

Pilotproject Integratie Terminologieserver (basis versie)

**Uitgave:
Stichting CBV**

September 2004

Aan deze uitgave werkten mee:

Dhr. R.E.P. Hamoen, MSc, Stichting CBV

Dhr. Dr. Z.J. de Langen, bestuurslid Stichting CBV

Dhr. J.M. Laurs, secretaris Stichting CBV

Mw. J. Pos, Registratie Advies Commissie Stichting CBV

Dhr. Drs. P.G.J.M. Spierings, Stichting CBV

Mw. L. Vandamme, Registratie Advies Commissie Stichting CBV

Mw. Prof. Dr. J.H.M. Zwetsloot-Schonk, voorzitter Stichting CBV

Auteur:

Bureau, Stichting CBV

© 2004 Stichting CBV,
Postbus 16070, 2301 GB LEIDEN,
www.cbv.nl
tel: (071)-521 5064
e-mail: info@cbv.nl

**Alle rechten voorbehouden. Niets uit deze uitgave mag worden
verveelvoudigd, opgeslagen in een geautomatiseerd
gegevensbestand en/of openbaar gemaakt, in enige vorm of
op enige wijze, elektronisch, mechanisch, door fotokopieën,
opnamen of op enige andere manier, zonder voorafgaande
schriftelijke toestemming van de uitgever**

**Voor het overnemen van (een) gedeelte(n) uit deze uitgave in
bloemlezingen, readers en andere compilatiewerken (artikel 16
Auteurswet 1912) dient met zich tot de uitgever te wenden.**

Inhoudsopgave

Voorwoord	4
1 Inleiding.....	5
1.1 Probleemstelling	5
1.2 Opdracht en doelstelling pilotproject	6
2 Werkwijze.....	7
3 Resultaten.....	9
3.1 Inrichten pilot-omgeving	9
3.2 Modelleren CBV- en CvV-verrichtingen.....	9
3.3 Het formeel modelleren van de CBV- en CvV-verrichtingen	11
3.3.1 Efforts	11
3.3.2 Analyse CBV-ontologie.....	11
4 Conclusie	17
4.1 Conclusie pilotproject en indicatoren	17
4.2 Meerwaarde eindresultaat	19
4.3 Aanbevelingen	20
5 Bijlagen	21
5.1 Een formele medische ontologie: theoretische achtergrond.....	21

Voorwoord

Het bestuur van de Stichting CBV heeft in haar beleidsplan (september 2001) de pijlpalen uitgezet voor de periode 2002-2006. In het beleidsplan wordt geanticipeerd op de vele veranderingen in de zorg. Hierbij valt ondermeer te denken aan het elektronische patiëntendossier (EPD), transmuralisatie van de zorg en de verandering in de bekostigingssystematiek van de zorg. Al deze veranderingen zorgen voor veranderende eisen aan de informatievoorziening in de zorg zoals de beschikbaarheid, toegankelijkheid, uitwisselbaarheid en vergelijkbaarheid van gegevens.

Eén van de speerpunten in het beleidsplan betreft het onderzoek naar de mogelijke toepassingsmogelijkheden van een terminologiesysteem binnen de curatieve zorgsector. De verwachting is dat dergelijke systemen de taal van het primaire proces, het medische jargon, kunnen vertalen naar de diverse (gestandaardiseerde) informatiebehoeften.

Een oriëntatie op beschikbare systemen leidde naar de firma Language & Computing (L&C) dat een geïntegreerd systeem kan leveren dat bestaat uit een meertalige ontologie (inclusief de Nederlandse taal) met diverse operationele toepassingen gericht op het beheer en gebruik in de zorgpraktijk. Eind 2002 heeft de Stichting CBV een opdracht verleend aan de firma L&C tot het uitvoeren van een pilotproject. De resultaten van dat pilotproject zijn door Dhr. Dr. W. Ceusters van L&C beschreven in het rapport *“Project Integratie terminologieserver - Formaliseren van het CBV bestand via een op realisme gestoelde ontologie”* d.d. 11-06-2003.

In het onderliggende rapport treft u de verwerking en evaluatie aan van de resultaten van dit pilotproject. Het grootste deel van de verwerking en evaluatie is uitgevoerd door dhr. drs. P. Spierings (Stichting CBV) in nauwe samenwerking met enkele leden van het bestuur en Registratie Advies Commissie van de Stichting CBV. Tijdens het project is gebleken dat de achterliggende theorieën en concepten zeer complex kunnen zijn die door een degelijke uitleg begrijpbaar gemaakt moeten worden.

De basisversie en de uitgebreide versie van het rapport met de bijlagen is via de website van de Stichting CBV, www.cbv.nl, te downloaden.

Technologische en methodologische ontwikkelingen hebben langere tijd nodig om betrouwbaar en op een doelmatige manier te kunnen worden ingezet. De resultaten van dit pilotproject bieden voldoende aanknopingspunten voor verdere ontwikkelingen binnen de ‘communicatie- en informatiewereld’ van de zorg. Hierbij zijn een breed gedragen aanpak door meerdere organisaties en de uitbreiding van de kennis m.b.t. terminologieserver belangrijke aandachtspunten.

