



Integrale aanpak voor digitalisering in de wetenschap

Uitvoeringsplan investeringen
digitale onderzoeksinfrastructuur





Integrale aanpak voor digitalisering in de wetenschap

Uitvoeringsplan investeringen digitale onderzoeksinfrastructuur

Inhoudsopgave

| | |
|---|-----------|
| 1 Samenvatting | 5 |
| 2 Achtergrond | 11 |
| 3 Doelstelling van de beoogde investeringen | 13 |
| 4 Uitvoeringsplannen | 15 |
| A Rekenfaciliteiten: Aanschaf nieuwe supercomputer | 16 |
| B Rekenfaciliteiten: Toekomstige investeringen in HPC & jaarlijkse gebruikersbijdrage reserveringen HPC | 17 |
| C Rekenfaciliteiten: Investerings in overige rekenfaciliteiten en opslag hardware | 19 |
| D Rekenfaciliteiten: Vernieuwing van de kennisbasis | 24 |
| E Rekenfaciliteiten: Expertise en ondersteuning van rekenfaciliteiten en datacentra door SURF | 29 |
| F Digitalisering: Impulsfinanciering lokale DCC's | 33 |
| G Digitalisering: Stimulering thematische DCC's | 36 |
| H Digitalisering: Ondersteuning DCC's door SURF | 39 |
| I Digitalisering: Investerings in eScience | 44 |
| 5 Bijlagen | 49 |
| Bijlage 1: Samenvatting adviezen uit Integrale aanpak voor digitalisering in de wetenschap | 49 |
| Bijlage 2: Definitie lokale DCC's | 53 |
| Bijlage 3: Invulling horizon-model binnen SURF | 54 |
| Bijlage 4: Toelichting voorziene onderwerpen innovatieagenda | 55 |
| Bijlage 5: Spreiding calls en projecten eScience Center 2014–2018 | 59 |
| Bijlage 6: Verklarende woordenlijst en afkortingen | 60 |

1 Samenvatting

De minister van Onderwijs, Cultuur en Wetenschap heeft de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) in 2018 gevraagd te adviseren over de groeiende digitalisering in de wetenschap en de benodigdheden met betrekking tot digitale infrastructuur en expertise om hiermee tevens bij te dragen aan de open science ambities geformuleerd in het Nationaal Plan Open Science (NPOS). NWO heeft hiertoe de Permanente Commissie Grootschalige Wetenschappelijke Infrastructuur (PC-GWI) om advies gevraagd. De ICT-subcommissie van de PC-GWI heeft in het voorjaar van 2019 een advies aan NWO uitgebracht, wat NWO heeft overgenomen in het advies *Integrale aanpak voor digitalisering in de wetenschap*. OCW heeft NWO vervolgens verzocht om de voorziene investeringen uit het advies nader uit te werken, inclusief een langetermijnstrategie voor *high performance computing*. Het voorliggende *Uitvoeringsplan investeringen digitale onderzoeksinfrastructuur* is hiervan het resultaat. Dit plan beschrijft voor alle onderdelen een voorstel op welke manier de middelen zullen worden ingezet. NWO zal eind 2022 evalueren in welke mate de extra investeringen tot dan toe het gewenste effect hebben bereikt en op grond daarvan de middelenverdeling eventueel herijken.

Deze samenvatting is bedoeld voor de algemene lezer. Vaktaal is zoveel mogelijk vermeden of uitgelegd. Naast de uitleg van bepaalde termen in onderstaande tekst is er een verklarende woorden- en afkortingenlijst te vinden in bijlage 6. In de uitwerking van de verschillende onderdelen wordt meer jargon gebruikt.

Grofweg zullen de extra ICT middelen op twee hoofdlijnen worden ingezet: 1. Rekenfaciliteiten en 2. Digitalisering, met ieder verschillende subonderdelen. Alle onderdelen zijn uitgewerkt in hoofdstuk 4 van dit uitvoeringsplan.

In onderstaande tabel staan alle onderdelen van het advies *Integrale aanpak voor digitalisering in de wetenschap* en de financiële implicaties samengevat, met verwijzingen naar de uitgewerkte plannen verderop in dit uitvoeringsplan.

| | Impuls-financiering 2018/2019 | Structurele financiering ≥ 2020 | Pagina |
|--|-------------------------------|---------------------------------|--------|
| Rekenfaciliteiten | | | |
| Aanschaf Supercomputer – door SURF | 18 (+2) M€ ¹ | | 16 |
| Jaarlijkse reserveringen toekomstige HPC behoeften – advies subcommissie ICT van de PC GWI | | 4 (+2) M€ ² | 17 |
| Grid computing en andere rekenfaciliteiten – door SURF | 3,0 M€ | 2,5 M€ | 19 |
| Opslag hardware – door SURF | 1,5 M€ | 1,0 M€ | 19 |
| Kennisvernieuwing rekenfaciliteiten – door SURF | 3,5 M€ | 3,5 M€ | 24 |
| Expertise en ondersteuning rekenfaciliteiten en datacentra – door SURF | 2,0 M€ | 2,0 M€ | 29 |
| Digitalisering | | | |
| Impulsfinanciering Lokale Digitale Competentie Centra – via een eenmalige call bij NWO | 4,5 M€ | | 33 |
| Stimulering Thematische Digitale Competentie Centra – via een doorlopende call bij NWO | 2,5 M€ ³ | 2,5 M€ | 36 |
| Inrichten en onderhouden gefedereerd netwerk, en ondersteuning lokale DCC's – door SURF | 2,0 M€ | 2,0 M€ | 39 |
| Netherlands e-Science Center – vergroten calls | 3,0 M€v | 2,5 M€ | 44 |
| Totaal | 40,0 M€ | 20,0 M€ | |

Rekenfaciliteiten

Om in de toekomst nog topwetenschap te blijven bedrijven is laagdrempelige toegang tot rekencapaciteit essentieel voor onze onderzoekers. Door de digitalisering van alle wetenschapsgebieden heeft elke wetenschapper behoefte aan rekencapaciteit. Deze vraag naar rekencapaciteit zal een grote impact hebben op het onderzoeksbudget.

In sommige wetenschapsgebieden is de digitalisering al eerder begonnen, denk maar aan het gebruik van supercomputers (*High Performance Computers*) en *Grid Computing* (*High Throughput Computing*). Recentelijk is *GPU-computing* erbij gekomen. Door de sterke groei in het verzamelen van data vindt er ook een explosieve groei in data-opslag plaats.

High Performance Computing (HPC) richt zich op met zeer grote snelheid uitvoeren van complexe berekeningen, zoals bijvoorbeeld klimaatberekeningen. Dit vereist supercomputers die voor individuele instellingen nauwelijks te financieren zijn. Vandaar dat SURF op nationaal niveau daar invulling aan geeft. Op Europees niveau concentreert EuroHPC zich op *pre-exascale* supercomputers als voorloper op de toekomstige *exascale* supercomputers.

High Throughput Computing (HTC) richt zich op zoveel mogelijk (vaak kleinere) rekentaken binnen een bepaalde tijd uit te voeren. Deze berekeningen worden uitgevoerd met relatief goedkopere computers die door middel van een computernetwerk met elkaar verbonden zijn. Dit wordt ook wel *Grid-computing* genoemd. Veel wetenschapsgebieden zullen door de digitalisering hier gebruik van maken. Voorbeelden

¹ Voor de vervanging van supercomputer Cartesius zal SURF een bedrag van 2 M€ bijdragen.

² SURF wordt gevraagd jaarlijks 2 M€ bij te dragen aan de reserveringen voor toekomstige HPC behoeften, te vergaren uit tariefinkomsten van academische gebruikers en commerciële exploitatie van de opvolger van supercomputer Cartesius.

³ De eerste call ter stimulering van thematische DCC's zal begin 2020 opengaan, het budget voor 2019 zal daarom worden doorgeschoven naar 2020.

hiervan zijn het gebruik van Grid computing om in datasets van wereldwijde beursinformatie te bestuderen hoe problemen op de financiële markten kunnen ontstaan bij een crisis; analyses van DNA-databases voor het beter begrijpen van ziektes; of het analyseren van grote hoeveelheden data die worden gegenereerd bij botsingen van deeltjes in de CERN deeltjesversneller.

GPU-computing maakt gebruik van de *Graphics Processor Unit* die gebruikt wordt om de schermen van computers aan te sturen. De essentie is dat in één operatie een hele dataverzameling bewerkt kan worden. Dit sluit heel goed aan bij het rekenmodel dat voor *Machine Learning* wordt gebruikt. *Machine Learning* is een onderzoeksgebied in de kunstmatige intelligentie waarbij algoritmes en technieken worden ontwikkeld waarmee computers kunnen leren. Veel wetenschapsgebieden zullen gebruik maken van *Machine Learning* om patronen te ontdekken in grote dataverzamelingen. *Machine Learning* wordt bijvoorbeeld gebruikt in spamfilters om automatisch patronen te herkennen en te bepalen of een e-mail ongewenst is; in spraakherkenning om gesproken taal te herkennen en om te zetten in een geschreven tekst; in gezondheidsmonitoring om risico's te kunnen voorspellen en voor te zijn.

Investing in rekencapaciteit

De nationale supercomputer is voor een aantal onderzoeksvelden een noodzakelijke voorwaarde om internationaal competitief te kunnen zijn. Om deze onderzoeksvelden te kunnen blijven bedienen en aansluiting te behouden met internationale computing initiatieven als PRACE, adviseert NWO een bedrag van 18 M€ beschikbaar te stellen voor het vervangen van de nationale supercomputer Cartesius. SURF verzorgt de aanschaf van de nieuwe supercomputer in 2020 en legt zelf 2 M€ in.

Gezien de groeiende behoefte aan rekenkracht zal deze supercomputer naar verwachting in 2025 ook weer vervangen moeten worden. NWO adviseert om vanaf 2020 jaarlijks 4 M€ uit de infrastructuurmiddelen te reserveren voor toekomstige high performance computing (HPC) behoeften. NWO adviseert verder om SURF om een jaarlijkse bijdrage van 2 M€ te vragen voor de toekomstige supercomputer behoeften. Deze reserveringen van totaal 6 M€ per jaar vullen een spaarpot waarmee, naast een uiteindelijke vervanger, ook een tussentijdse upgrade van de nieuwe supercomputer wordt gefinancierd.

Met het perspectief van investeringen in HPC is het belangrijk om ook een langetermijnstrategie voor HPC te hebben. Het Europese samenwerkingsproject EuroHPC en andere mogelijkheden voor internationale samenwerking op HPC gebied kunnen een waardevolle aanvulling op de nationale rekenfaciliteiten bieden. NWO heeft de ICT-subcommissie van de PC-GWI gevraagd om deze strategie op te stellen. Dit doet zij via interviews en een veldraadpleging en zal in november 2019 rapporteren aan NWO.

Verder adviseert NWO om vanaf 2020 jaarlijks 2,5 M€ ter beschikking te stellen voor de financiering van andere reken-systemen bij SURF, zoals Grid computing, met een eenmalige impuls van 3 M€ in 2019. Door te investeren in een breed spectrum van computing faciliteiten kan SURF ook gebruikers voor wie de nationale supercomputer niet geschikt is voorzien in hun behoeften. Voor dit onderdeel heeft SURF een investeringsplan gemaakt.

Daarnaast adviseert NWO om vanaf 2020 jaarlijks 1 M€ te investeren in uitbreiding van de opslag- en archiveringsmogelijkheden in de reken- en datacentra van SURF, om tegemoet te komen aan de snelle toename van datavolumes, met een eenmalige impuls van 1,5 M€ in 2019. Voor dit onderdeel heeft SURF een investeringsplan gemaakt.

Expertise en ondersteuning rekenfaciliteiten en datacentra – door SURF

SURF zorgt ervoor dat de Nederlandse kennisinstituten aangehaakt blijven bij de nieuwste technologische ontwikkelingen op ICT gebied. Daarom is het belangrijk te investeren in de vernieuwing en ontwikkeling van

hun kennis. NWO adviseert om vanaf 2019 jaarlijks 3,5 M€ te investeren in het opbouwen en vernieuwen van de benodigde kennis bij SURF. Samen met de Chief Information Officers (CIO's) van de universiteiten heeft SURF een plan opgesteld. Er zijn vijf thema's voorgesteld voor kennisopbouw. Een voorbeeld is machine learning. Door de ontwikkeling en de verspreiding van de nieuwste machine learning mogelijkheden aan onderzoekers en instellingen, beschikken onderzoekers in de toekomst over de beste mogelijkheden voor toepassing van machine learning in onderzoek. De thema's zullen jaarlijks worden geactualiseerd.

Om het potentieel van alle bovengenoemde hardware optimaal te benutten is het bovendien nodig dat SURF de gebruikers goed kan ondersteunen. Om de benodigde expertise voor gebruikersondersteuning aan te kunnen trekken en behouden, adviseert NWO om vanaf 2019 jaarlijks 2 M€ ter beschikking te stellen aan SURF. SURF breidt met deze middelen de menskracht uit de komende jaren.

Digitalisering

Hoe meer onderzoeksdata wordt gedeeld, des te belangrijker is het om deze goed te beschrijven en klaar te maken voor hergebruik. Daarom werken verschillende wetenschapsdomeinen samen en worden er afspraken gemaakt om onderzoeksdata FAIR te maken. FAIR- staat voor vindbaar (Findable), toegankelijk (Accessible), uitwisselbaar (Interoperable) en herbruikbaar (Reusable). Nederland doet onder andere mee aan het realiseren van de Europese Open Science Cloud (EOSC) om wetenschappelijke uitwisseling en vooruitgang in Europa te bevorderen. FAIR-data speelt hierin een sleutelrol. Het is ook een van de centrale pijlers in het Nationaal Plan Open Science (NPOS).

In Digitale Competentie Centra, of Digital Competence Centers (DCC's), vinden onderzoekers ondersteuning en technologische hulpmiddelen op het gebied van onderzoeksdata, onderzoekssoftware, en open- en FAIR-data. Lokale DCC's binnen kennisinstellingen vormen het eerste aanspreekpunt voor onderzoekers. Onderlinge uitwisseling van kennis en technologie kan worden gestimuleerd door middel van een veilig en gefedereerd netwerk en het oprichten van thematische DCC's. NWO wil de ontwikkeling van DCC's stimuleren door middel van het uitzetten van calls ter bevordering van lokale en thematische DCC's. Daarnaast worden er middelen beschikbaar gesteld aan SURF voor expertise en ondersteuning van de DCC's.

Het Netherlands eScience Center (NLeSC) zet via open *calls for proposals* zijn expertise en eScience Research Engineers in voor samenwerkingsprojecten met Nederlandse universiteiten en kennisinstellingen. Door het beschikbaar stellen van extra middelen aan het NLeSC zullen deze *calls for proposals* worden vergroot, waarmee ze een groter bereik zullen hebben en betere samenwerking en uitwisseling van kennis stimuleren.

Lokale Digitale Competentie Centra

Een lokale DCC is onderdeel van een wetenschappelijke instelling en goed toegankelijk voor onderzoeksgroepen. Lokale DCC's hebben als functie het faciliteren van data management van wetenschappelijke data. De functie betreft het faciliteren van data-opslag, -analyse, delen etc. zodat onderzoeksgroepen aan de Open Science/FAIR uitgangspunten kunnen voldoen. Meer concreet zijn DCC's: 1. landingsplaats voor projectsoftware voor onderhoud en beschikbaarstelling; 2. kennis- en adviescentrum voor FAIR data en software en lokale ICT-infrastructuur; 3. knooppunt in een gefedereerd netwerk voor data, computing en expertise.

Voor de oprichting en versterking van lokale Digitale Competentie Centra (DCC's) adviseert NWO éénmalig een bedrag van 4,5 M€ als impulsfinanciering beschikbaar te stellen. NWO ontwikkelt een eenvoudig financieringsinstrument om deze middelen bij DCC's in te zetten. Hiermee wordt voorzien in de versterking

van de capaciteit van datastewards en research software engineers aan kennisinstellingen, waarbij deze kennis en expertise vindbaar en laagdrempelig toegankelijk wordt voor onderzoekers. Belangrijk voorwaarde bij de toewijzing van middelen is de eigen bijdrage van de indienende instelling. Eind 2019 zal NWO de call openzetten voor aanvragen.

Thematische Digitale Competentie Centra

Een thematische DCC is een samenwerking tussen bestaande organisaties waarin onderzoeksgroepen rondom een specifiek thema besloten hebben om samen hun datamanagement in onder te brengen. Deze samenwerking wordt vaak medegefinancierd door de betreffende onderzoeksgroepen en hebben hun meerwaarde reeds bewezen. Een thematische DCC richt zich specifiek op het bij elkaar brengen van data van onderzoeksgroepen en het publiek beschikbaar stellen van data om analyse binnen het thema te vergemakkelijken.

NWO adviseert om voor thematische DCC's vanaf 2019 jaarlijks een bedrag van 2,5 M€ ter beschikking te stellen. Dit bedrag wordt vrijgegeven via een doorlopende call en op basis van gerichte projectvoorstellen die een eigen bijdrage vragen van de indienende instelling. In thematische DCC's worden bijvoorbeeld projecten uitgevoerd die FAIR bevorderen, of projecten die andere aspecten van de digitalisering ten goede komen, zoals het opzetten of koppelen van databanken, of het implementeren, breder beschikbaar maken en generaliseren van software.

Inrichten gefedereerd netwerk en ondersteuning van Digitale Competentie Centra

Lokale en thematische DCC's worden knooppunten in een gefedereerd netwerk. SURF speelt een sleutelrol in het technisch faciliteren en coördineren van het veilige, gefedereerde systeem dat de lokale en thematische DCC's onderling zal verbinden.

Voor het ontwikkelen van een veilig gefedereerd systeem en de ondersteuning vanuit SURF aan de lokale DCC's adviseert NWO om vanaf 2019 jaarlijks een bedrag van 2 M€ beschikbaar te stellen aan SURF. SURF heeft samen met de CIO's van universiteiten en onderzoeksinstituten zes onderwerpen vastgesteld waarop activiteiten worden ontwikkeld. Uitgangspunt hierbij is dat onderwerpen moeten bijdragen aan versterking van de landelijke IT- en data-infrastructuur voor wetenschappelijke onderzoekers in Nederland en de digitalisering van onderzoek. Hierbij zijn ook de instellingen zelf in de gelegenheid financiering aan te vragen bij SURF voor verdere ontwikkeling van het federatieve netwerk. Via een jaarlijkse evaluatie- en planningcyclus worden deze onderwerpen en zwaartepunten voor ondersteuning van DCC's tussen de kennisinstellingen en SURF afgestemd. Eén van de onderwerpen is bijvoorbeeld de ondersteuning door SURF van virtuele onderzoeksomgevingen, die meer en meer gebruikt worden.

De technische specificaties van het gefedereerde systeem zullen aan de NPOS stuurgroep worden voorgelegd, gezien het belang voor de ontwikkelingen in *open science*.

Netherlands eScience Center – vergroten calls

Het Netherlands eScience Center zet via bestaande calls for proposals zijn expertise en eScience Research Engineers in voor samenwerkingsprojecten met Nederlandse universiteiten en kennisinstellingen. Hierdoor wordt de relatie verstevigd tussen informatica-kennis en ICT-infrastructuur enerzijds en vragen en ontwikkelingen uit de onderzoeksdomeinen anderzijds. NWO adviseert om vanaf 2020 jaarlijks 2,5 M€ ter beschikking te stellen voor het uitbreiden van de bestaande calls, met een eenmalige impuls van 3 M€ in 2019, waarmee meer samenwerkingsprojecten kunnen worden opgestart.

2 Achtergrond

‘Op alle terreinen is de impact van digitalisering van de wetenschap merkbaar. De manier van inwinning, analyse, deling en presentatie van onderzoeksdata en -bronnen is sterk aan het veranderen: de omvang en complexiteit van datasets is fors toegenomen en internationale en discipline-doorkruisende samenwerkingen zijn belangrijker dan ooit. Bij sommige wetenschapsgebieden is de digitalisering jaren geleden al op gang gekomen en is het delen van data, software en reken capaciteit gemeengoed. Andere wetenschapsgebieden staan aan het begin van deze cruciale digitaliseringslag. De behoeften lopen dan ook sterk uiteen. Voor sommige wetenschappers is toegang tot *high-end computing* essentieel, anderen hebben behoefte aan het ontsluiten en/of combineren van publiek beschikbare data en weer anderen hebben behoefte aan het delen van software. In alle gevallen betreft het een combinatie van infrastructuur en expertise om de wetenschap beter te ondersteunen’⁴.

In het regeerakkoord Rutte III zijn middelen ter beschikking gesteld voor de versterking van de ICT-infrastructuur in Nederland. Deze extra investeringen in de nationale digitale infrastructuur maken het mogelijk op de beschreven ontwikkelingen in te spelen zodat Nederlandse onderzoekers kunnen blijven meedraaien in de wereldtop van data-intensief onderzoek. De minister van Onderwijs, Cultuur en Wetenschap (OCW) heeft NWO in 2018 gevraagd te adviseren over deze groeiende digitalisering in de wetenschap en de benodigheden met betrekking tot digitale infrastructuur en expertise, om hiermee tevens bij te dragen aan de ambitie die is geformuleerd in het Nationaal Plan Open Science (NPOS).

