

## Pilot: Verbinden van systemen voor gezondere peutermonden

Hogeschool: Hogeschool Utrecht

Lectoraten:

- *Innovaties in de Preventieve Zorg | Onderzoeksproject Gezonde peutermonden*
- *Kenniscentrum Gezond en Duurzaam Leven*
- *Contactpersoon: Margreet Riphagen | [margreet.riphagen@hu.nl](mailto:margreet.riphagen@hu.nl)*

Looptijd: februari 2021 – oktober 2021

### Beschrijving van de pilot

**Context** | Aan de Hogeschool Utrecht en aan andere hogescholen vindt in het domein van de zorg veel mens- gebonden onderzoek plaats. Dit onderzoek wordt uitgevoerd aan patiënten of cliënten of met gebruik van vrijwilligers. De betrokken onderzoekers zijn promovendi, docent-onderzoekers lectoren. Zij werken samen met externe zorgprofessionals en andere onderzoekers en belanghebbenden zoals gemeenten en patiëntenorganisaties. Veel van deze onderzoeksprojecten verzamelen een grote diversiteit aan (gevoelige) data. Data in de vorm van klinische uitslagen, experimentele en fysiologische parameters, vragenlijsten, patiëntgegevens, klinische voorgeschiedenis en anamnese rapporten, experimentele gegevens, dagboekjes, videobeelden, geluidsfragmenten etc.

**Casus** | onderzoeksproject ‘Gezonde Peutermonden’, wordt geleid vanuit het Lectoraat “Innovatie in de Preventieve Zorg” (lector: Dr. Katarina Jerkovic). In 2017 is gestart met het Gezonde Peutermonden onderzoek, waarbij een mondzorgcoach op het consultatiebureau wordt ingezet om tandbederf bij jonge kinderen te voorkomen. Binnen dit onderzoek wordt de (kosten)effectiviteit onderzocht van preventieve mondzorg vanaf de doorbraak van de eerste melktand. Kinderen worden hierbij vanaf de leeftijd van 6-8 maanden opgevolgd tot vierjarige leeftijd. Daarnaast worden belemmerende en bevorderende factoren bij de implementatie van Gezonde Peutermonden in een nieuwe omgeving onderzocht. En zo onderzocht wat de ‘beste’ implementatiestrategie is om zo de mondzorgcoach als reguliere zorg in te zetten. Het project kent vele uitdagingen op het gebied van: data verzamelen, autorisatie van toegang tot cliëntendata en persoonsgegevens, datamanagement, data-analyse, inrichting van elektronische werkprocessen om de communicatie tussen cliënt, ouders en zorgverleners te stroomlijnen en uit te voeren, archiveren van data, beschikbaar stellen voor hergebruik.

In deze pilot staat interoperabiliteit van verschillende services aangeboden door SURF en deels in-house bij HU centraal. In het onderzoeksproject Gezonde Peutermonden wordt er met veel verschillende externe partijen samengewerkt. Omdat het hier een onderzoek betreft met jonge kinderen ligt de privacy van de data gevoelig. Authenticatie van toegang tot verschillende databronnen is een complex vraagstuk in dit project. De onderzoeker heeft andere rechten dan een individuele zorgverlener (tandarts). Een ander vraagstuk in deze pilot is de koppeling tussen de patiëntgegevens en andere metadata. Voor een efficiënte data-analyse en ook voor hergebruik is goede meta- datering van de data van essentieel belang.

**Verandering** | Op dit moment ontbreekt een structurele implementatie van een Research Data Management (RDM) systeem in dit project en bij de HU. Stapsgewijs sluiten we in de pilot aan op de al bestaande basis SURF infrastructuur. Bovendien is het van belang voor de onderzoeker om op een robuuste en reproduceerbare manier toegang te kunnen krijgen tot (subsets van) de data met behulp van moderne analyse tools zoals R of Python. Dit kan ook worden vereenvoudigd door het inrichten van een

ETL (Extract, Transform & Load) proces. Hiermee kan een onderzoeker een database op reproduceerbare manier bevragen en data direct ophalen in de gewenste analyse- omgeving (VRE).

