

Bespreking van de verschillende mogelijkheden van dataextractie uit het elektronisch medisch dossier van de huisarts.

Auteurs: Koen Thomeer, [vul aan wie meewerkt]

Status: intern

Doel: dit document heeft als doel de verschillende vormen van dataextractie te bespreken met een oplijsting van de voor en nadelen van elke vorm. Dit is een technisch document. De mogelijk anonimisatie en codificatie na de extractie wordt niet in dit document besproken.

Introductie

Dataextractie uit het EMD van de huisarts is altijd een moeilijk onderwerp geweest. Aan de ene kant is het heel nuttig, het geeft immers epidemiologische info over de eerste lijn. Meer en meer komen er nu ook concepten dat de data nuttig kan zijn in het kader van de directe zorgverlening zelf: opvolging zorgpaden, eigen opvolging kwaliteit van zorg, preventieprojecten, ... Echter, er zijn heel wat obstakels, en dit op verschillende niveaus. Grootschalige dataextractie vereist een uniforme extractie. Uniforme extractie vraagt een uniforme registratie. Dit laatste is de zwakste schakel in de ketting en heeft verschillende oorzaken:

- bij de registrerende: in casu de huisarts. Deze is niet opgeleid in registreren, wat zijn taak niet makkelijk maken. Het zoeken naar de juiste code vergt tijd, wat in het begin demotiverend werkt, alsook de kwaliteit van de registratie is afhankelijk van de interpretatie van elke huisarts. Een goede opleiding en een makkelijk registratiesysteem zou dit probleem goed kunnen opvangen.
- Bij de registreersoftware: in casu het EMD van de huisarts. Deze EMD's zijn meestal ontstaan uit eigen initiatieven. Met de latere labeling georganiseerd door het FOD zijn er regels opgelegd. Een van deze regels was een uniform registratiesysteem volgens de DPRS 2 norm, echter deze zijn slecht geïnterpreteerd en bijgevoegd als een aanhangsel aan de bestaande programmakern.

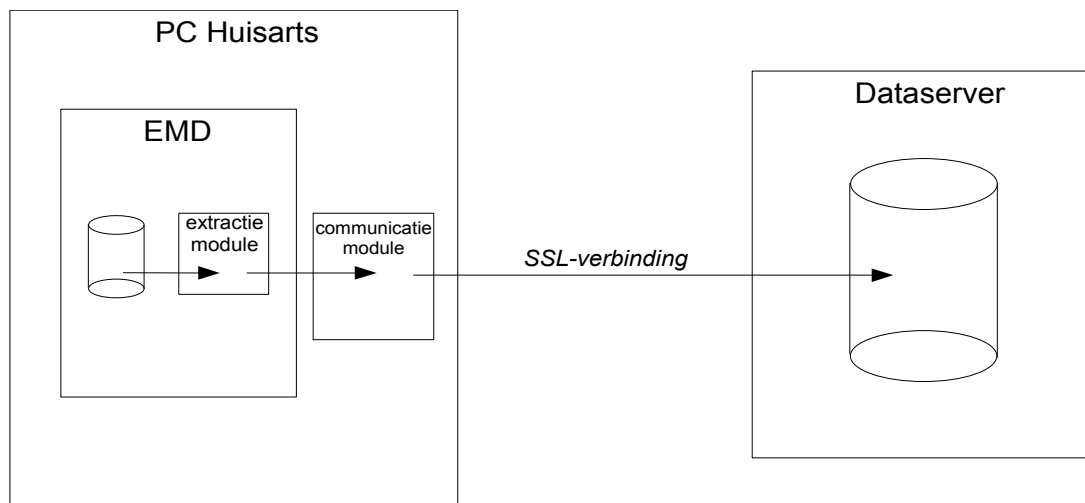
De dataextractie moet inspelen op deze problematiek. Daarom zijn er verschillende mogelijkheden besproken, met een opsomming van de voor- en nadelen. Later moeten deze afgetoetst worden bij de gebruikers in haalbaarheid en aanvaardbaarheid.