Over de inzet van de investeringen heeft NWO de Permanente Commissie Grootchalige Wetenschappelijke Infrastructuur (PC-GWI) om advies gevraagd. Een speciale ICT subcommissie van de PC-GWI, onder leiding van prof. dr. Apers, heeft na brede raadpleging van het wetenschappelijke veld en van de belangrijkste relevante bestuurlijke en faciliterende partijen in het voorjaar van 2019 een advies over de inzet van de extra middelen aan NWO opgesteld. Het citaat bovenaan deze pagina komt uit dit advies. NWO heeft het advies overgenomen en aan de minister van OCW voorgelegd. Een samenvatting van het advies staat in bijlage 1.

OCW heeft NWO verzocht om de voorziene investeringen nader uit te werken in samenspraak met de relevante partijen uit het veld zoals de universiteiten en SURF, inclusief een langetermijnstrategie voor *high performance computing*. Het voorliggende Uitvoeringsplan Investerings Digitale Infrastructuur is hiervan het resultaat. Dit beschrijft voor alle onderdelen een voorstel op welke manier de middelen zullen worden ingezet. In sommige gevallen is de uitwerking een proces dat nog gaande is, zoals het advies over toekomstige investeringen in HPC, waar een uitgebreide veldraadpleging nodig is.

⁴ Op 26 maart 2019 werd het advies “*Integrale aanpak voor digitalisering in de wetenschap*” door NWO gepubliceerd.

3 Doelstelling van de beoogde investeringen

Met de extra structurele investeringen en de impulsfinanciering, ingezet als een integraal coherent plan, heeft NWO als doel om effect bereiken op de volgende aspecten:

- Data delen volgens de principes van FAIR (*Findable, Accessible, Interoperable, Reusable*) en open science.
- Lokaal en via een op (inter)nationaal niveau gefedereerde digitale infrastructuur.
- Onderzoeksdata eenvoudiger analyseren door gebruik van krachtiger reken capaciteit, grotere opslagcapaciteit en nieuwe technologische ontwikkelingen.
- Ontwikkelde software meer benutten, doordat deze meer bekend en toegankelijk is.

De belangrijkste adviezen van NWO gaan in op het sterker toegankelijk, deelbaar en vindbaar maken van onderzoeksdata en -software, waardoor hergebruik van beide sterker kan plaatsvinden. Er wordt voorzien in de versterking van de capaciteit van datastewards en research software engineers aan kennisinstellingen met Digitale Competentie Centra (DCC's). NWO stimuleert hiermee de ontwikkeling van DCC's, teneinde deze expertise vindbaar en zichtbaar te positioneren. Deze DCC's zijn voor onderzoekers de lokale hub voor ondersteuning en technologische hulpmiddelen op het gebied van onderzoeksdata, open en FAIR-data. Onderlinge uitwisseling van kennis en technologie tussen DCC's wordt gestimuleerd middels de financiering van een veilig ('secure'), gefedereerd systeem, gecoördineerd door SURF. Naast lokale DCC's is er in het NWO advies ook sprake van thematische DCC's. Deze organisaties kunnen gezien worden als landelijke, thematisch georiënteerde DCC's.

Daarnaast is er in bijna alle wetenschapsgebieden een sterk groeiende behoefte aan meer reken capaciteit, snellere netwerken en grotere opslagcapaciteit. Snellere supercomputers en bijbehorende rekenkracht en ICT infrastructuur kunnen de groei mogelijk maken. Om dit alles voor onderzoekers breed toegankelijk te maken is tot slot expertise en ondersteuning nodig; laagdrempelig en lokaal waar mogelijk, specialistisch en gecentraliseerd waar nodig.

Evaluatie

Met bovenstaande doelstellingen in het achterhoofd zal NWO eind 2022 evalueren in welke mate de extra investeringen tot dan toe het gewenste effect hebben bereikt. Gezien de dynamiek en ontwikkelingen in het wetenschappelijke landschap, nationaal en internationaal, is de evaluatie mede gericht op eventuele herijking van beschikbare extra middelen. Zo kunnen de ontwikkelingen in analysemethoden een grote invloed hebben op specifieke behoeftes aan rekenkracht. NWO werkt de uitgangspunten van de evaluatie verder uit, samen met de betrokken partijen, inclusief succescriteria die kunnen worden meegenomen in de evaluatie.

4 Uitvoeringsplannen

Alle onderdelen van het uitvoeringsplan zijn uitgewerkt in nauw overleg met de uitvoerende partijen. Er is intensief contact geweest met SURF, NLeSC en de CIO's van de universiteiten; zij hebben in belangrijke mate bijgedragen aan de verschillende onderdelen van het uitvoeringsplan.

A Rekenfaciliteiten: Aanschaf nieuwe supercomputer

NWO heeft SURF gevraagd de aanbesteding uit te voeren voor een nieuwe supercomputer, ter waarde van 20 M€, waarvan SURF zelf 2,0 M€ bijdraagt. Na het besluit tot starten van het aanbestedingstraject zijn de volgende stappen voorzien:

- Het eerder door SURF in kaart gebrachte veld van kandidaat-leveranciers is door voorgenomen fusies en aangepaste strategieën van leveranciers significant gewijzigd. Hierdoor zal de aanbesteding starten met een nieuwe kwalificatie;
- Dialoogrondes met drie tot maximaal vijf leveranciers, afhankelijk van kwalificatie. Doel hiervan is te komen tot een definitief bestek, met criteria voor de levering van een zo geschikt mogelijk general-purpose systeem;
- Vaststelling en uitbrengen van het definitieve bestek;
- Offertefase voor leveranciers;
- Bepaling door de technische adviescommissie (opgenomen in het besturingsproces) van de voorkeursvolgorde van kandidaat-leveranciers op basis van de in het definitieve bestek vastgelegde weging van de beoordelingscriteria;
- De selectiecommissie (opgenomen in het besturingsproces), overwegende het advies van de technische adviescommissie, stelt de voorkeursvolgorde van leveranciers vast;
- Het SURF-bestuur bekrachtigt de voorkeursvolgorde van leveranciers;
- Na goedkeuring door selectiecommissie en SURF-bestuur meldt de technische adviescommissie de voorgenomen leverancier aan alle leveranciers en houdt, in verband met eventuele bezwaren, de wettelijke stand-still periode van 20 dagen in acht;
- In overleg tussen de NWO RvB en het SURF-bestuur, op initiatief van het SURF-bestuur en op basis van een notitie door de selectiecommissie, wordt de NWO RvB tijdens de stand-still periode geïnformeerd over het verloop van het aanbestedingstraject en het proces dat geleid heeft tot de keuze van de leverancier;
- Op basis hiervan worden de genoemde middelen direct na afloop van de stand-still periode door de NWO RvB vrijgegeven aan SURF en kan SURF/SURFsara BV het contract met de leverancier afsluiten. Hiermee vindt de definitieve gunning plaats.
- Dialoog met de voorgenomen leverancier om te komen tot een contract, na afloop van de stand-still periode;
- Levering, installatie, acceptatie en inproductie name worden vervolgens uitgevoerd conform de afspraken vastgelegd in het contract tussen SURF/SURFsara BV en de leverancier.
- Samenstelling selectiecommissie:
 - Erik Fledderus, voorzitter SURF
 - Kurt Jansen, Directeur operations SURFsara BV
 - Peter Michielse, CTO SURFsara BV
 - Lid van Commissie Wetenschappelijk gebruik Supercomputer (WGS)
- Doorlooptijd vanaf start aanbestedingstraject:
 - Tot aan afsluiten contract, afhankelijk van aantal benodigde dialoogrondes: 8–10 maanden;
 - Vanaf afsluiten contract tot levering, installatie en in productie nemen: 5 maanden.

B Rekenfaciliteiten: Toekomstige investeringen in HPC & jaarlijkse gebruikersbijdrage reserveringen HPC

Om deze behoeftevoorziening in HPC te verduurzamen is voorzien in een 'spaarpot', waarin jaarlijks, vanaf 2020, 4 M€ uit de nieuwe structurele ICT middelen wordt gereserveerd, aangevuld met jaarlijks 2 M€ afkomstig uit bijdragen van gebruikers van de faciliteit.

OCW heeft NWO verzocht een langetermijnstrategie (LTS) voor Nederlandse HPC investeringen te formuleren. NWO heeft de PC-GWI gevraagd hierover te adviseren, De subcommissie ICT is stelt dit advies op en voert hiervoor onder andere een veldraadpleging uit.

1 Gebruikersbijdrage van 2 M€ per jaar: evaluatie en tijdpad

Het realiseren van de jaarlijkse gebruikersbijdrage van 2 M€ is een voorwaarde voor de jaarlijkse bijdrage van OCW ter hoogte van 4 M€ aan de reserveringen. Deze reservering is in cash. SURF werkt in overleg met de ledenraad aan een voorstel hoe deze gebruikersbijdrage wordt gerealiseerd. Het volgend proces is afgesproken tussen NWO en SURF:

- De eerste periode van reserveringen zal worden ingezet op het moment dat de vervanger van supercomputer Cartesius in gebruik wordt genomen, en duurt twee jaar.
- Na het eerste jaar vindt er een tussentijdse evaluatie plaats, waarin SURF, NWO en OCW bezien of het innen van gebruikersbijdragen voor HPC het gewenste resultaat oplevert, en of de strategie voor het heffen van gebruikersbijdragen moet worden bijgesteld.
- NWO beheert de 'HPC spaarpot'.
- Twee jaar na ingang van het innen van gebruikersbijdragen voor HPC zal worden getoetst of de gebruikersbijdrage door SURF is gerealiseerd. In dat geval is de weg vrij voor OCW om een bijdrage toe te voegen aan de reserveringen voor HPC investeringen.
- De gebruikersbijdrage wordt onderwerp van de evaluatie en herijking van de nieuwe middelen voor digitale wetenschappelijke infrastructuur (jaarlijkse 20 M€) die in 2022 zal plaatsvinden.
- Indien SURF op het moment van evaluatie 4 M€ heeft gerealiseerd en het proces hieromtrent in de evaluatie positief wordt beoordeeld, volgt een tweede periode van reserveringen. Deze loopt van het derde tot en met het vijfde jaar, geteld vanaf de ingebruikname van de opvolger van supercomputer Cartesius. Hierbij vindt jaarlijks een tussentijdse evaluatie plaats, zoals na het eerste jaar.
- SURF rapporteert iedere zes maanden over de stand van zaken van het realiseren van de gebruiksbijdrage.

2 Kader advisering toekomstige investering HPC

Momenteel loopt de aanbestedingsprocedure voor de opvolger van de Cartesius, de supercomputer bij SURF. De Lange Termijn Strategie voor *High Performance Computing* (LTS HPC) betreft de opvolger van de opvolger van de Cartesius en de toegang tot *pre-exascale*, en later *exascale* supercomputers op Europees niveau.

Het primaire uitgangspunt bij het vormgeven van de strategie is de behoefte van de Nederlandse HPC gebruikers. Deze wordt in kaart gebracht door de ICT-subcommissie van de PC GWI via de volgende acties:

- gesprekken met verschillende partijen: – gebruikers van nationale en internationale supercomputers; SURF; leidinggevendenden van wetenschappelijke instellingen; nationale en internationale beleidsmakers.

- een veldraadpleging, uitgevoerd door de CIO's van de universiteiten waarbij een brede schare gebruikers van wetenschappelijke rekenfaciliteiten gevraagd wordt naar hun behoeften en voorkeuren hierin. De veldraadpleging is in gang gezet en maakt deel uit van de uiteindelijke langetermijnstrategie, die zal worden opgeleverd in oktober 2019.

Halverwege dit proces zijn de eerste observaties:

- de vraag naar HPC is een beperkt deel van de enorme vraag naar rekencapaciteit
- laagdrempelige toegang tot rekencapaciteit is essentieel voor onderzoekers
- excellentie-gedreven toegang tot EuroHPC is essentieel
- zowel wetenschapsgebieden als wetenschappelijk instellingen dienen een investeringsplan op te stellen en up-to-date te houden om de enorme vraag naar rekencapaciteit te plannen.
- om op Nederlandse onderzoekers in specialistische behoeften voor rekencapaciteit te kunnen voorzien is het noodzakelijk samenhang aan te brengen in de HPC voorzieningen die aanwezig zijn aan kennisinstellingen

Zodra de gesprekken afgrond zijn zal de langetermijn strategie HPC ter goedkeuring naar de RvB van NWO gaan. De planning is dat dat in november zal gebeuren.

C Rekenfaciliteiten: Investerings in overige rekenfaciliteiten en opslag hardware

Inleiding

Deze investeringen zijn bedoeld om de faciliteiten voor onderzoekers duurzaam aan te sluiten op de behoeftes van onderzoekers, en *state of the art* te houden in termen van capaciteit en functionaliteit. Uit diverse onderzoeken is naar voren gekomen dat de optimale vervanging van hardware vanuit duurzaamheidsoogpunt en performance ligt bij ongeveer vier jaar.

Dit bestedingsplan geeft een overzicht van de investeringen voor de domeinen Data Processing en Data Management. Het gaat om investeringen in 2019 van 3 M€ voor overige rekenfaciliteiten en 1,5 M€ voor opslag hardware en daarna vanaf 2020 jaarlijks 2,5 M€ voor overige rekenfaciliteiten en 1,0 M€ voor opslag hardware. Voor de jaren 2019-2020 is daarbij aangegeven wat de capaciteitsgroei is. Voor de jaren 2021–2022 is er een financiële doorkijk opgenomen, maar is mede door onzekerheid in ontwikkelingen in techno-logie en prijs een concrete groei van capaciteit niet opportuun. Dit bestedingsplan is opgesteld met de volgende uitgangspunten:

- Investerings in computing services vinden plaats in het kader van de vervanging van de nationale supercomputer. Naar verwachting zal dat een voortzetting betekenen van de huidige heterogene architectuur van de supercomputer (combinatie van CPU's en verschillende typen GPU's, evt. andere processoren);
- Investerings ten behoeve van innovatie en kennisopbouw/deling zijn ondergebracht in het SURF Open Innovatie Lab. Deze zijn voor 2019 en 2020 opgenomen in de bestedingen in het kader van vernieuwing van de kennisbasis.

Data Processing investeringen in 2019/2020 en doorkijk naar 2021/2022

De investeringen in het Data Processing domein betreffen rekenfaciliteiten anders dan de nationale supercomputer, bekend als de Nationale Grid infrastructuur, HPC Cloud platform en big data platformen. Voor ieder van deze faciliteiten is zowel sprake van vervanging (wegens eerder benoemde veroudering) als uitbreiding van reken capaciteit. Een belangrijke component in deze infrastructuur zijn specifieke gecombineerde reken-opslagfaciliteiten waarmee grote hoeveelheden data kunnen worden verwerkt.

Tabel 1: Toelichting faciliteiten Data Processing domein

| | |
|--------------------------------|--|
| Grid Service | De oorspronkelijke high throughput ("anders dan supercomputer") omgeving, geoptimaliseerd voor de verwerking van grote hoeveelheden data, oorspronkelijk opgezet voor hoge energie fysica, thans ook gebruikt door andere gemeenschappen (voornamelijk life sciences en in mindere mate astronomie); |
| HPC Cloud | Is de omgeving waarin een gebruiker de volledige controle over zijn rekeninfrastructuur, besturingssysteem en analysesoftware heeft. Deze omgeving is geschikt voor toepassingen (bijvoorbeeld complexe workflows of meer interactieve applicaties) waarbij de batch-scheduling zoals deze worden gebruikt in de reguliere HPC of Grid omgevingen niet gewenst zijn of andere besturingssystemen dan Linux vereisen. |
| Big data en data analytics | Hier worden met behulp van frameworks zoals Hadoop, Elastic Search en Jupyter de analyse van ongestructureerde datasets, streaming data en interactieve analyse mogelijk. Voorbeelden van toepassingen is de analyse van Twitter Data of analyse van sensordata die is verkregen uit Internet-of-Things. |
| Online Data Processing Storage | Dit omvat de opslag die vooral gebruikt wordt door de Grid-gebruikers, maar ook door HPC cloud en big data gebruikers; |

Voor de toekomst zien wij dat steeds meer hardware wordt geïntegreerd in één 'infrastructuurlaag' waarop vervolgens verschillende rekenplatforms kunnen worden aangeboden. In 2018 is gestart met de integratie van deze platformen in een single hardware platform, met daarbovenop een uniforme open source software laag. De faciliteiten Grid, HPC Cloud en big data zullen vanaf een enkele infrastructuur geleverd gaan worden. Dit geeft zogenaamde elasticiteit waardoor, als er tijdelijk minder vraag is naar het ene platform, deze infrastructuur niet leeg hoeft te staan maar kan worden gebruikt voor een ander platform waar schaarste is. Dit betekent een verbeterde bezettingsgraad van de infrastructuur (efficiëntie), waardoor er tevens meer gebruikers geholpen kunnen worden (effectiviteit). Dat levert de volgende voordelen op:

- Eenvoudigere en efficiëntere aanschaf van apparatuur;
- Eenvoudiger beheer: enkel platform in plaats van meerdere platformen;
- Minder capaciteitsbegrenzing meer als gevolg van separate platformen;
- Meer gebruiksvriendelijkheid in toegang;
- Snellere en efficiëntere deployment van nieuwe applicaties en toepassingen.

Daarnaast ligt het in het karakter van dit domein dat er voortdurend wordt gekeken naar de (door) ontwikkeling van reken- en analyseplatformen. Voorbeelden hiervan zijn de ontwikkeling van Research Cloud, een omgeving waarin onderzoekers zelf op een laagdrempelige manier zonder heel veel IT kennis een onderzoeksomgeving kunnen creëren met daarin reken-, opslag-, en applicatie faciliteiten. Een ander voorbeeld is een omgeving voor high throughput data processing die lijkt op Grid, maar dan veel eenvoudiger voor gemeenschappen die minder faciliteiten en geen wereldwijde opschaling nodig hebben. Deze platformen worden in beginsel ontwikkeld vanuit het onderdeel kennisvernieuwing en kunnen uiteindelijk gebruikmaken van de infrastructuur capaciteit.

Onderstaande tabel geeft de geschatte investeringen voor de genoemde faciliteiten aan:

Tabel 2: Overzicht investeringen Data Processing domein

| | 2019 | | | 2020 | | | 2021 | | | 2022 | |
|-----------------------------|--------------------|------------|-------------|---------------------|------------|-------------|---------------------|---------------|-------------|------------|-------------|
| | Huidige capaciteit | Vervanging | Uitbreiding | Capaciteit | Vervanging | Uitbreiding | Capaciteit | Vervanging | Uitbreiding | Vervanging | Uitbreiding |
| Data Processing Compute | Grid | 7280 cores | 830 k€ | | 7280 cores | 670 k€ | | | | | |
| | HPC Cloud | 1560 cores | 450 k€ | 200 k€ | 2360 cores | 25 k€ | 100 k€ | | | | |
| | Big Data | 1500 cores | | | 1500 cores | | | ca13000 cores | 570 k€ | 500 k€ | 842 k€ |
| | Ontwikkelplatform | 0 cores | | 225 k€ | 750 cores | 440 k€ | | | | | |
| Data Processing Storage | 14,5 PB | 465 k€ | 290 k€ | 16,7 PB | 450 k€ | 300 k€ | ca 19 PB | 720 k€ | 260 k€ | 557 k€ | 260 k€ |
| Capaciteit bij DNI partners | | | 500 k€ | 1800 cores 780TB | | 500 k€ | 1800 cores 780TB | | 500 k€ | | 500 k€ |
| Totaal | | | | 2960 k€ | | | 2485 k€ | | 2550 k€ | | 2499 k€ |

De prijs-prestatie index wordt met het voortschrijden van technologie over het algemeen gunstiger. Voor rekentoeepassingen betekent dit per investeringsbedrag een toename in prestatie per core (uitdrukt in Flop/s of HEPspec06, naar gelang de toepassing), maar dit resulteert niet in prijsdaling per core. Door de jaren heen zal het aantal cores daarom min of meer gelijk blijven. Voor opslag geldt zoals gezegd een prijsdaling per volume-eenheid, maar wordt deze meer dan opgeheven door de toename in behoefte. Naar verwachting zullen medio 2020 de verschillende infrastructuren geëvolueerd zijn tot het eerdergenoemde enkele platform.

SURFsara en Nikhef vormen samen de Tier-1 faciliteit ten behoeve van de LHC-experimenten bij CERN. De bij Nikhef geïnstalleerde apparatuur wordt niet alleen gebruikt door de hoge energie fysica community, maar ook door andere gebruikerscommunities. In die zin maakt de apparatuur van Nikhef deel uit van de nationale ICT-infrastructuur. Om die reden ontvangt Nikhef een bijdrage voor het gebruik van faciliteiten voor het deel dat niet binnen haar onderzoeksmissie hoort, zoals gebruik door life sciences.

