



Set van afspraken Personal Health Train Versie 0.5 –22 november 2021

*De **Personal Health Train (PHT)** is een verzamelbegrip voor technologieën en afspraken die decentrale data-analyse mogelijk maken. Deze voldoen daarbij aan de kaders van de PHT set van afspraken.*

Inhoudsopgave

1	Set van afspraken	5
1.1	Onderdelen	5
1.2	Achtergrond PHT	6
1.2.1	Definitie PHT concept	6
1.2.2	Doel van de PHT	7
2	Releasebeheer	8
3	Grondslagen	8
3.1	Achtergrond set van afspraken	8
3.1.1	Waarom een set van afspraken	8
3.1.2	Doelgroep beschrijving	10
3.1.3	Scope	10
3.2	Criteria	11
3.2.1	Doelen	11
3.2.2	Randvoorwaarden	11
3.3	Principes	12
3.3.1	Toelichting principes:	12
3.4	Opzet	15
3.5	Begrippenlijst	16
3.5.1	Definitie PHT en trein-metafoor	16
3.5.2	Trein	16
3.5.3	Spoor	16
3.5.4	Data-Station	16
3.5.5	Wat is een PHT-netwerk?	17
3.5.6	Wat is een 'PHT-dienstverlener' of 'PHT-leverancier'?	17
3.5.7	Waarvoor kan de PHT ingezet worden?	17
3.5.8	Kan de burger ook gebruiker van de PHT zijn?	17
4	Juridisch kader	18
4.1	Organisatie van het PHT-stelsel	18
4.2	Juridisch inhoudelijke ontwikkelingen	18
5	Normenkader	19
5.1	Normen	19
5.2	Toelichting normen	19
5.2.1	Duidelijk wie partijen zijn en wie toegang tot een PHT-netwerk heeft	19
5.2.2	Transparantie over kenmerken van datastation en treinen	19

5.2.3	Openheid over de berekeningen die op de data worden uitgevoerd en hoe de uitkomsten gebruikt gaan worden.....	20
5.2.4	Duidelijkheid over hoe een nieuwe vraag in het netwerk wordt beoordeeld en goedgekeurd	20
5.2.5	Data minimalisatie en doelbinding	20
5.2.6	Duidelijkheid over rollen in het kader van AVG	21
5.2.7	Afspraken over metadata en betekenis van data in een netwerk.....	21
5.2.8	Deelnemers in een PHT-netwerk maken gebruik van stations en treinen die het PHT-label mogen dragen	21
5.2.9	De data in een datastation wordt goed beheerd.....	21
5.3	Toetsing normen.....	22
5.4	Gebruik beeldmerk	22
5.4.1	Gebruik beeldmerk door netwerk	22
5.4.2	Gebruik beeldmerk door leverancier PHT-diensten	22
6	Technical Architecture and Agreements.....	23
6.1	Types of PHT-networks	23
6.1.1	Business roles and responsibilities.....	24
6.1.2	Interoperability Governance of the rail network.....	24
6.1.3	Operating trains on the railway	26
6.2	Managing semantic and technical interoperability in the rail network.....	26
6.2.1	Semantic Interoperability	27
6.2.2	Technical Interoperability	28
6.2.3	The capabilities of a station	29
6.2.4	The capabilities of a train.....	30
7	Beleid	31
7.1	Rollen binnen de PHT-governance.....	31
7.1.1	Bestuur	31
7.1.2	Management.....	32
7.1.3	Marktontwikkeling.....	32
7.1.4	Scientific community.....	32
7.1.5	Learning community	32
7.1.6	Formeel orgaan namens aanbieders van PHT-diensten, onderzoekers en gebruikers	32
7.1.7	Uitvoering (technische) afspraken	33
7.1.8	Certificering.....	33
8	Operationele processen.....	33
8.1	Operationele processen rondom ontwikkeling afsprakenet.....	33

8.1.1	Procesafspraken.....	33
8.1.2	Releases	33
8.2	Operationele processen rondom beeldmerk.....	33
8.3	Operationele processen rondom communicatie	34
8.4	Sturingsstructuur	34
9	Communicatie	34
Bijlage I – certificering PHT-netwerk.....		35
Checklist: voldoet ons netwerk aan de normen gesteld in het normenkader?.....		35

1 Set van afspraken

De Personal Health Train (PHT) is een verzamelbegrip van technologieën en afspraken die decentrale data-analyse mogelijk maken. De partijen in de PHT community werken in gezamenlijkheid aan een set van afspraken om gebruik te kunnen maken van decentrale data-analyse op basis van data aan de bron bij diverse organisaties.

1.1 Onderdelen

In deze set van afspraken worden afspraken gemaakt over inhoud, besturing, onderhoud en communicatie rondom de PHT.

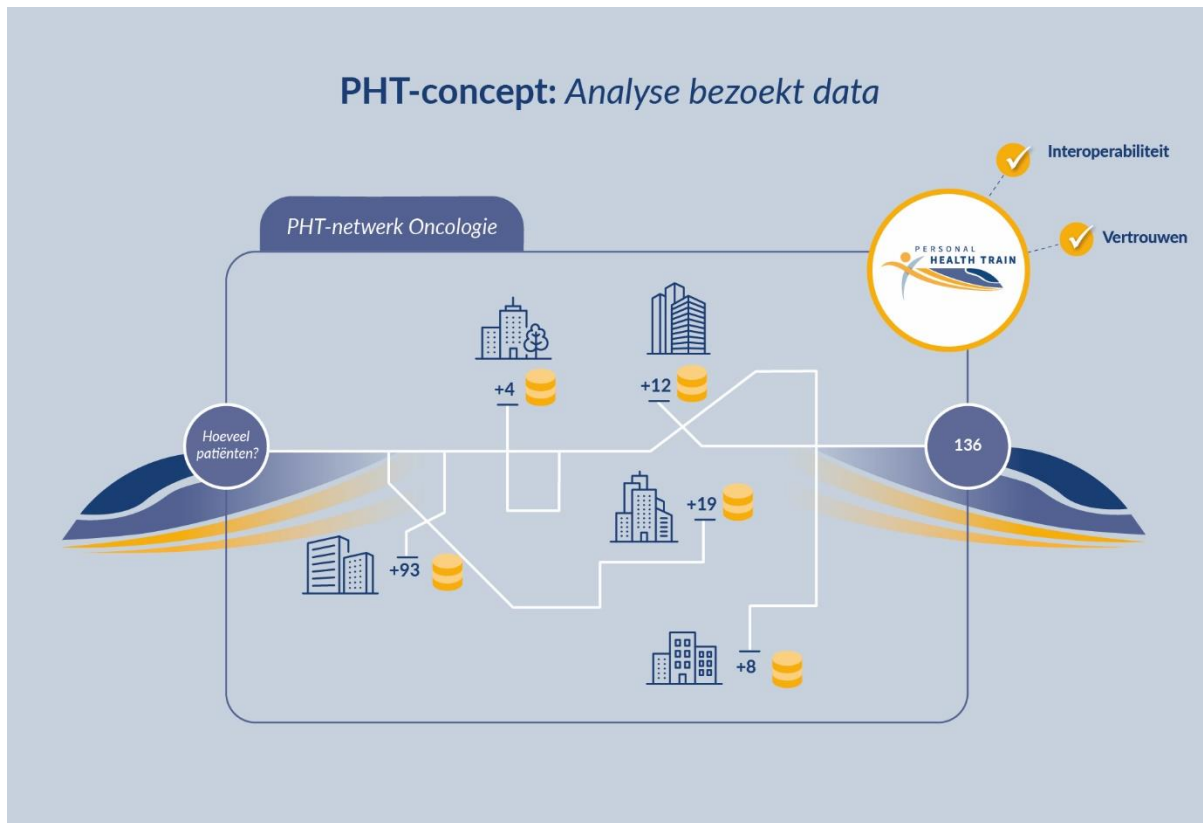
Het document beschrijft een samenhangende set van afspraken die bestaat uit de volgende onderdelen:

- Releasebeheer: in dit hoofdstuk wordt het releasebeheer van de set van afspraken beschreven. Het gaat dan om het bepalen van de huidige versie, hoe het wijzigingenbeheer verloopt en welke wijzigingen er per versie zijn doorgevoerd.
- Grondslagen: de grondslagen beschrijven het fundament van de set van afspraken. Dit hoofdstuk beschrijft waarom deze set van afspraken van belang is, de doelen die wij nastreven met de set van afspraken, de richtinggevende afspraken die we gezamenlijk maken, een beschrijving van de rollen en domeinen binnen de PHT en een beschrijving van de belangrijkste begrippen.
- Juridisch kader: dit hoofdstuk bevat een analyse van relevante wet- en regelgeving, en een aantal modelovereenkomsten.
- Normenkader: het normenkader beschrijft welke normen er vanuit de set van afspraken van de deelnemers worden verwacht. Daarnaast wordt ook ingegaan op eventuele certificering van deze afspraken.
- Gegevensuitwisseling: alle technische afspraken die noodzakelijk zijn om conform de PHT een netwerk te implementeren worden hier beschreven.
- Beleid: dit hoofdstuk gaat in op de vraag 'hoe om te gaan met besturingsvraagstukken'.
- Operationele processen: in dit hoofdstuk wordt besproken welke processen er nodig zijn voor een goede werking van de set van afspraken. Het gaat hier bijvoorbeeld om een beschrijving hoe een organisatie formeel deel kan gaan nemen aan de set van afspraken.
- Communicatie: in dit hoofdstuk beschrijft afspraken rondom communicatie over de set van afspraken en bijvoorbeeld het gebruik van logo's.

1.2 Achtergrond PHT

1.2.1 Definitie PHT concept

De Personal Health Train is een verzamelbegrip van technologieën en afspraken die decentrale data-analyse mogelijk maken. Het belangrijkste uitgangspunt van de PHT is dat de data niet naar de analyse gebracht wordt, maar de analyse aan de data wordt aangeboden. De PHT is een metafoor voor een trein die langs verschillende stations (databronnen) rijdt, daar algoritmes (analyses) uitvoert en de conclusies uit de analyses meeneemt. Deze manier van werken geeft invulling aan het *privacy-by-design* principe. Dit betekent dat in de analyse gedetailleerde gegevens gebruikt kunnen worden, terwijl in de conclusies enkel geaggregeerde gegevens worden getoond. Voor beheerders van data is het een effectieve manier om controle te houden over het gebruik van de beschikbare data. Voor gebruikers van data maakt de PHT het mogelijk om grote hoeveelheden data vanuit meerdere bronnen te analyseren, verbanden te leggen en complexe algoritmen uit te voeren.



Afbeelding 1.1. Het PHT-concept: De analyse bezoekt de data.

