



Praktijktoets FAIR data & Personal Health Train

Eindrapportage Personal Health Train, versie 1.0, 24 juni 2018

Management samenvatting

Dit document is de eindrapportage rondom de Personal Health Train van de praktijktoets FAIR Data & Personal Health Train door het Zorginstituut. In deze praktijktoets hebben we de FAIR principes toegepast op data van het Zorginstituut en is vervolgens de Personal Health Train in een praktijktoets geïmplementeerd. De activiteiten zijn uitgevoerd conform de afspraken tussen VWS en het Zorginstituut. De eindrapportage rondom FAIR Data is een separaat document.

Het belangrijkste concept van de Personal Health Train (PHT) is dat data niet naar de analyse wordt gebracht, maar de analyse naar de data wordt gebracht. De data blijft bij de bron (conform de outcome doelen van het informatieberaad) en het biedt de mogelijkheid om bij de bron complexe (zelflerende) algoritmes uit te voeren. Daarbij geeft het invulling aan het privacy-by-design principe door het mogelijk maken in analyses zelf gedetailleerde gegevens (persoonlijke gegevens) te gebruiken, terwijl in de conclusies die terug gaan naar de afnemer bijvoorbeeld enkel geaggregeerde gegevens zitten. Deze voordelen zagen wij terug in de implementatie. Deze voordelen maken een reeks aan data toepassingen mogelijk die op dit moment niet voor handen zijn.

Een andere belangrijke bevinding is dat wij de PHT in het licht zien van een breder verband van FAIR Data Services. Afhankelijk van de situatie en de organisatie die een station voor de PHT aanbiedt, moet gekeken worden welke vorm van de PHT passend is, mede omdat sommige implementatievormen van de PHT technisch en organisatorisch complexer zijn.

Voor een brede implementatie van de PHT moet er nader onderzoek en afstemming plaatsvinden. Denk hierbij aan vragen rondom techniek, privacy en organisatorische aspecten die eerst beantwoord moeten worden. In kleinere netwerken, waarbij er al een hogere mate van vertrouwen tussen partijen aanwezig is, kan de PHT op kortere termijn al een rol spelen. In de paragraaf Vervolg geven we een aantal suggesties voor vervolgonderzoek.



Inleiding

Van oktober 2017 tot en met mei 2018 heeft het Zorginstituut gewerkt aan een praktijktoets FAIR Data & Personal Health Train. FAIR Data en de Personal Health Train (PHT) zijn begrippen en concepten die in de afgelopen tijd veel aandacht hebben gekregen. Ze zijn ontwikkeld vanuit de academische wereld en worden vooral bekeken in het licht van data voor research & development. Het Zorginstituut heeft met deze praktijktoets de concepten verder onderzocht en toegepast op een meer operationele omgeving binnen de zorg. De praktijktoets is in twee fasen uitgevoerd. In de eerste fase heeft de focus gelegen op het leren toepassen van de FAIR data principes en het verkennen de implementatie van deze principes in de zorg. In de tweede fase is een simulatie van de Personal Health Train gerealiseerd. Deze eindrapportage geeft een overzicht van de resultaten en bevindingen van de tweede fase, de PHT.

De Personal Health Train

De Personal Health Train bouwt voort op de FAIR Data principes. De FAIR data principes beschrijven een nauwkeurige en meetbare set van kwaliteiten die aan iedere goede datapublicatie ten grondslag moeten liggen¹. Deze principes en de bijbehorende uitwerking helpen een antwoord te geven op de toenemende data uitdagingen (zoals de overvloed aan data, diversiteit aan standaarden, interoperabiliteit, maar ook fragmentatie van data) die we kennen. De FAIR data principes zijn algemene principes die van toepassing zijn op goed databeheer. FAIR Data gaat daarmee over de databron. De Personal Health Train is een concept dat gaat over het gebruiken (zoals het analyseren) van deze data.