Binnen Data Processing zien we, naast de verwachtingen van Lofar en LHC, de volgende toename van mogelijke gebruikerscommunities:

- aardobservatie: nieuwe satellieten naast Tropomi (bijvoorbeeld Sentinel-1);
- nieuwe sequencing/genomics project zoals MinE en andere *omics projecten;
- astronomie: ALMA-NL (Allegro), SKA;
- plantgenetica (WUR, UU);
- medisch gerelateerde life sciences, bv processing van MRI images voor brein analyse,
- kankeronderzoek, etc;
- geo-engineering projecten bv ten bate van risicosimulaties dijken;
- climate modeling groepen (o.a. bij de WUR).

Data Management investeringen in 2019/2020 en doorkijk naar 2021/2022

Investeringen voor het Data Management domein betreffen investeringen ten behoeve van het up-to-date houden en het faciliteren van de groei van de faciliteiten. Het betreft hier data-opslag die in beginsel niet aan een specifieke reken- of processing dienst is gekoppeld. Onderscheid wordt hier gemaakt tussen online opslag (waarbij data direct beschikbaar is voor gebruik vanaf traditionele disks en nieuwere media) en offline opslag (waarbij data op tapecartridges staat die door een robot worden weggezet en teruggehaald voor gebruik). Door de combinatie van offline storage en online storage wordt een kosten efficiënte dataopslag faciliteit gecreëerd die rekening houdt met toegangspatronen van data door gebruikers en applicaties (storage life cycle management).

Tabel 3: Toelichting faciliteiten Data Management domein

| | |
|-----------------|--|
| Offline storage | Bevat onderzoeksdata van afgerond onderzoek die zijn gearchiveerd, maar vooral zeer grootschalig experimenteel verkregen data die in een later stadium pas worden geanalyseerd en vanuit kosten oogpunt in de tussentijd offline worden opgeslagen. Dit laatste is bijvoorbeeld gangbaar in astronomie (bv LOFAR) en hoge energie fysica (bv van de Large Hadron Collider). |
| Online storage | Analyse van deze data kan plaatsvinden nadat de betreffende data van offline naar online opslag verplaatst is, of van data door de toegangspatronen permanent online opslag vereisen. Om die reden is er naast opslag op tape/cartridges altijd een online platform nodig waarop te analyseren data (tijdelijk) geplaatst kan worden en een snel netwerk daartussen. |
| Research Drive | Betreft een nieuwe online opslagfaciliteit waarmee onderzoekers in groepen eenvoudig data kunnen delen. Deze faciliteit koppelt bovendien goed met rekendiensten als Research Cloud waarmee de toegankelijkheid tot diensten en data aanzienlijk vereenvoudigd en verhoogd wordt, ook voor niet zeer IT technisch onderlegde onderzoekers. Deze faciliteit zal in overleg beschikbaar worden gemaakt als onderdeel van rekentijdaanvragen. |

Voor de jaren 2019–2022 wordt het volgende voorzien:

Tabel 4: Overzicht investeringen Data Management domein

| Huidige capaciteit | 2019 | | | 2020 | | | 2021 | | | 2022 | |
|-------------------------|------------|-------------|------------|------------|-------------|------------|------------|-------------|------------|-------------|--|
| | Vervanging | Uitbreiding | Capaciteit | Vervanging | Uitbreiding | Capaciteit | Vervanging | Uitbreiding | Vervanging | Uitbreiding | |
| Offline storage 60 PB | 630 k€ | 283 k€ | 70 PB | 480 k€ | 434 k€ | 90 PB | 890 k€ | 586 k€ | 57 k€ | 555 k€ | |
| Online storage 2 PB | 605 k€ | | 2 PB | | 250 k€ | 4 PB | | 250 k€ | | 250 k€ | |
| Research Drive platform | | | | | 70 k€ | | | | | 70 k€ | |
| Totaal | | | | 1518 k€ | | 1234 k€ | | 1726 k€ | | 932 k€ | |

De variabele kostencomponenten voor zowel offline als online storage bestaan voornamelijk uit tapeopslag media, disksystemen en licentiekosten. Grotere investeringspieken voor de offlineopslag zijn periodieke verversing van opslagmedia (veelal naar een nieuwe generatie cartridges met een grotere capaciteit) en diskcache. Voor online storage bestaan de investeringspieken uit de vervangingskosten van oudere storage-elementen die zijn afgeschreven en waarvoor de leverancier geen ondersteuning meer biedt die bovenop de reguliere uitbreidingen als gevolg van groei komen. Over het algemeen dalen door de tijd voor opslagmedia de kosten per opslagvolume. In de laatste jaren is echter een afvlakking waar te nemen. Tegelijkertijd zijn de kosten voor opslagmedia – naast bijvoorbeeld softwarelicenties en menskracht – slechts een beperkt aandeel van de integrale kosten voor opslag. Daarnaast neemt de jaarlijkse opslagbehoefte van de gebruikersgroepen die gebruik maken van de SURF voorzieningen exponentieel toe, met een factor 1,3.

In het huidige systeem wordt al het gebruik van offline storage gefinancierd uit middelen die SURF uit dit bestedingsplan ontvangt. Groeicijfers en de aankomende behoeften van voornamelijk grootschalige wetenschappelijke infrastructuur laten zien dat dit geen duurzaam bekostigingsmodel is. Daarbij laat dit geen ruimte voor het aanbieden van nieuwe faciliteiten die passen bij de opslagbehoefte van onderzoekers. Samen met de wetenschappelijke gemeenschappen werkt SURF naar een nieuw model waarin een gebruikersbijdrage vanuit deze gemeenschappen ontstaat. Naar gelang de succesvolle implementatie hiervan zal dit leiden tot een aangepast investeringsoverzicht en bijbehorende groeicijfers voor de jaren 2020–2022.

Succescriteria

Succes van investeringen kan gemeten worden over drie assen: financieel, community en gebruikers perspectief, en output criteria.

Allereerst de financiële toetsingscriteria:

| Toetsingscriterium financieel | Benodigd hiervoor |
|--|--|
| Prijs/prestatie verhouding van de investering | Goede prijs- en prestatievergelijking tussen aanbieders |
| Moment van investering | Inzicht in gevraagde capaciteit, inzicht in optimale vervangingsstrategie op gebied duurzaamheid en technologie roadmaps aanbieders. |
| Impact van de investering voor vergroting en diversiteit gebruikers (schaalvoordeel) | Inzicht in aantal potentiële gebruikers en gemeenschappen |

Daarnaast kan gekeken worden toetsingscriteria vanuit community en gebruikers perspectief:

| Toetsingscriterium community/gebruikers | Benodigd hiervoor |
|---|--|
| Aantal gebruikers en communities | Registratie bij NWO en bij SURF |
| Aantal aanvragen voor toegang | Registratie bij NWO en bij SURF |
| Tevredenheid | NPS monitoring door NWO/SURF |
| Diversiteitsratio, gebruik resources gemeten over instellingen en communities | Monitoring gebruik instellingen en communities |
| Mate van matching beschikbare capaciteit aan vraag in de tijd (utilisatie) | Monitoring gebruik vs beschikbare capaciteit |

Als laatste perspectief kan gekeken worden naar output criteria:

| Toetsingscriterium output | Benodigd hiervoor |
|----------------------------------|--|
| Output/publicaties/impact | Inzicht in NWO publicaties rapportages en de SURF output rapportages |
| Operationele performance | Meten van kwaliteitsparameters als beschikbaarheid, betrouwbaarheid, veiligheid etc. |

Periodieke registratie van deze indicatoren vindt deels al plaats. Met NWO zal besproken worden over de vorm van rapportage en de daarin te hanteren kwalificatie.

D Rekenfaciliteiten: Vernieuwing van de kennisbasis

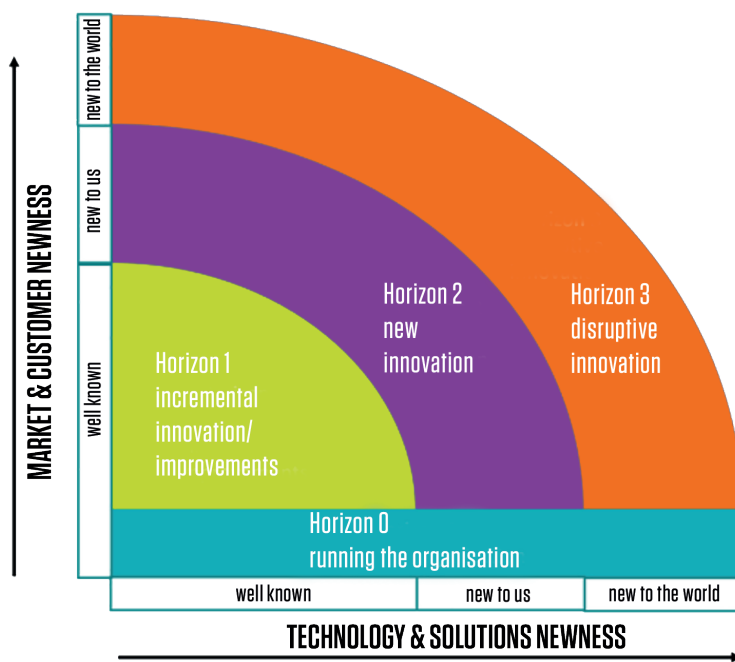
Inleiding

Dit hoofdstuk beschrijft de wijze waarop binnen SURF vernieuwingen tot stand komen, beschrijft de uitgangspunten voor de besturing van een innovatieagenda en doet een voorstel voor de hierin opgenomen onderwerpen. Hiervoor is jaarlijks 3,5 M€ beschikbaar vanaf 2019.

Positionering van vernieuwing en innovatie

Vernieuwing en innovatie zijn essentieel voor een infrastructuur en voorzieningen die aan blijven sluiten bij de behoeften van en toepassingen binnen het wetenschappelijk onderzoek.

Vernieuwing en innovatie zijn brede begrippen die kunnen reiken van kleine verbeteringen die deel zijn van de going concern tot aan grote disruptieve ontwikkelingen met een groot afbreukrisico. De verschillende delen van dit spectrum kennen hun eigen dynamiek en vragen om een andere sturing. Om verschillende vormen van vernieuwing te duiden gebruiken we hier het horizon-model (zie figuur 1) om een onderscheid te maken tussen vernieuwing van vandaag, van morgen en overmorgen. Bijlage 4 bevat een beschrijving van het horizon-model.



Figuur 1: Horizon-model voor vernieuwing en innovatie van Baghai, Coley et al. (2000)

De binnen dit uitwerkingsplan opgenomen agenda bestrijkt onderwerpen in horizon 2 die (nog) niet in het *LifeCycle Product Management*-proces vallen⁵ en enkele onderwerpen die in horizon 3 liggen.

De vernieuwingsactiviteiten bij SURF hebben tot doel om op lange termijn de noden en behoeften van het Nederlands wetenschappelijk onderzoek te kunnen blijven faciliteren. SURF heeft daarbij geen zelfstandige onderzoekstaak als kennisinstelling, maar gaat daar waar nodig de samenwerking aan met (informatica)onderzoek bij de kennisinstellingen. Door de laatste ontwikkelingen uit (informatica)onderzoek te kunnen vertalen in ondersteuning voor onderzoeksprojecten en eisen aan de infrastructuur, blijven de nationale reken- en datavoorzieningen geschikt voor het onderzoek van morgen en overmorgen. Via kennisproducten en community-buildingactiviteiten wordt de opgedane kennis ter beschikking gesteld aan de onderzoeks- en ICT-ondersteuning bij instellingen.

Veranderingen in het afstemmingskader

Op het moment van schrijven loopt een verandertraject voor de besturing van SURF. Hierin staat de vergroting van de (bestuurlijke) betrokkenheid vanuit de leden voorop. Dit is niet alleen noodzakelijk voor het merendeelcriterium waaraan SURF voor het behoud van de quasi-inbestedingspositie aan gehouden is, maar leden geven zelf ook aan dat de snelheid en impact waarmee digitalisering in het onderzoek plaatsvindt maakt dat zij zelf behoefte hebben om als coöperatie gezamenlijk invulling te geven aan innovatie voor onderwijs en onderzoek. Afhankelijk van besluitvorming zullen de gewenste veranderingen gedurende de tweede helft van 2019 worden uitgewerkt en in de ledenraad van 11 december 2019 worden vastgesteld. Dit geeft vanaf 2020 een nieuw kader voor de afstemming rondom vernieuwing en innovatie in aansluiting op het LCPM-proces.

Uitgangspunten voor kennisvernieuwing

Continue kennisvernieuwing door vroege innovatie in horizon-2 en -3 is cruciaal om duurzaam state-of-the-art ICT-infrastructuurdiensten en expertise te kunnen ontwikkelen en aan te kunnen bieden voor de ondersteuning van wetenschappelijk onderzoek in Nederland. Het NWO advies 'Integrale aanpak voor digitalisering in wetenschap' zegt daarover: 'SURF is de aangewezen partij om te waarborgen dat de Nederlandse kennisinstellingen aangehaakt blijven bij de nieuwste technologische ontwikkelingen; daartoe is het belangrijk te investeren in de vernieuwing en ontwikkeling van hun kennis. De subcommissie adviseert NWO vanaf 2019 jaarlijks 3,5 M€ te investeren in het opbouwen en vernieuwen van de voor de rekenfaciliteiten benodigde kennisbasis bij SURF. Deze middelen voorzien voor een deel in het verduurzamen van de financiering van de expertise die reeds aanwezig is bij SURF, en essentieel is voor de instandhouding van een ICT kenniscentrum van wereldniveau.'

SURF en de kennisinstellingen werken nauw samen in de kennisvernieuwing. SURF neemt daarbij de rol van voorloper en aanjager van vernieuwing in infrastructuur en toepassingen, maar zorgt vooral ook dat zoveel mogelijk kennis wordt geëxtraheerd, geborgd, en zo breed mogelijk wordt gedissemineerd en hergebruikt binnen de SURF instellingen. De agenda voor kennisvernieuwing wordt in nauwe samenspraak en afstemming met de SURF instellingen opgesteld en bijgesteld, zodat aansluiting bij de wensen behoeften van de instellingen wordt gewaarborgd. Relevante kennis en expertise binnen instellingen wordt ingezet om bij te dragen aan de kennisvernieuwing.

Vooruitlopend op de nieuwe afstemmingskaders, worden de volgende uitgangspunten voorgesteld:

- **Open innovatieprincipes:** in alle activiteiten in de Kennisvernieuwing wordt actief de samenwerking met de SURF-leden gezocht. Daarbij is ook duidelijk en nadrukkelijk ruimte open voor eigen inbreng van

⁵ Voor de ontwikkeling van een idee naar een levensvatbare dienst die aansluit op de behoefte, hanteert SURF het proces Life Cycle Product Management (LCPM). In het LCPM-proces worden stapsgewijs in afstemming met de SURF Portfolio Adviescommissie ideeën voor diensten en producten ontwikkeld. Deze commissie bestaat uit een afvaardiging vanuit de leden die ervoor moet zorgen dat de ideeën aansluiten bij de behoefte en diensten met een duurzaam businessmodel geïntroduceerd kunnen worden.

nieuwe ideeën door de leden en wordt aangesloten bij open innovatie van instellingen. Open participatie, open delen van kennis en technologie, en sturing op leren en het vergroten van kennis staan voorop.

- **Sturing vindt plaats op basis van de innovatieagenda.** Op deze agenda zijn onderwerpen opgenomen. Voor 2019 is de volgende agenda door SURF opgesteld, deze agenda⁶ wordt in samenspraak jaarlijks geactualiseerd (zie ook volgende punt):
 - A. Exascale Computing
 - B. Quantum Computing
 - C. Machine Learning
 - D. Data Management
 - E. Data Processing & Streaming Data
- **De innovatieagenda komt in samenspraak tot stand.** De bestaande afvaardiging vanuit het CSC-WO overleg waarmee dit voorstel is afgestemd wordt uitgebreid met vertegenwoordiging uit GWI's, WGS en experts uit het wetenschappelijk domeinonderzoek en het informaticaonderzoek. Deze groep zal een tijdelijke commissie vormen totdat het nieuwe besturingsmodel van SURF in werking treedt. Daarbij wordt jaarlijks de actualisering van de agenda, de verdeling van middelen per onderwerp en type project en de financiële bijdrage aan instellingen besproken. Vanuit goed gebruik wordt hierbij verhouding tussen Horizon 2 en 3 van ca 2:1 aangehouden.⁷ De commissie komt tenminste drie keer per jaar bij elkaar en bespreekt tussentijds op onderwerpniveau de voortgang en een eventuele bijgestelde aanpak. Ook kan zo tijdig ingespeeld worden op nieuwe ontwikkelingen. Per onderwerp wordt een commissielid verbonden voor eventueel benodigde tussentijdse afstemming;
- Binnen de kaders van de innovatieagenda heeft SURF **de ruimte om projecten te definiëren**, zodat snel ingespeeld kan worden op nieuwe ontwikkelingen, kansen en vraagsturing. SURF definieert projecten binnen de afgestemde innovatieagenda en rapporteert tijdens commissievergaderingen over voortgang en resultaten. Daarbij kan het gaan om drie type projecten:
 - i. *Technology scouting en assessments*: dit zijn kortdurende relatief kleine projecten rondom nieuwe ICT-technologieën die momenteel geen onderdeel uit maken van de voor onderzoekers beschikbare reken- of datavoorzieningen, maar er in de toekomst wel impact op kunnen hebben. Vaak gebeuren deze (vaak korte) verkenningen, assessments, prototyping en demonstrators in samenwerking met onderzoeksafdelingen of innovatielabs van universiteiten of hogescholen, met technologie-providers, start-ups, en industrie. Deze verkenningen worden gedaan vanuit de ICT-technologie, maar altijd met de toepassing in het vizier. Uitkomsten moeten inzicht geven in de toepasbaarheid van die technologieën, de impact op het onderzoek en de eisen aan toekomstige voorzieningen;
 - ii. *Strategische kennisinvestering*: lange-termijn projecten voor het behouden van een goede kennispositie op de kerngebieden van de nationale infrastructuur: computing, data processing, en datamanagement.⁸ Een significant deel van het onderzoekswerk wordt in nauwe samenwerking uitgevoerd bij onderzoeksgroepen bij Universiteiten, Hogescholen of onderzoeksinstellingen. Dit omvat ook deelname aan EC gefinancierde projecten voor bijvoorbeeld e-infrastructuur innovatie;
 - iii. *Samenwerkingsprojecten in onderzoek*: veelal meerjarige projecten waarin een brug wordt geslagen tussen wetenschappelijk onderzoek (bijvoorbeeld computer science of informatica-onderzoek) in combinatie met een toepassingsgebied. Vaak gaat het hier om inzet van SURF expertise bij projecten die door NWO of TKI (topconsortium voor kennis en innovatie) worden gefinancierd. Daar SURF niet als onafhankelijke kennisinstelling aangemerkt wordt, is zij voor participatie in dergelijke projecten aangewezen op aanwending van de gelden uit 'Vernieuwing van de Kennisbasis'.
- **Kennis binnen instellingen kan worden ingezet om bij te dragen aan de effectieve kennisvernieuwing op nationaal niveau.** SURF heeft hierbij niet de rol van financier, maar zet hierbij financiële middelen in om samenwerking rond nieuwe ontwikkelingen met instellingen te stimuleren en te bevorderen, en het mogelijk te maken unieke kennis bij instellingen in te zetten in het kader van de kennisvernieuwing. Voor projecten die met een financiële bijdrage aan instellingen worden uitgevoerd vindt een open inschrijving plaats. In overleg met de adviescommissie worden selectiecriteria opgesteld en een onafhankelijke beoordeling ingericht. Voor de bestedingen onder het voor 2019 beschikbaar gestelde

⁶ Zie bijlage 4 op een toelichting op de voorziene onderwerpen innovatieagenda

⁷ Zie ook: <https://www.pimcy.nl/the-3-horizons-drive-continuous-innovation/>

⁸ De andere kerngebieden networking, trust en identity worden uit andere bronnen bekostigd.

budget is 10%–15% van het totaalbudget gealloceerd voor activiteiten die onder regie van SURF bij instellingen worden uitgevoerd.

Vervolgstappen

Afhankelijk van de besluitvorming door NWO/OCW en afhankelijk van de door de SURF ledenraad vastgestelde afstemmingskaders zullen de volgende vervolgstappen worden gezet:

- Uitwerking van de innovatieagenda en onderwerpen in projecten incl. samenwerking met instellingen voor 2020. Bijlage 4 bevat de geïdentificeerde onderwerpen voor 2020.
- Opzetten van samenwerkingen.
- Verankering van de kennisvernieuwing binnen de afstemmingskaders.

Financiering van Kennisvernieuwing

Naast de 3,5 M€ geadviseerde middelen in het NWO-advies, heeft SURF in 2019 en 2020 nog de beschikking over een tweetal tijdelijke financieringen uit de Motie Van Meenen-Vos.