1.2.2 Doel van de PHT

De PHT kan voor verschillende doeleinden worden toegepast. Daarbij kan onder andere gedacht worden aan de inzet van de PHT voor medisch-wetenschappelijk onderzoek, de inzet van de PHT voor analyse voor individuele behandelingen of de inzet van de PHT voor analyse van kwaliteitsgegevens. De PHT geeft invulling aan verschillende vraagstukken en ontwikkelingen zoals;

- de maatschappelijke ontwikkeling van meer datagericht werken;
- een manier om kunstmatige intelligentie toe te passen in de zorg;
- de toenemende wens om data uit het zorgproces te hergebruiken voor onderzoek en analyse;
- de mogelijkheid om grote hoeveelheden aan data te gebruiken voor gepersonaliseerd advies;
- de ontwikkeling van technologieën als federated data, distributed learning en secure multi party computation;
- privacy-by-design principes 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;
- de ontwikkelingen om data machine leesbaar beschikbaar te hebben voor analyses.

Het PHT-concept wordt momenteel in verschillende netwerken, met verschillende doelen en op verschillende manieren toegepast (e.g. INKL, Health-RI, CoViDa en Zorginstituut Nederland). Veelal zijn deze gebaseerd op dezelfde basisprincipes, maar er worden ook contextafhankelijke afspraken en implementaties toegepast. Het is daarom wenselijk om gezamenlijk afspraken te maken over bijvoorbeeld ontwerp, toepassing en beleid rondom de PHT. De set van afspraken zorgt ervoor dat de deelnemende partijen op dezelfde wijze de PHT toepassen. Dit zorgt ervoor dat het gebruik van de PHT meer gemeengoed wordt en het eenvoudiger is om gebruik te maken van verschillende PHT-netwerken.

De afspraken in deze set van afspraken behelzen alle lagen van het European Interoperability Framework (EIF)¹ waar ook het interoperabiliteitsmodel van Nictiz (vijf lagen model) op gebaseerd is. Met deze set van afspraken wordt de diversiteit aan deelnemende partijen benadrukt. Deze partijen komen primair uit de Nederlandse zorg, maar de afspraken richten zich ook op interoperabiliteit in internationaal verband (e.g. bestaande internationale netwerken met Duitsland, Denemarken, China) en tussen sectoren (e.g. samenwerking met de Farm Data Train).

¹ https://ec.europa.eu/isa2/eif_en

2 Releasebeheer

Releasebeheer: Beschrijving van het releasebeheer van de set van afspraken. Het gaat dan om het bepalen van de huidige versie, hoe het wijzigingenbeheer verloopt en welke wijzigingen er per versie zijn doorgevoerd. Ook zal hier worden opgenomen of er meerdere versies tegelijkertijd actief kunnen zijn en hoe hier in de praktijk mee omgegaan wordt.

Hierin wordt ook opgenomen hoe het proces om te komen tot een nieuwe release verloopt (bv half open proces), wie daar bij betrokken is en hoe besluitvorming plaatsvindt.

Dit hoofdstuk wordt aangevuld.

3 Grondslagen

De grondslagen beschrijven het fundament van de set van afspraken. De grondslagen bestaan uit:

- **Achtergrond:** de achtergrond beschrijft waarom deze set van afspraken van belang is, wat de doelstelling is van de set van afspraken en op welk domein deze set van afspraken van toepassing is.
- **Criteria:** de doelen die wij nastreven met de set van afspraken en die helpen bij het toetsen van de diverse deelproducten aan deze doelen en eventuele randvoorwaarden.
- **Principes:** de richtinggevende afspraken over ontwerpkeuzes van de set van afspraken.
- **Opzet:** een beschrijving van de rollen en domeinen binnen de scope van de PHT set van afspraken.
- **Begrippenlijst:** een beschrijving van de belangrijkste begrippen voor de set van afspraken.

3.1 Achtergrond set van afspraken

3.1.1 Waarom een set van afspraken

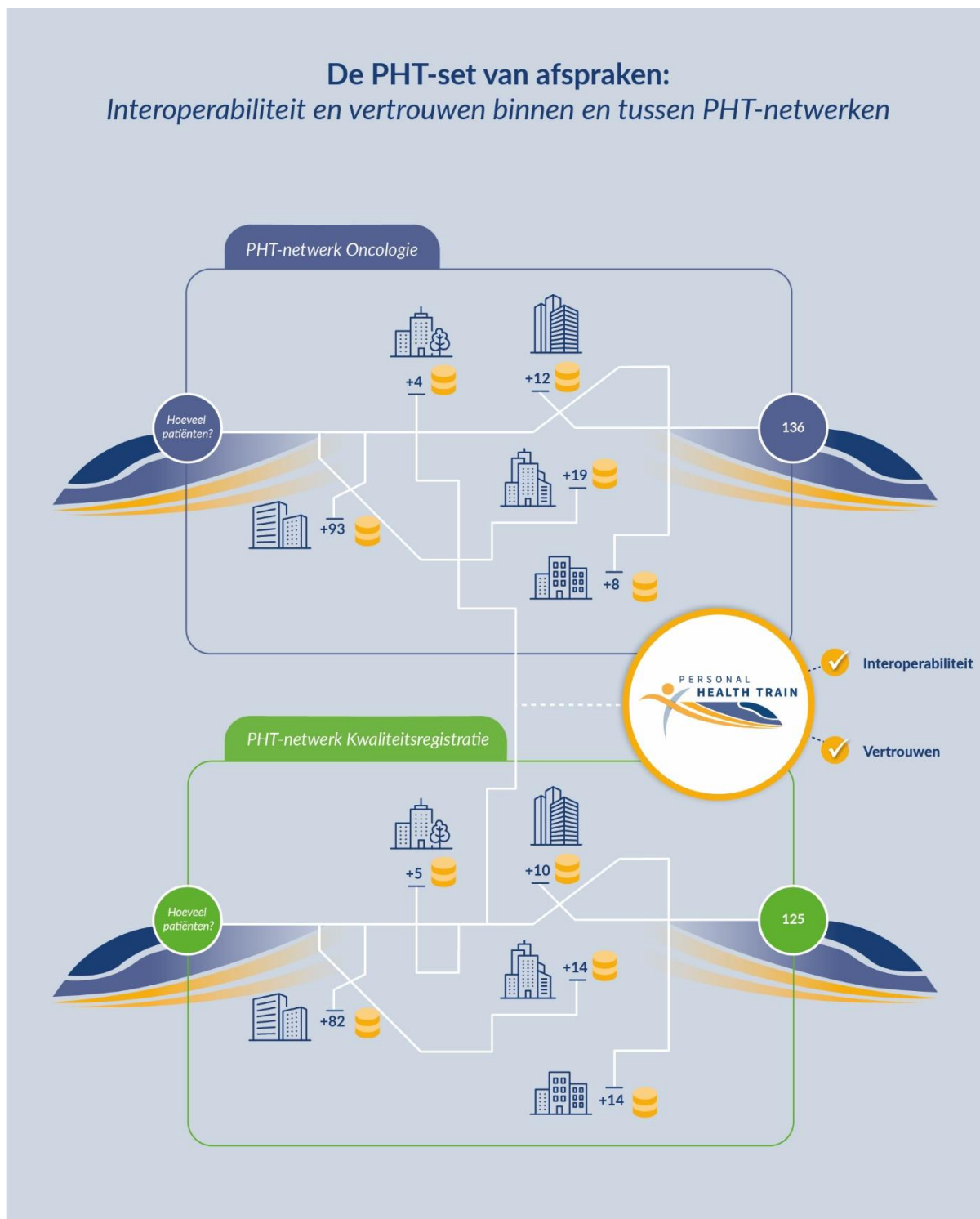
De PHT community werkt aan een set aan afspraken om het vertrouwen in en interoperabiliteit van decentrale gegevensanalyse te vergroten. Deze draagt bij aan de visie van de PHT community.

Deze afspraken hebben o.a. als doel om interoperabiliteit van de individuele PHT-implementaties in de verschillende netwerken te borgen. Deze interoperabiliteit is een essentieel en intrinsiek onderdeel van het PHT-concept. Een ander doel is het duidelijk benoemen van de randvoorwaarden voor de implementatie van een PHT netwerk.

Een gemeenschappelijke aanpak vanuit de PHT community is essentieel, omdat de mogelijkheden voor een goede decentrale analyse niet door één van de deelnemende partijen kan worden vormgegeven. Men heeft elkaar nodig. Kort samengevat betreft het afspraken om het vertrouwen tussen partijen met een nieuwe manier van werken te faciliteren, evenals de onderlinge interoperabiliteit (zowel technisch als semantisch) te bewerkstelligen.

De set van afspraken bestaat in de kern uit het in staat stellen van een netwerk van organisaties om het onderling toepassen van decentrale gegevensanalyse op basis van data aan de bron mogelijk te maken.

De PHT-set van afspraken: Interoperabiliteit en vertrouwen binnen en tussen PHT-netwerken



Afbeelding 1.2. De PHT-set van afspraken: Interoperabiliteit en vertrouwen binnen en tussen PHT-netwerken.

3.1.2 Doelgroep beschrijving

De set van afspraken richt zich op de leveranciers van PHT-diensten (treinen, stations en rails) en de organisaties in een PHT-netwerk.

3.1.3 Scope

De belangrijkste afbakening van het gebruik van de PHT is hieronder benoemd.

1. De PHT wordt ingezet in het domein van zorg en gezondheid. Het concept van de PHT wordt toegepast in meerdere domeinen (zoals bijvoorbeeld de biologie en landbouw), maar de scope van de set van afspraken is zorg en gezondheid.
2. De PHT is bedoeld voor secundair datagebruik, zoals het gebruik van data voor onderzoek, statistiek, beleid en analyse. De PHT richt zich dus niet op operationeel dataverkeer, zoals het delen van een laboratoriumuitslag tussen een ziekenhuis en een patiënt of het overdragen van een dossier tussen een ziekenhuis en een verpleeghuis.
3. De PHT gaat niet in op persoonlijke gezondheidsomgevingen (PGO), maar biedt wel de mogelijkheden om als burger data te gebruiken (een trein rond te sturen) of voor een PGO omgeving om als data station te opereren.
4. De PHT beschrijft de afspraken die nodig zijn om partijen gebruik te laten maken van data. Dit kan zijn voor zelflerende algoritmen of andere vormen van kunstmatige intelligentie (AI). De set van afspraken gaat niet in op de vraag of de inhoud van een trein (het algoritme) wenselijk is of een correcte vorm van AI is. Wel gaat de set van afspraken in op welke randvoorwaarden er zijn om een trein te kunnen vertrouwen.

3.2 Criteria

Criteria kunnen bestaan uit doelen (factoren waarbij gestreefd wordt naar een zo hoog mogelijke score) en randvoorwaarden (niet-onderhandelbare factoren). De doelen vormen een toetsingskader in de uitwerking van de afspraken in de diverse hoofdstukken.

