

## Samenwerking op service-infrastructuur en competenties voor data

### *Managementsamenvatting*

*Alle sectoren van de maatschappij, industrie, onderwijs en wetenschap ondergaan momenteel een digitale transformatie. De daarbij beschikbaar komende data en de mogelijkheden van artificiële intelligentie hebben een grote impact op alle mogelijke werkprocessen waar de huidige studenten later in terecht zullen komen.*

*Die digitale transformatie vergt een datageoriënteerde methodische service-aanpak om te blijven innoveren en voorop te kunnen lopen in het internationale speelveld. Onvermijdelijk betekent dit het maken van nieuwe strategische keuzes om innovatie te stimuleren.*

*SURF heeft jarenlange ervaring in het nemen van een leidende rol op de technische infrastructuur en de daarbij behorende services. Dit is het moment voor SURF om wederom een leidende rol op zich te nemen. Deze keer op het gebied van een datageoriënteerde service-infrastructuur en de competenties om die infrastructuur innovatief in te zetten voor onderwijs en onderzoek. SURF is daarvoor zowel nationaal als internationaal goed gepositioneerd.*

*De WTR adviseert de leden van SURF om deze visie verder uit te laten werken door SURF.*

### Veranderend landschap – nieuwe uitdagingen en nieuwe kansen

De maatschappij staat voor nieuwe uitdagingen die zich uitstrekken over alle sectoren en domeinen. Van de zorg voor het milieu, de economie, het onderwijs, de gezondheidszorg tot onderzoek. Hoe kan de kwaliteit omhoog? Hoe kan dat effectiever, efficiënter en met meer satisfactie voor de betrokkenen? De verregaande digitalisering van allerlei (keten)processen en de nieuwe data-analysmethoden vergen én produceren steeds meer en omvangrijkere hoeveelheden data. Dat levert nieuwe kansen en bedreigingen op. Data is het nieuwe goud en iedereen krijgt daar mee te maken.

Met name onderzoek is een belangrijke trendsetter in het gebruik van grote hoeveelheden data, breder bekend onder de noemer eScience - datageoriënteerde wetenschap. Het gaat daarbij niet alleen om grote hoeveelheden data, maar ook om inter- en cross-disciplinaire ontwikkeling van methoden en technieken voor analyse. Dit was een belangrijke drijfveer voor de oprichting van het NL eScience Center.

In het afgelopen decennium heeft deze ontwikkeling rond data zijn weg gevonden naar alle hiervoor genoemde gebieden in de maatschappij.

In het onderwijs gaat het bijvoorbeeld om verdere digitalisering, zoals toetsing en data-analyse, mogelijkheden om het onderwijs te flexibiliseren, het (her)gebruik van digitale leermaterialen (content), het benutten van studiedata, de faciliteiten die de markt biedt (of juist laat liggen) in de vorm van software, hardware en diensten aan studenten en medewerkers, etc.

In alle onderzoeksdomeinen speelt de vraag naar Research Data Management (dataopslag, -beheer, -verwerking en -analyse), rekencapaciteit en Open Science. Dé grote omwenteling tijdens het laatste decennium – ook in onderwijs en onderzoek - is de verschuiving van de aandacht voor de technische voorzieningen naar aandacht voor data ten behoeve van analyse en processen, bijv. met behulp van artificiële intelligentie. Het begrip 'data' wordt daarbij breed geformuleerd als 'objecten', van gegevens tot content en software, maar in ieder geval van een hogere orde dan bits en bytes.

Voor onze data worden we daarbij steeds afhankelijker van commerciële bedrijven, met name buiten Europa.<sup>1,2</sup> Dat leidt tot nieuwe Europese initiatieven, zoals GAIA-X<sup>3</sup>, dat de ontwikkeling beoogt van een competitieve, veilige en betrouwbare, federatieve en open data-infrastructuur, gebaseerd op Europese waarden.

<sup>1</sup> 'Digitalisering bedreigt onze universiteit. Het is tijd om een grens te trekken', <https://www.volkskrant.nl/columns-opinie/digitalisering-bedreigt-onze-universiteit-het-is-tijd-om-een-grens-te-trekken~bff87dc9/> Volkskrant, 22 december 2019.

<sup>2</sup> 'Investeer in de publieke kennisinfrastructuur voordat deze alleen nog in buitenlandse handen ligt', <https://www.volkskrant.nl/columns-opinie/investeer-in-de-publieke-kennisinfrastructuur-voordat-deze-alleen-nog-in-buitenlandse-handen-ligt~b7fc60c8/> Volkskrant, 23 december 2019.