De belangrijkste resultaten van dit pilotproject zijn de bevinding dat de kwaliteit van de ‘handmatige’ aanmaak van registratiecodes door de Stichting CBV van voldoende kwalitatief niveau zijn en daarnaast van hoog detailniveau zijn (de zgn. eindcodes). De meerwaarde van een terminologieserver voor de zorgsector ligt vooral bij de efficiency- en kwaliteitswinst in de vastlegging van zorginformatie tijdens de zorgverlening door de arts. Een bijkomend voordeel is de beschikbaarheid van een ‘technisch hulpmiddel’ bij het beheer en onderhoud van classificatie-, codestelsels en terminologieën. De aanbevelingen in dit rapport zijn vooral gericht op de aansluiting van het medische jargon van de arts op de classificatie- en codestelsels. Kleinschalige pragmatische projecten, bij voorkeur in samenwerking met andere instellingen, moeten de verdere gebruikstoepassingen en het nut van een terminologieserver in de zorgpraktijk inzichtelijk maken en bewijzen.

Directeur Stichting CBV,

R.E.P. Hamoen, MSc

1 Inleiding

Het bestuur van de Stichting CBV heeft in haar beleidsplan de visie voor de periode 2002-2006 uiteengezet¹. Eén van de speerpunten is de integratie van een breed gedragen en wetenschappelijk verantwoord terminologiesysteem waarmee de betekenis van zorginhoudelijke gegevens kan worden vastgelegd, de relaties tussen tabellen met zorginhoudelijke gegevens (diagnoses, verrichtingen, zorgvragen etc.) op een consistente en efficiënte wijze kunnen worden gelegd en een flexibele bewerking en presentatie van deze gegevens mogelijk wordt.

In dit rapport worden de resultaten en de conclusies van het pilotproject, dat in samenwerking met de Stichting CBV en L&C is uitgevoerd, besproken en bediscussieerd.²

Leeswijzer:

In Hoofdstuk 1 wordt de probleemstelling en doelstelling van dit pilotproject beschreven. In Hoofdstuk 2 worden de LinkFactory®-methodiek uitgelegd. In Hoofdstuk 3 worden de resultaten van het pilotproject weergegeven en de betekenis hiervan besproken. Vanuit de Stichting CBV zijn de door L&C uitgevoerde activiteiten steekproefsgewijs beoordeeld. Tot slot worden in Hoofdstuk 4 de toepassingsmogelijkheden voor het CBV-bestand weergegeven en de conclusies geformuleerd.

Voor de leesbaarheid is het rapport voorzien van een uitgebreide bijlage waarin de verdere technische uitwerking en een uitgebreide verantwoording van de analyse is weergegeven. M.n. bijlage 0 met daarin beschreven de achterliggende ontologische theorie wordt aanbevolen.

1.1 Probleemstelling

Het CBV-bestand is een gemeenschappelijk referentiebestand ontwikkeld om de eenmalige vastlegging van zorgactiviteiten binnen ziekenhuisinstellingen te faciliteren. De toepassingsmogelijkheden van het CBV-bestand zijn legio: elektronische medische documentatie, gegevensverstreking in het kader van landelijke registraties en facturatie zijn voorbeelden. Essentieel hierbij is de geïntegreerde én dynamische aanpak binnen het CBV-bestand. Door het leggen van onderlinge relaties tussen diverse codestelsels, referentielijsten en classificaties kunnen rapportages voor derden vanuit de basisregistratie worden afgeleid waardoor de administratieve werklast afneemt. Nieuwe technieken kunnen op verzoek van (beroepsbeoefenaren uit) de aangesloten instellingen direct binnen het bestand worden opgenomen, zodat aansluiting op de dagelijkse praktijk is gewaarborgd en een volledige en gedetailleerde vastlegging mogelijk is. Als gevolg van de invoering van prestatiebekostiging (DBC's) in de ziekenhuiszorg wordt de wens om de uitgevoerde zorgactiviteiten zo volledig mogelijk vast te leggen verder versterkt.

Het CBV onderkent dat het CBV-bestand in de huidige vorm een aantal tekortkomingen kent. Een van de belangrijkste is dat de betekenis van een term zoals gebruikt binnen de omschrijving van een verrichting niet altijd eenduidig is. Door het beperkte aantal beschikbare karakters voor de omschrijving van een verrichting is het niet altijd mogelijk de betekenis volledig vast te leggen. Dit leidt er toe dat een omschrijving op verschillende manieren geïnterpreteerd zou kunnen worden. Dit kan tot gevolg hebben dat de registratie foutief verloopt, maar ook dat de relaties tussen de binnen het CBV-bestand voorkomende codestelsels, referentielijsten en classificaties mogelijk foutief worden gelegd. Door het ontbreken van een expliciete formele betekenis van de gebruikte termen van omschrijvingen is het op dit moment moeilijk om in het kader van het afleggen van verantwoording over de verleende zorg binnen het geïntegreerde medisch specialistisch bedrijf de vastgelegde gegevens op een flexibele manier te bewerken (selecteren, aggregeren en ordenen). Dit omdat de afzonderlijke medische kenmerken van een verrichting niet expliciet beschikbaar zijn.

