELSA (ethical, legal, societal implications/aspects) van onderzoek.
Onderzoeksgegevens	Gegevens die in het kader van het uitvoeren van onderzoek ontstaan en/of worden gebruikt. Het begrip 'gegevens' dient hier zo breed mogelijk te worden opgevat, en omvat bijvoorbeeld ruwe meetpunten en analyseresultaten maar ook scripts en software, beschrijvingen van apparatuur en onderzoekopstellingen.
Onderzoeksontwikkeling	Het zorgen voor een geaccepteerd onderzoeksvoorstel. (Bron: HORA)
Onderzoekspublicatie	Het opstellen en verspreiden van publicaties rondom het onderzoek. (Bron: HORA)
Onderzoeksresultaten	Omvatten de antwoorden op onderzoeksvragen en alle bij de beantwoording van de onderzoeksvragen opgedane kennis en expertise, vastgelegd in welke vorm dan ook.
Onderzoeksuitvoering	Het daadwerkelijk uitvoeren van het onderzoek. (Bron: HORA)
Publiceren onderzoeksgegevens	Het beschikbaar stellen en ontsluiten van onderzoeksgegevens
Publiceren onderzoeksresultaten	Het aan een breder publiek beschikbaar stellen van resultaten uit een onderzoek.

¹⁵ <https://www.wikixl.nl/wiki/hora/>

Research Data Portal	Een platform dat centraal gegevens uit verschillende bronnen verzamelt en datatoegang biedt aan datawetenschappers / onderzoekers via een portal. Dit platform is beveiligd, controleerbaar en voldoet aan de regels en voorschriften. Waar bij de VRE de gegevensgebruiker centraal staat, draait het RDP om de gegevensleverancier / -eigenaar. (Bron: LCRDM Begrippenlijst)
RDM	Research Data Management
RDP	Research Data Portal
Uitbaten onderzoeksresultaten	Het gebruiken van onderzoeksresultaten om daar voordeel mee te behalen.
Uitvoeren onderzoek	Het uitvoeren van een zorgvuldige, verifieerbare en systematische studie van een bepaalde kwestie ter beantwoording van onderzoeksvragen.
Verwerken en analyseren onderzoeksgegevens	Het transformeren, combineren of anderszins manipuleren van onderzoeksgegevens.
Verzamelen onderzoeksgegevens	Het verzamelen van bestaande en nieuw gecreëerde onderzoeksgegevens.
Virtual Research Environment	Een omgeving waar een onderzoeker toegang heeft tot en werkt met al zijn relevante gegevens, analyses en tooling. Deze omgeving is beveiligd, is self-serviced, is in staat tot real-time samenwerking, levert data en proces audit trails en is controleerbaar conform alle regels en voorschriften. Waar bij de VRE de gegevensgebruiker centraal staat, draait het RDP om de gegevensleverancier / -eigenaar. (Bron: LCRDM Begrippenlijst)
Voorbereiden onderzoek	Het uitvoeren van voorbereidende handelingen die erop gericht zijn een onderzoek succesvol te kunnen uitvoeren.
VRE	Virtual Research Environment
VSNU	Vereniging van Universiteiten