1. dataextractie door de EMD-software en transfer via een extern programma naar server

Dit is een reeds gebruikt systeem in het ResoPrim project. Er wordt aan de EMD-software gevraagd om zelf een extractiemodule in te bouwen dat zorgt voor een extractie uit de EMD van de huisarts. De module dropt de data naar de communicatiemodule buiten het EMD. Deze encrypteert de data en stuurt deze naar de server waar het gedecrypteerd wordt.

Samengevat:

- extractiemodule in het EMD gemaakt door de softwarevendor van het EMD-pakket
- communicatiemodule geïnstalleerd op de PC van de huisarts, geleverd door de onderzoeksgroep



Voordelen:

- reeds bestaand systeem: minimale interventie nodig aan het extern programma
- vertrouwd door een groep gebruikers

Nadelen:

- ResoPrim project heeft heel wat problemen gehad inzake de kwaliteit van de toegeleverde data door de extractiemodule toegeleverd
- extractiemodule moet opnieuw gemaakt worden, aangezien de extractie anders is dan het ResoPrim project
- updateprobleem: bij toevoeging en/of aanpassingen aan de extractiemodule moet elke eindgebruiker zijn softwarepakket updaten

2. Dataextractie en transfer door een externe extractiemodule ingevoegd in het EMD van de huisarts

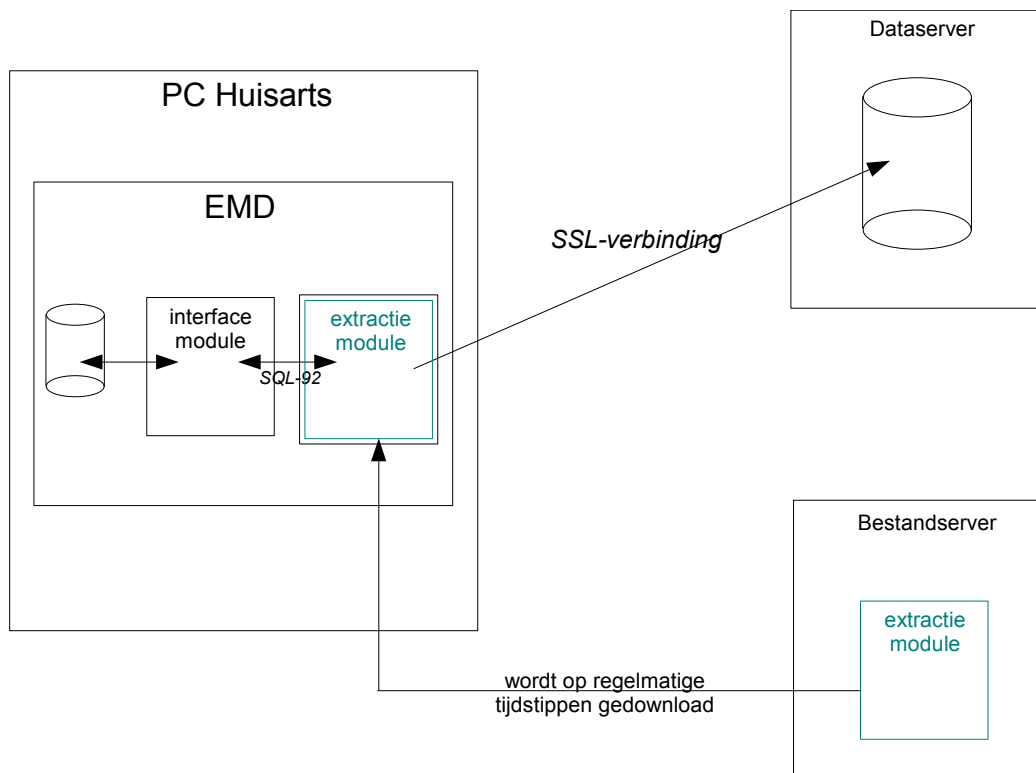
Op basis van de ervaringen in het ResoPrim project, is het volgende voorstel geformuleerd inzake de extractiemodule. De extractiemodule wordt geleverd door de onderzoeksgroep zelf. Deze staat steeds online beschikbaar op een server. De softwarevendor moet enkel voor de volgende modaliteiten zorgen:

- bij elke opstart controleren via de server of de laatste extractiemodule geïnstalleerd is
- een module inbouwen dat de interface verzorgt tussen de extractiemodule en de database van het softwarepakket. Er moet wel voor gezorgd worden dat de vragen die de externe extractiemodule stelt van generieke aard zijn. Het zullen SQL-92¹ statements zijn, echter de objecten moeten nog generiek gedefinieerd worden (patienid, diagnose, labowaarde, patientcontact, ...). Dit vergt het meeste werk!

De extractiemodule kan in dit geval ook de communicatie met de server zorgen op geëncrypteerde wijze.

Samengevat:

- externe extractiemodule die steeds online kan geupdated worden, verzorgt tevens voor de datatransfer naar de server
- interne interfacemodule door de softwarevendor gemaakt en die de communicatie verzorgt tussen de extractiemodule en de eigen database



Voordelen:

- systeem kan gebruikt worden voor meerdere projecten of voor een algemene registratie
- men is niet gebonden aan de softwarevendor voor veranderingen in de extractiemodule

Nadelen:

- interfacemodule zal veel werk vergen. Er moet een nationaal erkende standaard komen voor de definities van de objecten, opdat dit overal hetzelfde is. Men kan zich hierbij leiden aan de hand van de nationale DPRS norm.

3. Datatransfer voor thematische projecten door middel van een online invulformulier geactiveerd door een knop binnen het EMD.

Dit is eigenlijk geen vorm van extractie uit de database van het EMD, maar het registreren van data tijdens één contact tussen huisarts en patiënt in het kader van een zorgproject (zorgpaden, ...).