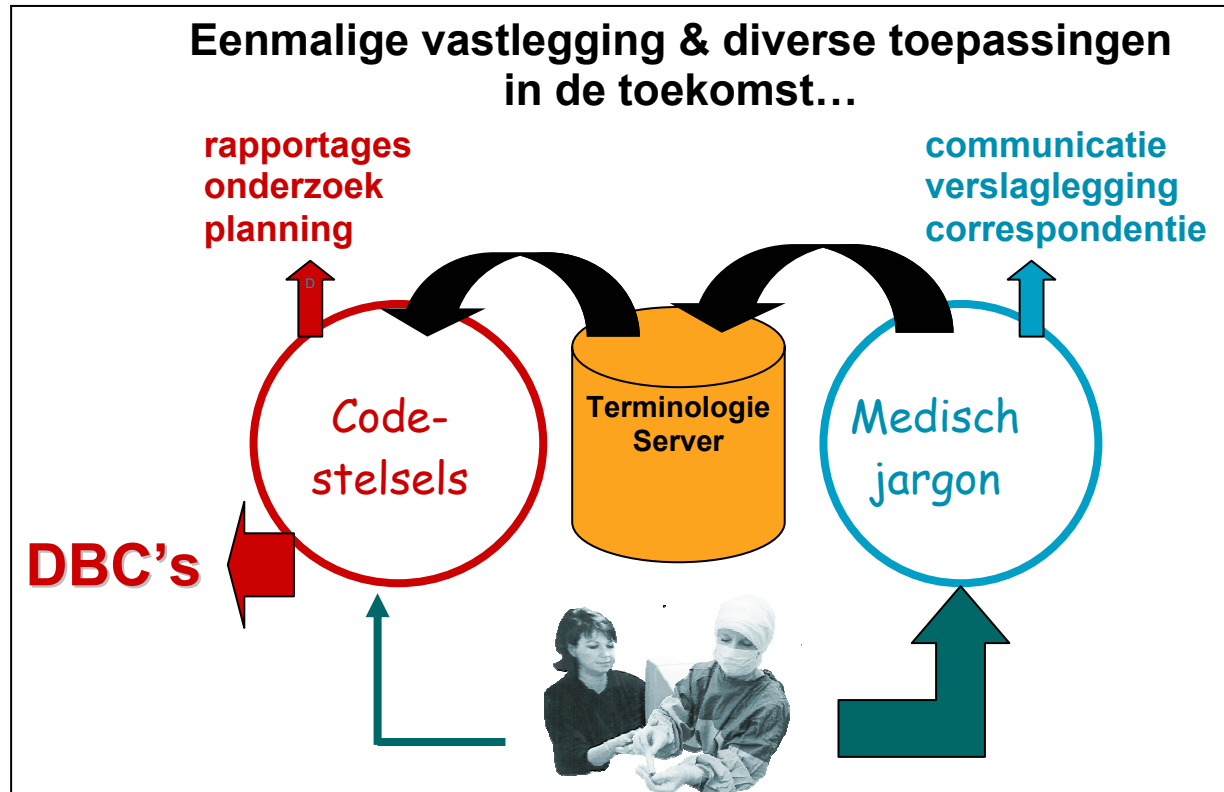
¹ Het CBV nu en in de toekomst, Uitgave Stichting CBV 2001

² Als onderlegger voor dit rapport is het L&C-rapport "Project Integratie Terminologieserver, Formaliseren van het CBV bestand via een op realisme gestoelde ontologie. Bevindingen inzake een pilootproject uitgevoerd op de deelgebieden traumatologie en urologie", dd. 11-06-2003 gebruikt.

Het voorgaande maakt duidelijk dat het voor zowel de eindgebruikers van het CBV-bestand als voor het CBV zélf van belang is te beschikken over een systeem waarmee:

1. de betekenis van de binnen het CBV-bestand voorkomende omschrijvingen formeel vastgelegd kan worden;
 2. de medische kenmerken van een verrichting expliciet beschikbaar gesteld kunnen worden.
- Een dergelijk systeem wordt een medische ontologie genoemd.

In onderstaande figuur is de plaats van de toepassing van de medische ontologie in de vorm van de terminologieserver binnen de zorgpraktijk schematisch weergegeven.



1.2 Opdracht en doelstelling pilotproject

Het Belgische bedrijf Language and Computing (L&C³) beschikt als enige commerciële organisatie binnen het Nederlands taalgebied over een meertalige ontologie en een beheersysteem waarbinnen de betekenis van medische begrippen inclusief de samenhang tussen die begrippen én de medische kenmerken van deze begrippen formeel is, of kan worden vastgelegd en gepresenteerd. Het systeem is wetenschappelijk theoretisch gefundeerd.⁴ Voor een inleiding op de achterliggende ontologische theorie wordt verwezen naar bijlage 5.1.

Het CBV heeft dan ook aan L&C de volgende drie opdrachten verstrekt:

1. Het modelleren van de verrichtingen behorende tot het domein van de specialismen Urologie en Traumatologie voor zowel het CBV- als het CvV-bestand (respectievelijk 1111 CBV- en 710 CvV-verrichtingen). Dit met als doel te komen tot een eenduidige formele omschrijving én een overzicht van de beschrijvende kenmerken per aangeleverde verrichting.

³ Meer informatie over Language and Computing treft u aan op www.LandCglobal.com

⁴ - Smith B, Ceusters W, *Towards industrial strength philosophy: how analytical ontology can help medical informatics*. Interdisciplinary Science Reviews, 2003, Vol 28, No. 2.

- Ceusters W, Martens P, Dhaen C, Terzic B, Linkfactory: *an Advanced Formal Ontology Management System*. Interactive Tools for Knowledge Capture Workshop, KCAP-2001, October 20, 2001, Victoria B.C. Canada (<http://sern.ucalgary.ca/ksi/K-CAP/K-CAP2001/>).

- Ceusters W, Smith, B, Flanagan J, *Ontology and Medical Terminology: Why Description Logics Are Not Enough*. Towards an Electronic Patient Record (TEPR 2003), San Antonio 10-14 May 2003, Boston, MA: Medical Records Institute (CD-rom publication)

2. Het leggen van een relatie tussen enerzijds de aangeleverde CBV- en anderzijds de CvV- verrichtingen uitgaande van de formele omschrijving van een verrichting op semantisch (betekenis) niveau.
3. Het beoordelen van de eenduidigheid van het CBV-bestand en de correctheid van de gelegde relaties door deze te toetsen aan de ontologie.