3.2.1 Doelen

1. De afspraken maken het mogelijk dat afnemers en bronhouders gegevens in onderling vertrouwen kunnen uitwisselen. Er is transparantie over doel van de analyse, het gebruik van de gegevens en wat er wel of niet mee kan of mag. De afspraken creëren vertrouwen bij deelnemers.
2. De afspraken stellen de partijen in staat om gezamenlijk een PHT-netwerk in te richten, met duidelijke voorwaarden en in onderling vertrouwen. Het volgen van de afspraken zorgt voor een gemeenschappelijke aanpak, wat essentieel is voor een goede decentrale analyse.
3. Met het naleven van de afspraken wordt een netwerk van organisaties in staat gesteld om op basis van informatie aan de bron een decentrale gegevensanalyse toe te passen.
4. De afspraken zorgen ervoor dat verschillende PHT-netwerken zo veel mogelijk met elkaar kunnen interopereren door transparantie in gemaakte keuzes.
5. De afspraken borgen vertrouwelijkheid van gegevens door *privacy by design*
6. De afspraken bieden toekomstvastheid van de PHT-oplossing, waarbij de afspraken zich richten op:
 - Strategische flexibiliteit voor de uitwisseling met nieuwe domeinen
 - Strategische flexibiliteit voor het gebruik van diverse informatiemodellen
 - Schaalbaarheid van de oplossing
7. Binnen de afspraken wordt een markt gecreëerd voor PHT-oplossingen met de juiste innovatie- en kwaliteitsprikkel en voldoende keuzemogelijkheden, daarbij richten de afspraken zich op:
 - Reële marktwerking voor dienstverlening op het gebied van datastations
 - Reële marktwerking voor dienstverlening op het gebied van treinen
 - Duidelijkheid over data gedreven businessmodellen

3.2.2 Randvoorwaarden

De uiteindelijke set van afspraken zal op elk moment in lijn moeten zijn met de Nederlandse wet- en regelgeving. Daarom moet de set van afspraken zo zijn opgezet dat partijen die betrokken zijn bij de uitvoering ervan in staat worden gesteld te voldoen aan deze wet- en regelgeving (AVG, gezondheidsrechtelijke wetgeving, Mededingingswetgeving). Dit betekent vooral dat een goede uitvoering van de afspraken niet mag vereisen dat partijen afwijken van wet- en regelgeving. Belangrijke juridische kaders, zoals de AVG, zullen voor dat doel getoetst en toegelicht worden.

3.3 Principes

Principes zijn richtinggevende afspraken over ontwerpkeuzes van de set van afspraken. De principes hebben een algemeen karakter en waar afspraken worden gemaakt over specifieke onderwerpen is dit opgenomen in het bijbehorende hoofdstuk.

- P1.** De outcome doelen van het informatieberaad worden gevolgd.
- P2.** De keuzes in de set van afspraken conformeren aan het Duurzaam Informatiestelsel Zorg Referentie Architectuur ([DIZRA](#)).
- P3.** De gegevens die gebruikt worden binnen de PHT-netwerken voldoen aan de FAIR data principes over vindbaarheid, toegankelijkheid, uitwisselbaarheid en herbruikbaarheid.
- P4.** De inrichting en het gebruik van de PHT-netwerken gebeurt conform de FAIR Referentiearchitectuur.
- P5.** Er kunnen meerdere PHT-netwerken zijn.
- P6.** De gebruiker van PHT-diensten heeft ten alle tijden een vrije keuze voor de leverancier.
- P7.** Data wordt gehaald uit operationele processen.
- P8.** Data is leesbaar voor machines.
- P9.** Data blijft bij de bron.
- P10.** De afspraken zijn niet vrijblijvend voor de deelnemende partijen.
- P11.** Centraal vertrouwen, decentrale operatie.
- P12.** Transparant waar dat kan, besloten waar het moet.

3.3.1 Toelichting principes:

P1 De outcomedoelen van het informatieberaad worden gevolgd

Het Informatieberaad Zorg (IB) is een bestuurlijke samenwerking tussen deelnemers uit het zorgveld en VWS. De set van afspraken draagt bij aan een duurzaam informatiestelsel en houdt rekening met de [outcomedoelen van het IB](#).

P2 De keuzes in de set van afspraken conformeren aan het Duurzaam Informatiestelsel Zorg Referentie Architectuur (DIZRA)

De [referentiearchitectuur DIZRA](#) staat voor een Duurzaam Informatiestelsel in de Zorg Referentiearchitectuur. Deze referentiearchitectuur geeft kaders mee voor de afsprakenstelsels voor gegevensuitwisseling. De deelnemers aan de PHT-set van afspraken conformeren zich aan de kaders van deze referentiearchitectuur.

P3 De gegevens die gebruikt worden binnen de PHT-netwerken voldoen aan de FAIR data principes over vindbaarheid, toegankelijkheid, uitwisselbaarheid en herbruikbaarheid

De internationale FAIR principes zijn richtlijnen voor gegevenshuishouding. De deelnemer aan de set van afspraken conformeert zich aan de richtlijnen zoals beschreven in de FAIR principes. Daarmee voldoen de gegevens die gebruikt worden aan de FAIR-dataprincipes over vindbaarheid, toegankelijkheid, uitwisselbaarheid en herbruikbaarheid.

P4 De inrichting en het gebruik van de PHT-netwerken gebeurt conform de FAIR Referentiearchitectuur

De FAIR referentiearchitectuur is ontwikkeld voor de PHT, en daarmee kader gevend voor elk netwerk wat in het kader van dit afsprakenstelsel ontstaat. Bij de ontwikkeling, inrichting en het gebruik van de PHT-netwerken wordt rekening gehouden met de afspraken en kaders die gesteld zijn binnen de FAIR Referentiearchitectuur. Dit zorgt ervoor dat de ontstane netwerken interoperabel blijven met FAIR Data Train netwerken in andere domeinen dan de zorg.

P5 Er kunnen meerdere PHT-netwerken zijn

De deelnemers aan de set van afspraken vormen niet gezamenlijk één groot PHT-netwerk, maar er zullen verschillende PHT-netwerken ontstaan. Deze netwerken kunnen van elkaar verschillen. Zij ontstaan vaak naar aanleiding van een bepaald doel. Bijvoorbeeld data-ondersteunde kwaliteitsverbetering of data-ondersteunde keuzehulp bij ingewikkelde aandoeningen. Daarom kunnen de netwerken verschillen in bijvoorbeeld de toegepaste technische implementatie voor decentrale data-analyse.

P6 Vrije keuze in leverancier van PHT-diensten

Om te voorkomen dat de zorgleverancier afhankelijk wordt van de leverancier van PHT-diensten (ontstaan van vendor lock-in) zullen deze diensten door meerdere partijen worden aangeboden. Toetreden tot de markt van PHT-diensten is voor iedere leverancier mogelijk.

P7 Data wordt gehaald uit operationele processen

De data wordt zo veel als mogelijk opgehaald uit het operationele (zorg) proces, waarbij registratie aan de bron het uitgangspunt is.

P8 Data is leesbaar voor machines

Het datastation moet de data leesbaar aanbieden aan de PHT-trein, de eigenaar van het station en/of de eigenaar van de data moet zorgen dat de data in het datastation leesbaar is voor deze toepassing. De technische afspraken voor de aanlevering van data zijn primair gericht op machine-naar-machine-uitwisseling. Daarnaast moeten de uiteindelijk uitgewisselde gegevens (ook) wel leesbaar zijn voor mensen.

P9 Data blijft bij de bron

De analyse die uitgevoerd wordt over de data binnen het datastation wordt bij de (zorg)leverancier opgehaald en geanalyseerd. De data zal daarbij niet worden getransporteerd, maar bij de bron blijven en bij de bron worden geanalyseerd.

P10 De afspraken zijn niet vrijblijvend voor de deelnemende partijen

Elke deelnemende partij conformeert zich aan de afspraken zoals opgenomen in de laatste versie van de set van afspraken.

P11 Centraal vertrouwen en decentrale operatie

Er zijn zo min mogelijk centrale componenten, er wordt zover mogelijk geen centrale infrastructuur ingericht. Binnen de PHT-netwerken zijn zo veel mogelijk componenten decentraal opgezet en ingericht. Dit komt het beheer van de individuele netwerken ten goede en bevordert de zelfredzaamheid van de netwerken.

P12 Transparant waar dat kan, besloten waar het moet

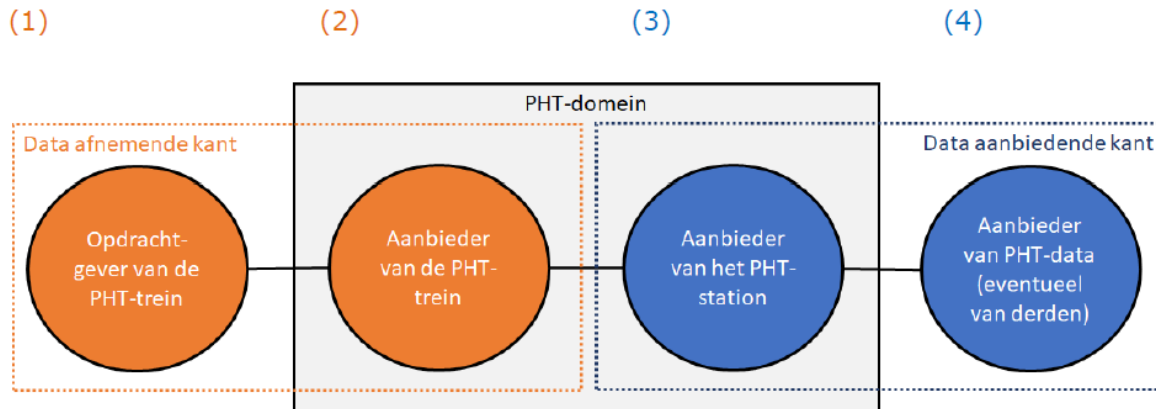
Om in vertrouwen met elkaar de PHT te kunnen gebruiken, is een grote mate van transparantie nodig. Binnen een PHT-netwerk moet helder zijn welke analyse de datastations uitvoeren, zodat het voor alle deelnemers helder is op welke wijze de data wordt gebruikt. Daar tegenover staat dat de data die gebruikt wordt gepseudonimiseerd is, dit vindt al plaats bij de data-aanbiedende kant binnen het datastation, zodat voor de data-afnemende kant persoonsgegevens zijn afgeschermd.

Principes die we zijn tegengekomen in andere set van afspraken, maar er nu voor gekozen hebben deze niet op te nemen:

- Ontwikkeling geschiedt in een half-open proces met verschillende stakeholders (het afsprakenstelsel wordt (door)ontwikkeld in samenspraak met de belangrijkste stakeholders, waaronder vertegenwoordigers van de deelnemers, de gebruikers en partijen met een belang bij het functioneren van het stelsel)
- De PHT set van afspraken spreekt alleen af wat nodig is (onderwerpen die al geregeld zijn in wet- en regelgeving of de facto technisch geen barrière vormen, worden niet opgenomen in het afsprakenstelsel)
- De afspraken zijn generiek waar mogelijk en specifiek waar nodig.