Het belangrijkste concept van de Personal Health Train is dat data niet naar de analyse wordt gebracht, maar de analyse naar de data wordt gebracht. De huidige manier van data-analyse begint in het algemeen met het verzamelen van de data op één locatie. Hier wordt de data vaak geschoond, getoetst op volledigheid en wanneer de data van verschillende bronnen komt, vertaald naar één definitie. De PHT gaat uit van het brengen van de analyse naar de data. De data blijft bij de bron en de eigenaar van de data stelt deze beschikbaar voor een langskomende 'trein' om te analyseren. Bij het gebruik van verschillende databronnen gaat de 'trein' langs verschillende 'stations' om de data te analyseren. De data blijft tijdens de analyses bij de bron en enkel de conclusies brengt de trein terug naar de analist. Qua privacy is dit een beter concept dan alle data verzamelen op één plek. Daarnaast voorkom je het dupliceren van data met alle risico's rondom actualiteit en volledigheid. Een derde voordeel van het concept is dat de eigenaar van de data veel meer controle houdt over de data omdat deze de eigen regels bepaalt voor het 'station'. En wanneer er voldoende draagvlak is voor het gebruik van de PHT is de gedachte dat er meer data beschikbaar komt voor analyse. Het concept van de PHT is ontwikkeld door het Dutch Techcentre for Life Sciences (DTL), samen met MAASTRO en het LUMC. Meer informatie is te vinden op hun site². Er zijn een aantal onderzoekstrajecten geweest die elke keer deelonderwerpen van het concept hebben onderzocht. Vanuit MAASTRO is onderzoek gedaan naar de validiteit van gedistribueerde zelflerende algoritmen³ en is er een eerste proefopstelling gemaakt voor de PHT⁴. Het Zorginstituut heeft deze ervaringen meegenomen in de implementatie van een simulatie van de PHT in een meer operationele omgeving.

¹ <https://www.nature.com/articles/sdata201618>

² <https://www.dtls.nl/fair-data/personal-health-train/>

³ <https://www.sciencedirect.com/science/article/pii/S2405630816300271>

⁴ <https://www.ncbi.nlm.nih.gov/pubmed/29678027>

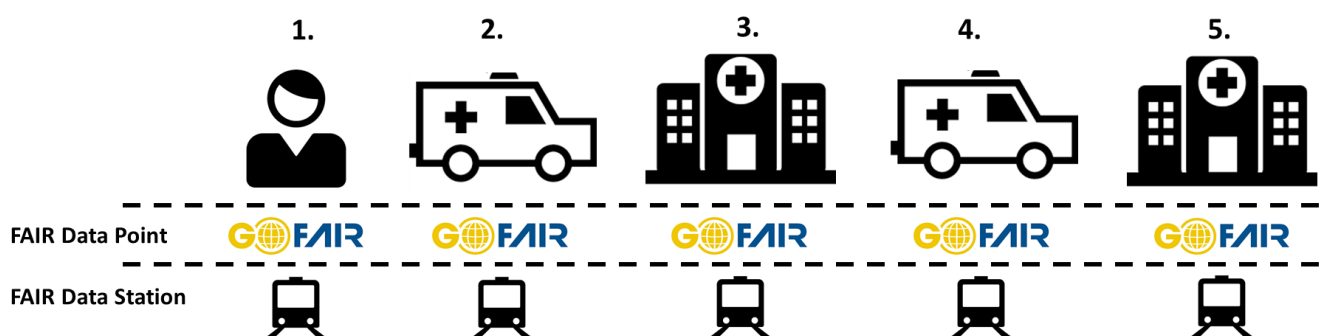
Implementatie Personal Health Train

Voor de praktijktoets PHT heeft het Zorginstituut een omgeving gerealiseerd waarbij er diverse datastations zijn waar een PHT zijn analyses kan doen. Hiervoor is een use case gekozen waarbij de tijdstippen en volgorde van processtappen een belangrijke rol hebben in het leveren van goede zorg.

