

RESEARCH SUPPORT IN NEDERLAND

DE STAND VAN ZAKEN BIJ RUG EN UMCG



SURF

Veel instellingen zijn bezig met research support: zij willen hun onderzoekers ondersteunen bij de toegang tot hoogwaardige ICT-infrastructuur voor het verwerken, analyseren, visualiseren, beheren en opslaan van onderzoeksdata. Elke instelling doet dat op haar eigen manier. SURF brengt in een serie in beeld hoe instellingen research support hebben georganiseerd en wat hun agenda is voor de toekomst. Deze serie komt voort uit het SURF innovatieproject Support4research. Dit project heeft als doel de SURF-dienstverlening aan onderzoekers beter te laten aansluiten bij de instellingen en haar onderzoekers. In deze aflevering is de RUG en het UMCG aan het woord.

Wilt u ook uw aanpak van research support delen, meld u dan aan via support4research@surf.nl.

SUPPORT VOOR RESEARCH RUG EN UMCG

RUG / UMCG

Aantal medewerkers RUG

5.900 fte (incl. UMCG
Onderwijs & Onderzoek)

Aantal medewerkers UMCG

9.064 fte

Wetenschappelijke staf

3.000 fte

Aantal hoogleraren

400 (100 vrouwen)

Aantal promovendi

2.000

Aantal studenten

30.000

Wetenschappelijke publicaties

6.000

Speerpunten onderzoek

- Energy
- Healthy Ageing
- Sustainable Society

Inleiding

De Rijksuniversiteit Groningen (RUG) heeft een rijke academische traditie die teruggaat tot 1614. De universiteit onderscheidt zich internationaal door de sterke koppeling tussen onderwijs en onderzoek, toegespitst op drie speerpunten: Energy, Healthy Ageing en Sustainable Society. Niet alleen heeft het onderwijs en onderzoek een direct maatschappelijke relevantie, ook werken de wetenschappers op grote schaal samen met partners uit het bedrijfsleven, maatschappelijke organisaties en de overheid.

Op het gebied van digitalisering heeft de universiteit drie speerpunten: Big Data, Smart Industry en Facets of Cultures - [Digital Humanities](#). Het Centrum voor Informatie Technologie (CIT) verzorgt alle IT-faciliteiten voor de universiteit, inclusief faciliteiten voor beheer en opslag van onderzoeksdata. Het CIT heeft met name expertise op het terrein van zeer grootschalige data en gevoelige data (zoals persoonsgebonden data en bedrijfsgevoelige gegevens).

Het Universitair Medisch Centrum Groningen (UMCG) is het tweede grootste ziekenhuis van Nederland en de grootste werkgever van Noord-Nederland. De ruim 10.000 medewerkers werken dagelijks aan de gemeenschappelijke doelstelling: bouwen aan de toekomst van gezondheid. Het UMCG heeft als kerntaken zorg, onderwijs en onderzoek. Als het gaat om onderzoek werkt het UMCG nauw samen met de RUG, en dan met name met het universitaire Centrum voor Informatie Technologie (CIT). Het onderzoek richt zich op nieuwe technieken en behandelingen, nieuwe medicijnen en nieuwe vormen van zorg waarbij de focus ligt op gezond en actief ouder worden, met andere woorden Healthy Ageing.

Visie op research support

Up-to-date datamanagement, training, IT-ondersteuning en -infrastructuur vormen de basis voor excellentie in onderzoek, onderwijs, maatschappelijke impact en talentontwikkeling. De RUG en het UMCG hebben een leidende positie op het gebied van High Performance Computing (HPC), visualisatie, geo-analyse en de virtuele (research) workspace. Naast het bieden van faciliteiten is er veel expertise opgebouwd op het gebied van o.a. data science, datamanagement, privacy en security. Centrale en decentrale expertise en faciliteiten bestaan naast elkaar, werken samen en vormen een levend en interactief netwerk dat is samengebracht in de [Data Federation Hub](#).

Het uitgangspunt is om onderzoekers tijdens de gehele research data life cycle ondersteuning te bieden. Onderzoekers hebben behoefte aan hulp bij het vinden van de juiste datasets, het bewerken van deze data, grootschalige opslagcapaciteit en rekenkracht. Veranderingen in wet- en regelgeving vergroten de behoefte aan ondersteuning.