3.4 Opzet

De opzet beschrijft op het hoogst mogelijke niveau welke rollen en domeinen er binnen de toepassing van de PHT actief zijn.



Afbeelding 3.1. De PHT-markt en de PHT-actoren.

Bij toepassing van de PHT zijn in de regel vier actoren actief. Aan de data-afnemende kant zijn dat (1) de opdrachtgever van de PHT-trein en (2) de aanbieder van de PHT. Aan de data-aanbiedende kant zijn dat (3) de aanbieder van het PHT-station en (4) de aanbieder van de PHT-data.

- De opdrachtgever van de PHT (1) kan bijvoorbeeld een medisch onderzoeker zijn die gegevens nodig heeft voor het doen van onderzoek, een overheidsorganisatie dat gegevens nodig heeft voor beleidsvorming, een arts of patiëntenvereniging die de beste behandeling wil onderzoeken voor een aandoening of een patiënt die informatie wil over haar/zijn behandelingen (hierna: de opdrachtgever).
- De aanbieder van de PHT-trein (2) is (de partij die) het algoritme (inzet) dat de data-analyse uitvoert. Dit is de trein (hierna: het algoritme). De opdrachtgever (1) kan het algoritme/de trein zelf inzetten of dat laten doen door een andere partij. De aanbieder van de PHT levert diensten op het gebied van het ontwikkelen, laten certificeren en uitsturen van algoritmen, logging en het verzamelen en toonbaar maken van conclusies.
- De aanbieder van het PHT-station (3) is (de partij achter) het datastation waar het algoritme draait (hierna: het datastation). De aanbieder van de PHT-data (4; zie de volgende bullet) kan zelf een datastation hebben of een datastation extern laten draaien. De aanbieder van het datastation levert diensten op het gebied van het beschikbaar stellen van data, het valideren van authenticatie, autorisatie en certificeringen, het bieden van koppelvlakken om de data te gebruiken en logging.
- De aanbieder van de PHT-data (4) is de data-eigenaar, zoals een ziekenhuis, onderzoeksinstituut of de patiënt zelf (hierna: de data aanbieder).

Binnen het PHT-domein (zie het middelste deel van bovenstaande figuur) vindt de daadwerkelijke analyse plaats door het algoritme (2) via het datastation (3).

Een groep van opdrachtgevers (actoren 1 en 2) en een groep van datastations (actoren 3 en 4) samen kunnen een PHT-netwerk vormen. Het kenmerk van een PHT-netwerk is dat zij gezamenlijk invulling

geven aan een deel van de afspraken. Dit zijn bijvoorbeeld afspraken over het gebruik van data tussen een groep ziekenhuizen of binnen het domein van de verpleeghuiszorg.

Als laatste actor zijn er de dienstverleners van rails, centrale componenten die nodig zijn om het PHT-netwerk te faciliteren. Enerzijds is dit de set van afspraken waarmee partijen werken en anderzijds zijn dit technische componenten die niet decentraal uitgevoerd kunnen worden.

3.5 Begrippenlijst

Definities van de belangrijkste begrippen.

3.5.1 Definitie PHT en trein-metafoor

De essentie van de metafoor “Personal Health Train” is dat de analyse (trein) op bezoek gaat bij de data (station) en vervolgens alleen de (tussen)resultaten van de analyse meeneemt. Hierdoor wordt het samenvoegen van data op een centrale plek t.b.v. data-analyse overbodig. Uiteraard is dit een streven en blijft in de praktijk het groei-principe “decentraal waar het kan, centraal waar het moet” bestaan. Zo zullen er steeds meer analyses decentraal uitgevoerd kunnen worden. Voor de PHT set van afspraken is het van belang om zo scherp mogelijk op te schrijven wat er onder de elementen trein, spoor, datastation en PHT-netwerk wordt verstaan. Waarbij het belangrijk is om te beseffen dat een metafoor nooit volledig overeenkomt met de werkelijkheid en alleen helpt de communicatie over de werkelijkheid te vergemakkelijken.

3.5.2 Trein

Onder de ‘trein’ in de PHT-metafoor wordt het volgende verstaan. De ‘trein’ staat voor de vraag die je wilt stellen aan de data. Dit wil zeggen de software en bijbehorende protocollen om de vraag te beantwoorden. De trein is de vraag zelf en het technische vehikel voor de vraag. De trein neemt géén (individuele) data mee. Voor de trein zijn er verschillende oplossingen beschikbaar. Deelnemers aan een netwerk zullen individueel een leverancier kiezen die de trein voor ze bouwt of bestaande oplossingen toepast. In de praktijk bestaat een “trein” vaak uit meerdere delen. Een deel dat de lokale analyse doet in een station (bijv. het aantal patiënten met diagnose X in station A) en een deel dat de lokale antwoorden verbindt tot een globaal antwoord (bijv. het totaal aantal patiënten met diagnose X van alle stations samen).

3.5.3 Spoor

Onder het ‘spoor’ in de PHT-metafoor wordt het volgende verstaan. Het ‘spoor’ is een verzamelbegrip voor de IT- voorzieningen die op een veilige en afgesproken manier de treinen en de stations met elkaar verbinden. Ook de juridische, technische en financiële verantwoordelijkheden en afspraken vallen hieronder. Dit wordt vertaald naar diensten voor autorisatie, identificatie, tools voor het vinden van stations en goedgekeurde treinen. Daarnaast worden onder het ‘spoor’ ook andere specificaties van afspraken en voorwaarden om treinen te laten rijden samengevat. Bijvoorbeeld de mate waarin privacy gewaarborgd kan worden, responstijden, query volume. Het spoor is géén inhoudelijke stakeholder, denk bijvoorbeeld aan een aanbieder van een IT-platform. Het spoor bestaat uit een lokale component waarmee het station aan het spoor wordt verbonden en een globale componenten die zorgt dat de treinen de stations kunnen bereiken en de communicatie tussen (delen van) de treinen afhandelt.

3.5.4 Data-Station

Onder het ‘data-station’ in de PHT-metafoor wordt het volgende verstaan. Het ‘data-station’ is de plek waar persoonsgegevens staan. Concreter gaat het daarbij om de beheerde omgeving van de data-

eigenaar, met daarbinnen de data. Het 'data-station' is de plek waar de trein de data kan vinden voor haar vraag. Daarvoor moet het 'data-station' vanuit technisch oogpunt verschillende soorten treinen kunnen toelaten. Dit wordt bereikt door middel van open standaarden waar treinen zich aan moeten conformeren. Het 'data-station' heeft een permanent karakter. Het 'data-station' is daarnaast ook de plek in de PHT-metafoor waar de analyse wordt uitgevoerd (lokale analyse). Deelnemers aan een netwerk zullen individueel een leverancier kiezen die het station voor ze bouwt of bestaande oplossingen toepast.

3.5.5 Wat is een PHT-netwerk?

Een PHT-netwerk beschrijft een groep van partijen die gezamenlijk data-analyses willen uitvoeren. Dit netwerk maakt hierbij gebruik van sporen, stations en treinen voor een specifiek doel. Deze combinatie van spoor, stations en treinen voor een specifiek doel kunnen in de technische zin bedoeld zijn met het 'PHT-netwerk'.

3.5.6 Wat is een 'PHT-dienstverlener' of 'PHT-leverancier'?

Een PHT-dienstverlener of -leverancier is een professionele aanbieder en/of verkoper van een PHT-station, trein en/of elementen van het spoor.

3.5.7 Waarvoor kan de PHT ingezet worden?

De PHT wordt ingezet voor statistiek en onderzoek. Dit kan onderzoek op het gebied van gezondheid en ziekte zijn maar ook t.b.v. van beleidsontwikkeling, kwaliteitsverbetering en voor verantwoording. Belangrijk is daarbij dat de PHT voornamelijk ingezet wordt voor een bepaald aggregatieniveau oftewel secundair data-gebruik. De PHT is dan ook niet bedoeld als vehikel voor primair data-verkeer zoals het primair vastleggen en overdragen van urgente gezondheidsinformatie. Hiervoor zijn er andere kaders (in ontwikkeling) zoals e-overdracht, MedMij etc. De PHT wordt vaak geassocieerd met personalised health care. De PHT wordt echter niet direct ingezet voor medische besluitvorming. Wel kan met behulp van de PHT een algoritme of model op basis van decentrale data ontwikkeld worden. Deze kan vervolgens ingezet worden voor de medische besluitvorming. Voorafgaande aan toepassing in medische besluitvorming moet echter een vorm van goedkeuring plaatsvinden. Deze zit in het medisch hulpmiddel-domein, betreft een inhoudelijke toetsing van de inhoud van de trein valt daarmee buiten de scope van de PHT-set van afspraken. De eigenaar van een algoritme is bijvoorbeeld verantwoordelijk om aan de medical device regulations te voldoen. Ook is er hiervoor een *veldnorm AI* in ontwikkeling. Een PHT-netwerk geeft geen garanties of de inhoud van de trein voldoende is voor medische besluitvorming.

3.5.8 Kan de burger ook gebruiker van de PHT zijn?

Op termijn zou dit kunnen en zeker in de ontwikkeling rondom *citizen science*. De PHT is ontwikkeld met de hier beschreven gebruiksdoeleinden in gedachten. Gebruik van de PHT buiten de hier beschreven scope is verantwoordelijkheid van de PHT-netwerken zelf. Hierover kunnen aanvullende afspraken gemaakt worden.

Dit hoofdstuk wordt verder aangevuld.

4 Juridisch kader

Juridisch kader: bevat een analyse van relevante wet- en regelgeving, en een verwijzing naar een aantal modelovereenkomsten.

Er zijn twee lagen van juridische kaders:

4.1 Organisatie van het PHT-stelsel

Hier wordt bijvoorbeeld vastgelegd waaraan partijen moeten voldoen om deel te kunnen nemen aan een PHT – netwerk en/of te voldoen aan de PHT-set van afspraken

4.2 Juridisch inhoudelijke ontwikkelingen

Juridisch inhoudelijke ontwikkelingen rondom de praktische implementatie van decentrale data-analyse. Zoals dataverwerkingsovereenkomsten, verantwoordelijkheden en aansprakelijkheden. Hier zijn al volop ontwikkelingen, werkwijzen en oplossingen in het PHT-netwerk te vinden. Deze zullen hier samengedragen moeten worden.

Dit hoofdstuk wordt verder aangevuld.

5 Normenkader

Normenkader: het normenkader beschrijft welke normen er vanuit de set van afspraken van de deelnemers worden verwacht. Daarnaast wordt ook ingegaan op eventuele certificering van deze afspraken.