Tabel 1: Financiering kennisvernieuwing domeinen Data Management, Data Processing en Computing

| | 2019 | 2020 | 2021 e.v. |
|--|--------|--------|-----------|
| Motie Van Meenen-Vos (OCW) | 1,0 M€ | – | – |
| Motie Van Meenen-Vos (EZK) – subtotaal Vernieuwing van reken- en data-infrastructuur en -diensten ⁹ | 1,0 M€ | 1,0 M€ | – |
| NWO-advies | 3,5 M€ | 3,5 M€ | 3,5 M€ |
| Totaal | 5,5 M€ | 4,5 M€ | 3,5 M€ |

Vanaf 2021 en verder zal alle innovatie voor reken- en datavoorzieningen vanuit de geadviseerde middelen gefinancierd worden. Voor de EZK-middelen is er een opzichzelfstaand projectplan. De OCW-middelen zijn bij NWO reeds aangegeven in het SURF werkplan 2019 dat op verzoek van NWO is opgesteld.

Tabel 2 geeft een eerste verdeling van de middelen uit het NWO-advies aan onderwerpen op de innovatie-agenda per horizon. Volledigheidshalve is hierbij de aansluiting gemaakt met de domeinen uit het genoemde werkplan.

⁹ De totale financiering vanuit EZK bedraagt 2,0 M€.

Tabel 2: Vernieuwing van de kennisbasis in domeinen met voorziene onderwerpen en omvang

| Domein | Onderwerp | Horizon | Omvang | | | | |
|--|----------------------------------|---------|------------------------|----------------------|-------------------|-----------------------|---------------------------------------|
| | | | T.l.v. NWO-advies (k€) | Totaal begrote fte's | Investerings (k€) | Overige kosten (k€) * | Totaal bijdrage aan instellingen (k€) |
| Data Management | Data Management | 2 | 100 | 0,5 | | | |
| | | 3 | 295 | 1,0 | 75 | 25 | |
| Data Processing | Data Processing & Streaming Data | 2 | 1604 | 9,3 | | 162 | |
| | | 3 | 227 | 0,8 | | 100 | |
| Computing | Exascale computing | 2 | 331 | 1,5 | 85 | 25 | |
| | Machine learning | 2 | 77 | 0,5 | | | |
| | Quantum Computing | 3 | 239 | 0,6 | | 150 | |
| Innovatiemanagement & Kennisdisseminatie | | | 220 | 0,6 | | 60 | |
| Lange-termijn behoefte onderzoekscommunities | | | 400 | 2,0 | | 100 | |
| Totaal menskracht, bijdragen en investeringen | | | 3493 | 16,7 | 160 | 222 | 400 |

* Overige kosten, waaronder operationele infrastructuurkosten en kosten kennisbenutting

E Rekenfaciliteiten: Expertise en ondersteuning van rekenfaciliteiten en datacentra door SURF

Inleiding

Dit hoofdstuk geeft een overzicht van de voorziene activiteiten van SURF die vallen in de categorie expertise en ondersteuning. Hiervoor is jaarlijks vanaf 2019 2 M€ beschikbaar. Het profiel van adviseurs die bij SURF werken combineert een diepgaande kennis van de nationale infrastructuur met kennis van het wetenschappelijke domein waar zij zelf vaak een achtergrond in hebben. Het stelt gebruikers daarmee in staat om het maximale uit de infrastructuur te halen, ten behoeve van hun wetenschappelijk onderzoek en publicaties. Het betreft hier dus bijvoorbeeld activiteiten die bijdragen aan het paralleliseren van gebruikersapplicaties, bijdragen aan het opzetten van *workflows* voor gebruikers en bijdragen aan het visualiseren van wetenschappelijke resultaten.

Overzicht bestedingen

De meeste activiteiten vinden plaats in de domeinen Data Processing en Computing. Het betreft hier personeel dat bij SURF in dienst is: Data Processing ca 7,5 fte, Computing ca 4,5 fte. Binnen het domein Data Management, gezien haar meer generieke opslagfaciliteiten, vindt minder directe ondersteuning van individuele gebruikers en gebruikersgroepen plaats.

Ondersteuning en expertise dragen sterk bij aan het effectief gebruik van de infrastructuur en de gebruikerstevredenheid en vormen hiermee naast de gebruikersvriendelijke infrastructuur en een licht en snel administratief proces, de belangrijkste pijler in de verlaging van de toegangsdrempel voor nieuwe gebruikersgroepen. Het is daarom van belang de inzet van expertise en ondersteuning in samenhang met een capaciteitsvraag op de harde infrastructuur te zien.

Jaarlijks wordt samen met NWO het actuele budget voor rekestijd op de supercomputer en – sinds 2019 – ook de andere rekenfaciliteiten vastgesteld. Onderzoekers kunnen vervolgens rekestijd (in core uren) op deze systemen aanvragen. Deze gebruikersvraag wordt door de commissie Wetenschappelijk Gebruik Supercomputer getoetst. Bij schaarste wordt er geprioriteerd of wordt er minder toegekend dan aangevraagd. Met ingang van 2020 zal in het budget voor deze call ook de beschikbare ondersteuning worden opgenomen. Onderzoekers zullen in hun rekestijdaanvraag dan ook de benodigde menskracht expliciet moeten opnemen.

7,5 fte overeenkomend met 11625 uur ondersteuning binnen het Data Processing domein

Dit betreft vooral:

- Het mogelijk maken van het gebruik van grid, HPC Cloud en de opschaling van de onderliggende applicatie. Dit vergt veel ondersteuning aan relatief onervaren gebruikers. Vaak voorkomende oplossingen worden generiek gemaakt, zodat nieuwe gebruikers hier meer zelfstandig gebruik van kunnen maken;
- Het optimaliseren van bestaande workflows – waar kan gebruik gemaakt worden van het feit dat de nationale infrastructuur zoveel groter is dan hetgene dat lokaal bij een instelling beschikbaar is;
- Het gebruik van specifieke applicaties en tools (Spark bijvoorbeeld) ten behoeve van big data simulaties toepassingen;
- Het gebruik van de SURF research Cloud, met als onderdeel het mogelijke gebruik van public cloud voorzieningen door onze gebruikers.

4,5 fte overeenkomend met 6975 uur ondersteuning binnen het Computing domein

Dit betreft vooral:

- Het optimaliseren en paralleliseren van gebruikersapplicaties die zo mogelijk breder gebruikt worden in de Nederlandse wetenschappelijke community. Het betreft hier vaak quantum-chemische en fluid dynamics applicaties;
- Huidige ondersteuning van gebruikers bij de migratie van applicaties van traditionele CPU's naar GPU's, zodat onderzoekers bij kunnen blijven bij de gangbare vorm van grootschalige simulaties;
- Migratie van big data applicaties naar de nationale supercomputer, waarmee meer mogelijkheden tot opschaling beschikbaar komen;
- Het visualiseren van grote hoeveelheden data en/of resultaten van wetenschappelijke simulaties.

Ontsluiting van rekensystemen

Een goede ondersteuning is een cruciaal onderdeel om de geavanceerde mogelijkheden van de landelijke rekensystemen beschikbaar te maken naar een brede groep onderzoekers. Dit gaat vaak verder dan het aanbieden van een account en het geven van instructies – om die reden wordt ondersteuning vanuit de adviseurs ingezet. SURF heeft een duidelijke strategie en aanpak ten aanzien van het ontsluiten van gebruikers op de verschillende systemen. Voor nieuwe gebruikers is het van belang dat men (1) bekend gemaakt wordt met de mogelijkheden van de nationale infrastructuur, en (2) nieuwe gebruikers op weg worden geholpen bij het gebruik van de rekenfaciliteiten die beschikbaar zijn binnen deze infrastructuur:

Ad.1: outreach

- Outreach in samenwerking met instellingen (via het Support4Research programma), waarbij via onderzoeksondersteuners bij instellingen onderzoekers worden geïnformeerd;
- Ontwikkelen van communicatiemateriaal over faciliteiten voor de long-tail doelgroep;
- (Poster)presentaties en demonstraties van diensten bij wetenschappelijke conferenties / partners onderzoekscommunities zoals Health-RI en ODISSEI;
- Demonstraties van diensten in samenwerking met centrale IT afdelingen bij instellingen aan onderzoekers en onderzoeksondersteuners. Als opvolging worden vervolgens pilots aangeboden.

Outreach is onderdeel van apart begrootte activiteiten vormt geen onderdeel van dit bestedingsplan. Meer specifiek zullen deze worden afgestemd op de ontwikkeling van DCC's.

Ad.2: training

- Uitgebreide algemene trainingsprogramma's voor nieuwe gebruikers;
- Samengestelde trainingsprogramma's met onderzoekscommunities en samenwerkingsorganisaties zoals DTL;
- SURF diensten aanbieden en toegankelijk maken via dienstportalen en websites van onderzoekscommunities en instellingen.

Voor gebruikers die de transitie maken van initieel gebruik naar geavanceerd gebruik is er een pakket aan trainingen beschikbaar gemaakt, zowel in nationaal verband als in internationale samenwerking (PRACE training centers, EUDAT, CompBioMed). Hierbij dient gedacht te worden aan bijvoorbeeld het opschalen van applicaties en de toepassing van geavanceerde machine learning technieken.

Een overzicht van alle trainingen van SURF op het gebied van reken- en datadiensten is te vinden op: <https://www.surf.nl/en/agenda/research-and-ict>.

Training voor individuele onderzoekers kunnen onderdeel zijn van rekentijdaanvragen. Trainingsaanvragen op instellingsniveau zijn onderdeel van aparte overeenkomsten.

Tot slot investeert SURF veel in afstemming met onderzoeksgemeenschappen, en co-creatie van diensten met onderzoeksgemeenschappen, om zo een optimale aansluiting te vinden. Dit dient tevens als input voor investeringen in faciliteiten en diensten. Dit valt buiten dit bestedingsplan.

Succescriteria

Uiteindelijk zijn het (internationale) succes en de impact van Nederlands onderzoek maatgevend voor het nut van investeringen en expertise. Het aantal ondersteuningsverzoeken is een indicator voor het nut van expertise en ondersteuning. Met de inbedding van de aanvragen van ondersteuningsuren in de NWO subsidie-instrumenten kunnen zowel registratie als verslaglegging vastgelegd worden. Toetsingscriteria kunnen dan zijn:

Tabel 1: Toetsingscriteria voor succes

| Toetsingscriterium | Benodigd hiervoor |
|--|--|
| Aantal aanvragen voor ondersteuning, aantal gevraagde uren | Opname van een call in NWO instrumentarium |
| Tevredenheid van gebruikers over ondersteuning | Opname van een call in NWO instrumentarium, inzicht in deze gebruikersrapportages met daarin een door de aanvrager aangegeven mate van succes. Aansluiting op NPS. |
| Aantal referenties naar ondersteuning in wetenschappelijke publicaties | Opname in subsidievoorwaarden in andere NWO calls |
| Aantal trainingen/bootcamps | Registratie door SURF |
| Kwantificering van outreach/kennisdeling | Registratie door SURF, in overleg met NWO |
| Samenwerkingen met gemeenschappen in dienstontwikkelingen | Aantal samenwerkingen in projecten |

Doorkijk naar 2021–2022

Om competitief onderzoek te kunnen blijven doen hebben onderzoekers in toenemende mate behoefte aan toegang tot ICT-infrastructuur en moeten zij op de hoogte te blijven van nieuwe technieken in het gebruik van die infrastructuur. Concrete voorbeelden van dit laatste zijn het gebruik van nieuwe hardware architecturen, de toepassing van *Artificial Intelligence* technieken als *machine learning* in data analyse en *processing*, en de mogelijkheden die *public clouds* van commerciële providers zoals Amazon Web Services en Microsoft Azure bieden.

Het betreft hier dus altijd de combinatie van architectuur, applicatiekennis en kennis van het wetenschappelijke domein. Uitbreiding en ontwikkeling van expertise in de juiste richting is continu aan de orde. Onderstaande tabel is een inschatting van fte's in 2020–2022, op basis van te verwachten nieuwe ontwikkelingen. Op basis van de werkelijke behoefte zal SURF jaarlijks met gebruikers en NWO in overleg treden over de samenstelling van het budget in de rekentijdcall en voor welke ondersteuning er naar alternatieve bekostigingsvormen zal moeten worden gezocht.

Tabel 2: Inschatting van aantal benodigde fte's in 2020–2022, op basis van te verwachten nieuwe ontwikkelingen per onderwerp en het aantal fte's dat met deze uitvoeringsplannen bekostigd kunnen worden

| Verwachtingen | 2019 (fte) | 2020 (fte) | 2021 (fte) | 2022 (fte) |
|---|-------------------|-------------------|-------------------|-------------------|
| Huidig niveau Data Processing | 7,5 | 7,5 | 7,5 | 7,5 |
| Huidig niveau Computing | 4,5 | 4,5 | 4,5 | 4,5 |
| Groei behoefte GPU ondersteuning | | 1,0 | 2,0 | 2,0 |
| Groei behoefte AI technieken in diverse applicaties | | 1,0 | 2,0 | 2,0 |
| Specifieke vraagstukken (waaronder mogelijke migratie van Hoge Energie Fysica gebruikers als gevolg van LHC upgrade, voorbereiding quantum computing,...) | | | 1,0 | 2,0 |
| Totaal | 12,0 | 14,0 | 17,0 | 18,0 |
| Beschikbare financiering vanuit dit bestedingsplan | 12,0 | 12,0 | 12,0 | 12,0 |
| Alternatieve bekostigingsvorm | 0,0 | 2,0 | 5,0 | 6,0 |

F Digitalisering: Impulsfinanciering lokale DCC's

Voor deze impuls is eenmalig 4,5 M€ beschikbaar.

Digitaal Competentie Centrum

In een Digitaal Competentie Centrum (DCC) wordt expertise op het gebied van data, software en computing op één plek gebundeld.¹⁰ De volgende functies zijn er voor de DCC's:

- landingsplaats voor projectsoftware voor onderhoud en beschikbaarstelling;
- kennis- en adviescentrum voor FAIR data en software en lokale ICT-infrastructuur;
- knooppunt in een gefedereerd netwerk voor data, computing en expertise.

Het doel van het financieringsinstrument voor de lokale DCC's is tweeledig:

1. versterken van ondersteuning van de onderzoekers op de kennisinstellingen op het gebied van data en software voor opslag en deling en het verduurzamen;
2. bundeling van expertise op het vlak van data en software om een betere toegankelijkheid en synergie te creëren.

Voor de verdere uitwerking van het instrument zal de door LCRDM opgestelde **definitie van een DCC** gehanteerd worden, zie bijlage 2.

Proces

Het doel is om een 'eenvoudig' financieringsinstrument in te richten, zonder een onderlinge competitie en met minimale administratieve lasten voor de aanvragers. Om dit te bereiken is het noodzakelijk om input op te halen bij diverse *stakeholders* over het financieringsinstrument.

Een punt van aandacht is dat er nog geen vaste definitie is voor een digitaal competentiecentrum (DCC), en wat de taken zijn van functies die actief zouden moeten zijn binnen het DCC, te weten de Data Steward en de Research Software Engineer. Om deze informatie te verzamelen worden de stakeholders gevraagd schriftelijk input te geven. Deze input kan dan worden meegenomen in de definitieve uitwerking van het financieringsinstrument.

Randvoorwaarden financieringsinstrument

Wie kan indienen?

- Nederlandse universiteiten
- Universitair Medische Centra
- KNAW instituutsorganisatie
- NWO instituutsorganisatie

¹⁰ "Integrale aanpak voor digitalisering in de wetenschap" p2.

Aanvullende eisen:

- De aanvragen van de Nederlandse Universiteiten worden ingediend door het CvB.
- De aanvragen van de Medische Centra worden ingediend door de RvB.
- Voor de KNAW en NWO instituten geldt dat de aanvraag ingediend wordt door de overkoepelende instituutsorganisatie. Het DCC dient ondersteuning te bieden aan alle onderzoeksinstituten behorend bij de betreffende instituutsorganisatie.
- Binnen deze call kunnen geen middelen aangevraagd worden vanuit de faculteiten of andere onderdelen van de organisaties.
- Per organisatie kan slechts één aanvraag ingediend worden.

Budget en aan te vragen kosten

Het beschikbare budget betreft een eenmalige impuls van 4,5 M€. Hiervan is 4,23 M€ beschikbaar voor financiering van aanvragen. NWO hanteert de standaard van 6% voor overheadkosten. Aanvragers kunnen maximaal 250 k€ aanvragen. Aan de subsidie is geen minimum verbonden.

Aanvragers kunnen financiering aanvragen voor:

- Personele kosten voor data stewards en research software engineers (op basis van VSNU akkoord NWP Academisch/HBO);
- Omscholingskosten voor de omscholing van bestaand personeel.

Aanvullende voorwaarden:

- Het is niet toegestaan om voor financiering personele kosten op te voeren voor personen die reeds via andere geldstromen gefinancierd worden;
- Omscholingskosten zijn enkel van toepassing op reeds bestaand personeel, gefinancierd vanuit de eerste geldstroom.

Cofinanciering

Er is een cofinanciering van 70% vereist.

Aanvullende eisen financiering

- Duur financiering is 2–4 jaar
- Plan voor verankering DCC binnen de eigen organisatie, dit moet bedrijfsmatig zijn ingestoken.
- DCC moet ondersteuning bieden aan de gehele organisatie. Plaatsing van het DCC binnen de organisatie wordt aan de instelling overgelaten.

Beoordeling

- Beoordeling vindt plaats op bureauniveau.
- Beoordeling op administratieve toelatingseisen, passendheid binnen de doelstellingen en verankering binnen de organisatie.

Besluitvorming

De Raad van Bestuur van NWO neemt het uiteindelijke besluit over de financiering van de ingediende aanvragen.

Tijlijn

| | |
|----------------|--------------------------------------|
| November 2019: | call open |
| Maart 2020: | sluiting call |
| April 2020: | besluitvorming Raad van Bestuur NWO. |

Monitoring en kennisdeling

Om de effectiviteit van de impulsfinanciering te meten zal NWO een tussentijdse evaluatie uitvoeren.

Hierbij zal worden gekeken naar:

- uitvoering van het opgestelde bedrijfsplan;
- verankering van het DCC binnen de organisatie;
- impact van het DCC op de organisatie;
- aansluiting van het DCC bij de verschillende onderzoekseenheden binnen de organisatie.

G Digitalisering: Stimulering thematische DCC's

Inleiding

NWO ontwikkelt een instrument voor financiering van projectvoorstellen van thematische DCC's. Dit vergt een brede, discipline- en domeinoverstijgende aanpak gezien het brede karakter van thematische DCC's. Aangevraagde voorstellen worden in competitie met elkaar beoordeeld. Voor dit onderdeel is jaarlijks 2,5 M€ beschikbaar (in 2020 is 5 M€ beschikbaar vanwege doorschuiven van middelen uit 2019).

Hieronder volgt een definitie van een thematische DCC. Vervolgens wordt het proces beschreven om tot een financieringsinstrument te komen. Als laatste worden een aantal randvoorwaarden van het financieringsinstrument beschreven.

Definitie thematische DCC's

De thematische DCC is onderdeel van een bestaande organisatie waar onderzoeksgroepen rondom een specifiek thema besloten hebben om hun datamanagement aan uit te besteden. Dit soort organisaties wordt vaak medegefinancierd door de onderzoeksgroepen en heeft zijn meerwaarde reeds bewezen. Dit soort organisaties richt zich specifiek op het bij elkaar brengen van data van onderzoeksgroepen en publiek beschikbare data om analyse binnen het thema te vergemakkelijken. Het zijn knooppunten waar kennis van data(stewardship), software en computing gebundeld wordt.

Proces

NWO gaat met verschillende stakeholders in overleg voor het uitwerken van een financieringsinstrument. Dit is in eerste instantie het NLeSC.

De randvoorwaarden zoals hieronder beschreven zullen aan de orde komen in de gesprekken. De hoogte van de cofinanciering staat niet ter discussie, wel de invulling ervan.

NWO voorziet in openstelling van de eerste call in mei 2020.

Omdat in 2019 het financieringsinstrument voor lokale DCC's wordt uitgerold, wordt met een latere invoering van het financieringsinstrument voor thematische DCC's voorkomen dat beide instrumenten met elkaar in competitie gaan. Beide instrumenten vereisen tenslotte een cofinanciering vanuit de kennisinstellingen.

Randvoorwaarden financieringsinstrument

Soort projecten

In het financieringsinstrument zal duidelijk moeten zijn wat voor projecten gefinancierd kunnen worden. Een randvoorwaarde voor de projecten is dat ze duurzaam zijn voor meerdere universiteiten of een domein. De eerste gedachten voor projecten zijn:

- Oprichting en onderhoud van discipline en subdiscipline specifieke data repositories en archieven

- Ontwikkeling en onderhoud van software tooling t.b.v. disciplinespecifieke data analyse en data management
- Interoperabiliteitsprojecten (koppelen van verschillende soorten data met het doel deze voor algemeen gebruik beschikbaar te stellen)
- Projecten die bijdragen aan de vindbaarheid en het hergebruik van bestaande onderzoeksdata
- Research Data Management en software carpentry workshops

Wie kan aanvragen

Aangezien het hier gaat om een samenwerking van een thematische DCC met meerdere kennisinstellingen zijn er meerdere aanvragers. De hoofdaanvrager komt vanuit een thematische DCC en medeaanvragers komen vanuit de kennisinstellingen. Het gaat hier om dezelfde instellingen als genoemd in het financieringsinstrument voor lokale DCC's.