De RUG en het UMCG hechten belang aan een transparante onderzoeksomgeving. Dit stimuleren we door de principes van Open Science te hanteren, zoals aangegeven in het [Strategisch Plan](#) van de RUG. Ook streven de RUG en het UMCG ernaar dat onderzoekers de [FAIR-principes](#) (Findable, Accessible, Interoperable and Reusable) hanteren bij het omgaan met en beheren van [onderzoeksdata](#).

Voor professionele support is het essentieel om mee te denken met de onderzoekers, complexe vragen te vertalen naar concrete oplossingen en in te spelen op de nieuwste ontwikkelingen. Daarom investeren we bij IT-support in innovatie en user-oriented service. De RUG en het UMCG participeren bijvoorbeeld met verschil-

lende private partners in de [proeftuinprojecten](#) van het Samenwerkingsverband Noord-Nederland, waar nieuwe, innovatieve technieken ontwikkeld worden zoals de proeftuinen *dHealth*, *5G* en *Mining Big Data*.

Samenwerking is zeer belangrijk voor optimale ondersteuning van onderzoekers. De verschillende expertiseonderdelen en faciliteiten moeten met elkaar in verbinding staan via korte lijnen. Belangrijk is dat de ondersteuning vindbaar en gebruikersgericht is en dat de onderzoeker verwezen wordt naar het juiste loket. De Data Federation Hub speelt een belangrijke rol in het verbinden van deze verschillende initiatieven en het bevorderen van de samenwerking.

Voorbeelden van internationale samenwerkingen (huidige en toekomstige) zijn grote sterrenkundige projecten zoals MUSE, MICADO, LOFAR, Euclid en KiDS. Het [Euclid Netherlands Science Data Centre](#) is in het CIT gehuisvest en daarmee speelt het CIT een grote rol in het ontwikkelen van het *Euclid Data Processing System* en het *Euclid Distributed Storage System*. De technologie die voor Euclid ontwikkeld wordt, zal ook het SNN-proeftuinproject Mining Big Data ondersteunen. De datavolumes zijn bij al deze projecten zeer groot: voor MUSE en KiDS wordt al op meerdere locaties een 1 petabyte aan data opgeslagen, en Euclid zal meer dan 50 petabytes aan data produceren.

Organisatie van research support

In Groningen is er een nauwe samenwerking tussen verschillende onderdelen van de RUG en het UMCG. De Data Federation Hub speelt hier een belangrijke rol als spin in het web. Bij de ondersteuning van onderzoekers geldt dat bij welk loket er ook wordt aangeklopt, onderzoekers altijd verder worden geholpen (no wrong door policy) en dat alle eerstelijns loketten beschikken over dezelfde informatie of weten waar de benodigde informatie te vinden is. De IT-services die aangeboden worden maken onderdeel uit van de NWO grootschalige nationale wetenschappelijke infrastructuur, vindbaar onder de naam CIT data warehouse. Bovendien werkt men vanuit RUG en het UMCG samen met verscheidene (inter)nationale partners.

Onderdelen die onderzoeksondersteuning bieden:

- **Centrum voor Informatie Technologie (CIT)** Het CIT is een ICT-expertisecentrum om grote datasets te analyseren en biedt ondersteuning bij het beheren en verwerken van onderzoeksdata. Het CIT heeft twee speerpunten: Big Data en gevoelige data.
- **Research and Innovation Support (RIS)** RIS biedt dienstverlening op het gebied van HPC, visualisatie (3D/AR/VR), Geodienst, data science en research data management/repositories.
- **Research Data Office (RDO)** Het RDO vervult een rol in eerstelijns support, research datamanagementplannen, en ondersteunt de onderzoekers vervolgens bij het selecteren van de juiste IT-oplossing.
- **UMCG - Service Desk Clinical Research Office (CRO)** De CRO is een service portal speciaal voor onderzoekers die klinisch onderzoek doen. Het CRO heeft een sterke link naar het RDO.
- **UMCG - Genomic Coordination Center (GCC)** Het GCC biedt ondersteuning voor de gehele research data cycle met een focus op bioinformatics, genomics en biobanking.
- **UMCG - Research Data Support (RDS)** Research Data Support ondersteunt onderzoekers bij het opzetten en uitvoeren van hoogwaardig mensgebonden onderzoek.

