

## Archetypes: sleutel voor semantische interoperabiliteit tussen EPD's?

Herman Lodder, Bertie Zwetsloot-Schonk  
Klinische Informatiekunde LUMC, Leiden

### Inleiding

Stel, het is 2010 en je bent clinicus in een academisch ziekenhuis. Eindelijk is een lang gekoesterde droom waarheid geworden: van al jouw patiënten zijn alle gegevens beschikbaar die elektronisch zijn opgeslagen, zowel in het ziekenhuis als daarbuiten. Naast dit zogeheten EPD (Elektronisch Patiënten Dossier) zijn ook alle benodigde guidelines, protocollen en classificatiesystemen online benaderbaar.

Security is geen discussiepunt meer en performanceproblemen behoren tot het verleden, alle gegevens staan met een druk op de knop op het scherm. Het lijkt een ideale situatie en je bent al zo aan de nieuwe situatie gewend dat het niet kunnen vinden van een papieren dossier al ver verleden tijd lijkt.

Maar zoals altijd, nieuwe mogelijkheden creëren nieuwe wensen en geven meestal ook aanleiding tot nieuwe vragen en problemen die daarvoor niet of nauwelijks aan de orde waren. Een paar voorbeelden:

- Hoe weet ik of de gegevens van elders betrouwbaar of volledig zijn?
- Ik zou het liefst gegevens laten valideren op het moment van data entry of data retrieval
- Is het mogelijk de vele aangeboden gegevens intelligent te filteren zodat ik alleen te zien krijg wat voor mij relevant is?
- Als ik toch eens vragen zou kunnen stellen aan het systeem op het niveau van klinische concepten in plaats van alleen over onderliggende gegevens, bijvoorbeeld: "heeft deze patiënt nier-insufficiëntie?"
- Mijn 'ideale' EPD moet zonder grote inspanningen aangepast kunnen worden aan de laatste medische inzichten. Bijvoorbeeld waar ik twee jaar eerder nog bij klacht X geen speciale aandacht besteedde aan de familie anamnese doe ik dat nu wel en vraag ik nu bovendien altijd een eiwit spectrum aan.

Hét probleem dat alle bovenstaande punten gemeenschappelijk hebben en waarvoor een oplossing moet worden gevonden is het realiseren van **semantische interoperabiliteit**. In andere woorden: hoe kunnen we garanderen dat gebruikers van een EPD de daarin samengebrachte gegevens 'veilig' kunnen gebruiken binnen hun eigen context? Immers de context waarin en het probleem waarvoor men het EPD raadpleegt hoeft helemaal niet dezelfde te zijn als de context en vraagstelling waarbinnen de gegevens primair zijn vastgelegd. De goede boodschap is dat er – al in 2007 – een Europese norm<sup>1</sup> beschikbaar is gekomen die mogelijk een oplossing biedt voor bovenstaande vragen en problemen. Sleutelwoord hierin is het begrip **archetype**, de achterliggende benadering heet **two-level modelling**.

In de volgende paragrafen wordt kort verteld wat onder een EPD (Engels: EHR – Electronic Health Record) wordt verstaan, aan welke toepassingen je kunt denken, welke eisen dit stelt aan de structuur en inhoud van een EPD, en hoe two-level modelling en archetypes semantische operabiliteit kunnen bewerkstelligen.

---

<sup>1</sup> EN 13606 "EHR Communication"

## **Wat is een EPD?**

Hoewel we in de literatuur vele definities tegen komen voor een EPD (EHR), hanteren we in dit document de op de ISO TR 20514 gebaseerde definitie:

**Een EPD is een verzameling persoons gegevens over de gezondheidstoestand van de patiënt en zijn / haar zorgproces die door ICT - onafhankelijk van tijd en plaats - benaderbaar zijn door daartoe geautoriseerde personen.**

Wat een EPD vooral onderscheidt van de gegevens in een ZIS, HIS of ander informatiesysteem is dat het alleen gegevens over de gezondheidstoestand van de patiënt en zijn zorgproces (met inbegrip van de overdenking van de zorgverleners) bevat, en dat het om een longitudinaal dossier gaat, wat impliceert dat het om gegevens gaat die afkomstig zijn van diverse bronnen (met elk een eigen context).