<sup>3</sup> Project GAIA-X, <https://www.bmwi.de/Redaktion/EN/Publikationen/Digitale-Welt/das-projekt-gaia-x-executive-summary.pdf>, Ministeries van Economische Zaken en Energie, en Onderwijs en onderzoek, Duitsland, 29 oktober 2019.

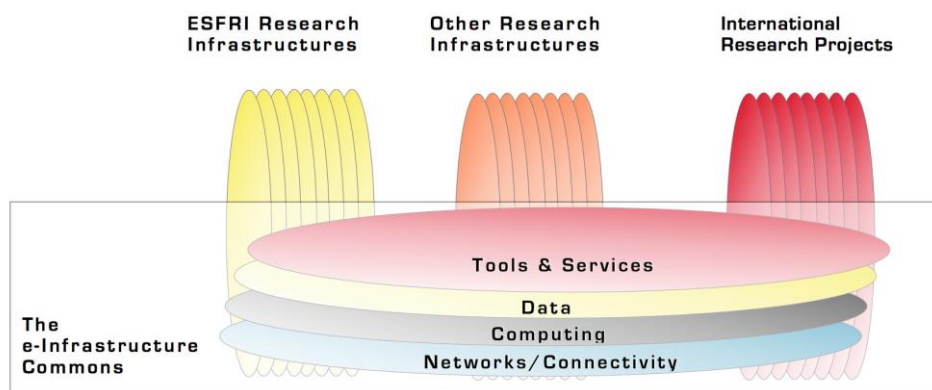
## Toekomstperspectief voor SURF en haar leden

### Verschuiving van aandacht naar research & educational datamanagement.

De grootste uitdaging van dit decennium is het innovatief, veilig en betrouwbaar ondersteunen van datastromen en de cyclus van data verzamelen, (selectief) opslaan, beschikbaar stellen, uitwisselen, selecteren, samenvoegen, verwerken en analyseren. Hoe zorgen we ervoor dat het uitgangspunt wordt gehonoreerd dat alle publiek gefinancierde databronnen en producten openbaar en transparant toegankelijk zijn, ofwel FAIR: *Findable, Accessible, Interoperable* en *Reusable*? Niet alleen voor onderzoek, maar ook voor bijv. onderwijs(leer)-materiaal.

Die grote uitdaging vergt een heroriëntatie in het maken van keuzes en de inzet van financiële middelen. Momenteel gaat het grootste deel van de middelen<sup>4</sup> waarover SURF beschikt naar de technische infrastructuur, het beheer en de gebruikersondersteuning: netwerk, computing en opslag. In de naaste toekomst zal (bij gelijkblijvende middelen) een groter deel nodig zijn voor de ontwikkeling van een nieuwe datageoriënteerde dienstenlaag. Daar zit een technische kant aan, maar vooral een omvangrijke organisatorische kant: het maken van afspraken over standaarden, over toegang en beveiliging, het verdelen van taken en verantwoordelijkheden, etc. Dat is niet de missie van de instellingen, maar het past wel bij SURF. Door hierin gezamenlijk te opereren in SURF kunnen de instellingen ontzorgd worden, zoals SURF dat in de beginjaren deed bij de opkomst van computers en netwerken.<sup>5</sup>

SURF en veel van haar leden en partners zijn, vanwege de urgentie, al diverse activiteiten op het gebied van data-management gestart. De aard, omvang en complexiteit van de voorliggende innovatietaak is echter van een hogere orde. Het vergt nationale samenwerking om internationaal aan te sluiten. Op Europees niveau wordt bijvoorbeeld verwacht dat de lidstaten nationale knooppunten vormen als onderdeel van de *e-Infrastructuur Commons* ten behoeve van de European Open Science Cloud.



Figuur 1 *e-Infrastructuur Commons* (EU) voor onderzoek. De nationale coördinerende taak van SURF breidt zich uit naar Data, Tools & Services

Om misverstanden te voorkomen, benadrukken we dat verdere missie-specifieke innovatie van de technische laag voor onderzoek en onderwijs hoogstnoodzakelijk blijft door het veranderende gebruik van data. De toepassing van artificiële intelligentietechnieken is daarin een dominante trend.

De technische uitdagingen zijn onverminderd groot, zeker ook voor Europa, om op wereldniveau een rol te kunnen spelen. De kosten daarvan overstijgen in toenemende mate de mogelijkheden van individuele organisaties en landen. Samenwerking is daarom een *conditio sine qua non*.

### Gemeenschappelijke voorzieningen voor diversiteit in gebruik

Hoewel er verschillen zijn tussen onderzoek en onderwijs in soorten data/objecten, hun omvang en toepassing, zijn de basisvoorwaarden voor een datageoriënteerde service-infrastructuur niet verschillend voor onderwijs en onderzoek. Integendeel, ze vertonen juist veel overeenkomsten. Voorkomen moet worden dat er meerdere service-infrastructuren los van elkaar ontstaan.