Huidige infrastructuur en bijbehorende diensten voor onderzoekers

Het CIT van de RUG levert een groot deel van de infrastructuur voor de onderzoekers. Data staat hierbij centraal: (big/complex/privacygevoelige) data is de kern van veel wetenschappelijk onderzoek. Naast opslagfaciliteiten (veilig en stabiel) biedt het CIT additionele IT-faciliteiten voor onderzoekers. De faciliteiten zijn in vier rekencentra opgesteld, met redundante netwerkkoppelingen intern en naar buiten. Ook maken de onderzoekers gebruik van externe dataservices zoals [DataverseNL](#), [DANS-Easy](#), [BBMRI-NL](#) en [CTMM/TRaIT](#).

REKENDIENSTEN	
Aantal clusters binnen de instelling	± 14 clusters
Totale omvang rekenkracht	5640 CPU cores (Peregrine HPC cluster PG), 200 nodes (PG), 30 TB RAM (PG), 600TB storage
3rd party clusters binnen de instelling	Genomic Coordination Center (GCC) clusters, Molecular Dynamics Group, Solid Matters Cluster, DAWN GPU cluster, LOFAR (Cobalt correlator/CEP4/CEP3) ASTRON (Aartfaac/Dragnet) OpenStack, VMware-rugcloud
Inkoop externe rekendiensten	SURFsara
OPSLAG	
Capaciteit centrale bulkstorage	<ul style="list-style-type: none"> • 3 PB Lustre; Datahandeling • 0.6 PB Lustre; HPC cluster • iRODS storage services • 0,5 PB incl dedupe data user storage Windows workspaces • 4 PB RUG-Cloud storage
Capaciteit centrale archief	<ul style="list-style-type: none"> • 2.5 PB Tape backup dubbel copy (2018) • 1 => PB disk backup met dedupe dubbel copy (2018) • 3.7 = > PB Tape Archive dubbel copy (2018) • 2 => PB Research archive (dcache - iRODS) (2018 - 2019)
NETWERK	
Routed capaciteit (extern)	2 x 10Gbit/s Campus IP 2 x 10Gbit/s HPC IP
MSP capaciteit	2 x 100Gbit/s MSP 1 x 10Gbit/s MSP
AUTHENTICATIE & AUTORISATIE-INFRASTRUCTUUR	
Aangesloten bij SURFconext federatie	Ja
Andere AAI-leverancier	Nee
DATACENTERS	
Datacenter Landleven (RUG)	Tier 2 vermogen 450 kW 1 x 100G netwerkkoppeling SURFnet 2 x 100G interconnect Smitsborg
Datacenter Smitsborg (RUG)	Tier 2 vermogen 300 kW 1 x 100G netwerkkoppeling SURFnet 2 x 100G interconnect Landleven
Datacenter Eemspoort UMCG (RUG deel)	Tier 3 vermogen 150 kW 2 x 40G netwerkkoppeling
Datacenter DUO (RUG - HPC)	Tier 1 vermogen 180 kW 2 x 100G netwerkkoppeling
Nieuwbouw datacenter RUG	Tier 3 vermogen 720 kW planning 2018-2019
EXPERTISE	
	Dataconsultants, datascientists, wetenschappelijk programmeurs, visualisatie, geodienst HPC-experts, IT-juristen

Virtual Research Workspace

De Research Workspace is een virtuele Windows ICT-omgeving, die de gebruiker deels zelf kan inrichten. Deze omgeving biedt toegang tot alle (research) ICT-diensten van de RUG en het UMCG. Deze diensten zijn 24/7 beschikbaar en zijn daar waar nodig voorzien van extra geheugen, extra krachtige processoren of zware grafische ondersteuning. De Research Workspace is zeer geschikt als de standaard universitaire werkplek niet toereikend is en wanneer de (Windows) applicaties niet geschikt zijn voor HPC-gebruik. Deze omgeving is ook geschikt bij strenge eisen aan de veiligheid van datasets of als extra privacy maatregelen nodig zijn.

Het portfolio Research Workspace bestaat uit de volgende workspaces, die een onderzoeker kan afnemen:

1. **Shared:** alle gebruikers delen alle componenten (geheugen/grafische kaart/processor).
2. **Dedicated Start:** componenten worden toegewezen aan een gebruiker, dus ze hoeven geen geheugen et cetera te delen.
3. **Dedicated Middle:** als Dedicated Start, maar met extra geheugen of rekenkracht.
4. **Dedicated Large:** op maat samen te stellen door opdrachtgever/onderzoeker.

Alle workspaces worden geleverd met een standaard basissoftwarepakket en standaard dataopslag. De universiteit zorgt voor het beheer (OS en applicaties). Alle andere dienstverlening is in overleg.