Hoofdaanvrager

In het rapport *Integrale aanpak voor digitalisering in de wetenschap* worden een aantal voorbeelden genoemd (DANS, 4TU Data Center, HealthRI, knooppunten bestaande Europese infrastructuren). Daarnaast is door het Netherlands eScience Center een korte lijst opgesteld, waar aanvullend Data Federation Hub, BISS institute, DataHub, Amsterdam Data Science, Science Data Centre en Netwerk Digitaal Erfgoed worden genoemd.

Aangezien er echter geen uitputtende lijst te geven is van partijen die in aanmerking komen om de rol van thematische DCC in te vullen wordt als startpunt gebruik gemaakt van een lijst van cumulatieve criteria. Een organisatie dient aan deze criteria te voldoen om in te mogen dienen bij dit financieringsinstrument:

- de organisatie dient een bestaand, zelfstandig orgaan te zijn met een juridische status;
- de organisatie mag geen winstoogmerk hebben;
- de organisatie is in de basis niet gelieerd aan één academische partij;
Aanvragers dienen aan te tonen dat de organisatie een duurzaam karakter heeft met meerjarige commitment van alle deelnemende partijen;
- de organisatie heeft geen onderzoeksdoelstelling;
- de organisatie heeft een software/data oogmerk.

Medeaanvragers

In dit financieringsinstrument is een cofinanciering vanuit kennisinstellingen verplicht. Voor de cofinanciering zijn medeaanvragers onderdeel van de aanvraag. Medeaanvragers zijn afkomstig van dezelfde kennisinstellingen als benoemd in het financieringsinstrument voor lokale DCC's.

Budget

Er is jaarlijks een bedrag van 2,5 M€ beschikbaar voor gerichte projectvoorstellen van thematische Digitale Competentie Centra. In de eerste ronde (call mei 2020) zal een budget van M€ 5,0 beschikbaar zijn, waarvan 4,7 M€ voor het financieren van projecten. NWO reserveert 6% van de beschikbare middelen voor gemaakte kosten, waaronder ook de organisatie van bijeenkomsten voor kennisuitwisseling tussen verschillende projecten en de samenhang met lokale DCC's.

Wat kan worden aangevraagd?

Het aan te vragen budget per project bedraagt maximaal k€ 500. Er is geen minimum weergegeven, vanwege de grote diversiteit in kosten van de projecten van de organisatie van een workshop tot een repository.

Thematische DCC's kunnen in de aanvraag kosten opvoeren voor:

- Personele kosten voor Data stewards en Research Software Engineers op basis van VSNU tarief NWP HBO of Academisch
- Materiële kosten (incl. kleine investeringen)
- Internationalisering
- Kosten voor kennisbenutting

Cofinanciering

De cofinanciering bedraagt minimaal 70% van het totaalbudget van het aangevraagd project voor het thematische DCC. Hiermee wordt bedoeld dat de NWO subsidie en de cofinanciering tezamen 100% vormen en 70% hiervan ingebracht moet worden door de samenwerkende partners in het thematische DCC in de vorm van cofinanciering. De cofinanciering kan zowel in cash als in kind worden bijgedragen.

Procedure

Beoordeling

De beoordeling zal plaatsvinden op door een commissie van experts. Er wordt hierbij gebruik gemaakt van de expertise van het Netherlands eScience Center. Er zal geen gebruik gemaakt worden van referenten. Het is de verwachting dat een commissie van experts voldoende kennis beschikt om te oordelen over de projecten. De beoordelingscriteria worden opgesteld in overleg met verschillende partijen.

Besluitvorming

De Raad van Bestuur van NWO neemt het uiteindelijke besluit over de financiering van de ingediende projecten.

Tijdslijn

| | |
|--------------------|--------------------------------------|
| Mei 2020: | financieringsinstrument open. |
| Eind oktober 2020: | sluiting indiening projecten. |
| December 2020: | besluitvorming Raad van Bestuur NWO. |

Monitoring en kennisdeling

De projecten worden afhankelijk van de omvang tussentijds geëvalueerd. Alle projecten dienen een eindrapportage in. Daarnaast kunnen bijeenkomsten worden georganiseerd om de kennisdeling tussen de thematische DCC's en de lokale DCC's te bevorderen, inclusief de projecten die bij de thematische DCC's gefinancierd worden.

H Digitalisering: Ondersteuning DCC's door SURF

Inleiding

De lokale DCC's worden de knooppunten in een gefedereerd netwerk. SURF speelt een sleutelrol in het technisch faciliteren en coördineren van het veilige, gefedereerde systeem dat de DCC's onderling zou moeten verbinden.

Dit document beschrijft de wijze waarop de centrale ondersteuning voor de DCC's tot stand komt, hoe de afstemming met de DCC's verloopt en welke activiteiten voor de komende periode zijn voorzien. Dit document is afgestemd met een representatieve vertegenwoordiging van universitaire CIO's, UB directeuren en CIO's van UMC's. Het hieruit voortkomende technische ontwerp zal uiterlijk in voorjaar van 2020 worden voorgelegd aan de NPOS stuurgroep.

Het DCC-landschap in het NWO advies

De lokale DCC's zijn dicht bij de onderzoekers gepositioneerd. Zij zijn het eerste aanspreekpunt voor onderzoekers. De onderzoekers vinden er ondersteuning en technologische hulpmiddelen op het gebied van onderzoeksdata, open en FAIR-data. Per instelling zal één lokale DCC opgezet worden. De manier waarop een lokale DCC binnen een instelling is georganiseerd kan per instelling verschillen. Logischerwijs specialiseren de lokale DCC's zich op grond van de wensen en eisen van het onderzoek binnen hun instelling. Deze specialistische lokale kennis wordt vervolgens vanuit het oogpunt van efficiency en interoperabiliteit ter beschikking gesteld aan de andere knooppunten binnen het DCC-netwerk.

In bijlage 2 is een uitgebreidere definitie van de lokale DCC overgenomen zoals die door de regiegroep RDM-voorzieningen van SURF is opgesteld.

Zoals de lokale DCC's hun kennis zullen gaan delen met anderen, zo zullen de thematische DCC's vanuit hun specifieke disciplinegerichtheid een specifieke bijdrage leveren. De centraal ondersteunende DCC (SURF) zorgt voor ondersteuning bij disciplineoverstijgende zaken.

De middelen voor ondersteuning van de lokale en thematische DCC's worden ingezet om de digitalisering van onderzoek een impuls te geven en het systeem te versterken. Digitalisering van onderzoek betekent – naast het FAIR maken van data - ook het zorgdragen voor duurzame archivering en permanente toegankelijkheid van onderzoeksprocessen en -resultaten. Transparantie en efficiency van het onderzoek en de resultaten moeten worden geborgd om te zorgen voor optimale reproduceerbaarheid en hergebruik. Centrale ondersteuning is hierbij essentieel.

Enerzijds gaat het om ontwikkeling en onderhoud van faciliteiten- en diensten ten behoeve van het landelijk netwerk, anderzijds om kennisopbouw en kennisdeling. Voorbeelden hiervan zijn de kennis rond het opbouwen, inrichten en onderhouden van computingfaciliteiten en de praktische *community building* binnen het Landelijk Coördinatiepunt Research Data Management (LCRDM).

SURF ondersteunt de lokale en thematische DCC's daarom op twee fronten. Ten eerste gebeurt dat via het technisch faciliteren en coördineren van het veilige, gefedereerd systeem dat de universitaire DCC's onderling gaat verbinden. Het gaat hierbij om technologieën en tools die de vroege innovatiefase al door

zijn en die in de uitrolfase een federatief karakter hebben. Hieronder valt ook ondersteuning bij voorzieningen die door de instellingen aan elkaar beschikbaar gesteld worden: SURF speelt hierin een coördinerende rol.

Ten tweede zet SURF zijn expertise in voor kennisdeling en coördinatie. Zo worden de DCC's ondersteund op het gebied van beleid, brengt SURF experts uit verschillende DCC's bij elkaar en worden onderzoeksondersteuners ondersteund. Deze elementen zijn onmisbaar voor het vormen van een veilig federatief netwerk.

Bekostiging van de operationele dienstverlening staat los van de ondersteuning van de DCC's. Hiervoor zijn binnen de coöperatie SURF en met NWO andere mechanismen afgesproken, zoals de NWO-call voor rekenfaciliteiten en dienstverlening tegen binnen de coöperatie afgesproken tarieven. Wel is het mogelijk dat de kosten in de gezamenlijke ontwikkel- en implementatiefase met de instellingen vanuit een onderliggend bestedingsplan gefinancierd worden. Vervolgens zullen, naarmate de voorzieningen dichter bij operationele dienstverlening komen, de instellingen meer gaan bijdragen in de gebruikskosten. Dit loopt totdat zij de uiteindelijke bekostigingssystematiek bereiken.

Manier van werken

Uitgangspunt is dat de activiteiten binnen dit plan moeten bijdragen aan versterking van de landelijke IT- en data-infrastructuur voor wetenschappelijke onderzoekers in Nederland en de digitalisering van onderzoek. Primair ligt de ondersteuningstaak van het gefedereerde netwerk van DCC's bij SURF. Er zal wel nadrukkelijke inzet vanuit DCC's zelf gevraagd en gefinancierd worden. In dit plan wordt uitgegaan van een verdeling van circa 50:50 tussen de activiteiten die door SURF, en de activiteiten door de instellingen zullen worden uitgevoerd. De zuiver lokale behoeften blijven een verantwoordelijkheid van de individuele instellingen en vallen buiten dit plan.

Voorstel is om in een reeds bestaand overleg tussen de SURF-directie en de kennisinstellingen¹¹ (waarin de CIO's van universiteiten, de CIO's van UMC's en directeurs van de universiteitsbibliotheken vertegenwoordigd zijn) gezamenlijk de behoefte van de instellingen vast te stellen, een plan te vormen en te sturen op onderliggende projecten. In 2019 wordt deze afstemmingsrol vervuld door de regiegroep-RDM-voorzieningen. Participatie van de instellingen in de concrete projecten zal vervolgens lopen via een systeem van tenders, waarbij een instelling zich op een onderwerp zich kan inschrijven.

Via een jaarlijkse evaluatie- en planningcyclus worden deze onderwerpen en zwaartepunten voor ondersteuning van DCC's tussen de kennisinstellingen en SURF afgestemd. Voor deze onderwerpen zullen de DCC's (lokaal en thematisch) uitgenodigd worden om voorstellen in te dienen. Hoewel SURF in beginsel niet als financier optreedt, is SURF een samenwerkingsorganisatie waarbij de mogelijkheid tot het indienen van voorstellen open staat. Openheid en transparantie van het allocatieproces zijn daarbij cruciale uitgangspunten. SURF heeft een bewezen aanpak om deze samenwerkingen te organiseren. Op hoofdlijnen is deze:

1. Agendering van onderwerpen, gezochte omvang en beoordelingscriteria voor voorstellen worden opgesteld in het regieoverleg tussen kennisinstellingen en SURF;
2. Beperkte, gerichte vraagstukken waarvoor een evidente gegadigde is voor de uitvoering, kunnen ook in gerichte opdrachtvorm worden gesubsidieerd. Beoordeling van voorstellen wordt gedaan in het regieoverleg;
3. Voor bedragen groter dan 50 k€ of kavels waarvoor meerdere voorstellen worden uit gezocht zullen in een beperkte tender worden uitgevoerd. Beoordeling van voorstellen groter dan 100 k€ door een onafhankelijke commissie, de SURF-Wetenschappelijke Technische Raad. Deze adviseert aan het SURF-bestuur.

¹¹ Dit bestaand overleg is onderdeel van de governance van SURF, op de SURF ledenraad 13/6/2019 is het proces gestart om een voorstel te maken om dit deel van de governance aan te passen aan de nieuwe bestuurlijke governance. Dit proces zal eind 2019 afgerond zijn.

4. Honorering van voorstellen wordt gedaan door het SURF-bestuur;
5. Voor projectomvang van groter dan 100 k€ kan een onafhankelijke commissie projectbewaking ingeregeld. Deze ziet toe op beheersaspecten van projecten: kwaliteit, tijd en geld. Adviezen van deze commissie kunnen tot inhouding van subsidies leiden. De commissie adviseert aan het SURF-bestuur;
6. Gedurende uitvoering draagt SURF zorg voor de afstemming tussen gelieerde projecten en de kennisdisseminatie voor impact buiten de directbetrokkenen.
7. Voor de eerste periode zijn de onderwerpen bij federatieve voorzieningen en beleidsondersteuning, kennisdeling en coördinatie in het volgende hoofdstuk aangegeven.

Specifieke activiteiten

Samenwerking technologieën federatieve voorzieningen RDM

Binnen onderstaande onderwerpen worden de volgende activiteiten ontplooid ter ondersteuning van de DCC's:

- Het faciliteren van kennisuitwisseling over de ontwikkelingen in binnen- en buitenland;
- Het maken van technische (architectuur)keuzes die aansluiten op de zich ontwikkelende standaarden van EOSC;
- Het ontwikkelen en beschikbaar stellen van softwarecomponenten voor de federatieve voorzieningen (door DCC, door SURF of betrokken van een andere partij) die technisch aansluiten op de EOSC waar relevant;
- Implementatie en uitrol van de verschillende componenten over, met en door de DCC's.

Binnen de huidige afstemming tussen de kennisinstellingen en SURF hebben de volgende concrete onderwerpen voor het komend jaar prioriteit gekregen. Dit zijn onderwerpen waarbij afstemming en ondersteuning op federatief niveau essentieel is.

Implementatie federatieve onderzoeksdatamanagementtechnologieën

Het opstellen van een gezamenlijke architectuur, plus de ontwikkeling en uitrol van federatieve RDM-componenten. Integratie van de oplossingen van SURF en de verschillende oplossingen van de DCC's worden hierbij bij elkaar gebracht om tot een efficiënte keuze voor een gezamenlijk systeem te komen.

Pseudonimisering/Trusted Third Party

Het delen of gebruiken van persoonsgebonden data in een secuur federatief systeem vergt op grond van de Gedragscode Wetenschappelijke Integriteit en de AVG extra maatregelen voor de privacybescherming van de onderzoeksobjecten. Pseudonimisering/Trusted Third Party zijn methodes die hieraan bijdragen. De centrale ondersteuning moet leiden tot een standaardisatie voor het delen van persoonsgebonden data voor onderzoek. Instellingen dragen de concrete gebruiksvoorbeelden aan en brengen de systemen in die zij zelf hiervoor ingericht hebben. SURF zal de rol op zich nemen van coördinator nemen.

Virtual Research Environments

Onderzoek geschiedt in toenemende mate digitaal en in samenwerkingsverbanden die de grenzen van één enkele onderzoeksinstelling overschrijden. Om veilig te kunnen werken met digitale data en in samenwerkingsverbanden ontwikkelen onderzoeksinstellingen een virtuele onderzoeksomgeving (VRE) voor hun onderzoekers. De centrale ondersteuning moet leiden tot het naadloos met elkaar laten samenwerken van nationale en lokale componenten, maar ook third party-diensten bij het vormgeven en inrichten van een VRE. Met de DCC's werkt SURF aan integratie van concrete toepassingen die bij onderzoekers in gebruik zijn, dan wel waar behoefte aan is binnen VRE/DRE-systemen die bij de verschillende partners beschikbaar zijn.

Sustainable Software

Om onderzoeksdata optimaal (her)bruikbaar te maken, een belangrijk aspect van het efficiënter en transparanter maken van onderzoek, is het niet alleen noodzakelijk de data, maar ook de bij het onderzoek gebruikte software te documenteren en duurzaam te archiveren en toegankelijk te houden. De centrale

ondersteuning op dit gebied moet leiden tot (het opzetten van) een gefedereerd systeem om het archiveren van verschillende soorten software uit te testen en mogelijk te maken. Hierbij wordt nauw samengewerkt met het Netherlands eScience Center, met name door gebruik van de *Research Software Directory*. De rol van de instellingen is het gezamenlijk opstellen van de architectuur, het aanbrengen van cases, en het op elkaar laten aansluiten van lokale en centrale componenten.

Science Collaboration Zone

Het organiseren en daarna beheren van toegang tot faciliteiten zoals als (cloud)applicaties, rekenclusters, datasets of generieke dataopslag is complex en tijdrovend. Science Collaboration Zone (SCZ) is een authenticatie- en autorisatieinfrastructuur gericht op onderzoekers. Inzet van SCZ maakt het voor onderzoekers en onderzoekssamenwerkingen mogelijk om gebruiksvriendelijk en veilig samen deze faciliteiten te gebruiken. SCZ bouwt voort op bestaande federatieve infrastructuren zoals SURFconext en eduGAIN. Het doel hiervan is om een schaalbare Authenticatie- en Autorisatiefaciliteit voor onderzoek in te richten. Deze maakt het mogelijk om data en andere voorzieningen tussen samenwerkende onderzoekers te delen, terwijl de toegang wel onder eigen controle wordt gehouden. De rol van de instellingen zal zijn om de lokale voorzieningen op dit systeem aan te sluiten.

Beleidsondersteuning, kennisdeling en coördinatie

Het **Landelijk Coördinatiepunt Research Data Management (LCRDM)**, faciliteert de koppeling tussen beleid en praktijk. 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. Het LCRDM brengt onderzoeksondersteuners, beleidsmakers, ICT-ers, bestuurders uit uiteenlopende onderzoeksinstellingen (WO, UMC, HBO) en onderzoeksfinanciers bij elkaar. Het netwerk bestaat inmiddels uit 180 personen die verbonden zijn aan 64 verschillende onderzoeksinstellingen. Keuze voor de aanpak van onderwerpen ligt bij de LCRDM-adviesgroep bestaande uit mensen met een leidende rol op RDM-gebied vanuit het perspectief van support, ICT of beleid. In de evolutie van het LCRDM zal na de opzet van DCC's gekeken worden hoe het systeem van onderwerpkeuze en uitkomstdisseminatie geoptimaliseerd kan worden, en hoe het LCRDM kan aansluiten op het systeem van DCC's. Op specifieke thema's kan het LCRDM expertise vanuit de instelling detacheren om de uitwerking van een onderwerp te versnellen.

Kennisdeling en Coördinatie, kernonderdeel van de ondersteuning van DCC's is de kennisdeling en facilitering die SURF organiseert voor de onderzoeksondersteuners rond de landelijke faciliteiten en diensten. Een aantal voorbeelden:

- Om het al bestaande netwerk van onderzoeksondersteuners te faciliteren worden Support-4-Research Masterclasses en Bootcamps georganiseerd samen met een lokale DCC voor het hele netwerk;
- Trainingen voor lokale en thematische DCC's;
- Het bij elkaar brengen van de verschillende DCC's om kennis te delen en inhoudelijke standpunten te bepalen rond onderwerpen die van gezamenlijk belang zijn. Vervolgens kunnen concrete uitkomsten daarvan in de reguliere governance van SURF ingebracht worden. Zo zal bijvoorbeeld een reguliere bijeenkomst ingericht worden om met de lokale DCC's informatie en beleid uit te wisselen (en te maken) rond High Performance Computing;
- De kennis en expertise die SURF in huis zal beschikbaar gemaakt worden voor de DCC's (bijv. ook rond het opbouwen, inrichten en onderhouden van computingfaciliteiten);
- De DCC's zijn het eerste aanspreekpunt van de onderzoekers. Onderzocht zal worden wat er additioneel nodig is om samenwerking tussen deze aanspreekpunten optimaal te faciliteren. Hierbij kan een verschil optreden tussen wat er nodig is voor lokale DCC's en thematische DCC's.

Naast de aansluiting van DCC's op het landelijk netwerk, kan vanuit dit bestedingsplan de afstemming over de aansluiting op de EOSC worden uitgevoerd. Wanneer de aansluiting van DCC's op de EOSC investeringen vragen, kan deze ook (deels) uit het dit bestedingsplan worden bekostigd. Deze middelen zullen dan via het tendermechanisme worden toegekend. Het geschikt maken van en investeren in de

Nationale Infrastructuur is onderdeel van de andere categorieën uit het NWO-advies (Kennisvernieuwing en Investerings in data- en computerfaciliteiten). Structurele bijdragen van de DCC's en de nationale infrastructuur aan de EOSC, zoals het gebruik van voorzieningen, valt buiten dit bestedingsplan.

Succescriteria

In samenwerking met de RDM-regiegroep stelt SURF voor eind 2019 een eerste set van succescriteria op. Succescriteria zullen gaan over het doel van het gefedereerde netwerk van DCC's: "Om FAIR te laten slagen is het noodzakelijk dat "lokaal gehuisveste" data en software toegankelijk worden vanaf andere onderzoeksinstellingen. De subcommissie beveelt de oprichting van een gefedereerd netwerk van lokale infrastructuren aan om dit te realiseren." Succescriteria zullen liggen op het gebied van adoptie (bv. hoeveel instelling maken gebruik van de technologieën die uitgerold worden), samenwerking (bv. hoeveel samenwerkings-, kennisdelings- en afstemmingsbijeenkomsten zijn gehouden) en standaarden (bv. welke standaarden zijn opgenomen en welke architectuurafspraken zijn door het netwerk van DCC's gemaakt).