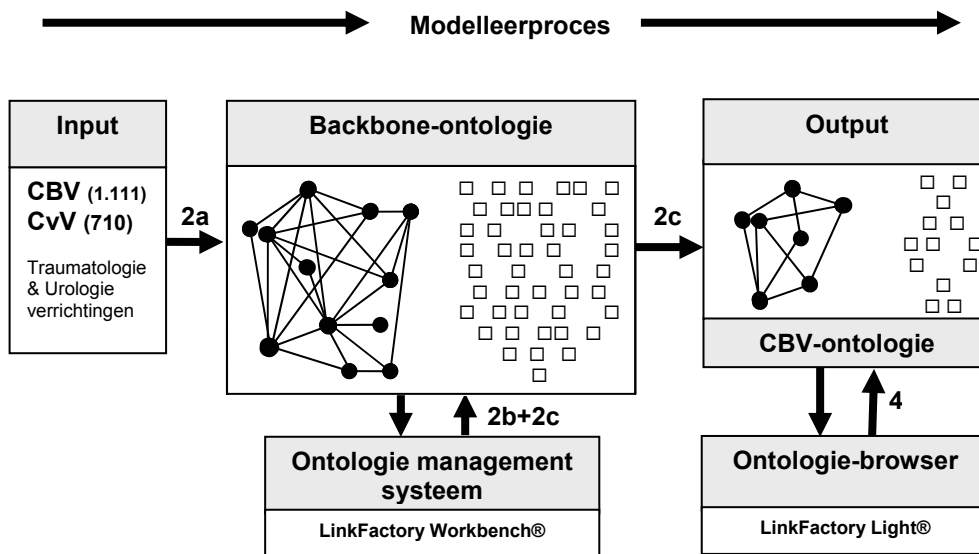
Nadat bovengenoemde specialismen zijn geformaliseerd, dient te worden bekeken in hoeverre het de registratie, ontsluiting en presentatie van zorggegevens in de praktijk verbeterd. Op grond daarvan kan de meerwaarde van het formaliseren van het volledige CBV-bestand worden afgezet tegen de kosten die gemaakt moeten worden om het gehele bestand te formaliseren.

2 Werkwijze

In het kader van het pilotproject werden de volgende deeltaken uitgevoerd en voorzieningen getroffen:

1. *Inrichten pilot-omgeving*
 - a. *Aanleveren CBV- en CvV-bestanden*
Door het CBV werden alle CBV-verrichtingen behorende tot het domein van de specialismen Urologie en Traumatologie geselecteerd. Tevens werden alle CvV-verrichtingen geselecteerd die een koppeling hadden met een van de CBV-codes. Afkortingen binnen de omschrijvingen van de selectie van CBV- en CvV-verrichtingen werden voluit geschreven en in tabelvorm aangeleverd aan L&C.
 - b. *Aanpassen bestaande L&C-software*
Voordat gestart kon worden met het inlezen van de door het CBV aangeleverde bestanden is de bestaande L&C-software aangepast. Dit om mogelijke problemen door het gebruik van diakritische tekens, speciale karakters en specifieke veldlengten te voorkomen.
 - c. *Opzetten discussieplatform*
Omdat in een reeds eerder uitgevoerde testfase bleek dat de omschrijvingen van de verrichtingen niet voldoende duidelijk waren om door derden ondubbelzinnig begrepen te kunnen worden, werd op de CBV-website een discussieplatform ingericht. Dit platform was toegankelijk voor L&C- en CBV-medewerkers en enkele (medisch-inhoudelijk)deskundigen.
 - d. *Realiseren remote-toegang*
Om inzicht te krijgen in de L&C-technologie én om de voortgang van het project te kunnen volgen werd er aan het CBV via Internet remote toegang verleend op de interne modelleromgeving van L&C.
2. *Modelleren CBV- en CvV-verrichtingen*
 - a. *Inlezen CBV- en CvV-bestanden en automatisch opzoeken semantische context*
Nadat de bestanden waren ingelezen en geverifieerd werd volledig automatisch de semantische context van alle aangeleverde verrichtingen bepaald. Dit gebeurde op basis van alle al binnen het L&C-systeem voorkomende verrichtingen, ongeacht bron en taal van herkomst (zoals Snomed CT, MEDCIN, ICD9-CM etc.).
 - b. *Modelleren backbone hiërarchie*
Om te achterhalen welke beschrijvende kenmerken nodig zijn voor het formaliseren van de CBV- en CvV-verrichtingen is uitgaande van de aangeleverde verrichtingen via de L&C-ontologie (LinkBase®) een eerste structuur bepaald: de backbone ontologie.
 - c. *Formeel modelleren CBV- en CvV-verrichtingen*
In deze fase werd elke individuele verrichting handmatig gemodelleerd met behulp van de uit de backbone ontologie afgeleide beschrijvende kenmerken met als uiteindelijk resultaat de CBV-ontologie.

De stappen 2a t/m c worden in onderstaande figuur schematisch weergegeven.