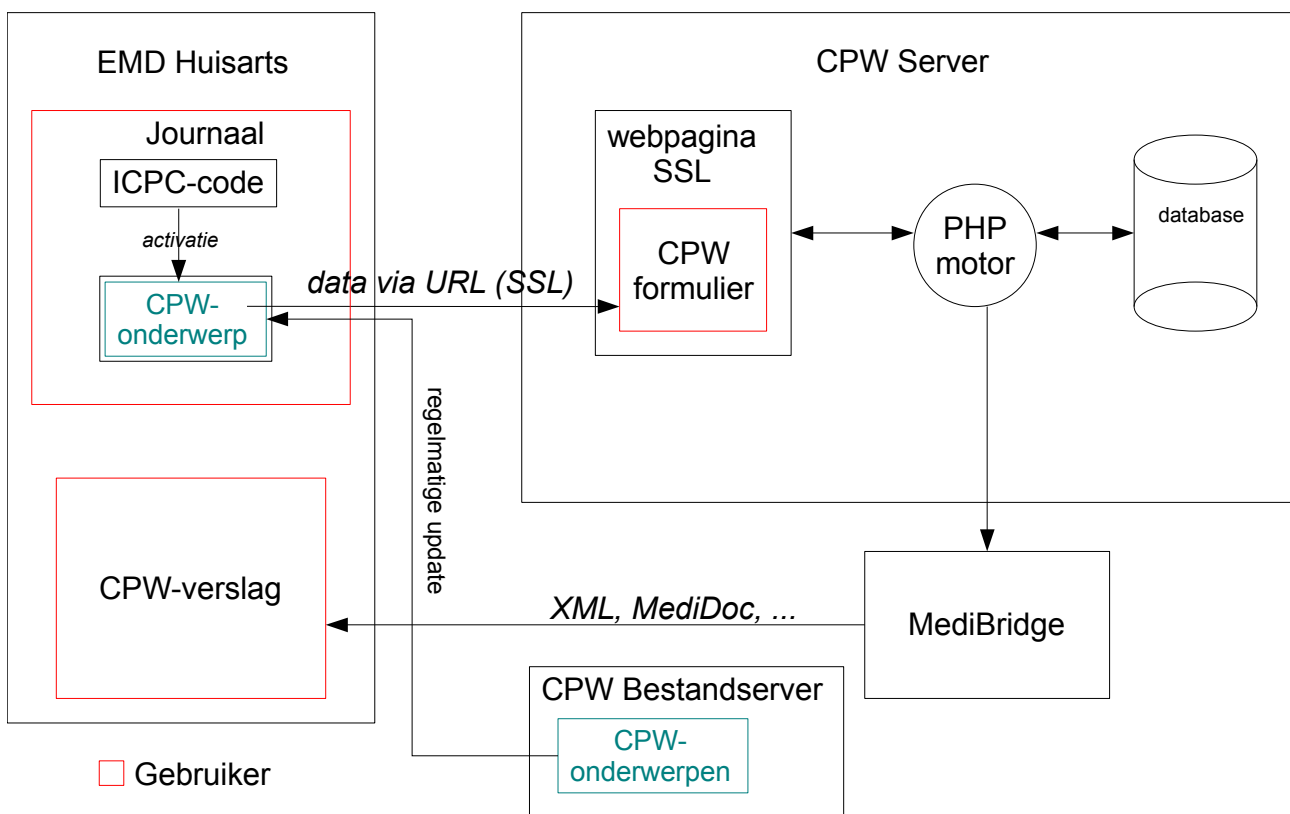
Indien tijdens een contact de huisarts merkt dat dit contact in aanmerking komt voor een bepaald zorgproject, kan hij in zijn EMD onder de noemer CPW (CarePathWay) het onderwerp kiezen dat in aanmerking komt. Dit kan ook automatisch geactiveerd worden aan de hand van de ICPC code die de arts geeft voor het consult. Bij het kiezen van het onderwerp, activeert het EMD een URL die de webbrowser van de huisarts opent. In deze URL zit al een reeks minimale data dat de webserver kan extraheren via de GET-methode. De benoeming van deze minimale data zal ook nationaal gestandaardiseerd moeten worden (patient-id, patientnaam, arts-id, ...). De huisarts vult dit formulier in en de server bewaart de noodzakelijke data dat nodig is voor het zorgproject. Tegelijk stuurt het een kopij van het invulformulier terug naar het EMD van de huisarts via de klassieke communicatieweg (MediBridge, ...). Het EMD heeft dus alle data.

De CPW-module moet door de softwarevendor ingebouwd worden in het EMD van de huisarts. Dit is redelijk eenvoudig. De CPW-module wordt bij elke opstart automatisch geupdated aan de hand van een XML-bestand dat beschikbaar is op het internet. Het XML-bestand bevat de volgende onderdelen:

- naam zorgproject
- URL van de server van het zorgproject
- minimale data die de URL moet bevatten
- reeks ICPC-codes wanneer CPW-onderwerp moet geactiveerd worden
- optioneel: standaardtekst dat in het EMD moet staan wanneer gebruik gemaakt wordt van zorgproject

Samengevat:

- softwarevendor van EMD maakt een CPW-module dat inhoudelijk geupdated wordt door middel van een extern XML-bestand
- bij een contact dat behoort tot een bepaald zorgproject, wordt de gebruiker via een URL (dat ook data bevat) verwezen naar een beveiligd extern webformulier. Daar wordt de data door de huisarts ingevuld.
- een kopij van de inhoud van het formulier wordt teruggestuurd naar het EMD van de huisarts via de klassieke kanalen (MediBridge, ...). Deze inhoud is al dan niet gestructureerd.
- de data wordt bewaard op de CPW-server, voor zoverre ze noodzakelijk zijn voor het zorgproject



Voordelen:

- weinig interventie nodig voor de softwarevendor van het EMD-pakket: CPW-module is makkelijk te maken.
- Inhoud van de CPW-module is flexibel en wordt geleverd extern.