Financiering Ondersteuning DCC's

Tabel 1: Financiering Ondersteuning DCC

| Onderwerp | Bijdrage aan instellingen (k€) | Bijdrage aan SURF (k€) | Externe Kosten (k€) | Totale kosten (k€) |
|--|--------------------------------|------------------------|---------------------|--------------------|
| federatieve RDM-tooling | | 150 | | |
| pseudonimisering | | 30 | | |
| Virtual research environments | 840 | 100 | 70 | |
| Sustainable software | | 30 | | |
| Science Collaboration Zone | | 140 | | |
| Subtotaal Technologieën Federatieve Voorzieningen RDM | 840 | 450 | 70 | 1360 |
| Landelijk Coördinatiepunt RDM | | 200 | | |
| Coördinatie RDM- en HPC-infrastructuur in NL | 100 | 140 | 60 | |
| Ondersteuning lokale en thematische DCC bij kennisopbouw en kennisdeling | | 140 | | |
| Subtotaal Beleidsontwikkeling, kennisdeling en coördinatie | 100 | 480 | 60 | 640 |
| Totaal Ondersteuning DCC | 940 | 930 | 130 | 2000 |

Toelichting:

- Bijdrage aan SURF bestaat uit menskracht en representeert 6,5 FTE.
- Externe kosten zijn kosten die niet bij SURF of bij de instellingen gemaakt worden. Het gaat hier om kosten die nodig zijn om het ondersteunen en faciliteren van een netwerk van DCC's mogelijk te maken, bijvoorbeeld om PR/communicatiekosten, website, events en catering, maar ook om inhuur van tijdelijk personeel voor projectmanagement, mochten deze binnen SURF niet beschikbaar zijn.

Digitalisering: Investerings in eScience

Voor dit onderdeel is vanaf 2020 jaarlijks 2,5 M€ beschikbaar en een eenmalige impulsfinanciering in 2019 van 3 M€. Het eScience Center zet de middelen in voor het uitbreiden van *open calls for proposals*. Deze calls zullen open staan voor alle wetenschappelijke disciplines. Het doel is:

1. Een substantiële toename van het aantal vernieuwende en collaboratieve onderzoeksprojecten;
2. Een substantiële toename van het honoreringspercentage van onderzoeksvorstellen ingediend in de open calls van het eScience Center.

Via de open calls voor proposals zet het eScience Center zijn expertise en eScience Research Engineers in voor samenwerkingsprojecten met Nederlandse universiteiten en kennisinstellingen. Hierdoor wordt de relatie verstevigd tussen informatica-kennis en ICT-infrastructuur (open onderzoekssoftware, big data-analyse, high-performance computing) enerzijds en vragen en ontwikkelingen uit de toepassingsdomeinen anderzijds. Naast de *in-kind* inzet van eScience-FTE's op projecten (in de vorm van eScience Research Engineers), bestaan de toekenningen in de open calls uit een *in-cash* component ten behoeve van de aanstelling van lokaal wetenschappelijk personeel op projectbasis (bv. AiO's, Postdocs). Aangezien de expertise van de eScience Research Engineers de voornaamste waarde van het eScience Center vertegenwoordigt, wordt met de extra financiering tevens het volgende doel nagestreefd:

3. Een substantiële toename van het totale aantal FTE eScience-inzet toegekend aan gehonoreerde onderzoeksvorstellen ingediend in de open calls van het eScience Center.¹²

Inzet middelen open calls 2019–2020

Op basis van de huidige call-strategie van het eScience Center worden de extra middelen per direct ingezet voor open calls met brede thematiek voor alle disciplines (zie Tabel 1). Hiermee worden de komende twee jaar minimaal 10 meerjarige samenwerkingsprojecten¹³ gestart, bovenop de projecten die het eScience Center op basis van het bestaande call-budget zal honoreren (zie Tabel 2). De middelen worden in gelijke mate verdeeld over de 5 hoofddomeinen van het eScience Center: *Environment & Sustainability; Humanities & Social Sciences; Life Sciences & eHealth; Physics & Beyond; Computer & Data Science*. Onderzoeksvorstellen zijn alleen *binnen* deze hoofddomeinen in competitie met elkaar; tussen de domeinen zal geen competitie bestaan.

De gepubliceerde open calls worden thematisch zo breed mogelijk opgezet. Dit past bij de missie en scope van het eScience Center, dat geen domein-inhoudelijke voorkeur heeft voor specifieke thema's of richtingen. Het eScience Center beoogt excellent onderzoek te ondersteunen dat voldoet aan de volgende uitgangspunten:

- a. Er moeten innovatieve wetenschappelijke vraagstellingen worden gefaciliteerd die zonder eScience niet of moeilijk realiseerbaar zijn;
- b. De 'stap voorwaarts' (o.a.: versnelling en verbetering van wetenschappelijke analyses) voor de wetenschappelijke disciplines moet worden gemaximaliseerd;
- c. De ontwikkeling van duurzame, herbruikbare en discipline-overstijgende onderzoekssoftware moet een kerndoel zijn.

¹² Het eScience Center heeft een langjarig steven de in-cash component in open calls for proposals niet uit de eigen middelen te financieren, maar deze te betrekken via samenwerkingen met externe financiers (waaronder NWO, ZonMW, ministeries, topsectoren, etc.). In de eerste open call (2019) zal de verhouding in-kind versus in-cash nog 50%-50% zijn, maar in de 2 opvolgende jaren zal deze via samenwerking met externe financiers stapsgewijs verhoogd worden naar tenminste 80%-20%.

¹³ Dit aantal is een minimum. Het eScience Center zal de komende jaren gezamenlijk en thematisch blijven programmeren met andere financiers, waardoor in totaal meer projecten toegekend kunnen worden. Enkele concrete voorbeelden van recente en aankomende co-funders voor thematische open calls zijn ZonMW, Topsector LSH, VWS, NWO (hart- en vaatzieken), NSO (aardobservatie en planetologie) en NWO, Topsector Energie en SIA (energietransitie).

Tabel 1: Besteding beschikbare middelen NWO-advies. Projectie minimaal extra budget, #projecten en #FTE-inzet in eScience Center calls 2019–2022

| Call | Domeinen | 2019 | | 2020 | | 2021*** | | 2022*** | |
|--|------------------------------|---------------|-----------|---------------|---------|---------------|---------|---------------|---------|
| | | In-kind* | In-cash** | In-kind | In-cash | In-kind | In-cash | In-kind | In-cash |
| Accelerating Scientific Discovery (ASDI) | Environment & Sustainability | 300 K€ | 300 K€ | 350 K€ | 150 K€ | 400 K€ | 100 K€ | 400 K€ | 100 K€ |
| | Humanities & Social Sciences | 300 K€ | 300 K€ | 350 K€ | 150 K€ | 400 K€ | 100 K€ | 400 K€ | 100 K€ |
| | Life Sciences & eHealth | 300 K€ | 300 K€ | 350 K€ | 150 K€ | 400 K€ | 100 K€ | 400 K€ | 100 K€ |
| | Physics & Beyond | 300 K€ | 300 K€ | 350 K€ | 150 K€ | 400 K€ | 100 K€ | 400 K€ | 100 K€ |
| Disruptive technologies (DTEC) | Computer & Data Science | 300 K€ | 300 K€ | 350 K€ | 150 K€ | 400 K€ | 100 K€ | 400 K€ | 100 K€ |
| Totaal | | 3,0 M€ | | 2,5 M€ | | 2,5 M€ | | 2,5 M€ | |
| # Extra projecten | | 5 | | 5 | | 9 | | 9 | |
| # Extra FTE eScience-inzet | | 12.5 | | 14.6 | | 16.7 | | 16.7 | |
| Verhouding in-kind/in-cash | | 50% | 50% | 70% | 30% | 80% | 20% | 80% | 20% |

* In-kind: bijdrage aan projecten in de vorm van eScience Research Engineers in dienst van het eScience Center.

** Bijdrage aan projecten voor aanstelling van lokaal wetenschappelijk personeel, met name AiO's en Postdocs.

*** Vanaf 2021 zal een nieuwe call-strategie ingevoerd worden. De calls vanaf dat moment zullen gebaseerd zijn op dezelfde uitgangspunten (zie volgende pagina) als de huidige calls (ASDI en DTEC).

Tabel 2: Projectie budgetten, #projecten en #FTE-inzet open calls 2019–2022 op basis van gecombineerde inzet van bestaande eScience Center financiering (3,0 M€ per jaar) en beschikbare middelen NWO-advies. Tevens vergelijking met minimale omvang op basis van het huidige eigen call budget (2018)

| | 2018 | 2019 | 2020 | 2021 | 2022 |
|---------------------------------|---------------|---------------|---------------|---------------|---------------|
| Environment & Sustainability | 0,6 M€ | 1,2M€ | 1,1 M€ | 1,1 M€ | 1,1 M€ |
| Humanities & Social Sciences | 0,6 M€ | 1,2M€ | 1,1 M€ | 1,1 M€ | 1,1 M€ |
| Life Sciences & eHealth | 0,6 M€ | 1,2M€ | 1,1 M€ | 1,1 M€ | 1,1 M€ |
| Physics & Beyond | 0,6 M€ | 1,2M€ | 1,1 M€ | 1,1 M€ | 1,1 M€ |
| Computer & Data Science | 0,6 M€ | 1,2M€ | 1,1 M€ | 1,1 M€ | 1,1 M€ |
| Totaal | 3,0 M€ | 6,0 M€ | 5,5 M€ | 5,5 M€ | 5,5 M€ |
| # Projecten* | 5 | 10 | 10 | 20 | 20 |
| # FTE eScience-inzet** | 12,5 | 25 | 32 | 37–46 | 37–46 |
| Honoreringspercentage*** | 5,8% | >10% | >10% | 20–25% | 20–25% |

* Voor 2018–2020 is hier het minimumaantal te honoreren projecten opgenomen, dat alleen ontstaat als er geen samenwerking met externe financiers is. Het daadwerkelijke aantal is maximaal 2x het genoemde aantal.

Voor 2021-e.v. is het verwachte aantal projecten opgenomen onder de beoogde nieuwe call-strategie. Dit bevat tevens ca. 12 jaarlijkse haalbaarheidsstudies van kleinere omvang dan de 'standaard' grote eScience projecten.

** Voor 2018–2019 is hier het minimumaantal toegekende FTEs bij een in-kind versus in-cash verdeling van 50%–50% opgenomen; voor 2020 bij een in-kind versus in-cash verdeling van 70%-30%. Voor 2021-e.v. (onder de nieuwe call- strategie) is hier een range aangehouden voor een in-kind versus in-cash verdeling tussen 80%–20% en 100%–0%.

*** Voor 2019–2020 wordt direct en minimaal een verdubbeling van het honoreringspercentage verwacht t.o.v. 2018 op basis van de extra beschikbare middelen. Voor 2021-e.v. is een directe vergelijking lastiger i.v.m. de geheel nieuwe strategie; desalniettemin wordt een honoreringspercentage van minimaal 20–25% nagestreefd voor grote projecten.

Hoewel de open calls thematisch breed zullen worden ingestoken, kan worden besloten om specifieke disciplines of thema's extra aandacht te geven als deze in voorgaande jaren onderbelicht zijn gebleven.¹⁴ Op technologisch vlak is een gerichte focus ook mogelijk, bijvoorbeeld rondom veelgevraagde en innovatieve technologische thema's als *explainable AI*, *remote sensing*, robotica, multi-model simulaties, *Internet of Things*, etcetera. De directie van het eScience Center stelt dergelijke thematische richtingen voor aan het Bestuur op basis van directe contacten met en expliciete vragen uit de disciplines en op basis van o.a. de meest recente Kennis- en Innovatieagenda's. Het Bestuur van het eScience Center besluit uiteindelijk over de domein-inhoudelijke en thematische speerpunten van de calls, gebaseerd op de eerdergenoemde uitgangspunten (a-c).

Het beschikbare extra budget voor de periode 2019-2020 zal direct leiden tot een significant groter aantal samenwerkingsprojecten, zelfs tot een verdubbeling ten opzichte van het aantal projecten uit het huidige eigen call budget.¹⁵ Op basis van de extra financiering zal daarom op korte termijn een bredere en duurzamere impact van eScience binnen en tussen alle wetenschappelijke domeinen verwacht kunnen worden. Tevens verwachten we in deze periode een verdubbeling van de honoreringspercentages bij onze calls for proposals, die in 2018 tot ongeveer 5% waren gedaald.

Inzet middelen open calls 2021-e.v.

In reactie op de uitkomst van de externe evaluatie van het eScience Center (Q2 2019), en in aanloop naar een nieuwe samenwerkingsovereenkomst met NWO en SURF over het eScience Center (start in 2021), wordt een nieuwe organisatiestrategie en call-strategie ontwikkeld. De voorbereiding op de nieuwe call-strategie zijn in 2018 al gestart door de directie, het bestuur en de externe adviescommissie van het eScience Center.¹⁶ De drie hoofddoelstellingen van deze nieuwe call-strategie zijn: (1) verdere verhoging van de mate van inclusiviteit voor alle disciplines, (2) verdere verhoging van de honoreringspercentages, en (3) versterkte aandacht voor softwaregeneralisatie- en verduurzaming.

Behalve deze hoofddoelstellingen zal de nieuwe call-strategie erop gericht zijn innovatief en excellent onderzoek nog intensiever te ondersteunen door middel van eScience-expertise. Dit wordt dusdanig ingericht dat de gehele keten van vroege ideevorming, diepgaand wetenschappelijk onderzoek, en opschaling en productontwikkeling op een geïntegreerde wijze wordt bediend. De strategie beoogt zo een nog grotere groep onderzoekers vooruit te helpen, variërend van kleinere innovatieve projecten (high risk, high gain) tot grotere infrastructurele en discipline-brede initiatieven. De strategie zal bestaan uit meerdere fasen, elk met een open call waarin projecten geselecteerd worden via onafhankelijke externe review. De kleinere bottom-up calls in de eerste fase staan breed open voor alle domeinen en onderzoekers en hebben het karakter van een haalbaarheidsstudie. De focus in de latere fasen ligt op het creëren van wetenschappelijke doorbraken door innovatieve toepassingen van onderzoekssoftware en digitale methoden.¹⁷ Door het meerjarige karakter van de meeste eScience Center-projecten zullen ieder jaar minimaal 50 projecten draaien, met grote spreiding over alle wetenschappelijke disciplines.

¹⁴ Een voorbeeld is het Social Sciences & Humanities domein. Onderzoeksvoorstellen uit de hoek van de Social Sciences zijn de afgelopen jaren in de open calls van het eScience Center minder vaak gehonoreerd dan voorstellen uit de Humanities. Wij overwegen daarom in de eerstkomende call voorrang te verlenen aan de Social Sciences. Deze wijze van opereren past het eScience Center al langer toe. In 2018/2019 is bijvoorbeeld extra aandacht geschonken aan de zogenaamde 'Big Science' disciplines, waaronder High-Energy Physics en Radioastronomie.

¹⁵ Binnen het huidige, geheel eigen beschikbare budget voor open calls (ca. 3M€ per jaar) heeft het eScience Center ruimte voor tenminste 5 nieuwe grote onderzoeksprojecten per jaar, 1 per discipline.

¹⁶ De definitieve call-strategie (en dus precieze aantal en typen open calls en projecten) zal in 2020 vorm krijgen, in afstemming met en na goedkeuring van NWO, SURF en het bestuur van het eScience Center en hangt direct samen met de nieuwe beoogde samenwerkingsovereenkomst (vanaf 2021) tussen NWO, SURF en het eScience Center.

¹⁷ Eventuele thematische programmering is flexibel en hangt mede af van programmering van overige funders (vnl. NWO), nieuwe ICT-ontwikkelingen en actuele vragen naar eScience vanuit de wetenschappelijke domeinen. Jaarlijkse afstemming met stakeholders binnen de disciplines, zoals programma-managers bij NWO, de ENW-adviestafels, SGW-panels en programmaclusters (binnen ZonMW) maakt hier ook onderdeel van uit.

Proces totstandkoming open calls

De inzet van de call-middelen vindt plaats middels een gestructureerd proces. Jaarlijks is er contact met verschillende *stakeholders* binnen de vijf domeinen van het eScience Center (grotendeels overeenkomend met de NWO-domeinen en ZonMW). De afgelopen jaren heeft dat geresulteerd in calls en projecten over de hele breedte van de wetenschap, zoals weergegeven in de bijlage 6. De processtappen tot het komen van zowel joint calls met externe financiers als geheel eigen calls die tezamen de volle breedte van de wetenschap bedienen, staan beschreven in Tabel 3. De maximale tijd vanaf het eerste moment van afstemming tot de publicatie van calls is ongeveer een jaar.

Tabel 3: Jaarlijks proces totstandkoming open calls

| 1 ^o kwartaal | 2 ^o kwartaal | 3 ^o kwartaal | 4 ^o kwartaal |
|---|---|--|--|
| Joint calls: Programmatische en concrete afstemming met externe funders, o.a. NWO-domeinen en program officers | Joint calls: Uitwerking op hoofdlijnen concept samenwerkingen in thematische joint calls | Afstemming en besluit eScience Center Bestuur; keuze uit eigen en joint (concept) calls, met gelijke verdeling over alle domeinen en eScience technologieën; tevens maximale inclusiviteit | Definitieve calls, besluit eScience Center Bestuur. Eerste aankondiging en publicatie in Q1/Q2 volgende jaar. |
| Eigen calls: Inventarisatie en gerichte uitvragen bij wetenschappelijke stakeholders, KIA's, etc. | Eigen calls: Bij ontbreken joint calls met externe funders en/of wens tot specifieke focus: concept-uitwerking eigen calls | Op basis besluit bestuur: finale uitwerking en afronding call-teksten | |

Gedurende het hele jaar:

- Uitleg en introductie van eScience Center (expertise, calls, technologie) bij potentiële funding-partners (NWO, ZonMW, topsectoren, etc.)
- Inhoudelijke consultaties met wetenschappelijke veld (domeinen, netwerken en onderzoeksgremia)

Joint calls gedefinieerd in samenwerking met externe financiers zijn vaak minder breed gedefinieerd dan de eigen calls. Om toch een goede langjarige spreiding te garanderen over alle relevante onderwerpen en thema's zullen door het Bestuur van het eScience Center alle nieuw te publiceren calls domein-inhoudelijk en technologisch gewogen worden ten opzichte van de in de meest recente jaren gehonoreerde projectvoorstellen.

Naast domein-inhoudelijke en technologische variaties, kunnen ook de doelgroepen van de open calls over de jaren variëren. Normaliter zal 'maximale inclusiviteit' voorop staan, zodat zoveel mogelijk wetenschappers bediend kunnen worden. Om de honoreringspercentages beheersbaar te houden kan het bestuur echter variaties in de beoogde doelgroepen aanbrengen, om zo het aantal indieningen af te remmen dan wel te stimuleren. Voor de hand liggende voorbeelden van doelgroepen zijn als volgt:

- Individuele excellente (gevestigde) wetenschappers;
- Individuele jonge wetenschappers die een grote verandering teweeg kunnen brengen in hun disciplines met behulp van onderzoekssoftware, ICT en data, maar zelf nog niet 'gevestigd' zijn;
- Grote, goed verbonden (inter)nationale consortia, bij voorkeur met eigen expertise en community op gebied van ontwikkeling van duurzame software;
- Samenwerkingen tussen wetenschap en bedrijfsleven.

Verbinding eScience-activiteiten bij universiteiten

Het rapport *Integrale aanpak voor digitalisering in de wetenschap* benoemt onder het kopje eScience de verduurzaming van software; de verantwoordelijkheid hiervoor ligt bij lokale digital competence centers (DCC's). Het eScience Center biedt aan een rol te spelen bij het verspreiden, vertalen en overdragen van

kennis en tooling, bijvoorbeeld door training, naar lokale DCC's rondom verduurzaming van software en aanpalende aspecten van open science. Samenwerking met SURF (rondom LCRDM, Support4Research, Bootcamps) ligt voor hand en heeft afgelopen jaren al in beperkte mate plaatsgevonden. De onderscheidende expertise van het eScience Center ligt hoofdzakelijk rondom de onderwerpen FAIR data en software, software sustainability en tools voor reproduceerbaarheid van onderzoek, zoals containerization en workflows. Zoals de externe evaluatiecommissie eerder dit jaar over de rol van het eScience Center stelde in haar eindrapport: "Dissemination is key for the relevance of the centre".

5 Bijlagen

Bijlage 1: Samenvatting adviezen uit Integrale aanpak voor digitalisering in de wetenschap

Dit is een samenvatting van het NWO adviesrapport. Adviezen zijn opgesteld door de ICT-commissie ('commissie Apers') van de PC-GWI.

Lokale Digitale Competentie Centra

Voor de oprichting van (en versterking van bestaande) DCC's adviseert de commissie NWO tijdelijk middelen beschikbaar te stellen. De commissie adviseert NWO hiervoor éénmalig een bedrag van 4,5 M€ als impulsfinanciering beschikbaar te stellen. Hiervoor zal een eenvoudig financieringsinstrument worden ingericht door NWO, waarbij de administratieve lasten voor aanvragers worden geminimaliseerd.