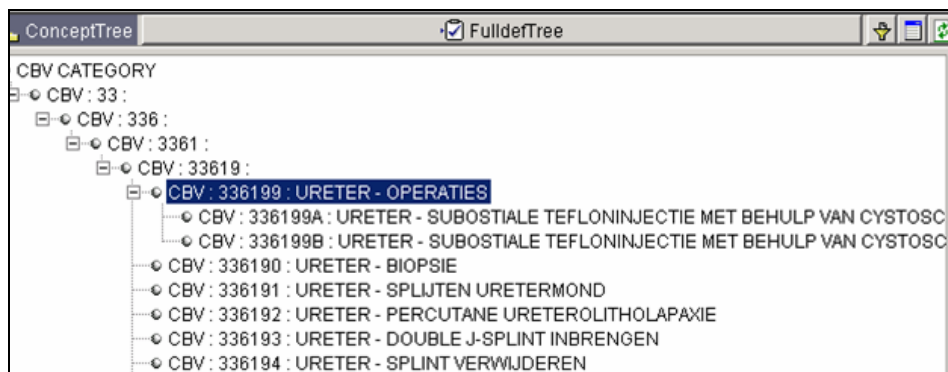
Figuur 1: Modellerproces CBV en CvV-verrichtingen

3. *Evaluatie door L&C*
De CBV-ontologie werd door L&C onderzocht op pragmatische tekortkomingen.
4. *Presentatie resultaten in demonstratie-browser*
De CBV-ontologie werd afgeleverd in een demonstratie-browser, LinkFactory Lite®, waarmee het eindresultaat op verschillende manieren ontsloten en gepresenteerd kan worden (zie figuur 1).
5. *Inzicht voor CBV-vertegenwoordigers*
Middels bijeenkomsten is aan de CBV-medewerkers, inhoudelijke deskundigen, afgevaardigden vanuit het CBV-Bestuur en de Registratie Advies Commissie (RAC) inzicht gegeven in de toegepaste methodiek en de theoretische achtergronden.
6. *Evaluatie resultaten door CBV-vertegenwoordigers*
Om te kunnen beoordelen in welke mate het formaliseren bijdraagt aan een kwalitatieve verbetering van het bestand zijn de volgende indicatoren bepaald:
 - a) Het aantal verrichtingen waarbij sprake is van doublures, spellingsfouten en niet reguliere afkortingen;
 - b) Het aantal verrichtingen waarvan de betekenis niet rechtstreeks is af te leiden uit de omschrijving en die dus voor meerdere uitleg vatbaar zijn;
 - c) Het aantal verrichtingen dat ontbreekt in het bestand wanneer gelet wordt op de uiteindelijke semantische structuur. Dit kan worden gezien als een maat voor de volledigheid van het CBV- en CvV-bestand;
 - d) Het aantal niet correct gelegde relaties tussen het CBV- en het CvV-bestand wanneer de door het CBV gelegde relatie wordt vergeleken met de relaties gelegd bij het formeel modelleren.

3 Resultaten

3.1 Inrichten pilot-omgeving

Door het CBV zijn een tweetal subsets van verrichtingen als aparte bestanden in databaseformaat aan L&C geleverd. De eerste subset bestond uit alle CBV-verrichtingen binnen het CBV-bestand waarvan bekend was dat deze behoorden tot de specialismen Traumatologie en Urologie. Dit betrof in totaal 1111 verrichtingen. De tweede subset die werd ingelezen bestond uit alle CvV-codes die gekoppeld waren aan de CBV-verrichtingen uit de eerste subset. Om de verrichtingen niet als een platte lijst te presenteren binnen de LinkFactory Workbench® (het ontologie management systeem van L&C) werd er bij het inlezen op basis van de CBV- en CvV-codering voor beide subsets van verrichtingen een aparte pseudo-hiërarchie gecreëerd (zie figuur 5).



Figuur 5: Voorbeeld pseudo-hiërarchie CBV-subset

Tijdens het automatisch inlezen van beide bestanden en de daarop volgende steekproefsgewijze verificatie door L&C werden geen problemen geconstateerd.

3.2 Modelleren CBV- en CvV-verrichtingen

Als tweede stap werd volledig automatisch de semantische context van alle aangeleverde verrichtingen bepaald. Dit gebeurde op basis van alle reeds binnen de L&C-ontologie geformaliseerde verrichtingen, ongeacht bron en taal van herkomst. Om te achterhalen welke concept-criteria nodig waren voor het formaliseren van de CBV- en CvV-verrichtingen werd in een cyclisch proces een hiërarchische structuur voor de aangeleverde verrichtingen bepaald: de backbone ontologie. Figuur 6 geeft een voorbeeld van de structuur zoals deze aan de aangeleverde CBV- en CvV-verrichtingen via de L&C-ontologie kan worden toegekend.

- 1 ENTITY
- 2 DOMAIN-ENTITY
- 3 TIME BEARING ENTITY
- 4 PROCESS
- 5 PHYSICAL PROCESS
- 6 PROCEDURE
- 7 VERRICHTINGEN INZAKE CONTROLE VAN URINEBLAAS
- 8 VOORSCHRIJVEN VAN CONDOOMKATHETER
- 9 -> HAS-CCC CVV : 9-364 : VOORSCHRIJVEN VAN CONDOOMKATHETER
- 10 VOORSCHRIJVEN VAN ENURESISWEKKER
- 11 -> HAS-CCC CVV : 9-362 : VOORSCHRIJVEN VAN ENURESISWEKKER
- 12 VOORSCHRIJVEN VAN PENISKLEM
- 13 -> HAS-CCC CVV : 9-365 : VOORSCHRIJVEN VAN PENISKLEM
- ...
- 86 MARSHALL-MARCHETTI OPERATION
- 87 -> HAS-CCC CBV : 336370 : URINEWEGEN - URETROVESICALE
- OPHANGOPERATIE - MARSHALL-MARCHETTI
- 88 -> HAS-CCC CVV : 5-595.02 : RETROPUBISCHE URETROSUSPENSIE VOLGENS
- MARSHALL-MARCHETTI-KRANTZ
- 89 URETHROPEXY