High Performance Computing

Voor grote aantallen of zware berekeningen kunnen onderzoekers terecht bij het centrale Peregrine Linux rekencluster van de universiteit. Honderden onderzoekers uit een breed scala aan disciplines gebruiken het rekencluster, van taalwetenschap tot moleculaire dynamica en van robotica tot genomics. Het CIT heeft het tweede nationale HPC-center in Nederland (na SURFsara). Het Genomics Coordination Center (GCC) biedt daarnaast HPC-faciliteiten voor mensgebonden onderzoek binnen het UMCG. Ook huisvest en beheert het CIT drie clusters voor de LOFAR-radiotelescoop.

Reality Center

Het Reality Center van het CIT beschikt over zeer geavanceerde virtual reality (VR)-faciliteiten en voorziet wetenschappers van hypermoderne faciliteiten op het gebied van (3D)-visualisatie. Door deze dienst centraal aan te bieden heeft het Reality Center optimale expertise kunnen opbouwen en maakt het hergebruik van kennis en software mogelijk. Onderzoekers kunnen bij het Reality Center terecht voor datavisualisatie, simulaties, 3D-design, cursussen en workshops. Het Reality Center maakt het mogelijk om nieuwe informatie te verkrijgen uit complexe data, complexe concepten uit te leggen, constructiefouten te voorkomen, tijd te besparen in het ontwerpproces en besluitvormingsprocessen te bevorderen.

Netwerkinfrastructuur

De universiteit beschikt over een state-of-the-art netwerk onder beheer van het CIT. Het CIT werkt continu aan de vernieuwing en verbetering van het netwerk. Daarbij richt het CIT zich op een betere beveiliging van het netwerk, vergroting van de capaciteit en virtualisatie van het netwerk (software defined networking). Daarnaast zorgen maatregelen als zoning van het netwerk voor optimale security.

(Big) data handling

In de afgelopen jaren is er veel expertise opgedaan in data handling, data delen en hergebruik van data, met name op het gebied van astronomie, energie en de life sciences. Voorbeelden van zulke grote researchinfrastructuur en projecten zijn [LifeLines](#), [Energysense](#), [GLIMPS](#) LOFAR en Euclid. Deze Big Data expertise is

vastgelegd in de bestaande research data support organisaties, bijvoorbeeld RIS en GCC, zodat deze expertise blijven kan worden aangeboden aan onderzoekers.

Data science

Om de toenemende hoeveelheid data voor onderzoek te benutten, is een goede en slimme analyse van de data nodig. Een datascientist kan helpen om uit de grote bulk aan gegevens informatie te halen. De datascientists werken in multidisciplinaire teams samen met experts van bijvoorbeeld het Center for Data Science and Systems Complexity (DSSC) en het Genomic Coordination Center (GCC). Ook worden er verschillende cursussen aangeboden, variërend van datawarehousing, programmeren in Python en R, gebruik van HPC- en visualisatiefaciliteiten tot machine- en deep learning. Daarnaast is het CBS Academic Data Center gevestigd in het CIT. De Data Federation Hub ondersteunt bij het samenbrengen van de verschillende expertises en hergebruik van best practices op het gebied van data science.

Geodienst

De Geodienst is hét ruimtelijk expertisecentrum van de RUG dat onderzoeksondersteuning en innovatieve diensten levert op het gebied van geodata, geo-analyse en GIS-software. De Geodienst onderhoudt een open data portal en biedt cursussen op maat.

Toolbox UMCG en RDMP (Research Data Management Plan) webtool

De toolbox clinical research bevat een breed scala aan Standard Operating Procedures (SOPs), templates en best practices om mensgebonden onderzoek te ondersteunen. Onderzoekers kunnen de toolbox gebruiken als hulpmiddel om gedragscodes, richtlijnen en wet- en regelgeving te vertalen naar de onderzoekspraktijk. Bovendien kunnen onderzoekers specificeren van welke faciliteiten (intern en extern) zij gebruik maken in hun onderzoeksprojecten. Voor research datamanagement is daarnaast de RDMP-webtool beschikbaar. De RDMP-webtool is per onderzoeksinstituut ingericht en helpt de onderzoeker de meest passende faciliteiten in te zetten bij het onderzoek.