**Resultaat** | Deze pilot is erop gericht om de data-cyclus binnen het project ‘Gezonde Peutermonden’ te professionaliseren en daarbij de onderzoeker(s) en alle betrokkenen centraal te stellen. Good practices uit research datamanagement, softwareontwikkeling en database ontwerp hebben daarbij een uitgangspunt gevormd voor het ontwerp van de VRE. Dit ontwerp is stapsgewijs geïmplementeerd binnen het programma Digitale Onderzoeksomgeving tijdens het uitvoeren van deze pilot. In eerste plaats is er gekeken naar de bijdrage aan de casus van de pilot en daarna waar mogelijk naar de overige onderzoekers. Zo is er een VRE gerealiseerd die hergebruikt kan worden door andere casussen. Samen met onderzoekers, ondersteuners en andere betrokkenen is er steeds een stukje van de ontwerpschets geconcretiseerd en geïmplementeerd, om kleine werkende producten op te leveren, die geëvalueerd konden worden en meteen een aspect van de casus oplossen.

Veel van de taken rond management van data, opslag, archiveren, regelen van toegang etc. ligt bij de onderzoekers zelf. Hiervoor gebruiken zij op dit moment een veelheid aan verschillende systemen. Dit leidt ertoe dat data verspreid staat, er een schaduwboekhouding ontstaat, fouten sluipen in communicatie en dataverwerking, etc. Dit met alle gevolgen voor het onderzoeksproces en de rapportage daarvan. Onderzoekers zijn hier veel tijd aan kwijt en het lijdt hen onnodig af van hun kerntaken. In deze pilot is een bestaand onderzoeksproject als casus gebruikt om deze processen te harmoniseren, te automatiseren en toekomstbestendig te maken.

## **Resultaten & impact**

### **WEB-APPLICATIE**

Samen met de onderzoeker is gekeken naar alle te verzamelen data, reeds verzamelde data, administratieve documenten en overzichten die gebruikt werden om het onderzoek uit te voeren. Aan de hand hiervan is een webapplicatie gemaakt voor alle professionals die de data moeten gaan verzamelen voor de onderzoekers. Enkele belangrijke punten van deze webapplicatie:

- Onderzoekers kunnen zelf peuters toevoegen aan het systeem, deze gegevens worden automatisch gepseudonimiseerd;
- De webapplicatie, in de vorm van een dashboard, biedt een overzichtelijke weergave per peuter: welke data er wanneer verzameld moet worden;
- Het invoeren van gebitsgegevens is volledig gedigitaliseerd waardoor fouten worden voorkomen en er geen meetgegevens verloren gaan;
- Alle vragen die tijdens een meting beantwoord moeten worden zijn afgestemd met de onderzoeker. Antwoorden worden gestandaardiseerd ingevoerd.

Er is veel mogelijk als je software op maat maakt, processen kun je automatiseren en laten aansluiten op jouw onderzoeksvraag. Daarmee bespaar je tijd gedurende je hele onderzoek. De ‘winst’ opgesomd voor dit project:

- Tijds winst, zowel bij metingen als administratief;
- Overzicht bij alle onderzoekmedewerkers (dashboard);
- Processen binnen het onderzoek worden geautomatiseerd;
- Hogere kwaliteit van je data, compleet, overzichtelijk, daardoor:
- Geen/nauwelijks dataverlies voor je analyse:
  - Alles is ingericht op en gekoppeld aan het nieuwe platform van HU (SURF) Research Cloud:

- AVG proof;
- Alle data wordt veilig gearhiveerd voor analyse;
- Makkelijk een virtuele onderzoeksomgeving in de cloud aan te koppelen, waar alle verzamelde onderzoeksdata geanalyseerd kan worden met Python of R;
- Opslaglocatie HU Research Drive en virtuele onderzoeksomgeving toegankelijk voor alle onderzoekers van het onderzoeksproject

