

## **FAIR DATA: Wereldwijd data delen blijkt zeer relevant in bestrijding coronavirus**

*Eind juli werd in Uganda het eerste COVID-19 FAIR Data Point ingericht. Een doorbraak, aangezien een COVID-19 FAIR Data Point gegevens van coronapatiënten beschikbaar maakt voor zowel mensen als computers. Door middel van COVID FAIR Data Points kunnen gegevens van coronapatiënten anoniem met elkaar vergeleken en gecombineerd worden, waardoor nieuwe wetenschappelijke inzichten kunnen ontstaan. De Nederlandse onderzoeks-financiers ZonMw en de Philips Foundation zien de enorme waarde van deze ontwikkeling en ondersteunen FAIR Data via verschillende projecten.*

### **FAIR data kan ons dus helpen een virus te bestrijden, of een volgende uitbraak voorkomen. Maar hoe werkt dit precies ?**

Terwijl ziekenhuizen wereldwijd dag en nacht bezig zijn om de gevolgen van het SARS-COV 2 virus te beteugelen en onderzoekinstellingen met man en macht proberen de exacte werking van het virus te begrijpen om medicijnen en vaccins te ontwikkelen, is het essentieel dat we grote hoeveelheden betrouwbare patiënt- en virus-gerelateerde onderzoeksgegevens verzamelen, publiceren, delen en analyseren.

Het probleem binnen veel onderzoeksterreinen is dat belangrijke gegevens vaak niet vindbaar, niet toegankelijk, niet interoperabel, dus niet herbruikbaar zijn voor andere onderzoekers. Dit maakt wereldwijde wetenschappelijke samenwerking lastig. Het probleem speelt niet alleen in de medische wetenschap, we zien dit op een groot aantal terreinen. FAIR biedt hier een oplossing voor.

### **FAIR data herbruiken**

Het concept van FAIR is simpel. FAIR staat voor Findable, Accessible, Interoperable en Reusable. FAIR Principles vormen de basis om onderzoekers beter met elkaar te laten samenwerken, wereldwijd en onafhankelijk van het onderzoeksonderwerp. Recente studies hebben aangetoond dat meer dan 80 % van alle gepubliceerde wetenschappelijke gegevens nóóit wordt hergebruikt! Hierdoor is wereldwijd een enorme verspilling op het gebied van publiek en privaat gefinancierd onderzoek ontstaan.

De FAIR principes geven aan dat gegevens/data moet worden verrijkt met metadata, die de data dusdanig goed omschrijft dat deze data door een machine of algoritme volledig en correct geïnterpreteerd en verwerkt kan worden. Op deze manier kunnen grote hoeveelheden data door een computer worden verwerkt.

### **Data kunnen blijven waar ze zijn geproduceerd**

De FAIR Principles geven aan dat het niet persé nodig is om gegevens op een centrale locatie op te slaan. Sterker nog, vaak zien we dat data zo gevoelig of persoonsgebonden is dat gegevens zeer zeker níét op een centrale plek mogen worden opgeslagen. Het FAIR concept van data-visite is dan een mooie, veilige en betere oplossing om veel decentraal opgeslagen data te verwerken. Een slim algoritme wordt losgelaten op diverse geanonimiseerde databronnen die op verschillende instituten in de wereld opgeslagen staan in zogenaamde FAIR DATA POINTS, waarbij de data-eigenaar toestemming heeft gegeven deze data onder duidelijke, van te voren bepaalde voorwaarden te “bezoeken” voor onderzoeksdoeleinden.

Zo kan zeer veel decentraal opgeslagen data in enkele seconden worden verwerkt door computers en kunnen hypothesen en mogelijke correlaties razendsnel en efficiënt worden getoetst.

Essentieel is dat de data op locatie blijft, dat FAIR het data-eigendom volledig bij de databron houdt en daarmee impliciet voldoet aan de Algemene Verordening Gegevensbescherming (AVG). Dit in tegenstelling tot veel andere oplossingen waarbij data vaak wordt overgedragen aan buitenlandse partijen, zonder dat er duidelijkheid is over het eigendom en het gebruik van de overgedragen data.

### **FAIR aanpak heeft zich bewezen voor COVID-onderzoek**

Een pilotproject van een FAIR data-infrastructuur is recentelijk door samenwerking tussen de GO FAIR Foundation, de Philips Foundation en de Nederlandse financier van medisch wetenschappelijk onderzoek ZonMw tot stand gebracht. Dit VODAN-project (Virus Outbreak Data Network) is volledig volgens de FAIR principes geïmplementeerd. Het project stoelt op het wereldwijd gestandaardiseerde Case Report Form (CRF) van de WHO, is gebaseerd op het daarvan afgeleide semantische model, werkt met geanonimiseerde patiënt gegevens gecombineerd met relevante gegevens over het virus en wordt momenteel succesvol uitgerold in ziekenhuizen, instituten en universiteiten in meerdere landen, op meerdere continenten. Voor onderzoekers in de medische wereld een enorme stap voorwaarts omdat hierdoor sneller en efficiënter veel data kan worden verwerkt.

### **Bruikbaar voor de hele wetenschap**

De FAIR infrastructuur is niet uitsluitend opgezet voor het bestrijden van het coronavirus en voor het voorkomen van de uitbraak van een volgend virus. Dezelfde infrastructuur kan ook worden gebruikt bij onderzoek op het gebied van bijvoorbeeld oncologie, cardiologie, dementie of bepaalde zeldzame ziekten. De FAIR principes zijn generiek en algemeen toepasbaar op allerlei situaties waar veel databronnen moeten worden gedeeld, gecombineerd of bezocht. Hierbij kan ook gedacht worden aan data in de wereld van de agrarische sector, de financiële wereld of de energiesector.

FAIR verandert de wereld van ons wetenschappelijk wereldwijde onderzoek fundamenteel en deze ontwikkeling zal leiden tot een hausse aan wetenschappelijke doorbraken.



Meer informatie:

*Drs. Bert Meerman, Director GO FAIR Foundation*

<https://gofairfoundation.org>

*E-mail: [b.meerman@gofairfoundation.org](mailto:b.meerman@gofairfoundation.org)*

*Mobile: +31 6 53661044*