Van de deelnemers van de set van afspraken wordt verwacht dat ze zich houden aan de principes van de set van afspraken. De principes zijn breed opgezet, om dit concreet te maken is er een normenkader uitgewerkt.

5.1 Normen

Binnen de set van afspraken hanteren wij een aantal normen. Binnen dit kader opereren netwerken en organisaties met elkaar. Er wordt verwacht van de deelnemer van de set van afspraken dat deze normen ten allen tijden worden nageleefd.

5.2 Toelichting normen

Hieronder volgt een toelichting per norm. Er wordt ingegaan op wat er concreet van de organisatie wordt verwacht.

5.2.1 Duidelijk wie partijen zijn en wie toegang tot een PHT-netwerk heeft

De organisaties binnen het netwerk zorgen ervoor dat ten alle tijden duidelijk is welke organisaties lid zijn van het netwerk en hoe toegang tot het netwerk is geregeld.

Hierover maken we de volgende afspraken:

- Alle betrokken partijen hebben publiek gemaakt wie er in het netwerk zitten
- Er is duidelijk met welk doel het PHT-netwerk is opgericht
- Het netwerk maakt onderling afspraken over autorisatie en identificatie van de deelnemers
- Alle betrokken organisaties binnen één netwerk hanteren dezelfde autorisatie en identificatie aanpak
- De gekozen aanpak rondom autorisatie en identificatie is transparant voor alle deelnemers en transparant voor potentiële gebruikers van de data, deze is beschreven en beschikbaar

5.2.2 Transparantie over kenmerken van datastation en treinen

De organisaties binnen het netwerk zorgen ervoor dat alle metadata over de PHT-producten transparant is voor elkaar.

Hierover maken we de volgende afspraken:

De organisaties binnen het netwerk dragen er zorg voor dat in de metadata van de trein ten minste vermeld is;

- Door wie trein is gebouwd
- Datum van de uitvraag
- Analyse die wordt uitgevoerd (encryptie van de berekening)

- Welke data gebruikt wordt voor de analyse

De organisaties binnen het netwerk dragen er zorg voor dat in de metadata van het datastation ten minste vermeld is;

- Door wie het datastation is gebouwd
- Welke data er in het datastation aanwezig is
- Op welke datum de laatste update van de data is geweest
- Hoe de data in het datastation geraadpleegd kan worden
- Wie de eigenaar is van de data
- De totstandkoming van de data

5.2.3 Openheid over de berekeningen die op de data worden uitgevoerd en hoe de uitkomsten gebruikt gaan worden

De organisaties in het PHT-netwerk zorgen ervoor dat er bij het uitvragen van data openheid is over het doel van onderzoek en de berekeningen die op de data worden uitgevoerd.

In de procedure die het netwerk hiervoor opstelt wordt ten minste vastgelegd:

- Grondslag uitvraag (gebaseerd op wettelijk kader of onderlinge afspraken)
- Doel uitvraag (wat is de onderzoeksvraag)
- Welke data er nodig is (en tevens gebruikt zal worden)
- Welke berekening er wordt uitgevoerd op de data
- Waar de uitkomst van de berekening naartoe gaat
- Hoe de uitkomst van de berekening gebruik gaat worden
- Hoe er duiding wordt gegeven aan de uitkomst

5.2.4 Duidelijkheid over hoe een nieuwe vraag in het netwerk wordt beoordeeld en goedgekeurd

De organisaties in het netwerk zorgen ervoor dat inkomende analysevragen voldoen aan de eisen van het netwerk.

De organisaties in het netwerk stellen een procedure op om te beoordelen of een inkomende analysevraag voldoet aan de eisen van het netwerk. Denk hierbij aan eisen op het gebied van:

- Juridische en organisatorische aspecten. Zoals doelbinding, juiste data beschikbaar, dataminimalisatie etc.
- Maatschappelijke en ethische aspecten. Zoals beschreven in de norm Medische inhoudelijke kwaliteit AI (in ontwikkeling¹).

5.2.5 Data minimalisatie en doelbinding

De organisaties in het netwerk beschrijven hoe er wordt gezorgd dat er niet meer gegevens worden gebruikt dan wat er minimaal nodig is voor het doen van het onderzoek.

De organisaties in het netwerk stellen procedures op om te controleren of de berekeningen die worden uitgevoerd passen binnen de eisen van dataminimalisatie. Zo controleert het netwerk bijvoorbeeld of een onderzoeker zo min mogelijk data gebruikt en of de berekening die wordt uitgevoerd past bij de data die wordt geraadpleegd.

5.2.6 Duidelijkheid over rollen in het kader van AVG

De organisaties binnen het PHT-netwerk zorgen dat duidelijk is welke partij welke rol heeft in het kader van de AVG.

Voor de verschillende rollen binnen de AVG moeten afspraken gemaakt worden. De organisaties binnen het netwerk zullen besluiten of er gebruik wordt gemaakt van een *super user* en wie *verwerkersverantwoordelijke* zijn. Tevens zal er een verwerkingsovereenkomst opgesteld worden voor de *verwerker* van de data.

We maken hierover de volgende afspraken:

- De verwerkersverantwoordelijke partijen zorgen dat er een 'onderlinge regeling' wordt vastgesteld (hierin wordt o.a. zaken vastgesteld rondom privacyverklaringen, DPIA en handeling rondom datalek).
- De organisaties binnen het PHT-netwerk zorgen dat de bevoegdheden van de verwerker zijn vastgelegd in een verwerkersovereenkomst.

5.2.7 Afspraken over metadata en betekenis van data in een netwerk

Door de organisaties binnen het netwerk worden afspraken gemaakt over de metadata die beschikbaar moet zijn, en het is voor alle organisaties binnen het netwerk duidelijk welke data waar staat. Hierbij worden de FAIR principes gevolgd.

De metadata bevat bijvoorbeeld elementen zoals:

- Eigenaar van de data
- Laatste update van de data
- Op welke momenten en door wie de data is verwerkt
- Welke informatie er in de data staat

5.2.8 Deelnemers in een PHT-netwerk maken gebruik van stations en treinen die het PHT-label mogen dragen

De organisaties in het netwerk zorgen ervoor dat er alleen stations en treinen gebruikt worden die een PHT-label gebruiken.

Hierover maken we de volgende afspraken:

- De organisaties in het netwerk maken gezamenlijk een afspraak over hoe gecontroleerd wordt dat de gebruikte PHT-diensten gecertificeerd zijn (denk hierbij aan gebruik maken van verifiable credentials of dit juist meer centraal opzetten)
- Als een netwerk een PHT-leverancier benadert voor het bouwen van een data- station, legt het netwerk vast dat het om een gecertificeerde leverancier gaat
- Als een netwerk gebruik maakt van een trein, controleert het netwerk of deze trein afkomstig is van een gecertificeerde leverancier van PHT-diensten

5.2.9 De data in een datastation wordt goed beheerd

De eigenaar van de data zorgt ervoor dat de data in het datastation op een correcte wijze wordt geplaatst, beheerd en verwijderd. Tevens blijft de dataeigenaar ten alle tijden verantwoordelijk voor het juiste gebruik van de data.

Hierover maken we de volgende afspraken:

- De data in het datastation is voldoende beveiligd (voldoet b.v. aan de NEN7510)

- De data in het datastation is FAIR
- Bij een wettelijk bewaartermijn voor de data moet deze gehandhaafd worden (als de gegevens niet onder een wettelijke bewaartermijn vallen moet de eigenaar van de data zelf bepalen welke behaartermijn voor de gegevens gerechtvaardigd is)

5.3 Toetsing normen

De normen binnen dit kader worden getoetst middels een self assessment. Daarvoor kan een organisatie gebruik maken van de checklist (zie bijlage). Op het moment dat de organisaties binnen het netwerk het self assessment hebben gedaan kunnen ze een certificaat aanvragen. Na ontvangst mogen zij gebruik maken van het beeldmerk PHT en mogen ze zich een PHT-netwerk noemen.

5.4 Gebruik beeldmerk

Er zijn twee vormen van het gebruik van het beeldmerk. De eerste betreft het gebruik van het beeldmerk door een netwerk, wat zich dan een PHT-netwerk mag noemen. De tweede betreft de diensten die een leverancier aanbiedt (datastation/trein), welke dan het PHT-beeldmerk mogen dragen.

Het beeldmerk PHT is een beschermd beeldmerk, dat betekent dat als een lid van de set van afspraken zich niet aan de normen of technische eisen houdt deze het beeldmerk niet (meer) mag gebruiken. De beheerder van de set van afspraken behoudt zich ten alle tijden het recht om deelnemers hun lidmaatschap te ontzien en het recht op gebruik van het beeldmerk in te trekken.

Stappen die de organisatie kan zetten om het beeldmerk te ontvangen en de procedure bij het eventuele intrekken van het beeldmerk door de beheerder van de set van afspraken zijn opgenomen in het hoofdstuk 'Operationele processen'.

5.4.1 Gebruik beeldmerk door netwerk

Als netwerk mag je het PHT-beeldmerk gebruiken als voldaan wordt aan de normen zoals gesteld in het normenkader. Hiervoor kan gebruik worden gemaakt van een checklist, opgenomen in Bijlage I.

Als binnen een PHT-netwerk een organisatie zich niet aan de set van afspraken houdt, heeft de beheerder van de set van afspraken het recht om het PHT-netwerk hierop aan te spreken en het beeldmerk in te trekken.

5.4.2 Gebruik beeldmerk door leverancier PHT-diensten

Als leverancier mag je het PHT-beeldmerk gebruiken als je producten voldoen aan de technische eisen zoals beschreven in het hoofdstuk 'Gegevensuitwisseling'.

Als een leverancier van PHT-diensten zich niet aan de afspraken houdt, heeft de beheerder van de set van afspraken het recht om de leverancier hierop aan te spreken en het gebruik van het beeldmerk in te trekken.

6 Technical Architecture and Agreements

In dit hoofdstuk zijn de technische architectuur en afspraken van de PHT beschreven.

The Personal Health Train is a metaphor for a system that allows decentralized analysis of (record-level) data, where, by design, record-level data remains at the source as much as possible and only aggregated statistics are shared. Examples of this type of analysis range from retrieving counts or ratios to full-fledged federated analysis. Using the latter, more complex mathematical/statistical models are fitted to the data. Many of today's use cases are using sensitive data for analysis, but the Personal Health Train is not limited to use cases with sensitive data only. For instance, in some use cases, a physician may need to combine personal health data (sensitive) with a publicly available list of medicines (non sensitive). Here we define sensitive data as *confidential information that must be kept safe and out of reach from outsiders unless they have permission to access it*.

6.1 Types of PHT-networks

The base of the system is made up of **dispatchers**, **trains**, and **stations**. These components are connected through a network of rails and are collectively known as the infrastructure. The rails represent the terms and conditions to access the data in the network. Data controllers are the controllers of the stations where data is made available. The terms and conditions are jointly agreed and maintained by the data controllers and monitored in an organizational role called the **operational control center rail**. The network can be public, private, permissionless, or permissioned.