90 -> HAS-CCC CVV : 5-597.1 : URETROPEXIE, NIET ELDERS GECLASSIFICEERD
 91 BURCH PROCEDURE
 92 -> HAS-CCC CBV : 337352 : VAGINA - COLPOSUSPENSIE - VOLGENS BURCH
 93 URINEWEGEN - URETROVESICALE OPHANGOPERATIE VOLGENS BURCH -
 LAPAROSCOPIE
 94 -> HAS-CCC CBV : 336362 : URINEWEGEN - URETROVESICALE
 OPHANGOPERATIE VOLGENS BURCH-LAPAROSCOPIE
 95 -> HAS-CCC CBV : 337352A : VAGINA - COLPOSUSPENSIE - VOLGENS BURCH -
 VIA LAPAROSCOPIE
 ...
 1634 LITHOTRIPSY OF KIDNEY (see line : 1589)
 1635 LITHOTRIPSY OF URINARY BLADDER
 1636 BLAAS - LITHOTRIPSIE - CHEMOLYSE (see line : 1529)
 1637 BLAAS - LITHOTRIPSIE - LASER
 1638 -> HAS-CCC CBV : 336262B : BLAAS - LITHOTRIPSIE - LASER

Figuur 6: Voorbeeld hiërarchische structuur backbone hiërarchie

Deze hiërarchische structuur moet als volgt worden geïnterpreteerd:

- Iedere regel, voorafgegaan door een incrementeel nummer, bevat een concept
 - De hiërarchie worden binnen de structuur aangegeven door voorloopspaties
 - De omschrijvingen van de concepten komen overeen met:
 - óf een concept uit de LinkBase domeinontologie, bijvoorbeeld het concept “*VERRICHTINGEN INZAKE CONTROLE VAN URINEBLAAS*” (zie regelnummer 7);
 - óf een concept uit het metadomein, bijvoorbeeld het concept “*RETROPUBISCHE URETROSUSPENSIE VOLGENS MARSHALL-MARCHETTI-KRANTZ*” (zie regelnummer 88). Deze metaconcepten zijn herkenbaar binnen deze structuur omdat ze voorafgegaan worden door het linktype “HAS_CCC”. Dit linktype staat voor “has a Coding & Classification Category” en geeft de directe verwijzende relatie weer vanuit het domeinconcept naar het metaconcept. Deze groep omvat dus de aangeleverde CBV- en CvV-omschrijvingen.
- De betekenis van de metaconcepten is binnen deze structuur gelijk aan het direct erboven voorkomende domeinconcept. Bijvoorbeeld: het metaconcept “*RETROPUBISCHE URETROSUSPENSIE VOLGENS MARSHALL-MARCHETTI-KRANTZ*” (regelnr. 88) heeft eenzelfde betekenis als het domeinconcept “*MARSHALL-MARCHETTI OPERATION*” (regelnr. 86). Indien meerdere CBV en/of CvV-omschrijvingen rechtstreeks onder eenzelfde domeinconcept voorkomen, wijst dit op synoniemen of doublures binnen respectievelijk het CBV- en/of CvV-bestand (zie hiervoor bijvoorbeeld de regelnr. 94 en 95).
- Omdat het mogelijk is om binnen de hiërarchische structuur op verschillende manieren tot een bepaald concept te komen (figuur 4, blz. 8 illustreert dit) komen sommige concepten meerdere keren voor. Daar waar dit het geval is wordt de omschrijving van het concept gevolgd door de tekst “see line : xxx”, waarbij xxx doorverwijst naar een ander regelnummer binnen de structuur (zie bijvoorbeeld regelnr. 1636).

Een analyse van de hiërarchische structuur levert de volgende resultaten op:

Groep	Soort concept	Aantal	Gemiddelde diepte	Maximale diepte
1	Concepten	7465		
2	Concepten die doorverwijzen	2203		
3	CBV-concepten	1111	13,33	19
4	CvV-concepten	744	12,70	18
5	Backbone-concepten	1764		
6	Vanuit CBV/CvV bereikte domeinconcepten	1643		
7	CBV/CvV als diepste concept	727		

Tabel 1: Resultaten backbone hiërarchie

Groep 1 vertegenwoordigt alle concepten, 7465 in totaal, die binnen de hiërarchische structuur voorkomen.

Groep 2 wordt gevormd door die concepten binnen de structuur waarbij een doorverwijzing (“see line: xxx”) is opgenomen en die dus meerdere keren voorkomen. Dit zijn er in totaal 2203.

Groep 3 en 4 bevatten respectievelijk de CBV-concepten (1111) en CvV-concepten (744).

Dat er 34 CvV-concepten meer binnen de structuur voorkomen dan dat er oorspronkelijk zijn aangeleverd (744-710) wordt veroorzaakt doordat meerdere, in betekenis verschillende, domeinconcepten verwijzen naar eenzelfde CvV-metaconcept. Een voorbeeld: het metaconcept "ENKELZIJDIGE ORCHIDECTOMIE D.M.V. INCISIE VAN SCROTUM INCLUSIE: MET MEENEMEN VAN EPIDIDYMIS" met CvV-code 5-622.0 komt tweemaal in de structuur voor. Zowel onder het domeinconcept "ENKELZIJDIGE ORCHIDECTOMIE D.M.V. INCISIE VAN SCROTUM" als onder het domeinconcept "ENKELZIJDIGE ORCHIDECTOMIE D.M.V. INCISIE VAN SCROTUM MET MEENEMEN VAN EPIDIDYMIS".

Kijkend naar de gemiddelde diepte van de CBV- en CvV-concepten (het gemiddelde niveau binnen de hiërarchische structuur waarop een metaconcept voorkomt) kan geconcludeerd worden dat de CBV-concepten gemiddeld meer detail kennen dan de CvV-concepten (gemiddelde diepte van respectievelijk 13,33 en 12,70).