Met name binnen onderzoek zijn al veel initiatieven ontstaan vanuit verschillende onderzoeksdomeinen en -disciplines om te komen tot een datageoriënteerde service-infrastructuur. Enkele voorbeelden in

<sup>4</sup> Exclusief de middelen voor content, met name tijdschriftenlicenties.

<sup>5</sup> *Samenwerking...? Reken maar!* Stuurgroep Samenwerking Universitaire Rekencentra, 1985.

## Toekomstperspectief voor SURF en haar leden

---

Nederland zijn Health RI<sup>6</sup>, CLARIAH<sup>7</sup> en ODISSEI<sup>8</sup>. Zij werken samen met verschillende partners – leden van SURF, overheidsorganisaties en bedrijfsleven - aan een eigen onderzoeksinfrastructuur (voorzieningen en diensten) voor grote datacollecties. Het zijn belangrijke initiatieven, maar overkoepelende samenwerking is noodzakelijk om fragmentatie te voorkomen van basisvoorzieningen (techniek en services) die in hun aard veel overeenkomsten vertonen.

Onderwijs is minder ver dan onderzoek als het gaat om initiatieven voor datageoriënteerde service-infrastructuren, maar kent meer samenwerking op overkoepelend niveau, zoals in het Versnellingsplan Onderwijsinnovatie met ICT. Het voordeel daarvan is dat onderwijs - in vergelijking met onderzoek - relatief gezien een *greenfield* is op dit gebied. Het nadeel is dat er minder ervaring is met datageoriënteerde service-infrastructuren, maar door die verschillen kunnen onderwijs en onderzoek elkaar juist versterken. Zij kunnen elkaar ook versterken in het scholen van de nieuwe aanwas aan onderzoekers, docenten en medewerkers voor het gebruik van die nieuwe service-infrastructuur. Innovatie ontstaat niet vanzelf door het bieden van nieuwe voorzieningen, mensen moeten het vooral innovatief benutten.

### Eén overkoepelend plan voor SURF: een nationale service-infrastructuur en competenties voor data

Op dit moment is het datalandschap in Nederland, evenals in andere landen, gefragmenteerd. Data zitten momenteel voor een belangrijk deel 'opgesloten' in silo's van en voor specifieke domeinen, instellingen, bedrijven en overheidsorganisaties.

Om te komen tot een nationaal knooppunt is (federatieve) samenwerking noodzakelijk. SURF is bij uitstek geschikt om hierin het voortouw te nemen als neutrale ICT-organisatie van al haar leden. Dat past ook bij de aanpak van de grootschalige wetenschappelijke infrastructuur in Nederland waarin aan indieners van voorstellen wordt gevraagd om aan SURF duidelijk te maken welke voorzieningen zij nodig hebben. Het gaat daarbij niet alleen om de technische en servicevoorzieningen van SURF, maar ook om die van andere instellingen en organisaties. Alle partijen met data, dus ook organisaties zoals CBS en KB en andere niet-academische partners, moeten meegenomen worden in deze plannen, zodat een technisch en organisatorisch ecosysteem ontstaat voor zowel onderzoek als onderwijs.

Dit vergt een overkoepelend plan waar de activiteiten van SURF en de deelnemers onder geschaard kunnen worden. De realisatie - ontwikkeling, onderhoud en aanpassing - van dit overkoepelende plan zal zeker het huidige decennium in beslag nemen. Het vergt enorm veel coördinatie, aansturing en communicatie om bestaande datageoriënteerde service-infrastructuren te integreren en daarnaast nieuwe services en de benodigde gebruikerscompetenties te ontwikkelen.

### Tot slot

SURF heeft meer dan dertig jaar ervaring als het gaat om complexe samenwerking. En met succes. Juist daarom is SURF de meest geëigende organisatie om gezamenlijk met de leden en partners in te zetten op een nieuwe nationale infrastructuur, nu voor datageoriënteerde services en de daarvoor benodigde competenties. Er is nog een lange en complexe weg te gaan, maar de WTR heeft er vertrouwen in dat SURF die nieuwe datageoriënteerde infrastructuur gedurende de komende jaren zal kunnen realiseren samen met haar leden en partners.

De WTR adviseert de leden van SURF om deze visie verder uit te laten werken door SURF.

---

<sup>6</sup> Health Research Infrastructure – to establish a state-of-the-art infrastructure for data, samples and images that will be key to facilitating excellent health research and sustainably bridging the health innovation gap.

<sup>7</sup> Common Lab Research Infrastructure for the Arts and Humanities.

<sup>8</sup> Open Data Infrastructure for Social Science and Economic Innovations.