- **Public networks** are exposed to the public, any data controller can join whenever they want. More so, anyone can see the network and dispatch trains in the network. A public network does not mean that there are no terms and conditions, only that everyone is allowed to join who meets these terms and conditions. Joining a network means that a data controller agrees to the terms and conditions of the network's interoperability framework. The operational control center rail ensures that everyone complies.
- **Private networks** are not exposed, it's not open for the public to join in. They are only accessible to a certain group of people or organizations. The data controllers in the network will use a train operating company to operate the trains and implement access control. Train dispatchers need to use a train operating company to dispatch trains.
- **Permissionless networks** are networks that require no permission to join and interact with. An example is a station with open data. It's public and permissionless. Open data is data that can be freely used, reused and redistributed by anyone - subject only, in the extreme, to the requirement of attribution and sharing.
- **Permissioned networks** are networks that require permission from data controllers to use data, mostly (sensitive) personal data. Data controllers need to have full control on the processing of personal data and apply the highest security standards to protect the data. All components in the network are controlled and must comply with the terms and conditions. The operational control center rail ensures that everyone complies. Most permissioned networks are also private.
- **Hybrid networks** which combine public, private, permissionless, and permissioned stations in a network.

In practice, several Personal Health Train implementations are currently emerging. Different design decisions have been taken based on the requirements of the use cases addressed.

6.1.1 Business roles and responsibilities

A **network** represents a collaboration between one or more (clinical) data controllers that wish to make data available and provide services to use the data they control. It is within this collaboration that agreement is sought on which data is made available, which vocabularies are used, and the way it can be accessed.

The data controllers jointly decide on the terms and conditions of data usage, perhaps limiting the complexity of the analyses or detailing how results should be disseminated. The governance of the network is organized in a role called the operational control center rail.

Data controller

As defined in the GDPR Article 4 (7): the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data.

The data controller provides data through a station provided by a station provider. The data controller may be a station provider itself.

For example, within a network it is agreed who is allowed to drive on the rails (i.e., who's allowed to run analyses), what type of usage is allowed (e.g., a specific query or any custom Docker image) and whether (financial) compensation is required.

Networks are usually formalized through (legally binding) contracts that additionally define how liabilities are handled.

Agreement

The railway network:

- represents a collaboration where one or more data controllers wish to make data available and provide services to use the data they control.
- decide upon the terms and conditions of the network, and the data usage within the network.
- decide upon the network to be public, private, permissionless, or permissioned.

6.1.2 Interoperability Governance of the rail network

The operational control center rail is an organization that acts as an authority of the rails/tracks under its control. As such it can control:

- certification of train operating companies and station providers,
- validation of purpose, and ethics of analyses,
- validation/certification of train implementations,
- certification of train operating companies,
- and identification of railway users such as train dispatchers and data controllers.

The network will have to decide what is applicable.

The authority may be a joint responsibility of all data controllers in a network but may also be a single responsibility of the data controller itself. This monitoring- and validation - function could be centrally implemented (e.g., on a central server) yet a decentralized implementation could also work.

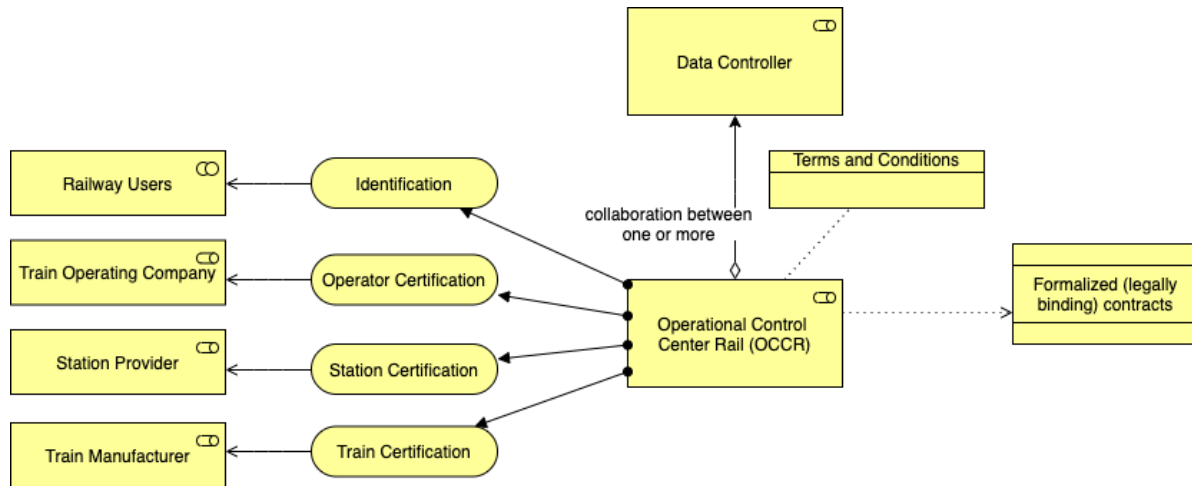


Figure 6.1: Responsibilities of the Operational Control Center Rail

Actors in the governance framework of the operational control center rail:

- **Train operating company;** the stakeholder (individual or organization) who provides access to the rails of the network.
- **Train manufacturer;** the stakeholder (individual or organization) who builds the train and stores it in a depot.
- **Train dispatcher;** the train dispatcher represents the role of a stakeholder that is responsible for the dispatch of a given train to stations. Examples include the physician who sends a train to perform a treatment analysis on a hospital's station or a researcher who dispatched a statistical analysis train to a given station.
- **Station provider;** the data controller makes data available through a station provided by the station provider. The station may be a managed service of a supplier, or the data controller may have the role of station provider itself. When the station provider is a supplier, then it acts as a data processor of the data controller.
- **Railway users (not a role, but a collective);** the users of the infrastructure - the train dispatchers and data controllers providing data availability.

Example:

The National Health Care Institute wants to research the effectiveness of medication. Together with the controllers of patient and quality registries, it forms a governance framework with agreements, terms, and conditions for conducting research. The National Health Care Institute will act as the operational control center rail to ensure compliance. The network will be a private permissioned network.

In this collaboration, agreements have been made to regulate access and usage of the Netherlands Cancer Registry data. IKNL is the train operating company in this case. Train dispatchers must use Vantage6 from IKNL. Because IKNL also is a data processor of the Netherlands Cancer Registry, it acts in two different roles.

6.1.3 Operating trains on the railway

The **train operating company** is the stakeholder (individual or organization) who provides access to the rails of the network. The train operating company is allowed access to the rail network under the terms and conditions of the network. The train operating company **provides services** to train dispatchers to run trains in the network and provide train depots for storage and reuse of trains.

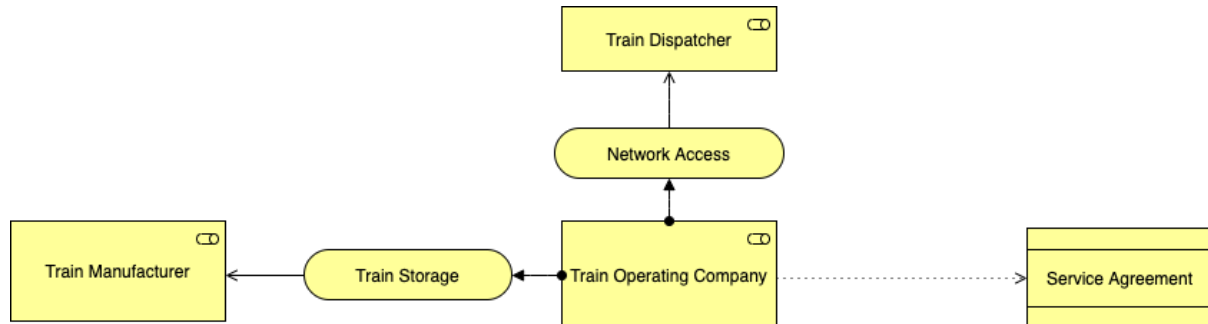


Figure 6.2: Responsibilities of the Train Operating Company

The train operating company provides the following services:

Storage of trains in a **train depot**; the depot hosts a collection of readily available trains that can be deployed within a network. For example, a Docker registry holding images that can be run by a station, may be considered a depot.

Access to the railway (e.g., a portal may be provided) including **authentication** of train dispatchers and stations, and secure transportation of trains to the station.

For participants in a PHT network, it should be transparent who the participating parties are and what rights they have. No others should be allowed, and participants should only be able to perform operations (e.g. access data, send trains) as agreed upon. The operational control center rail controls the usage and the traffic in the network and is responsible for doing so under the accountability of the data controllers in the network. Access is granted to certified train operating companies which deliver the machinery to train dispatchers to dispatch trains.

The data stewards of the data controllers should be enabled to have oversight and control of their station, such that they comply with their data protection standards.

6.2 Managing semantic and technical interoperability in the rail network

Probably, an organization controlling data (the data controller) will use a managed service for a station offered by a vendor, a station provider. The data controller will probably be part of a network with other data controllers operating in the same domain. So they can agree together and join forces to define the terms and conditions of data availability and usage. Also they can jointly implement an operational control center rail.

Aspects of technical interoperability include interface specifications, interconnection services, data integration services, data presentation and exchange, and secure communication protocols. Technical interoperability is not realized by the implementation of a single standard. It is implemented with advertising the capabilities it supports (expressed in the station's metadata, the data catalog).

Capabilities are a mix of standards developed by a Standard Development Organization (SDO) and usage requirements to gain the capability.