Ook van belang is om te benadrukken dat niet alle gegevens die worden verzameld in een EPD terecht dienen te komen. Zo zullen van een MRI-scan uit bijvoorbeeld alle 2000 gemaakte beelden na selectie en interpretatie slechts een paar illustratieve beelden uiteindelijk in het EPD terecht komen.

Om met (de gegevens in) een EPD te kunnen werken zijn applicaties nodig, die veelal nog lokaal per zorgproces of discipline worden ontwikkeld.

Op nationale schaal wordt gewerkt aan het stapsgewijs realiseren een landelijk EPD. Het betreft hier een virtueel dossier, waaruit een zorgverlener via het zgn. landelijk schakelpunt (LSP) gegevens kan opvragen uit de dossiers van andere zorgverleners. De eerste toepassingen die onder regie van VWS worden gerealiseerd zijn het elektronisch medicatiedossier (EMD) en het waarneemdossier voor huisartsen (WDH). De voorbereidingen voor realisatie van andere toepassingen<sup>2</sup> zijn reeds in gang gezet.

## **Van one-level naar two-level modelling**

Net zoals elk (medisch) informatiesysteem dient ook een EPD te zijn gebaseerd op een informatiemodel dat alle voor registratie en communicatie relevante concepten in hun samenhang beschrijft. Het gaat daarbij om concepten uit het medische domein (zoals systolische bloeddruk), om niet-medische concepten (zoals adres), en om meer generieke concepten (zoals tekstveld of folder). Het model beschrijft bijvoorbeeld dat een adres bestaat uit straatnaam, huisnummer, postcode en woonplaats en dat deze gegevens zowel bij patiënten als bij zorgverleners kunnen worden geregistreerd. Alle concepten van het informatiemodel zijn tijdens de ontwikkelfase hard gecodeerd in de software en de databases van het informatiesysteem. Dit werkt prima zolang het informatiemodel stabiel is. Verandert er een concept of komt er een concept bij dan moet – om de nieuwe situatie optimaal te kunnen blijven ondersteunen – zowel de software als de databases worden aangepast. Zoals bekend betekent het opleveren van een nieuwe softwareversie bij grote systemen meestal een grote inspanning, zowel qua tijd, geld als menskracht.

---

<sup>2</sup> Zie [www.nictiz.nl](http://www.nictiz.nl)

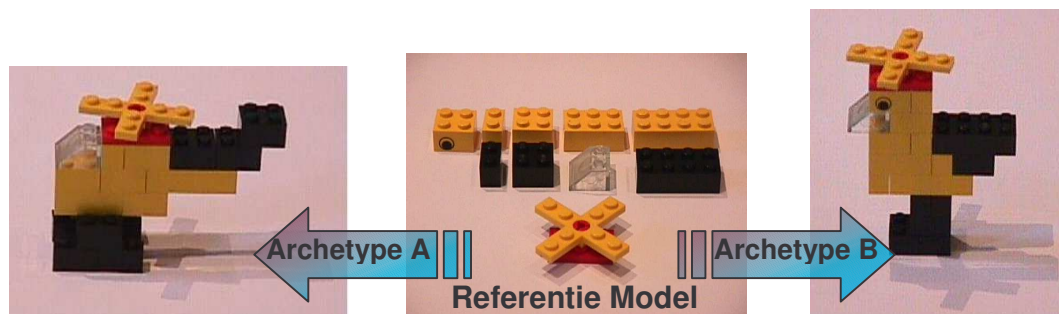
Nu hebben we bij een EPD te maken bij twee schijnbaar tegenstrijdige eisen. Aan de ene kant moeten door het longitudinale karakter van het EPD de datastructuren stabiel zijn zodat zowel gegevens uit verleden, heden als toekomst in hetzelfde EPD kunnen worden vastgelegd. Tegelijkertijd is flexibiliteit noodzakelijk in een zo dynamisch domein als dat van de gezondheidszorg waarin medische kennis, zorgprocessen en 'best practices' juist sterk aan verandering onderhevig zijn.