De Use Case: Intra Arteriële Trombectomie

Intra Arteriële Trombectomie (IAT) is een relatief nieuwe behandeling van een herseninfarct (CVA) waarbij de arts het stolsel in het bloedvat verwijdert middels een katheter die via de lies wordt ingebracht. In de volksmond wordt dit ook dotteren in de hersenen genoemd. Dit is een hoog-complexe behandeling waarbij er in Nederland twaalf centra zijn aangewezen om deze zorg te leveren. De behandeling moet binnen een aantal uren na het ontstaan van het infarct plaatsvinden en de tijd tussen de diverse stappen in het patiënt proces zijn daarmee van cruciaal belang.

De PHT biedt mogelijkheden om informatie achteraf te verzamelen. Een patiënt, wetenschappelijke vereniging, een onderzoeker of een kwaliteitsorganisatie is geïnteresseerd in de doorlooptijden van



Figuur 1: Proces use case IAT

IAT behandelingen. Welke stappen heeft de patiënt doorlopen en hoeveel tijd heeft elk van deze stappen geduurd. Waar in het proces kunnen we verbeteringen doorvoeren? Om deze vraag te kunnen stellen en beantwoorden met behulp van een PHT voorzien we het volgende proces waarbij de elementen van een PHT zijn toegevoegd.

1. Bij de patiënt treedt een herseninfarct op. Hij belt met 112 om met spoed naar het ziekenhuis gebracht te worden.
2. De ambulance (dienst Friesland) haalt de patiënt op en brengt deze naar het ziekenhuis in Leeuwarden. Relevante informatie uit het eigen registratiesysteem publiceert de ambulance dienst in een FAIR Data Point.
3. De artsen in het regionale ziekenhuis van Leeuwarden diagnosticeren dat de patiënt een herseninfarct heeft gehad en hij krijgt intraveneuze trombolysie (IVT). Daarnaast vermoedt het ziekenhuis dat de patiënt in aanmerking komt voor een IAT behandeling en een tweede ambulance wordt geregeld. Ook het ziekenhuis publiceert relevante informatie uit het eigen registratiesysteem in een FAIR Data Point.



4. De tweede ambulance (dienst Groningen) brengt de patiënt van het ziekenhuis naar het UMCG (een IAT behandelend ziekenhuis). De ambulance dienst publiceert de relevante informatie uit het eigen registratiesysteem in een FAIR Data Point.
5. Het UMCG voert de IAT behandeling uit en publiceert enkel zijn eigen relevante informatie in een FAIR Data Point.

Om uiteindelijk informatie uit deze keten te krijgen rondom de doorlooptijden van het proces is voor de praktijktoets gebruik gemaakt van de Personal Health Train. Een trein gaat rijden vanuit de patiënt die deze informatie wil hebben. Deze trein gaat langs de verschillende stations, analyseert de daar aanwezige data, bepaalt zijn route en stuurt uiteindelijk de resultaten van de analyse terug aan de patiënt. In de simulatie gaat het om de indicatoren rondom de doorlooptijd.

Resultaten tweede fase

Voor de tweede fase is de use case uitgewerkt in een simulatie omgeving. Hierbij zijn de volgende onderdelen gerealiseerd.

FAIR Data Point; de plek waar de eigenaar van de data zijn data neerzet. Voor de simulatie zijn er een viertal datapoints gerealiseerd met relevante data; een voor elke organisatie in de use case. Voor elk datapoint zijn een ontologie⁵ en voorbeelddata opgesteld. In het algemeen zullen organisaties een selectie van operationele data beschikbaar stellen in hun data point en deze, conform de FAIR Data principes, voorzien van een ontologie. Deze data wordt dan in het data point geplaatst. In de praktijktoets gebruiken we de open source editie van OpenLink Virtuoso voor het Data Point. Virtuoso is een van de meest gebruikte databases voor de opslag van zogenaamde Linked Data triples (object – predicaat – subject). Deze databases worden Triple-stores of RDF-stores genoemd.