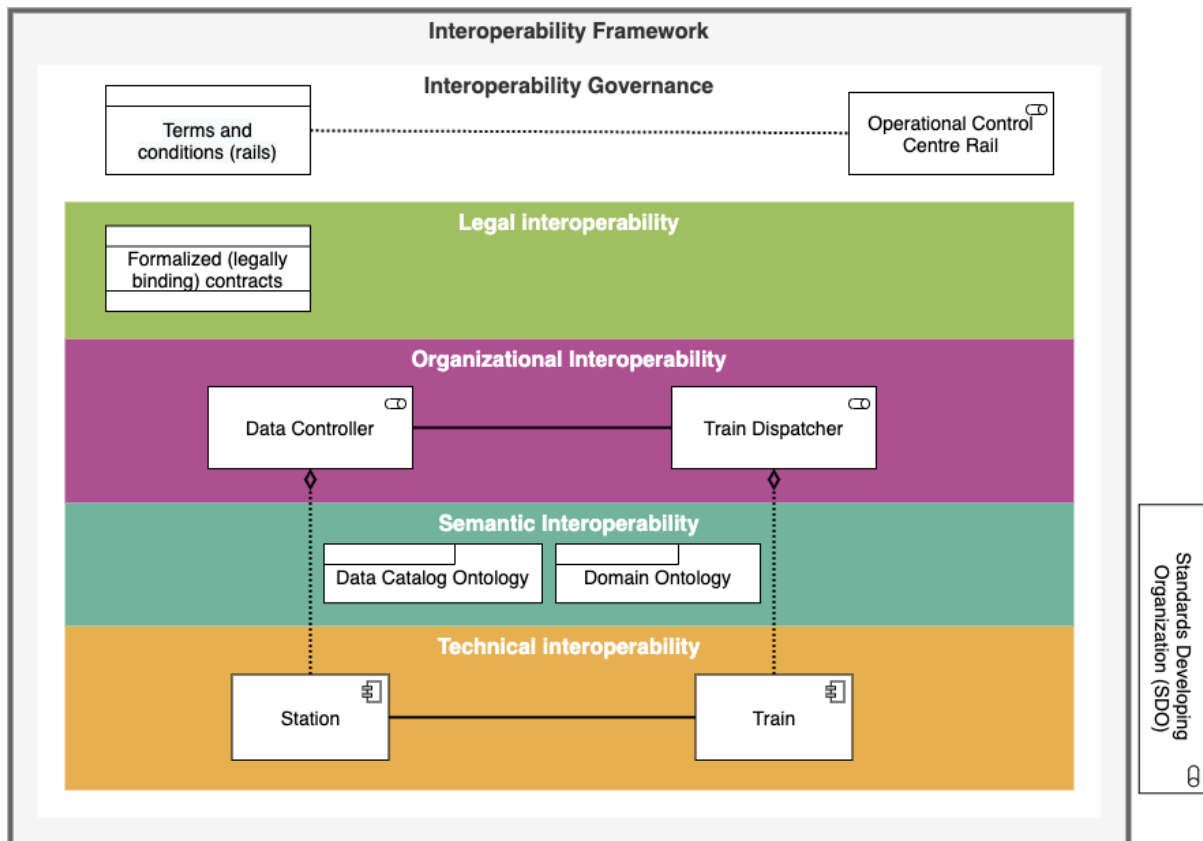


Figure 6.3: Levels of interoperability based on the European Interoperability Framework (EIF)

6.2.1 Semantic Interoperability

Semantic interoperability ensures that a train dispatcher and the train have knowledge of the precise format and meaning of the data available in a station, in other words 'what is *accessed and analysed* is what is *understood*'. Data must use a formal, accessible, shared, and broadly applicable language for knowledge representation.

In the EIF, semantic interoperability covers both semantic and syntactic aspects:

- The semantic aspect refers to the meaning of data elements and the relationship between them. It includes developing ontologies, vocabularies and schemata to describe data exchanges, and ensures that data elements are understood in the same way by all communicating parties.
- The syntactic aspect refers to describing the exact format of the information to be exchanged in terms of grammar and format.

This is implemented with advertising the ontologies, vocabularies and formats it supports through a capability statement (expressed in the station's metadata). Advertisement must be done using the latest version of the W3C Data Catalog Vocabulary (DCAT). Metadata includes data about the station and the content of the data made available in the station.

Examples of station metadata:

- Vendor / creator of the data station software, version
- Common interface to Share Metadata with Network
- Usage license for the Station
- Data interaction capability of the station, i.e., which type of trains the station supports.
- A minimal metadata schema for the station **SHOULD** be defined and agreed by the community.

Example of content metadata:

- Date of last update of dataset
- Requirements / specifications / license for accessing the data.
- Provenance information of the content, e.g., data controller, creation/update dates, version (when applicable), etc.
- A minimal metadata schema for the content **SHOULD** be defined and agreed by the community.

Agreement

Data controllers must implement **semantic interoperability** through a capability statement of the ontologies, vocabularies, and schema's it supports. The statements are expressed in the station's metadata using the latest version of the W3C Data Catalog Vocabulary (DCAT).

Data controllers must follow FAIR principles.

6.2.2 Technical Interoperability

Trains and stations are the technical components on the railway. Data controllers agree to host a station on premise or agree to a managed service from a service provider.

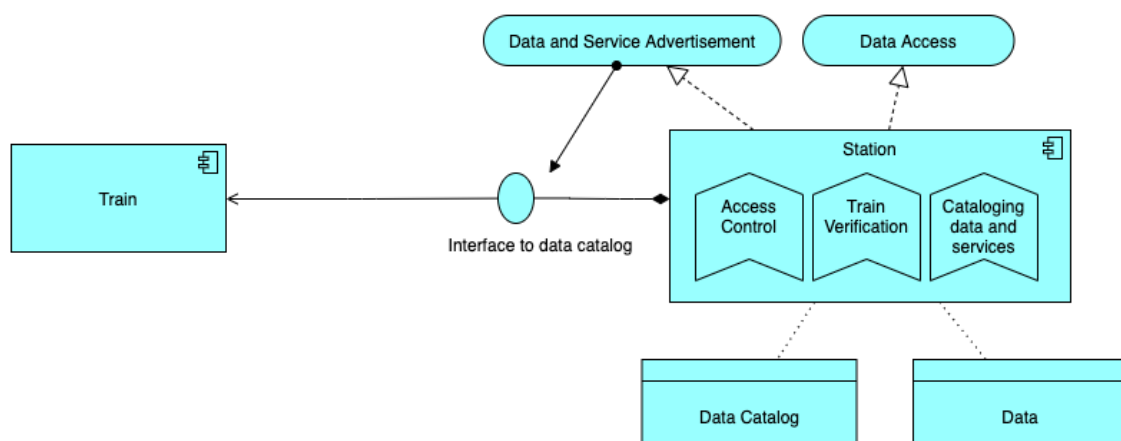


Figure 6.4: Trains must have the capability to interface with the stations data catalog (metadata)

The technical components in the interoperability framework:

- **Train** represents the method by which information is conveyed between dispatchers and stations. They may be considered shuttling messages back and forth (or going round in circles).

- **Station** represents an information system or application having access to (personal) data, typically a specialized application to represent knowledge of the data controller.

6.2.3 The capabilities of a station

A station is an information system or application, typically having access to (sensitive) record-level (patient) data. Its primary task is to receive trains and process their contents. To do this safely, stations also have something akin to a customs service: they have the ability to check if a train is allowed to enter the station and be processed. Data controllers are accountable for every processing of the data they control. Measures need to be in place to protect personal data.

Considering the variety of train types (different data interaction mechanisms), a station must advertise the types it supports through a capability statement (expressed in the station's metadata).

Agreement

- In most situations, stations need access to (personal) record-level data.
- Stations must implement technical interoperability through a capability statement of the interfaces, and train types, it supports. The statements are expressed in the station's metadata using the latest version of the W3C Data Catalog Vocabulary (DCAT).
- Stations must handle incoming trains (i.e., computation requests) and return the result of the train.
- Stations must perform access control, if the network is permissioned, to ensure privacy when using personal data.
- Stations must be hosted with an information security management system (ISMS) implemented. The ISMS must be an ISO/IEC 27001 certified system.
- Stations must follow the rules of law applicable to the geographical region (e.g. NEN 7510:2017 would be applicable in the Netherlands).

6.2.4 The capabilities of a train

The contents of trains are flexible and, depending on the use case, may range from a simple message to a (database) query, (the URL) of a containerized executable environment (e.g., a docker container) or a return value. It might be as complex as a federated machine learning model with hyper-parameter tuning. Different train types are envisioned, but not all will be supported by each station:

The train manufacturer, source, destination, and type must be clearly written on the message envelope (by means of the metadata). The payload may be hidden from view.

Agreement

Trains must implement a mechanism to interact with the data made available in a given station.

The metadata of a train must express:

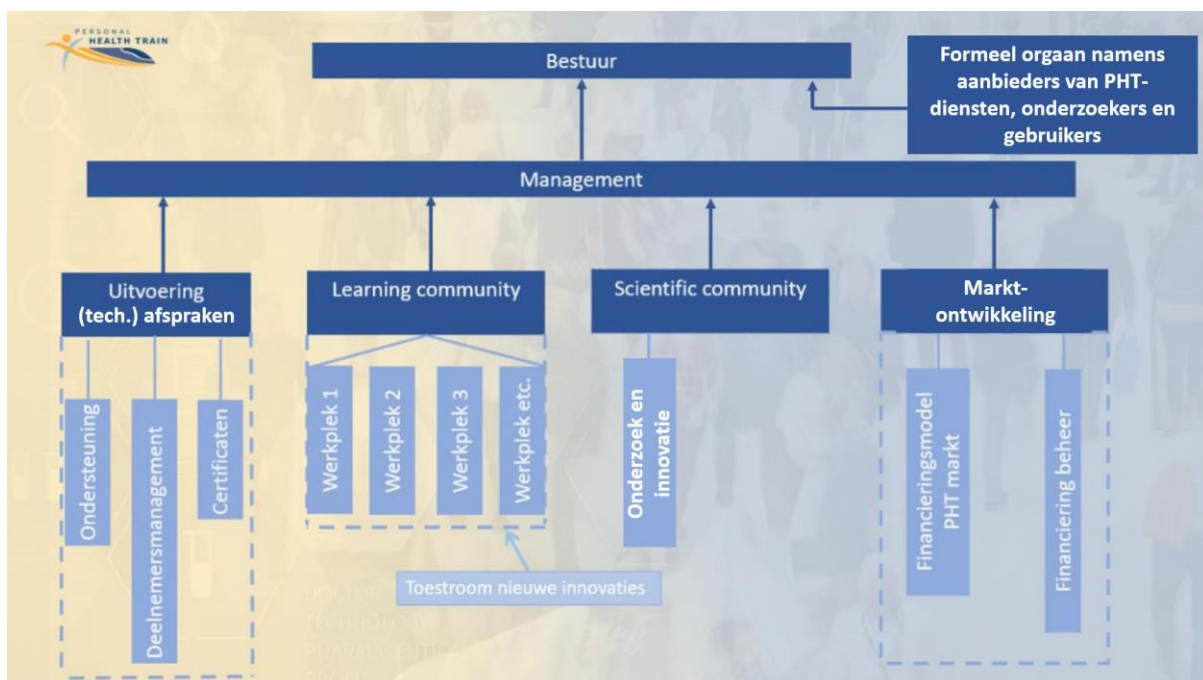
- The train type.
- The origin of the train, the train dispatcher, and the manufacturer.
- The route and destination of the train.
- The goal it tries to achieve; stations may choose which types they support.
- Information about certification of the train, interoperability/compatibility with data, the review/ approval process and potential privacy risks
- Proof of integrity and origin; trains must prove the origin and integrity of the train, so that evidence exists that the train is the same train as it was built by the train manufacturer.

The metadata must use a W3C Resource Description Framework (RDF) format.