Digitale Competentie Centra in een gefedereerd systeem

Voor de ontwikkeling, het onderhoud en het leveren van ondersteuning adviseert de commissie NWO jaarlijks een bedrag van 2 M€ beschikbaar te stellen aan SURF. De commissie beveelt aan dat SURF samen met de CIO's van universiteiten en onderzoeksinstituten een plan opstelt waarin de functionaliteit van het gefedereerde systeem wordt beschreven, en hoe het zal worden gerealiseerd. Om te waarborgen dat het gefedereerde systeem in voldoende mate bijdraagt aan het realiseren van de ambities gesteld in het Nationaal Plan Open Science (NPOS) zouden SURF en de het landelijk CIO overleg, alvorens het plan uit te werken, de hoofdlijnen moeten afstemmen met de stuurgroep NPOS.

Thematische Digitale Competentie Centra

De commissie adviseert NWO voor deze categorie DCC's jaarlijks een bedrag van 2,5 M€ ter beschikking te stellen vanaf 2019. Dit bedrag wordt vrijgegeven op basis van gerichte projectvoorstellen die matching vereisen van academische instituten; op deze manier kunnen de instellingen zelf keuzes maken in het prioriteren van deze DCC's.

Het prioriteren van investeringen in thematische DCC's vereist een brede, discipline- en domeinoverstijgende visie op digitalisering. De commissie adviseert gebruik te maken van de kennis van het NLeSC om de middelen die hiervoor worden gereserveerd te verdelen. De commissie beveelt aan dat NLeSC in samenwerking met NWO een financieringsinstrument ontwerpt voor de verdeling van deze middelen. Zij zullen voorstellen welke partijen voor financiering in aanmerking komen en hoe een evenredige verdeling tussen wetenschappelijke domeinen kan worden bewerkstelligd. Gezien de ambities van DANS en SURF ligt een sterkere samenwerking tussen die partijen voor de hand.

Voorwaarden financiering DCC's

Het behoeft nadruk dat de kosten van de digitalisering van de wetenschap uiteindelijk vanuit de wetenschapsbudgetten gedragen zullen worden. Data stewards en research software engineers zullen langzamerhand een vast onderdeel van het ondersteunend personeel gaan uitmaken. Kennisinstellingen zullen zich hieraan moeten committeren wanneer zij gebruik maken van bovenstaande financiering. Voor de bekostiging van versterking van lokale DCC's adviseert de commissie het vragen van cofinanciering; naar mening van de commissie kan 30% van de kosten voor maximaal drie jaar worden gefinancierd uit de 1,5 M€ die hiervoor jaarlijks ter beschikking wordt gesteld, de overige 70% zou voor rekening van de kennisinstellingen moeten komen. Ook voor financiering van de thematische DCC's adviseert de commissie deze 30/70 verhouding, en er moeten waarborgen voor duurzame financiering door de kennisinstellingen worden afgegeven. De universiteiten zullen moeten zorgen voor dat de impulsfinanciering voor de lokale DCC's ten goede komt aan alle wetenschappelijke domeinen, en NLeSC en NWO zullen de gelijkmatige spreiding van de financiering voor thematische DCC's over de domeinen moeten waarborgen. Daarnaast doet de commissie de aanbeveling dat de VSNU gevraagd wordt te onderzoeken hoe de betrekkelijk nieuwe categorieën ondersteunend personeel die bij lokale DCC's worden gehuisvest een aantrekkelijk carrièreperspectief gegeven kan worden.

Rekenfaciliteiten

Hardware

De nationale supercomputer is een belangrijke faciliteit, en voor een aantal onderzoeksvelden een noodzakelijke voorwaarde om internationaal competitief te kunnen zijn. De commissie adviseert NWO een bedrag van 18 M€ beschikbaar te stellen voor het vervangen van de nationale supercomputer Cartesius, om deze onderzoeksvelden ook in te toekomst te kunnen bedienen, en aansluiting te behouden met internationale computing initiatieven als PRACE. De commissie adviseert NWO een éénmalige bijdrage ter hoogte van 2 M€ van SURF zelf te vragen om het gewenste ambitieniveau voor de supercomputer te kunnen realiseren. De commissie adviseert dat de middelen die hiervoor beschikbaar komen in beheer van NWO blijven totdat SURF overgaat tot aanschaf van de nieuwe supercomputer.

Ook adviseert de commissie NWO vanaf 2020 jaarlijks 2,5 M€ ter beschikking te stellen voor de financiering van andere rekensystemen bij SURF, met een eenmalige impuls van 3 M€ in 2019. Door te investeren in een breed spectrum van computing faciliteiten kan SURF ook gebruikers voor wie de supercomputer niet geschikt is voorzien in hun behoeften. Daarnaast adviseert de commissie NWO vanaf 2020 jaarlijks 1 M€ te investeren in uitbreiding van de opslag- en archiveringsmogelijkheden in de reken-datacentra van SURF, om tegemoet te komen aan de rappe toename van datavolumes, met een eenmalige impuls van 1,5 M€ in 2019. SURF is de aangewezen partij om te waarborgen dat de Nederlandse kennisinstellingen aangehaakt blijven bij de nieuwste technologische ontwikkelingen; daartoe is het belangrijk te investeren in de vernieuwing en ontwikkeling van hun kennis. De commissie adviseert NWO vanaf 2019 jaarlijks 3,5 M€ te investeren in het opbouwen en vernieuwen van de voor de rekenfaciliteiten benodigde kennisbasis bij SURF. Deze middelen voorzien voor een deel in het verduurzamen van de financiering van de expertise die reeds aanwezig is bij SURF, en essentieel is voor de instandhouding van een ICT kenniscentrum van wereldniveau.

Het is belangrijk om te waarborgen dat de infrastructuurvernieuwing van SURF nauw aansluit bij de wensen en noden van de instituten die zij bedienen. De commissie beveelt daarom aan dat SURF de CIO's van de universiteiten en onderzoeksinstituten een rol geeft in het sturen van de vernieuwing bij SURF, om de doelmatigheid te waarborgen. De commissie vraagt SURF en het landelijk CIO-overleg te beschrijven hoe zij hun samenwerking op het front van infrastructuurvernieuwing gaan intensiveren in het plan voor het gefedereerde systeem.

Gezien de groeiende behoefte aan rekenkracht zal de opvolger van Cartesius naar verwachting in 2025 vervangen moeten worden. De commissie adviseert NWO vanaf 2020 jaarlijks 4 M€ uit deze infrastructuurmiddelen te reserveren voor toekomstige high performance computing (HPC) behoeften. Deze reserveringen zouden moeten worden beheerd door NWO. Om ook in de toekomst, wanneer de nieuwe supercomputer end-of-life is, in de behoeften van gebruikers van de supercomputer te kunnen voorzien, adviseert de commissie NWO vanaf 2020 een jaarlijkse bijdrage van 2 M€ aan deze reserveringen van SURF te vragen, te realiseren door tariefinkomsten van academische gebruikers en commerciële exploitatie van de machine. Uit deze reserveringen kan, naast een uiteindelijke vervanger, ook een mid-life upgrade voor de opvolger van Cartesius worden gefinancierd.

De verwachting is dat de opgebouwde reserveringen in 2025 niet toereikend zullen zijn om een nieuwe nationale supercomputer van een redelijk ambitieniveau te financieren. Mogelijk kunnen internationale HPC samenwerkingen, zoals het ambitieuze EuroHPC, maar mogelijk ook minder grootschalige “lokale” internationale samenwerkingen, kosteneffectieve en strategisch opportune aanvullingen op de Nederlandse HPC ambitie bewerkstelligen. Het behoeft nadruk dat een eventuele deelname aan EuroHPC ten koste zal gaan van mogelijkheden om in nationale HPC capaciteit te investeren. De kosten en baten van deze keuze dienen zorgvuldig te worden gewogen.

Expertise

Om het potentieel van alle bovengenoemde hardware optimaal te benutten is het nodig dat SURF de expertise die benodigd is voor het ondersteunen van de gebruikers kan aantrekken en behouden. Hiervoor adviseert de commissie vanaf 2019 jaarlijks 2 M€ ter beschikking te stellen. Gebruikersondersteuning is een noodzakelijke voorwaarde voor het ontsluiten en laagdrempelig toegankelijk maken van de digitale infrastructuur. Daarnaast adviseert de commissie, zoals eerder genoemd in de paragraaf Digitale Competentie Centra in een gefedereerd systeem, dat ondersteuning vanuit SURF aan lokale DCC's wordt bekostigd uit de 2 M€ die volgens de eerder genoemde aanbeveling jaarlijks ter beschikking gesteld zouden moeten worden voor het gefedereerde systeem.

Investering in supercomputer doorzetten

Gezien de capaciteitsproblematiek op de huidige nationale supercomputer beveelt NWO aan de in het bestedingsplan voorgestelde investering in de vervanging van de supercomputer (eenmalig 18 M€ uit overheidsmiddelen en 2 M€ cofinanciering door SURF) onverwijld door te zetten, zodat het vervangings-traject kan worden ingezet. De wachttijden voor het gebruik van de nationale supercomputer lopen snel op, en gezien de tijd die het aanbestedingsproces en de installatie van de nieuwe faciliteit zullen vergen, is het belangrijk dat de hiervoor benodigde middelen spoedig kunnen worden vrijgegeven.

Andere investeringen vergen nadere uitwerking met relevante partijen, resulterend in een uitvoeringsplan van bestedingen. De raad van bestuur van NWO is van mening dat de andere investeringen die de PC-GWI in het bestedingsplan aanbeveelt nog nadere uitwerking vereisen om deze te kunnen effectueren. Dit geldt ook voor de aanbevolen reservering van middelen ten behoeve van toekomstige HPC investeringen. Hiertoe zal NWO in samenspraak met de relevante partijen (waaronder in elk geval de universiteiten en SURF), de details van het bestedingsplan uitwerken in een uitvoeringsplan.

Langetermijnstrategie & samenhang met Europees samenwerkingsproject EuroHPC

Het bestedingsplan benadrukt het belang van het vormen van een langetermijnstrategie voor HPC. Het Europese samenwerkingsproject EuroHPC en andere mogelijkheden voor internationale samenwerking op HPC gebied kunnen een waardevolle aanvulling op de nationale rekenfaciliteiten bieden. In EuroHPC bundelen Europese landen hun kennis en middelen om HPC faciliteiten te bouwen die krachtiger zijn dan

wat binnen het bereik van de individuele lidstaten ligt, waarbij investeringen van lidstaten worden gematcht door de Europese Unie.

Deze mogelijkheden moeten zorgvuldig worden gewogen tegen het belang van investeringen in nationale rekenfaciliteiten, met oog voor de beperkte beschikbare middelen, de behoeften van Nederlandse onderzoekers en strategische nationale en internationale belangen.

Investeringen middelen -indien nodig- herijken in 2022

Tot slot wordt door de RvB van NWO gewezen op, gezien de snelle ontwikkelingen in de digitalisering, het belang van evalueren van de effecten van de door de PC-GWI aanbevolen investeringen in 2022 en op grond daarvan de middelenverdeling -indien nodig- te herijken.

Bijlage 2: Definitie lokale DCC's

Onderstaande is opgesteld door de regiegroep RDM-voorzieningen en is in dit document opgenomen ter referentie.

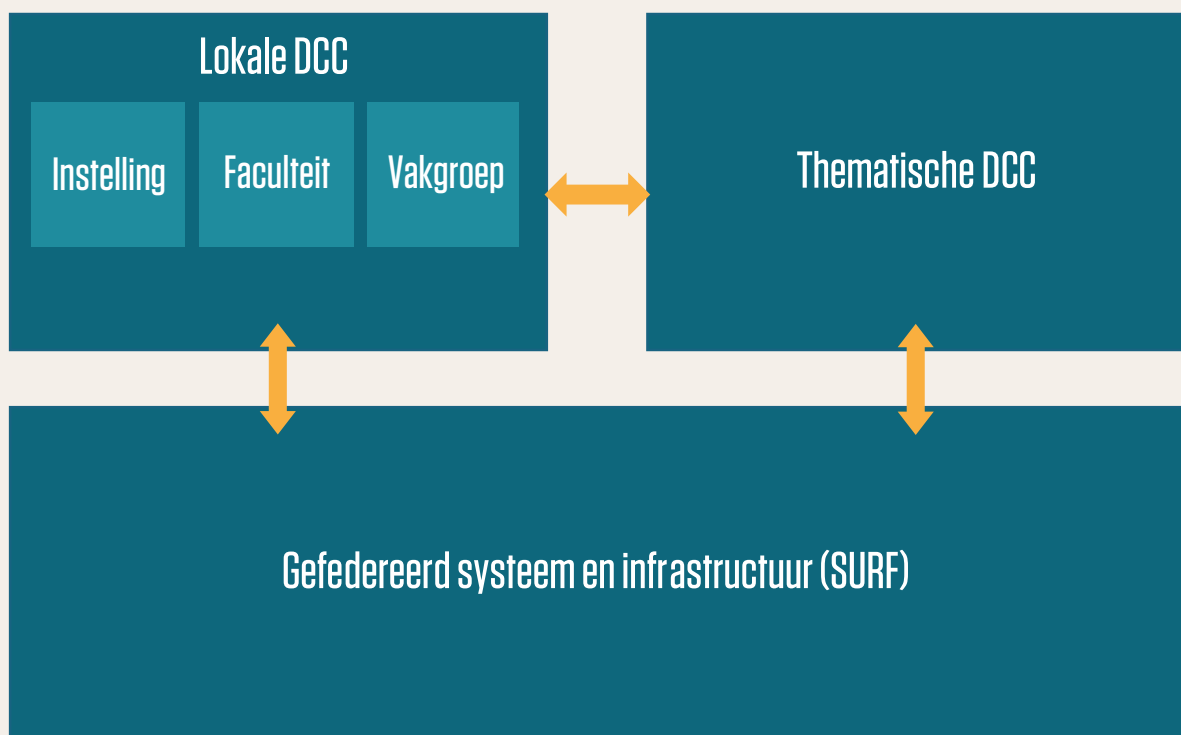
Definitie lokale DCC's

De lokale digitale competentie centra zijn dicht bij de onderzoekers gepositioneerd en zijn het eerste aanspreekpunt of loket voor onderzoekers. De onderzoekers vinden er ondersteuning en tooling op het gebied van onderzoeksdata, open- en FAIR-data, ook als de diensten feitelijk door horizontale of verticale thematische DCC's worden geleverd.

De ondersteuning en tooling van lokale DCC's omvat voorzieningen voor dataopslag, rekenen en bijbehorende netwerken, zowel voor tijdens het onderzoek als daarna. De ondersteuning sluit aan bij de taken en rollen zoals deze door het LCRDM zijn vastgelegd in de publicatie "Datastewardship op de kaart"¹⁸. Opleiding en training gaan naast Open Science en FAIRdata ook over tooling.

Binnen universiteiten en UMC's is er veelal sprake van een (virtuele) organisatie, waarbij tooling en ondersteuning binnen vakgroepen, faculteiten en/of op instellingsniveau geleverd worden. Financieringsvoorstellen worden op instellingsniveau gecoördineerd en ingediend. Voorgesteld wordt bij de impulsfinanciering het aantal onderzoekers als uitgangspunt te nemen bij de verdeling. De randvoorwaarde van cofinanciering is een ander uitgangspunt.

Thematische DCC's zijn tweedelijns ondersteuning voor de lokale DCC's voor domein- of vakgebiedspecifieke voorzieningen. SURF is een andere tweede lijn voor het gefedereerde systeem, de generieke infrastructuur en de landelijke kennisdeling. De lokale DCC's vervullen een coördinatie- en dispatchfunctie naar de tweede lijn.



¹⁸ <https://zenodo.org/record/2642066#.XRE4MY-xV1M>

Bijlage 3: Invulling horizon-model binnen SURF

Een nadere invulling van het horizon-model binnen SURF is in onderstaande tabel weergegeven. Horizon 1 innovatie wordt ingevuld via het LCPM-model dat binnen SURF gehanteerd wordt.

| | Horizon 1: 75-100% kans van slagen | Horizon 2: 25-75% kans van slagen | Horizon 3: < 25% kans van slagen |
|--------------|--|--|---|
| Drivers | Verbetering, uitbreidingen, varianten, kostenbesparingen. | Volgende generatie oplossingen voor belangrijkste gebruikers/doelgroepen (bestaande oplossingen voor nieuwe gebruikers, of bestaande oplossingen die we nog niet gebruiken). | Nieuwe oplossingen voor nieuwe vragen. |
| Vraagsturing | Sturing per dienst via Life Cycle Product Management (LCPM). Afstemming via SURF Portfolio Advies Commissie en stuurgroepen | | Sturing per onderwerp via jaarlijkse innovatieagenda. Afstemming nog niet geïnstitutionaliseerd. Contouren voor afstemming in dit document, een en ander afhankelijk van governance op SURF. |
| Uitvoering | Continuous improvement en standaardisatie binnen de diensten in domeinen. | Exploratie van haalbaarheid en potentiële waarde met instellingen en gebruikers. | Definitie van nieuwe ideeën en concepten in samenwerking met experts van buiten. |

Bijlage 4: Toelichting voorziene onderwerpen innovatieagenda

Exascale computing

Exascale computing refereert aan computersystemen met een performance van minimaal 1 exaflop/s, of 10^{18} (een miljard miljard) floating point berekeningen per seconde. Exascale computing biedt enerzijds mogelijkheden voor veel nauwkeurigere simulaties, modelvorming en data-analyse, maar anderzijds is daar het effectief gebruik van nieuwe hardware (GPUs en toekomstige ontwikkelingen) en software (bijvoorbeeld rekening houdend met hardware failures) voor nodig.

- **Lange termijn doelstelling/uitdaging:** De ondersteuning van onderzoekers bij het gebruik van nieuwe ICT-infrastructuur op exascale niveau en bij het omgaan met navenant grotere databestanden.
- **Urgentie:** Wetenschappelijk succes wordt in toenemende mate bepaald door beschikbaarheid van en toegang tot grootschalige wetenschappelijke infrastructuren. De ICT-ondersteuning voor onderzoeksprojecten en vakgroepen bij het gebruik van deze infrastructuren is daarbij cruciaal voor het effectief gebruik. In Europa komen deze voorzieningen in de aankomende 2–4 jaar beschikbaar. Om de wetenschappelijke concurrentie te kunnen blijven aangaan is een Nederlandse competentie om van dergelijke voorzieningen gebruik te kunnen maken noodzakelijk.
- **Rol van SURF:** Kennisopbouw en hands-on ervaring bij het gebruik van dergelijke systemen stelt SURF in staat onderzoekers te ondersteunen om bij hun onderzoek zo goed mogelijk gebruik te maken van de beschikbare technologie. SURF houdt verbinding met internationale ontwikkelingen zoals PRACE, EuroHPC en HPC Competence centers en werkt nationaal samen met onderzoeksgroepen en instellingen om deze kennis nationaal in te bedden.
- **Aanpak:**
 - Assessment van technologieën die naar verwachting onderdeel zullen uitmaken van exascale systemen (bijv. GPU's, low-precision GPU's, FPGA's, etc.);
 - Assessment van big data applicaties en toepassingen op HPC/supercomputer omgevingen;
 - Assessment van nieuwe architectuurconcepten als unified computing voor het flexibeler en breder gebruik van multipepetascale en exascale systemen;
 - Assessment van methodieken voor energy-efficient computing;
 - Deelname in Europese projecten zoals het European Processor Initiative die gericht zijn op een onafhankelijke positie van Europa op het gebied van ICT technologie;
 - Deelname in Europese projecten die zich richten op het geschikt maken van applicaties voor het exascale tijdperk.

Machine learning

Machine learning wordt in sterk toenemende mate ingezet in wetenschappelijk onderzoek, gedreven door de exponentieel toenemende hoeveelheden wetenschappelijke data, en de machine learning capabilities van high performance computers (convergence of HPC en Big Data Analytics). Machine learning is nu reeds een significant onderdeel geworden van vele wetenschappelijke workflows.

- **Lange termijn doelstelling/uitdaging:** Optimale inzet van machine learning in wetenschappelijke workloads (zowel aan applicatie-kant als aan infrastructuurkant)
- **Urgentie:** Machine learning biedt grote en vernieuwende mogelijkheden in wetenschappelijk onderzoek, waar kennis en infrastructuur nodig is ter ondersteuning voor competitief onderzoek door Nederlandse onderzoekers. Daarnaast zijn nieuwe benaderingen van numerieke simulaties nodig door eindigheid aan parallelisme van applicaties en eindigheid van Moore's Law.
- **Rol van SURF:** Vernieuwing van kennis en faciliteiten om onderzoekers in Nederland optimaal te ondersteunen. Door de ontwikkeling en de verspreiding van de nieuwste ML-mogelijkheden aan

onderzoekers en instellingen, beschikken onderzoekers in de toekomst over de beste mogelijkheden voor toepassing van machine learning in onderzoek. SURF houdt ook hier verbinding met internationale ontwikkelingen zoals PRACE, EuroHPC en HPC Competence centers.