FAIR Data Station; het station dat diensten aanbiedt om bij de data te komen. Bij de implementatie zijn wij tot de conclusie gekomen dat er meerdere typen diensten zijn en daarmee meerdere technische implementaties. Afhankelijk van de gebruikstoepassing zal de dienst meer of minder beveiligd moeten zijn en afhankelijk van de complexiteit van de analyse zal de dienst meer of minder mogelijkheden moeten bieden. Technisch gezien hebben wij drie type diensten gerealiseerd die een station kan aanbieden.

1. SPARQL endpoint voor vragen waarbij privacy minder van belang is (denk aan Open Data), maar wel veel mogelijkheden wordt geboden aan de analist om zelf zijn vragen te definiëren. Deze dienst is relatief eenvoudig om te implementeren en biedt de gelegenheid om nieuwe gebruikstoepassingen voor data te ontwikkelen. In de implementatie van de use case heeft elk station een SPARQL endpoint.
2. RESTful interface om privacy goed te kunnen waarborgen en vragen vooraf zijn gedefinieerd. Deze dienst biedt weinig flexibiliteit voor de analist, is iets complexer om te implementeren dan een SPARQL endpoint, maar geeft veel zekerheid vanuit beveiliging en privacy. In de implementatie van de use case bieden de stations van de ambulance diensten deze interface, naast het SPARQL endpoint.
3. Docker container omgeving waarbij de analist een algoritme in een docker container kan draaien bij de eigenaar van de data. Deze dienst komt het dichtste bij de metafoer van de trein. In de praktijktoets is voor Docker gekozen, als container van de algoritmes, omdat deze

⁵ Een ontologie, binnen het semantische web wordt omschreven als een door computers interpreteerbare beschrijving van de werkelijkheid (kennisrepresentatie).

technologie de software onafhankelijk en autonoom maakt van de omgeving waarin het moet opereren. Er zijn goede waarborgen mogelijk voor privacy én de analist heeft veel mogelijkheden om de analyse vorm te geven. Deze dienst is technisch gezien complexer om te implementeren dan de eerste twee diensten. Ook zijn er afspraken nodig tussen de gebruikers van deze diensten om de juiste waarborgen te creëren op het gebied van privacy, techniek en organisatie. In de implementatie van de use case bieden de stations van beide ziekenhuizen deze interface.

Trein; de trein vormt de logica van de analyse. De vorm van de trein verschilt afhankelijk van het type dienst. Deze kan in de vorm van een query zijn, een aanroep van een interface of een volledig geprogrammeerd (zelflerend) algoritme. Bij de docker container als trein wordt er altijd eerst een verzoek verstuurd naar de data eigenaar om toestemming te krijgen de container ter plaatse de analyse uit te laten voeren.

Rails; de 'rails' zijn de afspraken, waarborgen en interfaces van de PHT. De rails zijn geen fysieke infrastructuren zoals VPN-verbindingen. Organisaties kunnen eenvoudig met bestaande middelen aan de slag en zodoende is er geen behoefte aan centrale infrastructuur. De rails zijn de afspraken over de type diensten die een station aanbiedt, over welke logging en audit trails we met zijn allen inregelen, welke authenticatie we gebruiken, etc. De vorm van deze afspraken kan uiteindelijk een afsprakenstelsel zijn of een vergelijkbare manier om de waarborgen op te stellen en te bewaken. Uiteindelijk is er bij de PHT een bepaald niveau van vertrouwen nodig. Welke waarborgen en criteria je precies nodig hebt om dit vertrouwen te krijgen zal zich moeten uitwijzen. Hier is het naar onze mening nodig om vanuit koplopers op het gebied van de PHT voorbeelden te geven en richtingen te bepalen.







Dashboard; een plek waaruit een monitor de voortgang van de 'trein' kan bijhouden. Vanuit dit dashboard wordt de analyse gestart en worden de resultaten van de analyse getoond. In het onderstaande figuur wordt het dashboard getoond.