7 Beleid

Beleid: dit hoofdstuk gaat in op de vraag 'hoe om te gaan met besturingsvraagstukken' en vormt de basis voor de operationele processen.

Het is nog niet bekend in welke rechtsvorm beheer en doorontwikkeling van de PHT-set van afspraken georganiseerd zal worden. Hiervan hangen organisatievragen af zoals de benoeming van het bestuur. Onderstaande weergave van de rollen binnen de PHT-governance geeft alvast een beeld van de benodigde activiteiten ten behoeve van PHT-beheer en doorontwikkeling.



Afbeelding 7.1.: PHT governance structuur.

7.1 Rollen binnen de PHT-governance

7.1.1 Bestuur

In het bestuur ligt de eindverantwoordelijkheid voor het functioneren van het afsprakenstelsel. Taken zijn o.a.:

- Strategische besluitvorming
- Visie bepalen (meerjarenplan)
- Afstemming externe overlegorganen (zoals het Informatieberaad, de AI coalitie, de Architectuur Board etc.)
- Beslissingen nemen over de financiering
- Kaders meegeven richting management

7.1.2 Management

Het management is verantwoordelijk voor de tactisch en operationele besluitvorming, taken zijn o.a.:

- Verantwoording richting bestuur over voortgang
- Stakeholdermanagement

7.1.3 Markontwikkeling

In de markontwikkeling wordt gekeken naar financiële randvoorwaarden voor de (lange termijn) governance en naar modellen over hoe een PHT-netwerk financieel kan opereren. Ten behoeve van financieel duurzaam groeien en implementeren van de PHT in de praktijk. Binnen de markontwikkeling liggen de volgende twee kerntaken, daarbij krijgen ze kaders mee vanuit het bestuur:

- Ontwikkelen financiële modellen voor 'PHT markt'
- Ontwikkelen financieringsmodel voor beheer PHT

7.1.4 Scientific community

De scientific community richt zich o.a. op:

- (Technologische) innovatie binnen de PHT
- Nadenken over wat nieuwe technologieën voor de set van afspraken betekenen
- Advies richting bestuur over innovatieve ontwikkelingen en nieuwe technologieën
- Onderzoek rondom PHT ondersteunen en in kaart brengen

7.1.5 Learning community

Binnen de learning community komen alle partijen bij elkaar die bezig zijn met PHT of willen starten met een implementatie van de PHT. De learning community is in de basis een open community, met eventueel besloten vergaderingen voor alleen leden. Taken van de community zijn o.a.:

- Nieuwe initiatieven helpen en aantrekken
- Signalen opsporen in het veld en terugleggen bij het bestuur
- Ervaringen delen ten behoeve van kennisdeling
- Documenten opleveren zoals lessons learned, best practices, templates etc.
- Klankbord voor bestuur

7.1.6 Formeel orgaan namens aanbieders van PHT-diensten, onderzoekers en gebruikers

Er wordt een formeel orgaan opgericht namens aanbieders van PHT-diensten, onderzoekers en gebruikers. Dit is er op het moment nog niet. Een dergelijk orgaan is nodig om feedback te krijgen vanuit de gebruiker van de afspraken. Deelnemers aan dit orgaan kunnen ook deelnemen aan de learning community. Het verschil is, dat dit orgaan een goede afspiegeling is van alle gebruikers en daarbij een formele rol krijgt rondom de doorontwikkeling van de set van afspraken. De precieze vorm en rol moet nader bepaald worden, en ook moet erover nagedacht worden wie er de 'gebruikers' zijn. Taken kunnen zijn:

- Meebeslissen over strategische besluiten
- Gevraagd advies geven richting bestuur

7.1.7 Uitvoering (technische) afspraken

De laatste rol binnen de governance is de uitvoering. Hier ligt de verantwoordelijkheid van de volledige operationele uitvoering van de set van afspraken. In de kern heeft de uitvoering minimaal de volgende taken:

- Beheer afsprakenset (doorvoeren van wijzigingen)
- Deelnemersmanagement zoals in- en uittreding, lijsten bijhouden, contactgegevens etc.
- Certificering (beheer van certificaten, aanvragen verwerken etc.)

7.1.8 Certificering

Dit hoofdstuk wordt later aangevuld.

8 Operationele processen

Operationele processen: In de operationele processen wordt beschreven welke processen er nodig zijn voor een goede werking van de set van afspraken.

In dit hoofdstuk worden de operationele processen beschreven die nodig zijn voor een goede werking van de set van afspraken. Dit betreft operationele processen rondom de ontwikkeling van de set van afspraken zelf en het beeldmerk van PHT, en processen rondom communicatie en sturing.

8.1 Operationele processen rondom ontwikkeling afsprakenset

In deze paragraaf komen processen rondom het opstellen, wijzigen, publiceren en implementeren van de afspraken. Dit hoofdstuk wordt later aangevuld.

8.1.1 Procesafspraken

(bijvoorbeeld; deelnemers en de beheerder kan wijzigingsvoorstellen indienen, de beheerder onderhoudt de afspraken, waaruit bestaat een wijzigingsverzoek, wie besluit over het doorvoeren van een wijzigingsverzoek)

8.1.2 Releases

Dit hoofdstuk wordt later aangevuld.

8.2 Operationele processen rondom beeldmerk

Procedure ontvangen beeldmerk:

Procedure intrekken beeldmerk PHT-netwerk:

- De beheerder van de afsprakenset meldt bij de technische werkgroep dat er een tekortkoming is geconstateerd bij een leverancier
- De technische werkgroep beoordeelt de tekortkoming en besluit over het al dan niet verwijderen van de leverancier uit de set van afspraken, de werkgroep zet de volgende stappen bij besluit tot verwijderen:
 - de werkgroep geeft bij de leverancier aan welke afspraken zijn geschonden en wat er minimaal moet worden aangepast om aan de set van afspraken te voldoen

- de werkgroep geeft bij de beheerder van de set van afspraken aan dat de leverancier het beeldmerk niet meer mag gebruiken
- De beheerder van de set van afspraken verwijderd de leverancier van de lijst met deelnemende organisaties, ten minste tot de leverancier weer aan de afspraken voldoet

Procedure intrekken beeldmerk leverancier PHT-diensten:

- De beheerder van de afspraken set meldt bij de technische werkgroep dat er een tekortkoming is geconstateerd bij een leverancier
- De technische werkgroep beoordeelt de tekortkoming en besluit over het al dan niet verwijderen van de leverancier uit de set van afspraken, de werkgroep zet de volgende stappen bij besluit tot verwijderen:
 - de werkgroep geeft bij de leverancier aan welke afspraken zijn geschonden en wat er minimaal moet worden aangepast om aan de set van afspraken te voldoen
 - de werkgroep geeft bij de beheerder van de set van afspraken aan dat de leverancier het beeldmerk niet meer mag gebruiken
- De beheerder van de set van afspraken verwijderd de leverancier van de lijst met deelnemende organisaties, ten minste tot de leverancier weer aan de afspraken voldoet

8.3 Operationele processen rondom communicatie

Voorbeelden zijn het melden van een nieuw netwerk of het stoppen van PHT-netwerk.

Dit hoofdstuk wordt later aangevuld.

8.4 Sturingsstructuur

Dit hoofdstuk wordt later aangevuld.

9 Communicatie

Communicatie, Gebruik van logo, etc. etc.

Dit hoofdstuk wordt later aangevuld.

Bijlage I – certificering PHT-netwerk

Voordat organisaties samen een PHT-netwerk kunnen vormen, moeten een aantal zaken helder zijn. De volgende vragen zijn het startpunt voor het vormen van een nieuw PHT-netwerk

- Wat is het **doel** van het PHT-netwerk initiatief? Waarom willen jullie gezamenlijk een nieuw netwerk vormen?

Na het besluit om gezamenlijk een nieuw PHT-netwerk te willen vormen moeten een aantal zaken geregeld worden. Zorg ervoor dat:

- Jullie zijn **aangemeld** bij de beheerder van de set van afspraken (jullie worden opgenomen in de lijst met PHT-netwerken)
- Jullie een **certificaat aanvragen** bij de beheerder van de set van afspraken

Checklist: voldoet ons netwerk aan de normen gesteld in het normenkader?

Als de organisaties binnen het netwerk voldoen aan alle afspraken in onderstaande lijst, mogen de organisaties zichzelf onderdeel van een PHT-netwerk noemen en het certificaat gebruiken

- ☐ Alle betrokken partijen hebben publiek gemaakt wie er in het netwerk zitten
- ☐ Er is duidelijk met welk doel het PHT-netwerk is opgericht
- ☐ Er zijn onderling afspraken gemaakt over autorisatie en identificatie van de deelnemers
- ☐ Alle betrokken organisaties binnen het netwerk hanteren dezelfde autorisatie en identificatie aanpak
- ☐ De gekozen aanpak rondom autorisatie en identificatie is beschreven en beschikbaar voor alle deelnemers
- ☐ Bij een data-uitvraag is een procedure opgesteld om het doel van de uitvraag vast te leggen, de volgende informatie wordt minimaal vastgelegd:
 - o Grondslag uitvraag (gebaseerd op wettelijk kader of onderlinge afspraken)
 - o Doel uitvraag (wat is de onderzoeksvraag)
 - o Welke data er nodig is (en tevens gebruikt zal worden)
 - o Hoe de uitkomst van de berekening gebruik gaat worden
 - o Hoe er duiding wordt gegeven aan de uitkomst
 - o Waar de uitkomst van de berekening naartoe gaat
- ☐ Bij een data-uitvraag is een procedure opgesteld om de berekening die wordt uitgevoerd vast te leggen
 - o Welke berekening er wordt uitgevoerd op de data
- ☐ Er is een lijst met criteria opgesteld om een inkomende analysevraag te beoordelen (deze criteria kunnen zaken bevatten rondom organisatorische en juridische aspecten zoals doelbinding, juiste data beschikbaar, dataminimalisatie)
- ☐ Er is een procedure opgesteld om te controleren of een inkomende analysevraag voldoet aan de door het netwerk opgestelde eisen
- ☐ Er is een procedure opgesteld om te controleren of de berekeningen die worden uitgevoerd passen binnen de eisen van dataminimalisatie

- ☐ Voor alle organisaties binnen het netwerk is duidelijk wie *super user* en wie *verwerkersverantwoordelijke* zijn.
- ☐ Er is een verwerkingsovereenkomst opgesteld voor de *verwerker* van de data waarin de bevoegdheden van de verwerker zijn vastgelegd
- ☐ Er is een procedure opgesteld om ervoor te zorgen dat een verwerker voordat deze gebruik maakt van de data de verwerkersovereenkomst ontvangt en ondertekent en archiveert
- ☐ De verwerkersverantwoordelijke partijen zorgen dat er een 'onderlinge regeling' wordt vastgesteld (hierin wordt o.a. zaken vastgesteld rondom privacyverklaringen, DPIA en handeling rondom datalek).