Nadelen:

- is enkel per zorgproject mogelijk. Het is immers een contactregistratie en geen dataextractie uit het EMD. De data is aldus beperkt.

4. Dataextractie uit het EMD van de huisarts op basis van een gestandaardiseerde XML-database in het EMD van de huisarts.

4.1. maken van een gestandaardiseerde XML-database (generiek XML dossier)

In theorie moet elk gelabeld EMD episodegerigistreerd kunnen werken. Elk contact zou ook een SOEP-registratie moeten hebben, alsook verschillende van deze elementen zouden moeten kunnen geclassificeerd worden in ICPC en ICD. We weten dat in de praktijk deze onderdelen in lagen boven de vroegere basislaag is gebouwd van het EMD en dit daardoor niet goed werkt.

Nu, in het kader van eenduidige gegevensuitwisseling voor de continuïteit van zorgen, vragen we de softwareontwikkelaars:

- niet dat ze hun EMD-layout éénduidig maken
- maar wel dat de dataextractie uit hun EMD op een generiek wijze kan uitgevoerd worden.

Het formaat moet nog verder uitgedacht worden, maar minimaal moeten de SOEP-elementen in, alsook de mogelijk coderingen, in de structuur van het 'health care episode'.

4.2. transformatie van het generiek XML dossier voor extractie

Het hele XML dossier is niet nodig voor elk project/onderzoek/samenwerking om getransfereerd te worden naar de buitenwereld.

Voor elk project/onderzoek/samenwerking krijgt de gebruiker een generiek **Hermes-bestand**. Dat pakket is een zip gecomprimeerd bestand, dat verplicht de volgende elementen bevat:

- XSLT transformatie bestand
- Public Key van de ontvanger
- URL-adres van de ontvanger
- elektronisch getekend hash-bestand van de hele inhoud

Het XSLT-bestand extraheert de gegevens uit het generiek XML-dossier die nodig zijn voor het project/onderzoek/samenwerking. Men krijgt dus een mini-dossier dat klaar staat om verzonden te worden.ⁱ

Er zijn nog vragen die nog verder uitgeklaard moeten worden voor deze extractie:

- welke patienten moeten geselecteerd worden voor extractie?
 - automatisch volgen algoritme in Hermes-bestand? (vb: alle diabetes patiënten)
 - manueel door de huisarts (bij integreren van Hermes-pakket in EMD, komt er automatisch de optie voor de huisarts om patient te includeren in het dossier van de patient)
- moet er in algoritme in het Hermes-bestand zitten dat er automatisch (om de zoveel tijd, of bij elke verandering in het EMD) een extractie gedaan wordt?
- ...

ⁱ In dezelfde lijn kan het XSLT bestand een SumEHR maken.

Het is de bedoeling dat hetzelfde Hermes-bestand in elke 'merk' van EMD kan geïntegreerd worden, alsook dat een EMD verschillende Hermes-bestanden kan bevatten (zorgpaden, wacht dienst, profielen, ...)