The screenshot displays a dashboard with two main panels. The left panel, titled 'Task', contains two dropdown menus: the first is set to 'UMCG' and the second to 'Robert'. Below these is a blue 'RUN TASK' button. The right panel, titled 'Result', shows a list of data points for a patient. The first entry is for 'MCL' (Marlee MacCosty) with details: Burger Service Nummer - 494127328, Naam - Marlee MacCosty, Geboortedatum - 1960-09-04T00:00:00, Patientnummer - 164505, and a timestamp of Wednesday, May 23rd, 2018 13:44:58. The second entry is for 'AZG' (Medical Center Leeuwarden) with details: Aankomstlokatie - Universitair_Medisch_Centrum_Groningen, Aankomsttijd - 2017-02-01T22:17:38, Vertrektijd - 2017-02-01T22:01:36, Vervoer van patient - patient_795838, and a timestamp of Wednesday, May 23rd, 2018 13:44:48. Below the task configuration, there is a 'Live log' section showing a message: 'Result successfully fetch from the ds_umcg data station' with a green checkmark icon. Below the log are 'FINISH' and 'RUN TASK ?' buttons.

Figuur 2: Dashboard praktijkomgeving PHT

Beschrijving Processtappen

Omdat de metafoor van de trein bij de verder detaillering van proces implementatie knelt, laten we deze in dit hoofdstuk los. Het starten van de analyse gebeurt in het dashboard.

-  1. Hiervoor wordt vanuit het dashboard een verzoek naar het UMCG gestuurd om een analyse te draaien⁶. In dit verzoek wordt opgenomen wie de analyse wil draaien en waar de container die deze analyse kan uitvoeren te downloaden is.
-  2. Het UMCG kan (geautomatiseerd) dit verzoek beoordelen en haalt (na goedkeuring) de docker container met de analyse op. In de omgeving van het UMCG wordt de eerste stap van de analyse uitgevoerd. Hierbij krijgt de analyse de behandeltijden (IAT) van het UMCG en de informatie dat de patiënt met de ambulance is binnengebracht.
-  3. Op basis van deze informatie vraagt de container aan het data station van de ambulancedienst Groningen op wat de tijden en de herkomst waren van het vervoer. Op basis van deze informatie stuurt de container een verzoek naar het data station van MCL om een docker container op te halen voor de analyse. In dit verzoek wordt opgenomen wie de analyse wil draaien en waar de container die deze analyse kan uitvoeren te downloaden is.
-  4. Het MCL kan (geautomatiseerd) dit verzoek beoordelen en haalt (na goedkeuring) de docker container met de analyse op. In de omgeving van het MCL wordt wederom een deel van de analyse uitgevoerd. Hierbij krijgt de container de behandeltijden (IVT) van het MCL en de informatie dat de patiënt met de ambulance is binnengebracht.
-  5. Op basis van deze informatie vraagt de container aan het data station van de ambulancedienst Friesland op wat de tijden en de herkomst waren van het vervoer.
-  6. Alle deelresultaten worden nu teruggestuurd naar het dashboard waar de processtappen van de behandeling nu allemaal in kaart zijn gebracht.

Dit proces van de PHT kan naast voor een individuele patiënt ook toegepast worden op een groep patiënten.

Implementatie aspecten PHT

Binnen de PHT staat de rails voor de afspraken die gemaakt moeten worden om de PHT succesvol in een netwerk te implementeren. Daarbij heeft een FAIR Data Service in de vorm van een Docker container voor een organisatie de grootste impact. Een externe partij plaatst software in de vertrouwde omgeving van de organisatie en krijgt toegang en rechten om software uit te voeren op een dataset. Dit vormt een inbreuk op de maatregelen voor gegevensbescherming van de organisatie. Een inbreuk dat alleen door vertrouwen, en door borging van vertrouwen, beheersbaar kan worden gemaakt. De rails zijn daarom van essentieel belang voor het functioneren van de PHT. Afspraken die gemaakt moeten worden gaan bijvoorbeeld over de technische standaarden van een FAIR Data Station. Welke selfservice-interface⁷ pas je toe om treinen uit te rollen en uit te voeren op een FAIR Data Station? In de praktijktoets is hiervoor een taskmanager geïmplementeerd, maar hier bestaan diverse varianten op. Ook moeten er afspraken komen op het terrein van workflow en database-interfaces. Verwacht mag worden dat deze afspraken zullen evolueren tot een volledige

⁶ De analyse had ook gestart kunnen worden vanuit de andere FAIR Data Stations, maar wij hebben in de praktijktoets ervoor gekozen om op de plek waar de patiënt het laatst is geweest te starten met de analyse.