- **Aanpak:**
 - Innovatie van nieuwe efficiënte en schaalbare infrastructuur en algoritmes (High Performance Deep Learning) voor machine learning toepassingen op grootschalige en complexe data, in samenwerking met key users (casestudies), internationale partners en technologie providers (o.a. Intel).
 - Assessment van potentie van verbeteringen van numerieke simulaties door toepassing van machine learning technieken, in nauwe samenwerking met key users uit verschillende wetenschappelijke disciplines (Deep learning enhanced HPC), resulterend in inzicht in potentieel en impact, en best practices.
 - Big Science gedreven innovatie van Machine learning technologieën ism NLeSC (eTEC-BIG call)
 - Bijdrage aan grote wetenschappelijke projecten zoals Cortex (Centrum voor Onderzoek in Real Time naar het Explosieve Universum)

Quantum computing

Grootschalige quantum computers zouden theoretisch in staat zijn om bepaalde klassen problemen veel sneller op te lossen dan een klassieke computer, bijvoorbeeld integer factorisatie of simulatie van many-body systemen. Bovendien kunnen quantum computers problemen oplossen die op klassieke computers niet haalbaar zijn vanwege het superpositiekenmerk. De potentiële impact van quantum computing op de huidige ‘klassieke’ high performance computing kan in de toekomst enorm zijn (ref.: The next decade in quantum computing and how to play, Philipp Gerbert (BCG), November 2018).

- **Lange termijn doelstelling/uitdaging:** Het Nederlandse wetenschappelijk onderzoeksveld is vroegtijdig voorbereid en kan concurrentievoordeel te halen uit Quantum Computing ontwikkelingen en faciliteiten zodra deze beschikbaar komen
- **Urgentie:** Quantum revolutie gaat razendsnel, en Nederland heeft een voorloperspositie in technologieontwikkeling. We komen nu in de ‘quantum ready phase’ (Gerbert 2018) en moeten nu handelen om als sector goed voorbereid zijn op de ‘quantum advantage phase’ (3–5 jaar) waarin quantum technologie toegepast gaat worden (richting mainstream).
- **Rol van SURF:** SURF stimuleert en faciliteert daarin als nationaal HPC expertisecentrum met een groot aantal wetenschappelijk HPC gebruikers de benodigde kennisontwikkeling: wat vraagt het (kennis, tools, tijd) om bestaande wetenschappelijke applicaties te transformeren tot quantum applicaties. Daarmee wordt inzichtelijk wat de mogelijke impact is op de bestaande workload van HPC voorzieningen in Nederland. De lange-termijn HPC strategie kan hierop worden afgestemd.
- **Aanpak:**
 - Ontwikkeling en executie van wetenschappelijke quantum algoritmes door onderzoekers wordt door SURF ondersteund door middel van expertise, samenwerking (o.a. met QuTech) en faciliteiten. Zo werken SURF en QuTech samen aan de ontwikkeling van een schaalbare quantum simulator, die beschikbaar wordt gesteld aan onderzoekers op de nationale supercomputer (via het Quantum Inspire platform van QuTech) ten behoeve van het ontwikkelen en runnen van quantum algoritme.
 - Twee concrete quantum applicatie use cases worden ontwikkeld in samenwerking met het wetenschappelijke onderzoeksveld, technology providers (zoals QuTech) en quantum algoritme-ontwikkelaars (zoals QuSoft): 1) in quantum chemie: implementatie van een interface tussen quantum chemie applicatie en quantum computer/simulator, en implementatie en verbetering van quantum variational eigensolvers) en 2) in quantum machine learning: vele use cases mogelijk. Voor deze toepassingsgebieden wordt het snelst en het grootste quantum voordeel verwacht, bovendien zijn dit zeer belangrijke toepassingen op de huidige nationale supercomputer. Best practices, best examples van 1^e praktische implementaties worden zo ontwikkeld.

Data processing & streaming data

Onderzoekers combineren hebben steeds meer behoefte aan schaalbare oplossingen voor data verwerking en data analyse, zodat enorme hoeveelheden data zoals die geproduceerd worden in bijvoorbeeld astronomie (data van bv telescopen), hoge energie fysica (data van bv de deeltjesversneller van CERN), sociale wetenschappen (data van Twitter) en life sciences (data uit DNA studies) kunnen worden verwerkt. Analysemethoden worden veilig toegepast in grote en complexe workflows, in toenemende mate via Virtual Research Environments (VREs) voor onderzoek, gekoppeld aan geavanceerde en gedistribueerde clouddiensten (public en private clouds).

- **Lange termijn doelstelling/uitdaging:** Ondersteunen van het Nederlandse wetenschappelijk onderzoeksveld en wetenschappelijke instellingen voor bij het uitvoeren van (real-time) effectieve verwerking en analyse van grote en complexe onderzoeksdata.
- **Urgentie:** Uitdagingen in multidisciplinair onderzoek worden steeds groter en nieuwe onderzoeksmethodes ontwikkelen zich razendsnel. Data wordt steeds meer real-time gegenereerd en verwerkt. Ook de technologie (vb. machine learning, IoT) ontwikkelt zich, evenals de markt met snel ontwikkelende commerciële cloud diensten. Dit vraagt nieuwe benaderingen en 'future architectures' voor ICT-infrastructuren voor onderzoek, zodat ook in de toekomst wetenschappelijk onderzoek optimaal ondersteunt wordt.
- **Rol van SURF:** Vernieuwing van kennis en faciliteiten om onderzoekers in Nederland optimaal te ondersteunen, door innovatie op het gebied van onder andere virtual research environments, data processing en clouds voor onderzoek, in nauwe samenwerking met instellingen, grootgebruikers en research infrastructures, in nationaal en internationaal verband
- **Aanpak:**
 - Assessment van (cloud) technologieën voor toepassing in ondersteuning van (big science) wetenschappelijk onderzoek, ism technology providers en wetenschappelijke domeinen (o.a. via ESCAPE (European Science Cluster of Astronomy & Particle physics ESFRI research infrastructures))
 - Bijdrage aan de ontwikkeling van federatieve VRE componenten voor instellingen/communities, met bijzondere aandacht voor integratie met federatief access and identity management, integratie met research data management en containertechnologie en de interfacing via een federatieve en interactieve catalog voor data en tools. Oriëntatie op Europese ontwikkelingen
 - Big Science gedreven innovatie van Processing van streaming data ism NLeSC (eTEC-BIG call)
 - Ontwikkeling van high-throughput data processing cloud platform met flexible service deployment
 - Ontwikkeling van een open self-service platform voor federatie van gedistribueerde cloud diensten en data (Research Cloud)

Data management

Goed research data management is een van de belangrijkste uitdagingen van onderzoekers, gedreven door Open Science ontwikkelingen (onder andere FAIR data eisen), door de exponentiele toename van de hoeveelheid en complexiteit van research data, alsook de steeds grotere samenwerkingsverbanden waarin data wordt gedeeld en gebruikt.

- **Lange termijn doelstelling/uitdaging:** Ondersteunen van het Nederlandse wetenschappelijk onderzoeksveld en wetenschappelijke instellingen voor optimaal (her)gebruik en management van research data
- **Urgentie:** de huidige eisen van funders en instellingen aan FAIR data management, exponentiele groei van research data
- **Rol van SURF:** Vernieuwing van kennis en faciliteiten om onderzoekers in Nederland optimaal te ondersteunen, door innovatie op het gebied van nieuwe federatieve architectures, diensten en kennis voor optimaal laagdrempelig research data management in nauwe samenwerking met instellingen, grootgebruikers en research infrastructures, in nationaal en internationaal verband

- **Aanpak:**
 - Innovatie van een nieuwe generatie federatieve RDM stack, via nationale samenwerkingsverbanden en via internationale samenwerkingen met e-infrastructuren (EUDAT, EGI) en ESFRIs
 - Onderzoek naar mogelijkheden van het ondersteunen van data delen met behoud van data soevereiniteit van data eigenaren (digitale data exchange)
 - Big Science gedreven innovatie van Large-scale (Distributed) Data Organization, Management & Semantics ism NLeSC (eTEC-BIG call)

Lange-termijn behoefte onderzoekscommunities

- **Lange termijn doelstelling/uitdaging:** de kennis en voorzieningen van de nationale infrastructuur aansluiten bij de behoeften van de verschillende wetenschapsdomeinen door samenwerking met georganiseerde onderzoeksgemeenschappen in kennisvernieuwing. Dit draagt bij aan de ontwikkeling van generieke nationale dienstverlening voor onderzoek en een betere aansluiting van Research Infrastructures op de nationale infrastructuur.
- **Urgentie:** Aansluiting op de ontwikkelingen bij goed georganiseerde onderzoeksgemeenschappen is essentieel om fragmentatie in onderzoeksdienstverlening te voorkomen, en de positie en zichtbaarheid van nationale RI's internationaal te vergroten. De snelheid van de (internationale) ontwikkelingen vereist snelheid en coördinatie van handelen.
- **Rol van SURF:** SURF bundelt de vraag van georganiseerde onderzoekscommunity's in de nationale e-infrastructuur
- **Aanpak:** in samenwerking met georganiseerde community's worden relevante roadmaps gemaakt die de ontwikkeling van huidige en toekomstig Research-Infrastructures ondersteunen. Daartoe benodigde Proof-of-Concepts worden gedefinieerd en uitgevoerd.

Bijlage 5: Spreiding calls en projecten eScience Center 2014–2018

Table 1: Submitted and accepted proposals in ASDI and DTEC calls

| Domain | 2014 | | 2015 | | 2016 | | 2017 | | 2018 | | Totals | |
|--------------------------------|--------------|----------|--------------|----------|--------------|----------|--------------|-----------|-------------|----------|--------------|-----------|
| | Subm | Acc | Subm | Acc | Subm | Acc | Subm | Acc | Subm | Acc | Subm | Acc |
| Environment & sustainability | 0 | 0 | 13 | 1 | 15 | 1 | 8 | 1 | 27 | 2 | 63 | 5 |
| Humanities & social sciences | 4 | 2 | 16 | 2 | 9 | 1 | 20 | 3 | 45 | 2 | 94 | 10 |
| Life sciences & eHealth | 16 | 4 | 19 | 1 | 20 | 2 | 19 | 1 | 0 | 0 | 74 | 8 |
| Physics & beyond | 0 | 0 | 8 | 3 | 8 | 1 | 51 | 8 | 0 | 0 | 67 | 12 |
| Computer and Data Sciences | 0 | 0 | 20 | 2 | 17 | 3 | 14 | 4 | 0 | 0 | 51 | 9 |
| Total | 20 | 6 | 76 | 9 | 69 | 8 | 112 | 17 | 72 | 4 | 349 | 44 |
| Average acceptance rate | 30.0% | | 11.8% | | 11.6% | | 15.2% | | 5.6% | | 12.6% | |

Bron: Self-assessment eScience Center 2014 – mid 2018



Figure: Map of projects 2014-2018 (Bron: Self-assessment eScience Center 2014 – mid 2018)

Bijlage 6: Verklarende woordenlijst en afkortingen

| | |
|---------------------|--|
| CIO | Chief Information Officer – binnen een organisatie de hoogst verantwoordelijke op het gebied van ICT |
| CPU | Central Processing Unit is een processor (of ‘chip’), een stuk hardware in een computer die instaat voor basisbewerkingen en – controle bij uitvoeren van programmacode. |
| Data steward | Een Data steward houdt zich bezig met het data beleid (data governance) van de organisatie en zorgt voor de juistheid, volledigheid, integriteit en kwaliteit van de gegevens in de organisatie. Datastewardship staat voor het geheel van verschillende taken en verantwoordelijkheden dat betrekking heeft op de zorg voor data, gedurende de verschillende fases van de gehele onderzoekslevenscyclus. |
| DCC | In Digitale Competentie Centra, of Digital Competence Centers (DCC’s), vinden onderzoekers ondersteuning en technologische hulpmiddelen op het gebied van onderzoeksdata, onderzoekssoftware, en open- en FAIR-data. Lokale DCC’s binnen kennisinstellingen vormen het eerste aanspreekpunt voor onderzoekers. Onderlinge uitwisseling van kennis en technologie kan worden gestimuleerd door middel van een veilig en gefedereerd netwerk en het oprichten van thematische DCC’s. NWO wil de ontwikkeling van DCC’s stimuleren. |
| EOSC | European Open Science Cloud – in 2016 voorgesteld door de Europese Commissie ter bevordering van een competitieve data- en kenniseconomie in Europa. EOSC biedt Europese onderzoekers een open en betrouwbare virtuele omgeving voor opslag van onderzoeksdata via een gefedereerd netwerk van bestaande en toekomstige digitale infrastructuren. |
| EuroHPC | De EuroHPC Joint Undertaking is een samenwerking tussen de EU en Europese landen om samen een wereldklasse supercomputer infrastructuur te ontwikkelen en onderhouden. Gebruikers in de deelnemende landen kunnen daardoor gebruik maken van een groter scala aan High Performance Computing mogelijkheden. |
| Exascale | Exascale computing is het gebruik van een nieuwe generatie geavanceerde computersystemen, waarin de totale capaciteit van alle processoren de waarde van 1 exaFLOPS overschrijdt (1 exaFLOPS staat gelijk aan 10^{18} (een miljard miljard) floating point operations per second, een eenheid die wordt gebruikt om de rekenkracht van CPU’s aan te duiden), maar waarin ook andere componenten, zoals geheugengrootte, geheugensnelheid, dataopslagruimte, bandbreedte en latentietijd, proportioneel zijn. |
| FAIR | FAIR staat voor Findable, Accessible, Interoperable, Reusable; om data en andere onderzoeksgegevens open toegankelijk te maken voor iedereen volgens het Open Science beleid is het noodzakelijk dat deze data zoveel mogelijk Vindbaar (Findable), Toegankelijk (Accessible), Interoperabel (Interoperable) en Herbruikbaar (Reusable) worden opgeslagen op een daarvoor geschikt systeem. |

| | |
|----------------------------------|---|
| Gefedereerd netwerk DCC's | Een gefedereerd netwerk is een digitaal netwerk waarin Digitale Competentie Centra samenwerken. Een lokale DCC blijft zelfstandig opereren, maar vormt tevens een knooppunt in het digitale netwerk, zoals bijvoorbeeld in het netwerk van een thematische DCC, waarin de lokale DCC samenwerkt met andere lokale DCC's. |
| GPU | Graphics Processor Unit is een processor (of 'chip') die snel wiskundige berekeningen kan uitvoeren, met name om beelden en video te genereren. Een GPU processor kan de taak van een CPU versnellen en/of overnemen. Een GPU bevindt zich op een videokaart. |
| Grid service | Grid bestaat uit een groot aantal computerclusters, verspreid over de hele wereld, die verbonden zijn via een snel netwerk. Met Grid kunnen grote hoeveelheden data gelijktijdig worden verwerkt in een korte tijd (via parallel processing) en opgeslagen. |
| HPC | High Performance Computing – richt zich op met zeer grote snelheid uitvoeren van complexe berekeningen, zoals bijvoorbeeld klimaatberekeningen. Dit vereist supercomputers die voor individuele instellingen nauwelijks te financieren zijn. |
| HPC cloud | High Performance Computing Cloud – een virtueel systeem waarmee de gebruiker toegang en controle heeft over zijn/haar eigen HPC computer of computer cluster en gebruik kan maken van zijn/haar eigen aansturingssysteem en software. |
| HTC | High Throughput Computing – richt zich op zoveel mogelijk (vaak kleinere) rekentaken binnen een bepaalde tijd uit te voeren. Deze berekeningen worden uitgevoerd met relatief goedkopere computers die door middel van een computernetwerk met elkaar verbonden zijn. Dit wordt ook wel Grid-computing genoemd. |
| KNAW | Koninklijke Nederlandse Akademie van Wetenschappen – adviesorgaan van de regering, genootschap van excellente wetenschappers en institutenorganisatie; de KNAW is verantwoordelijk voor tien nationale onderzoeksinstituten en drie instituten die infrastructuur bieden voor onderzoek. |
| LCRDM | Landelijke Coördinatiepunt Research Data Management – landelijk netwerk van experts of het gebied van research data management. Het LCRDM maakt de koppeling tussen beleid en oplossing op landelijk niveau. Het LCRDM brengt onderzoeksondersteuners, beleidsmakers, ICT-ers, bestuurders uit onderzoeksinstituten en onderzoeksfinanciers bij elkaar. |
| Lokale DCC | Lokale Digitale Competentie Centra bevinden zich binnen kennisinstellingen en vormen het eerste aanspreekpunt voor onderzoekers voor ondersteuning en toegang tot technologische hulpmiddelen op het gebied van onderzoeksdata, onderzoekssoftware, en open- en FAIR-data |
| NLeSC | Netherlands eScience Center – in 2011 opgericht door NWO en SURF; ondersteunt en versterkt multidisciplinair en data-intensief wetenschappelijk onderzoek waarin effectief en innovatief ICT wordt gebruikt. |

| | |
|---------------------|--|
| NPOS | Nationaal Plan Open Science of Nationaal Platform Open Science – Voor alle Nederlandse kennisinstellingen en onderzoeksdomeinen worden in 2020 open access publiceren en optimaal hergebruik van onderzoeksdata de norm. Het motto van het Plan hierbij is: open als het kan, gesloten als het moet. In het Platform komen de betrokken partijen uit het Plan bijeen en zorgt ervoor dat Nederland vordert met de realisatie van de ambities ten aanzien van Open Science en houdt zicht op de ontwikkelingen |
| NWO | Nederlandse Organisatie voor Wetenschappelijk Onderzoek – zelfstandig bestuursorgaan, wetenschapsfinancier en institutenorganisatie; NWO is verantwoordelijk voor negen nationale onderzoeksinstituten. |
| Open Science | Open Science houdt in dat onderzoeksresultaten vrij toegankelijk zijn voor eenieder die deze wil inzien. Dit geldt voor zowel wetenschappelijke publicaties als voor andere vormen van wetenschappelijk output, zoals ook onderzoeksgegevens en -data. Op die manier kan waardevolle kennis benut worden door andere onderzoekers, bedrijven en maatschappelijke instellingen. |
| PC GWI | Permanente Commissie voor Grootchalige Wetenschappelijke Infrastructuur – deze commissie is in 2015 benoemd door NWO in opdracht van het Ministerie van Onderwijs, Cultuur en Wetenschap. De commissie brengt alle grootchalige onderzoeksfaciliteiten die voor Nederlandse onderzoekers toegankelijk zijn in kaart en stelt vast welke faciliteiten nog ontbreken. De commissie adviseert NWO strategisch over toekomstige investeringen in grootchalige wetenschappelijke infrastructuur. De PC GWI commissie bestaat uit 12 leden met een breed overzicht over de Nederlandse wetenschap, onder leiding van commissievoorzitter Prof. dr. ir. C.J. van Duijn. De ICT-subcommissie (commissie Apers) richt zich op ICT-vragen. |
| PRACE | Partnership for Advanced Computing in Europe – een internationale vereniging, gevestigd in Brussel. De PRACE Research Infrastructure biedt een wereldklasse High Performance Computing dienst voor onderzoekers in wetenschap en industrie in Europa. De computersystemen en diensten die toegankelijk zijn via PRACE worden beschikbaar gesteld door vijf PRACE leden, die Spanje, Italië, Zwitserland, Duitsland en Frankrijk vertegenwoordigen. |
| Pre-exascale | Bij Pre-exascale computing wordt gebruik gemaakt van een nieuwe generatie geavanceerde computersystemen, die sneller zijn dan de huidige computersystemen, maar nog niet zo snel als Exascale computing. De pre-exascale computersystemen die momenteel in Europa in ontwikkeling zijn, zijn met 150 petaFLOPS (1 petaFLOPS staat gelijk aan 10^{15} (een miljoen miljard) floating point operations per second, een eenheid die wordt gebruikt om de rekenkracht van CPU's aan te duiden) naar verwachting 4-5 keer sneller dan het huidige top supercomputersysteem van PRACE. |
| SURF | Coöperatieve vereniging van Nederlandse onderwijs- en onderzoeksinstituten, die samen zorgen dat de onderwijs- en onderzoeksgemeenschap kan beschikken over de nieuwste en beste ICT-voorzieningen voor toponderzoek en talentontwikkeling. De SURF coöperatie bestaat uit ruim 100 leden, waaronder universiteiten, hogescholen, middelbaar beroepsonderwijs-scholen, universitair medisch centra en andere onderzoeks- en onderwijsinstellingen. |

Thematische DCC

Onderzoeksinstituten en lokale Digitale Competentie Centra kunnen samenwerkingen aangaan via thematische Digitale Competentie Centra, waardoor een onderlinge uitwisseling van kennis en technologie wordt gestimuleerd via een veilig en gefedereerd netwerk.

**Nederlandse Organisatie
voor Wetenschappelijk Onderzoek**

Laan van Nieuw Oost-Indië 300
2593 CE Den Haag
Postbus 93138
2509 AC Den Haag

+31 (0)70 344 06 40
www.nwo.nl