In de eerder genoemde Europese standaard wordt "two-level modelling" beschreven als manier om aan beide tegenstrijdige eisen tegemoet te komen. In plaats van één informatiemodel dat alle concepten beschrijft, wordt gebruik gemaakt van twee informatiemodellen:

1. een stabiel zogeheten **Referentie Model (RM)** en
2. een flexibel model dat **Archetype Model (AM)** wordt genoemd.

Het - relatief kleine - RM beschrijft alle basis concepten die in een EPD kunnen voorkomen. Dit model kan relatief klein zijn omdat het geen invulling geeft aan de medische domein concepten, daarom is het niet of nauwelijks aan verandering onderhevig en hoeft het in principe maar éénmaal te worden geïmplementeerd. Het AM bestaat uit zogenoemde Archetypes, die medisch relevante concepten definiëren door vast te leggen uit welke RM basis concepten een dergelijk medisch concept is opgebouwd en welke waarden die concepten mogen aannemen. Een voorbeeld van een eenvoudig archetype is een bloeddrukmeting, voorbeelden van meer complexe archetypes zijn medicatie order of familie anamnese.

Om de relatie tussen beide modellen duidelijk te maken wordt als analogon wel Lego gebruikt, zie figuur 2. Het RM is dan de verzameling basis Lego steentjes, het AM bevat de bouwbeschrijvingen om van de basis steentjes zinvolle constructies te maken. Een bouwbeschrijving (archetype) kan worden gewijzigd of een nieuwe kan worden toegevoegd zonder de doos met basis Lego steentjes te hoeven wijzigen.



**Figuur 1: Archetype als bouwbeschrijving in het Lego analogon**

Sterk punt van deze opzet is dat bouwbeschrijvingen kunnen worden opgesteld door diegenen die deskundig zijn in het betreffende domein. Dus bijvoorbeeld in het geval van concepten over vogels kunnen de archetypes het beste door ornithologen worden gemaakt (en beheerd) en ingeval van medische concepten door artsen en niet door de ingenieurs die de basis steentjes hebben ontworpen.

## Archetypes nader bekeken

Een meer formele definitie<sup>3</sup> van wat we hiervoor vergeleken met een Lego bouwbeschrijving luidt:

**An archetype is a computable expression of a domain content model in the form of structured constraint statements, based on some reference model.**

**Main features of an archetype are its re-usability and context-neutrality.**

In deze definitie gaat het dus om:

- Medisch relevante concepten, die niet context-gebonden zijn en daarom breed toepasbaar zijn. Om deze reden wordt per archetype ook gezocht naar de maximale data set (bouwstenen) die van toepassing is op het bijbehorende medische concept. Via specialisatie kunnen indien lokaal noodzakelijk bepaalde data worden weggelaten of worden toegevoegd.
- Een referentiemodel waarin de concepten (en hun onderlinge relaties) worden beschreven en waar het archetype gebruik van maakt.
- Constraints, een soort rules die bepalen uit welke bouwstenen (RM concepten) een archetype is opgebouwd, of deze verplicht of optioneel zijn, welke waarden deze concepten mogen aannemen, of deze herhaald mogen voorkomen, en welke terminologie term ermee is geassocieerd.
- Een representatie die door de computer verwerkt kan worden en daarom gebruik maakt van een formele taal<sup>4</sup>.