Training en scholing

Voor het gebruik van applicaties en IT-faciliteiten zijn op maat trainingen en workshops beschikbaar over data-analyse, HPC, data science, geo software/tools en (3D-)visualisatie. Ook bieden zowel het UMCG als de RUG verschillende trainingen/workshops aan over research datamanagement en privacyaspecten van onderzoeksdata. Een website met een overzicht van alle trainingen is in aanbouw.

De agenda voor de nabije toekomst

Voor de universiteit is het van belang dat data 24/7 laagdrempelig, betrouwbaar en veilig beschikbaar is vanaf meerdere locaties. Ook de veilige samenwerking op basis van federatief identity management is zeer gewenst. De data moet toegankelijk zijn vanaf werkplek, research-werkplek, rekencluster en visualisatiesystemen. Hiertoe is een landelijke federatieve infrastructuur nodig met voldoende bandbreedte. Daarnaast is laagdrempelige toegang tot een verscheidenheid aan expertise belangrijk.

Data Federation Hub

Om de bestaande faciliteiten te coördineren en zichtbaarder te maken werken we aan intensivering van de samenwerking en bundeling van kennis. De Data Federation Hub is een gedeelde resource op het gebied van datamanagement en data science voor alle onderzoeksdomeinen. Primaire aandachtsgebieden zijn training, FAIR datacatalogi, kennisuitwisseling en data science.

De onderzoeker centraal

De onderzoeker komt nog meer centraal te staan. Hiervoor is het allereerst belangrijk dat de onderzoeker de ondersteuning weet te vinden. Daarnaast werken we toe naar een centraal aanspreekpunt. De Data Federation Hub dient daarbij als gateway naar alle beschikbare data expertise.

Virtual Research Environment (VRE)

De komende jaren werken we aan het realiseren van een Virtual Research Environment (VRE). De ondersteuning bij de gehele data life cycle staat hierbij centraal. Ook zetten we in op het verbinden met andere VRE's in Nederland. Met een goed werkend VRE kunnen we ook externe partijen toegang geven tot de verschillende interne systemen.

Human Subject Research (HSR) Programma

In januari 2017 zijn de RUG en het UMCG het universiteitsbrede *Human Subject Research (HSR) Programma* gestart voor data uit mensgebonden onderzoek. Het HSR heeft tot doel om een up-to-date infrastructuur en hulpmiddelen te leveren om (privacy)gevoelige onderzoeksdata op te slaan en te bewerken. Een infrastructuur die innovatief onderzoek stimuleert in lijn met het huidige beleid op het terrein van research datamanagement, Open Science en FAIR data.

(Inter)nationale positionering

De komende periode versterken we de (inter)nationale positie van onze IT-diensten en nemen we actief deel aan de (inter)nationale e-infrastructuur. We werken mee aan landelijke federatieve diensten, in samenwerking met onder andere SURF, Netherlands eScience Center en NWO. Ook op het gebied van health participeren we in verschillende (inter)nationale initiatieven, zoals BBMRI, ELIXIR, Health-RI en DTL. Daarnaast zetten we blijvend in op de (internationale) samenwerkingen op het gebied van astronomie. Tot slot nemen we vanuit Groningen actief deel aan het European Open Science Cloud initiatief.

Voor contact met de Data Federation Hub, mail naar: dfh@rug.nl

COLOFON

Auteur

Wietske Degen, Rijksuniversiteit Groningen
Salome Scholtens, Universitair Medisch Centrum Groningen

Projectleiding

Jan Bot, SURFsara

Redactie

Erik van der Spek, Hendrixx van der Spek
Jocelyn Manderveld, SURFnet
Judith Pijnacker, SURFnet

Eindredactie

Judith Pijnacker, SURFnet

Ontwerp

Crasborn Communicatie Vormgevers
www.crasborn.nl

Datum

Januari 2018

Copyright

Beschikbaar onder de licentie Creative Commons
Naamvermelding 4.0 Nederland.
www.creativecommons.org/licenses/by/4.0/nl

Meer weten?

SURF

Moreelsepark 48
3511 EP Utrecht

Postbus 19035
3501 DA Utrecht

+31 (0)20 88 787 30 00
info@surf.nl
support4research@surf.nl
www.surf.nl/s4r

SURF is de ICT-samenwerkingsorganisatie van het onderwijs en onderzoek in Nederland. Dankzij SURF beschikken onderzoekers, studenten en docenten in Nederland over de best mogelijke ICT-voorzieningen voor toponderzoek en talentontwikkeling.

The SURF logo consists of the word "SURF" in a bold, white, sans-serif font, centered within a black speech bubble shape that points downwards and to the right.