Groep 5, de backbone-concepten zijn die domeinconcepten die geen directe relatie hebben met een CBV- of CvV-concept, maar die wel op het pad liggen van een CBV- of CvV-concept naar de top van de hiërarchische structuur. Dit aantal, 1764, is zeer groot en is een maat voor het aantal conceptcriteria (eigenschappen) van belang voor de detailmodellering van de CBV- en CvV-omschrijvingen. Daarnaast kan het zicht geven op het aantal mogelijke "gaten" in hiërarchisch opgebouwde classificaties zoals de CvV. Wat betreft dit laatste is het dan wel van belang dat niet zoals nu het geval is alleen de eindklassen (de CvV-codes waarop geregistreerd mag worden), maar ook de hogere klassen binnen de hiërarchie gemodelleerd worden.

Groep 6 geeft het aantal domeinconcepten weer waarnaar, vanuit de CBV- en CvV-metaconcepten, direct wordt verwezen (middels het "IS-CCC-OF" linktype). Dit is het geval voor 1643 domeinconcepten. Meerdere CBV- en CvV-metaconcepten verwijzen dus naar één domeinconcept en moeten dus eenzelfde betekenis hebben. Met andere woorden van de 1821 (1111+710) voorkomende CBV- en CvV-concepten zijn er 178 (1821-1643) niet uniek. Voor een groot gedeelte wordt dit veroorzaakt door CvV-concepten die exact hetzelfde betekenen als CBV-concepten, maar ook doublures tussen CBV-concepten onderling komen voor (zie hiervoor ook paragraaf 3.3.2.2).

Groep 7 geeft het aantal CBV- en CvV-concepten aan die eindbladeren zijn in de hiërarchische structuur. Dit wil zeggen dat binnen de L&C-ontologie nergens een meer gedetailleerd concept is terug te vinden. Hierbij is van belang te weten dat een classificatiesysteem als SNOMED RT hier al wel in zit. Rekening houdend met de aanwezige doublures blijkt dat circa 30% van de aangeleverde CBV- en CvV-concepten nog niet voor te komen binnen de L&C-ontologie.

Het totaal van 7465 concepten (groep 1) is samengesteld uit diverse soorten concepten uit de groepen 2 t/m 6 (2203+1111+744+1764+1643).

3.3 Het formeel modelleren van de CBV- en CvV-verrichtingen

3.3.1 Efforts

Als derde stap werd elke individuele verrichting gemodelleerd en gedefinieerd met behulp van de uit de backbone ontologie afleidbare conceptcriteria met als uiteindelijk resultaat de CBV-ontologie. Het modelleren door L&C kostte in totaal 3 maanden. Deze inspanning zou volgens L&C met 50 tot 75% verminderd kunnen worden door aan de ene kant de CBV- en CvV-concepten niet te relateren aan andere classificatiesystemen (SNOMED e.d.) en aan de andere kant alleen de aangeleverde CBV- en CvV-verrichtingen te modelleren, waarbij aan de concepten geen volledige formele definitie wordt toegekend en geen gebruik wordt gemaakt van de reeds bestaande L&C-ontologie. Dit laatste betekent wel dat er een extra inspanning moet worden geleverd omdat er geen gebruik gemaakt kan worden van de al binnen LinkBase bestaande concepten en conceptcriteria.

3.3.2 Analyse CBV-ontologie

3.3.2.1 Conceptcriteria

De CBV-ontologie bevat in tegenstelling tot de backbone-ontologie naast de CBV- en CvV-concepten en bijbehorende linktypes alleen die concepten van de domeinontologie die een *rechtstreekse* relatie hebben met één van de CBV- en CvV-concepten. Dit betekent dat indien zich de situatie voordoet dat geldt dat concept A "is-een" concept B "is-een" concept C en indien hierbij concept B niet gerelateerd

is aan een CBV- of CvV-concept, concept B niet in de uiteindelijke CBV-ontologie is terug te vinden. Deze situatie wordt binnen de CBV-ontologie weergegeven als concept A “is-een” concept C. M.a.w. de backbone-concepten (1764, zie groep 5 tabel 1) uit de backbone hiërarchie zijn binnen de CBV-ontologie *niet* opgenomen.

Uit de CBV-ontologie werden alle conceptcriteria (de eigenschappen) geëxtraheerd die deel uit maakten van de CBV en CvV-modellering. Een conceptcriterium bestaat hierbij uit een bronconcept, het linktype en het doelconcept: een voorbeeld, bij het conceptcriterium “overige osteotomie van femur en patella is_a incision of bone” vormt “overige osteotomie van femur en patella” het bronconcept, “is_a” het linktype en “incision of bone” het doelconcept. In totaal werden 7987 conceptcriteria geëxtraheerd uit de CBV-ontologie. Bij de start van het project bestonden er hiervan 2366 (bijna 30%) reeds binnen LinkBase, 2 bestaande conceptcriteria werden in de loop van het project gewijzigd en er werden er 5619 tijdens de CBV- en CvV-modellering nieuw aangemaakt.

De volgende tabel geeft een verdeling van de inspanning van het automatisch en handmatig toevoegen van conceptcriteria. Dit voor zowel de conceptcriteria die voor het project al bestonden als voor de nieuw aangemaakte conceptcriteria.