Om ervoor te zorgen dat niet overal hetzelfde wiel hoeft te worden uitgevonden wordt er gewerkt aan een **bibliotheek** van standaard archetypes, die voldoen aan de volgende ontwerpeisen:

- Volledigheid ('wholeness'); de informatie in een archetype moet zó volledig zijn dat interpretatie ervan context-neutraal kan plaatsvinden. Om deze reden bestaat een archetype uit een maximale data set, waarvan overigens het merendeel van de data optioneel is.
- Eén-op-één relatie met een klinisch concept ('discreteness'); ten behoeve van eenduidigheid en de mogelijkheid tot hergebruik moet zoveel mogelijk worden voorkomen dat archetypes overlappen. Er zijn mogelijkheden om in een lokale situatie te werken met samengestelde archetypes (zie hieronder) en met specialisaties van archetypes.

Om ook echt te kunnen voldoen aan deze ontwerpeisen moeten standaard archetypes via een "quality assurance and change control" proces door de (internationale) medische beroepsgroep worden ontwikkeld en beheerd.

## Templates

Omdat medische gegevens in de dagelijkse praktijk niet op archetype niveau worden uitgewisseld maar door middel van invulformulieren, rapporten of berichten is het begrip **template** geïntroduceerd. Een template bundelt één of meer archetypes, en kan

<sup>3</sup> Uit: "Archetype Definitions and Principles" – Thomas Beale, Sam Heard

<sup>4</sup> Hiervoor is de zogeheten ADL (Archetype Development Language) ontwikkeld.

het gebruik ervan geheel toesnijden op de lokale situatie, bijvoorbeeld door default waarden toe te kennen, door extra verplichte velden toe te voegen of door te definiëren welke terminologie (subset) mag worden gebruikt.

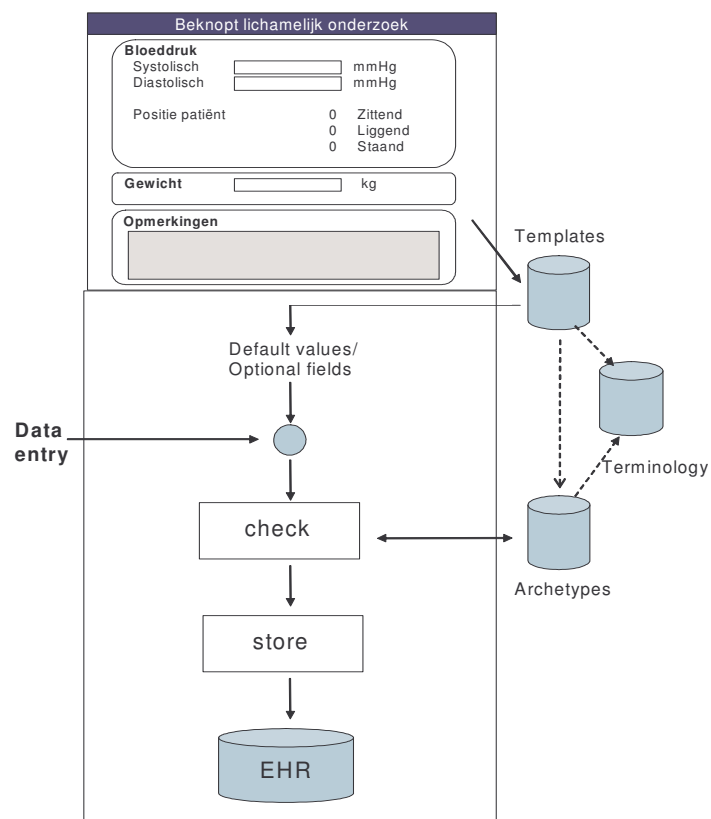
Templates worden daarom dus lokaal ontwikkeld, gebruikt en beheerd in tegenstelling tot de eerder genoemde standaard archetypes.