### REPRODUCEERBARE DATA-ANALYSE

Doordat de data op een gecentraliseerde en gestroomlijnde manier verder verzameld kon worden met de webapplicatie zoals hierboven beschreven, kwam er in het project ook de behoefte naar voren om de al verzamelde data conform de nieuw verzamelde data op te schonen en daarmee samen te brengen. Hiervoor zijn we samen met de onderzoeker grondig door de data gegaan en hebben we gezamenlijk een set R scripts en een workflow opgeleverd die het mogelijk maakt de data gestructureerd op te schonen en te fuseren met de nieuw verzamelde data. Deze scripts zijn tot stand gekomen doordat de onderzoeker gebruik kon maken van de combinatie van HU Research Drive, RStudio Server op HU Research Cloud en een HU GitHub Organization. De code is opgesteld onder versie beheer (git) en wordt opgeslagen in een private GitHub-repo. Dit benadert het uitgangsprincipe (infrastructure by code) waar we in deze pilot mee hebben kunnen werken.

### CONCLUSIE

Veel praktijkgericht onderzoek aan onze eigen hogeschool maar ook bij andere hogescholen (en ook universiteiten) hebben in onderzoeksprojecten te maken met de aspecten die we in deze pilot hebben willen oplossen. Deze pilot is daarom dan ook een mooie good practice op zichzelf voor andere instellingen.

De ontwikkelde VRE kan hergebruikt worden door andere onderzoekers, zowel binnen als buiten de HU. De opgedane kennis wordt hiermee gedissemineerd. Er is zoveel mogelijk ingezet op open source ontwikkelingen van toepassingen en de bijbehorende scripts zodat anderen hierop verder kunnen bouwen. Uiteraard zijn hierin de privacy en veiligheid als randvoorwaarden in meegenomen. Scripts, handleidingen, en andere materialen zijn gedeeld in de HU GitHub repository waar mogelijk. Externen kunnen toegang aanvragen. De lessons learned, het ontwerp, de code base en de good practices worden zoveel mogelijk gedeeld in de interne organisatie, volgens het bij jullie wel bekende motto “good practice is to share practice”.

Er is een template gemaakt van de applicatie met CI/CD pipelines voor andere onderzoekers:

<https://github.com/uashogeschoolutrecht/django-app-template>

### Betrokken SURF-diensten

#### Welke SURF-diensten waren onderdeel van de pilot?

Voormalig SURF HPC Cloud, de vervangende dienstverlening HU (SURF) Research Cloud, & HU Research Drive.

#### Hoe zijn deze diensten ingezet of gebruikt binnen de pilot?

SURF Research Cloud is ingezet als OTA omgeving voor het ontwikkelen en hosten van de webapplicatie. Er is gebruik gemaakt van de static ip's om domein namen te koppelen aan de applicatie.

SURF HPC Cloud is ingezet als analyse omgeving voor data: Een dedicated VM (UBUNTU 18.04) met daarop RStudio Server en een koppeling met HU Research Drive, maakte het mogelijk om op een plek de data te beheren en de analyse uit te voeren

HU Research Drive is ingezet als de centrale data store in dit project (*single source of truth*)

### **Betrokkenen en inzet**

#### **Welk lectoraat(en) of kenniscentra waren betrokken bij de pilot?**

Kenniscentrum Gezond en Duurzaam leven met lectoraat Innovaties in de Preventie Zorg waren betrokken bij deze pilot. De ontwikkelde toepassing wordt breed gedeeld binnen de HU onder andere onderzoekers en lectoraten aan de hand van de daarvoor ingerichte kennissessies.

#### **Welke inzet en expertise is vanuit SURF geleverd?**

SURF heeft bijgedragen aan diverse ontwerp sessie en als eerste in kaart gebracht op welke manieren er data werd verzameld binnen het pilotproject. Tevens hebben zij continue als klankbord gefungeerd richting de projectgroep. Ook heeft SURF een platform geboden om de ontwikkelde webapplicatie te tonen aan de buitenwereld.