⁷ Interfaces die erop gericht zijn geautoriseerde gebruikers (afnemer van data) zoveel mogelijk geautomatiseerd te ondersteunen.



stack voor het ontwikkelen, uitrollen, uitvoeren en beheren van FAIR Data Stations en bijbehorende FAIR Data Services. Dit betekent een stack zoals deze nu door cloud-leveranciers (bijvoorbeeld Amazon, Google, Microsoft, IBM maar ook diverse kleinere marktpartijen) wordt aangeboden. Deze partijen bieden dan bijvoorbeeld FAIR Data Services aan waarmee individuele zorgaanbieders een FAIR Data Station kunnen realiseren. Eveneens mag verwacht worden dat PHT gekoppeld zal worden aan een prijsmodel om de kosten van processorgebruik en infrastructuur te ondervangen.

Voor kleinschaligere netwerken, die al een grotere mate van vertrouwen kennen, zijn er andere oplossingsvarianten mogelijk. In de meest eenvoudige vorm worden er bilaterale afspraken gemaakt over de uitrol en uitvoer van een PHT in de omgeving van een databron. Bij dit type netwerken kan je bijvoorbeeld denken aan een groep ziekenhuizen, een regionaal verband van verpleeghuizen, een aantal onderzoeksinstituten, een aantal PGO leveranciers of een aantal publieke organisaties. Binnen deze netwerken kan met al beginnen aan het implementeren van een PHT.

Juridische aspecten

Voor het juridische perspectief hebben we samen gewerkt met Pels Rijcken voor een eerste analyse. Daarbij is het belangrijkste vraagstuk bij het werken met persoonsgegevens wie de gegevensverantwoordelijke is. Wanneer de analyse die de PHT uitvoert geen persoonsgegevens gebruikt zijn er geen bijzonderheden, maar wanneer dit wel gebeurt zijn afspraken nodig. In het onderstaande scenario wordt dit toegelicht.

Een onderzoeker wil een analyse bij een ziekenhuis draaien. Hij is daarbij geïnteresseerd in geaggregeerde data, maar heeft tijdens de analyse ook toegang tot persoonsgegevens nodig. Conform de PHT is er met privacy-by-design rekening gehouden en krijgt dus enkel de analyse (container) toegang tot de persoonsgegevens. De uiteindelijke resultaten die de onderzoeker ziet zijn enkel totalen en bevatten geen persoonsgegevens meer. De vraag hierbij is wie conform de wetgeving van de AVG gegevensverantwoordelijke is en wie gegevensverwerker.

Na analyse vanuit juridisch perspectief blijkt echter dat beide partijen in dit scenario een gedeelde gegevensverantwoordelijkheid hebben. Alleen is het niet direct duidelijk hoe je aan de eisen daarvan voldoet als onderzoeker. Je weet namelijk alleen het soort data dat door de analyse is gebruikt, maar niet de daadwerkelijk data. De invulling van bijvoorbeeld een inzage verzoek van een persoon kan een onderzoeker dus enkel in samenwerking met de data eigenaar organiseren. Hiervoor moet je afspraken maken in het stelsel van de PHT. In een apart analyse is een volledig overzicht van de juridische vragen opgenomen⁸.