## Hoe ‘werken’ archetypes en templates?

In het voorgaande is aangegeven wat een archetype is en dat archetypes via templates kunnen worden samengevoegd en aangepast voor gebruik in de lokale situatie. In deze paragraaf gaan we nader in op drie belangrijke toepassingen van two-level modellering: A. data-entry validatie, B. archetyped-based (intelligent) querying, en C. mogelijk maken van semantische interoperabiliteit

### A. Data entry validatie

We gaan als voorbeeld uit van een situatie waarin de arts een aantal bevindingen van een beknopt lichamenlijk onderzoek wil invoeren (zie figuur 3). Hij maakt daarbij gebruik van een invoerscherm waarop gegevens over bloeddruk, gewicht en eventuele aanvullende opmerkingen kunnen worden ingevuld. In een systeem dat is gebaseerd op archetypes zal het invoerscherm corresponderen met een template ‘beknopt lichamenlijk onderzoek’. Deze template bevat en is een samenvoeging van de archetypes ‘bloeddruk’, ‘gewicht’ en ‘opmerkingen’. Het invoeren



**Figuur 3: gebruik templates/archetypes bij data entry**

van de bloeddruk gegevens gebeurt onder controle van het archetype ‘bloeddruk’, wat als groot voordeel heeft dat de gegevens die worden opgeslagen zowel qua structuur als ook qua mogelijke waarden en geassocieerde terminologie geheel voldoen aan wat dat archetype voorschrijft. Deze domeinkennis (in templates, archetypes en terminology repositories) wordt buiten de programmatuur ontwikkeld en onderhouden, en tijdens run-time aangeroepen. De medici kunnen deze domein kennis zelfstandig, dus zonder tussenkomst van automatiseerders aanpassen en onderhouden.

*B. Intelligent zoeken in data (archetyped-based querying)*

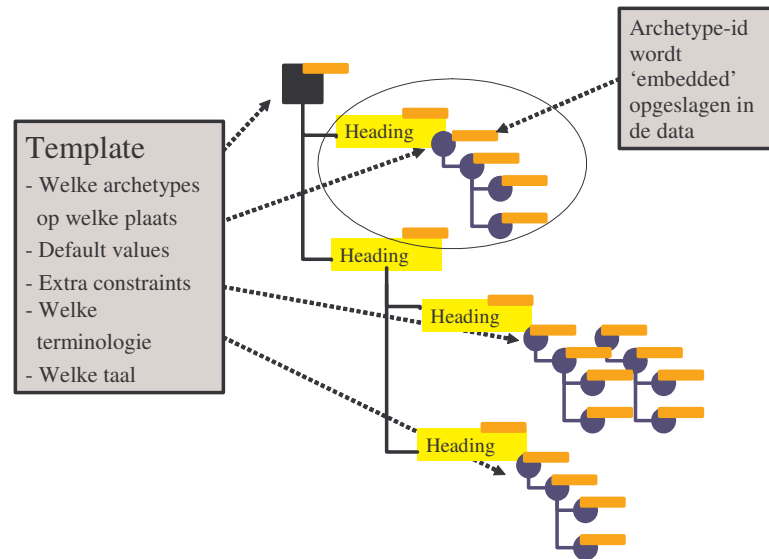
Op het moment dat datastructuren worden gecreëerd volgens de specificatie van een template (of archetype) wordt van elke data-item ‘embedded’ opgeslagen welk deel van welke archetype binnen het template hierop betrekking heeft (zie figuur 4). Zo is bij de systolische bloeddruk in het eerder gebruikte voorbeeld vastgelegd dat het gaat om het eerste item in archetype ‘bloeddruk’ en dat het om een verplicht item gaat.

Door van deze kennis

gebruik te maken kan vooral bij het zoeken in grote hoeveelheden data grote winst worden behaald in vergelijking met zoeken volgens de “brute force” methode.