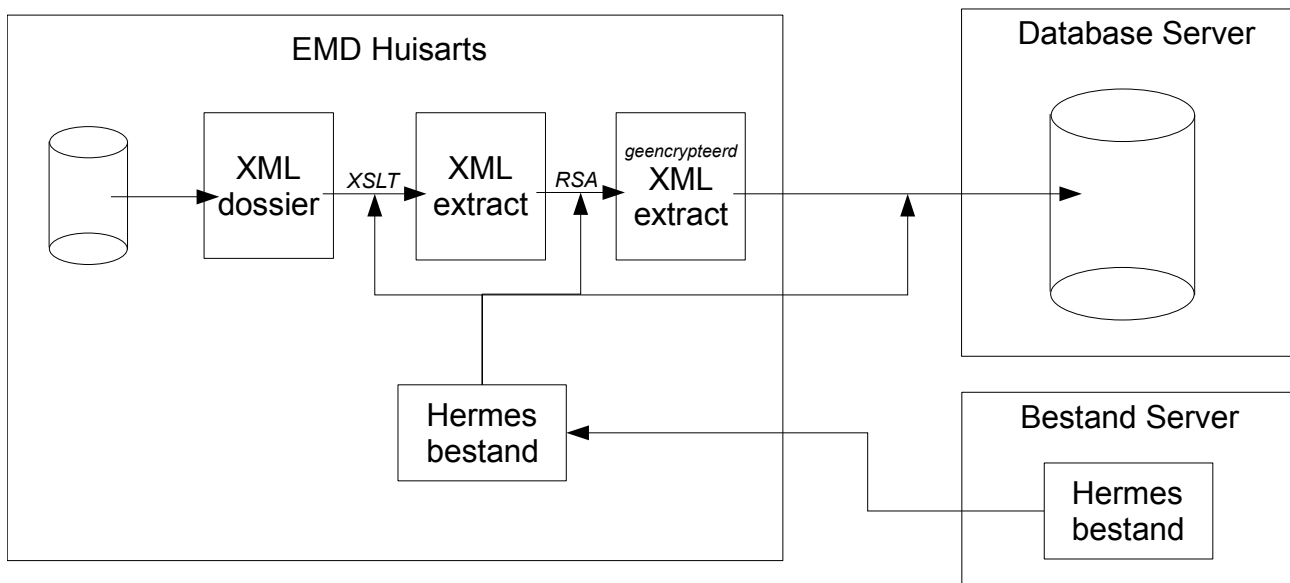
4.3. generieke verzendmodule

Deze module verzorgt de verzending naar de ontvanger van het project/onderzoek/samenwerking. Dit gebeurt via een generiek webservice protocol dat de nationale richtlijnen volgt. Het EMD zelf zorgt voor de verzending.

De verzendmodule encrypteert de extractie met de public key uit het Hermes-bestand en verzend dit naar het URL-adres dat ook in het Hermes-bestand staat.

Samengevat:

- softwarevendor zorgt dat hele dossier kan geëxtraheerd worden in een XML-database. Deze database volgt een nationaal erkend xls-schema. De structuur van de XML-database heeft als leidraad de DPRS-norm.
- softwarevendor zorgt voor een module dat een Hermes-bestand kan integreren. Het moet mogelijk zijn verschillende Hermes-bestanden te integreren, elk volgens hun eigen project: algemene registratie, projectregistratie, ... De module maakt een extract uit XML-database volgens de richtlijn in het Hermes-bestand en verzend deze naar de databaseserver volgens de modaliteiten in hetzelfde Hermes-bestand



Voordelen:

- generieke en zeer eenvoudige vorm van extractie
- zeer flexibel: mogelijk voor projectregistratie (zorgpaden), alsook algemene extractie;
- type extractie hangt alleen af van Hermes-bestand. Geen andere veranderingen noodzakelijk
- aangezien het XML-dossier de DPRS norm volgt, zal dit de softwarevenders meer verplichten om hun systeem meer DPRS compatiebel te maken, wat vermoedelijk een betere registratie mogelijk maakt

Nadelen:

- de nationale normen van het XML-dossier moeten gemaakt worden: dit kan zeer uitgebreid (tot BD-resultaten, gewicht, laboresultaten, ...), maar aangeraden is om eenvoudig te beginnen (episodegeregistreerd SOEP)
- aangezien de huidige EMD's de verplichte DPRS-norm aangepast hebben aan hun eigen interpretatie, zal het heel moeilijk zijn om een degelijk XML-dossier te kunnen extraheren.

1 <http://www.contrib.andrew.cmu.edu/~shadow/sql/sql1992.txt>