⁸ Zie Privacyaspecten van Personal Health Train van Pels Rijcken, auteur Nina Bontje



Conclusies

FAIR Data en de PHT kennen een oorsprong in het werkveld van research & development en het Zorginstituut heeft in de praktijktoets gekeken naar de toepasbaarheid in brede zin, zoals meer operationele omgevingen (e.g. procesinformatie, kwaliteitsinformatie). Naar aanleiding van de praktijktoets kunnen we de volgende conclusies trekken.

Algemeen

Ondanks dat de PHT in ontwikkeling minder ver is als FAIR Data is het Zorginstituut overtuigd van de kansen die het biedt voor data-analyse en de benodigde interoperabiliteit die nu niet of moeizaam tot stand komen. Daarbij is van belang om de PHT te zien als onderdeel van FAIR Data Services. Er is namelijk niet één PHT. De PHT kent verschillende verschijningsvormen (diensten) afhankelijk van de kenmerken van de data eigenaar en de complexiteit van de analyse. Een docker container implementeren voor een eenvoudiger vraag is een vorm van 'over-engineering'. Voor complexe algoritmen (denk onder andere aan AI, machine learning, intelligent agents), biedt deze vorm van implementatie echte meerwaarde t.o.v. huidige methodes. Voor minder complexe vragen zijn andere FAIR Data Services (andere verschijningsvormen van de PHT), zoals een SPARQL endpoint of een RESTful interface, zeer bruikbare technieken. Bij deze bredere kijk op FAIR Data Services gaat de metafoer van de trein scheef, maar blijft het concept van de analyse naar de data brengen onverminderd sterk.

Technische achtergrond

Zoals eerder aangegeven is er niet één vorm van de PHT. Er zijn verschillende verschijningsvormen (implementaties) mogelijk. Een SPARQL endpoint als implementatie werkt goed voor bijvoorbeeld Open Data. Een RESTful interface voor een toepassing waarbij input en output van een interface nauwkeurig gedefinieerd (en daarmee geoptimaliseerd) kan worden. De implementatie van de PHT met behulp van docker containers werkt goed voor complexe analyses waarbij de variatie van data (veel diverse locaties of veel diversiteit aan definities) groter is. In de praktijktoets hebben wij met weinig moeite mooie toepassingen kunnen maken die de kracht van deze diensten toont.

De PHT in een Internet of FAIR Data & Services kent geen grote centrale infrastructuur, maar gaat uit van een aantal diensten die het mogelijk te maken data beschikbaar te stellen. Het gaat dan om FAIR Data Points en FAIR Data stations. Dit zijn diensten die organisaties die data delen (en willen gebruiken) realiseren. De 'rails' die dan nodig is om deze diensten met elkaar te laten praten zijn voornamelijk de afspraken die nodig zijn over technische interfaces, implementatie van autorisatie en authenticatie, en waarborgen als audit trails.

Privacy & waarborgen

De PHT kent een aantal aspecten die 'Privacy by Design' mogelijk maken. Data eigenaren blijven in controle van hun data en de waarborgen in het systeem kunnen de balans tussen flexibiliteit en privacy bewaken. Je kan bijvoorbeeld analyses uitvoeren op gecombineerde data die op meerdere plekken is verzameld. Het resultaat van de analyse is echter het enige dat de analist uiteindelijk ontvangt. De achterliggende data blijft bij de eigenaar. Voor deze privacy waarborgen (het vertrouwen) zijn zoals aangegeven afspraken nodig. Dit zit niet enkel in de gebruikte techniek. De



vraag is wel in hoeverre deze afspraken schaalbaar zijn, mede vanwege het potentiële internationale karakter van de PHT.

Transparantie

FAIR data werkt aan transparantie. De principes draaien voor een groot deel om het aangeven welke data je in huis hebt, deze rijkelijk van definities te voorzien en duidelijk te zijn hoe iemand toegang kan krijgen tot deze data. Aan de andere kant is er ook een mate van vertrouwen. De PHT kent een aanvraag mechanisme waarbij de analist aan de data eigenaar toestemming vraagt om een analyse uit te voeren. Daarmee kan je diverse waarborgen voor transparantie introduceren. En wanneer je met de PHT een analyse bij een databron uitvoert is het voor de data eigenaar duidelijk welke data je gebruikt. Daarmee draagt het bij aan een transparanter informatiestelsel.