Zo zal bij het zoeken naar “de laatst vastgelegde systolische bloeddruk in een EHR transactie” de volgende stappen worden doorlopen:

- Bepaal alle archetypes waarin ‘systolische bloeddruk’ voorkomt via een query op de archetype repository
- Doorloop alle EHR transacties te beginnen met de meest recente en stop wanneer er een match is met de eerder gevonden archetype(s)
- Haal binnen deze EHR transactie het item op dat binnen het archetype is gedefinieerd als ‘systolische bloeddruk’.



**Figuur 4: data elementen direct geassocieerd met bijbehorend archetype-path (archetype-id + plaats binnen archetype)**

*C. mogelijk maken van semantische interoperabiliteit*

Informatiesystemen gebaseerd op een referentiemodel en waarin de domein concepten expliciet zijn gedefinieerd m.b.v. archetypes, bieden niet alleen voor menselijke gebruikers duidelijke meerwaarde ten opzichte van hun voorgangers, ze kunnen ook worden ingezet ten behoeve van bijvoorbeeld decision support en statistische toepassingen, juist omdat de betekenis van de medische gegevens op expliciete wijze is vastgelegd (en dus terug te vinden is) in het informatiesysteem. De onder A genoemde validatie tijdens data-entry is dan ook van toepassing bij validatie tijdens data-retrieval.

Er zijn verschillende scenario's denkbaar om tussen twee EPD systemen gegevens uit te wisselen met behulp van archetypes<sup>5</sup> (zie onderstaande tabel):

<sup>5</sup> Gebaseerd op een presentatie door Dipak Kalra



Systeem 1 Archetype enabled j/n	EHR_Extract (EN 13606)	Systeem 2 Archetype enabled j/n	Mate van Semantische interoperabiliteit
j	Conform archetype A	j	Maximaal *
j	Conform archetype A	n	Beperkt
n	Conform local clinical schema	j	Minimaal, echter via mapping naar een bestaand of nieuw archetype is verdere verwerking mogelijk
n	Conform local clinical schema	n	Minimaal

\* Systeem 2 kan archetype A al lokaal beschikbaar hebben, bij een centrale archetype repository opvragen, of direct bij systeem 1 opvragen (welke dan in feite dienst doet als centrale archetype repository).

### Wat is de huidige stand van zaken?

Vooral binnen de openEHR Community (in 2006 al 648 leden in 63 landen) wordt hard verder ontwikkeld aan de standaards en (in samenwerking met industriële partners) aan tools om grootschalige implementaties te faciliteren. Het feit dat Europa gekozen heeft voor de EN13606 norm is een grote stap vooruit. Ook het HL7 consortium lijkt steeds meer geneigd om archetypes een plaats te geven in hun methodologie. Software bedrijf Ocean Informatics (van Thomas Beale en Sam Heard, de geestelijke vaders van het two-level paradigma) heeft commerciële en research activiteiten in tal van landen waaronder Nederland. Afgelopen april is in het LUMC een tweedaagse workshop gehouden waarin onder andere Thomas Beale docent was. In Spanje wordt in het kader van het LinkEHR project gewerkt aan een tool met behulp waarvan gegevens uit bestaande informatiesystemen kunnen worden beschreven en getransformeerd met behulp van archetypes.

Van groot belang binnen Nederland is de door de OIZ op 20 juni 2007 georganiseerde conferentie waarin voor het eerst openlijk een discussie plaatsvond over de te kiezen EPD-norm (HL7 of EN13606). Recent is in dit kader door NICTIZ een opdracht aan TNO verstrekt om een vergelijking tussen beide normen in kaart te brengen.

### Literatuur

- Beale T. *Archetypes: Constraint-based Domain Models for Future-proof Information Systems*. OOPSLA workshop on behavioural semantics, 2002.
- David Moner et al. *Archetype-Based Semantic Integration and Standardization of Clinical Data*, In: Proc. Of the 28<sup>th</sup> IEEE EMBS Annual Int. Conf, NY City, Aug 30-Sept 3, 2006.
- Eichelberg M. *EHR Standards and Semantic Interoperability*. E-Health Conference, Berlin, April 17-19, 2007.