Implementatie

De implementatie van de PHT heeft een aantal maanden tijd in beslag genomen, met name omdat er een aantal componenten gerealiseerd zijn waarbij er weinig voorbeeld implementaties zijn. Ook zijn er diverse architecturale vraagstukken nog niet beantwoord die wel nodig zijn voor de implementatie van de PHT. Hierbij zijn in de praktijktoets keuzes gemaakt, zoals het gebruik van Docker container of de meta data die bij een aanvraag gebruikt wordt. In de verdere doorontwikkeling van de PHT zullen deze keuzes besproken moeten worden. Daarnaast zal de implementatie veelal in netwerken plaatsvinden, zoals in de use case in een netwerk van ketenzorg. Bij deze implementaties zal altijd afstemming nodig zijn.

FAIR Data Services

Zoals aangegeven gaat de metafoer van de trein bij een nadere uitwerking van het concept verloren. Er komt een moment dat een toepassing van de PHT als intelligent agents zou leiden tot een metafoer van stilstaande treinen. Daarnaast biedt de PHT veel mogelijkheden voor PGO's om data analyse uit te voeren en om patiënt data beschikbaar te maken voor bijvoorbeeld onderzoek, maar kan een toepassing van de PHT ook niet op 'Personal' niveau liggen. Gedurende de praktijktoets zijn we daarom steeds meer de term FAIR Data Services gaan gebruiken. Dit geeft een breder beeld van de potentie die FAIR Data en Services biedt voor een netwerk van organisaties die door ontologieën verbonden zijn. Een netwerk waarin microservices aangeboden worden om interpretaties mogelijk te maken er nieuwe waardevolle diensten gerealiseerd kunnen worden. Het is dus belangrijk om de PHT te zien in het grotere geheel van FAIR Data Services.

Vervolg

De praktijktoets heeft inzicht gegeven in de werking en implementatie van de PHT. Zoals aangegeven staat het Zorginstituut positief tegenover deze ontwikkelingen. Er zijn vooral bij de implementatie van FAIR Data Services (PHT) wel nog verschillende aspecten die nog niet volledig zijn uitgewerkt. Het gaat dan om de schaalbaarheid van de PHT en diverse vragen op architectureel en juridisch gebied. Deze zijn hieronder benoemd.



- De haalbaarheid in de praktijk en het toetsen van de operationele inzetbaarheid. Vanuit de praktijktoets hebben we vertrouwen dat het operationeel inzetbaar is, mede omdat de PHT veelal gebruik maakt van bestaande technieken. De toetsing hiervan kan door de implementatie van de PHT op te nemen in het simulatiemodel van het actieprogramma iWlz en daarmee te operationaliseren.
- Openstaande vragen op het juridische gebied met als hoofdvraag wie gegevensverwerker is en wie gegevensverantwoordelijke in de situaties van de PHT. Dit geeft mede input voor de waarborgen (afspraken) die bij de PHT nodig zijn. Een eerste verkenning hiervan is gemaakt samen met juristen van Pels Rijcken.
- Openstaande technische vragen die zich met name richten op de beveiliging van data binnen de diverse diensten van een data station. Een voorbeeld is de mate waarin RDF stores autorisatie kunnen regelen op elementair niveau van de data. De praktijktoets geeft voldoende vertrouwen dat dit kan, maar de beste manier om dit technisch te realiseren is een onderdeel van vervolg onderzoek.
- Het komen tot de benodigde afspraken om een Internet of FAIR Data & Services binnen de zorg tot stand te laten komen. De rails (set aan afspraken) van de PHT dient zodanig vorm te krijgen dat meerdere partijen de PHT kunnen implementeren. Hierbij kan de overheid een coördinerende rol oppakken totdat de denkwijze en technologie een zodanig gemeengoed zijn geworden dat veldpartijen het zelf oppakken.