	Voor project	Tijdens project	Totalen
Automatisch	1192	979	2171
Handmatig	1174	4640	5814
Totalen	2368	5619	7987

Tabel 2: Automatisch versus handmatig toekennen conceptcriteria

Wat in deze tabel opvalt, is dat tijdens het project sprake is van een veel groter aandeel van het handmatig toekennen van conceptcriteria dan voor het project. Tijdens het project blijkt 17% van het modellerwerk automatisch te verlopen, 83% handmatig. Voor het project betrof dit voor iedere component ongeveer een gelijk deel. Dit wordt voornamelijk veroorzaakt doordat de omschrijvingen van de CBV- en CvV-codes:

- 1) niet de grammaticale regels van het Nederlands volgen;
- 2) spellingsfouten en niet-conventionele afkortingen bevatten en
- 3) een groot deel van de omschrijvingen vaag zijn beschreven.

Hierdoor kunnen ze binnen LinkBase niet automatisch worden herkend als een bestaand domeinconcept.

Een aantal voorbeelden van spellingsfouten en niet-conventionele afkortingen:

BIMALLEOLAIREENKELFRACTUUR	VERR
CONSERVATIOEVE	VESICULASEMINALIS
CONSSERVATIEVE	ZOALSBLADDERNECK
ENFALANGEN	ZONDERVERDERE
EVTUEEL	BEKKENB
GESPECIFICEERDEFRACTUREN	BEKKENBODEMSP
METFIXATIE	BLAASSENS
METULNA	CONSERVATIEV
OPERATIEVEFIXATIE	DETRUSORREFL
OVERIGEGESPECIFICEERDE	ELEKTROMYOGR
PERATIES	PIJNBESTRIJD
TESTSEXCL	PROVOC
URETROVESICALEOVERGANG	REG
VANBESTAANDE	REGI
VANRETROPERITONEAAL	ZELFST

Een voorbeeld van een vage omschrijving is de volgende:

338384L HAND - HECHTEN EXTENSORPEZEN - LINKS

Welke extensorpezen worden hier bedoeld? Er zijn immers twee mogelijkheden: of het betreft de extensoren van de hand (en dus gelegen in de pols), of het betreft de extensoren van de vingers (en dus gelegen in de hand). Overigens is het niet zo dat iedere vaagheid in een omschrijving vermeden zou moeten worden. Immers, binnen een ontologie is dit onmogelijk omdat, per definitie, ieder ouderconcept vager is dan de kindconcepten die eronder voorkomen.

Verder kwam naar voren dat sommige woorden, vanuit ontologisch standpunt, binnen de omschrijvingen van verrichtingen geen betekenis hebben of beter niet gehanteerd kunnen worden.

Enkele voorbeelden:

Omdat het CBV-bestand niet hiërarchisch is gestructureerd, kunnen omschrijvingen met het woord "overige" nergens naar verwijzen. Het woord "overige" voegt geen eigenschap toe aan het concept zelf. Dit in tegenstelling tot de CvV-verrichtingen waar duidelijk sprake is van een hiërarchische classificatie. Bij onderstaande CBV-verrichtingen is dit het geval:

336239	BLAAS	- EXCISIE - OVERIGE - NIET TRANSURETRAAL
336349	BLAAS	- OVERIGE PLASTIEKEN
339238L	BUIK	- LAPAROSCOPIE- DIAGNOSTISCH - OVERIGE
338510R	FEMUR	- OVERIGE OSTEOTOMIE - RECHTS
339004B	HUIDTRANSPLANTATIE	- THIERSCH-PLASTIEK NAAR OVERIGE LOKALISATIES
336769	MANNELIJK GESLACHTSORGAAN	- REANASTOMOSE OPERATIES - OVERIGE
336059	NIER	- PLASTISCHE OPERATIES - OVERIGE
338612D	ONDERBEEEN	- OVERIGE OSTEOTOMIEEN TIBIA-FIBULA - RECHTS
336846C	PENIS	- OVERIGE BEHANDELINGEN BIJ PRIAPISME
336479A	URETHRA	- DILATATIE - OVERIGE
336429	URETHRA	- EXCISIE OVERIGE PATHOLOGISCHE AFWIJKINGEN
336449	URETHRA	- OVERIGE PLASTISCHE OPERATIES

Bovenstaande geldt ook voor woorden als "niet gespecificeerd", "NNO" etc.

Ook wordt het gebruik van "en" niet aangeraden, tenzij daadwerkelijk aan de volledige omschrijving van de betreffende verrichting moet worden voldaan zoals het geval is bij bijvoorbeeld de verrichting "PENIS - FRENULUMPLASTIEK EN PREPUTIUMPLASTIEK". Hetzelfde geldt voor het gebruik van "of", "en/of". In omschrijvingen als "URETER - ANTEGRADE SONDAGE EN DILATATIE EN-OF ENDOPROTHESE" of "BLAAS - DRAINAGE OF EXPLORATIE PERIVESICALE WEEFSEL" worden meerdere verrichtingen binnen één omschrijving ondergebracht. Beter zou zijn deze als aparte concrete verrichtingen binnen het bestand op te nemen.

Tijdens het modelleren gaven van de 1821 aangeleverde verrichtingenomschrijvingen er 44 voor L&C aanleiding tot het stellen van een vraag op het discussieplatform. Op zichzelf is dit geen direct relevant cijfer. Bepaalde formuleringen zoals "en", "of", maar ook "-" kwamen regelmatig voor in verschillende omschrijvingen. L&C ging er hierbij vanuit dat een eerder gegeven antwoord op een vraag betreffende de betekenis van deze woorden waarschijnlijk bij een nieuwe vraag op dezelfde manier zou worden beantwoord. Daarnaast kunnen L&C-medewerkers aan bepaalde omschrijvingen een specifieke interpretatie hebben gegeven, zonder zich bewust te zijn van mogelijke andere (bij CBV-medewerkers en/of medische professionals bekende) interpretaties. Dit laatste punt zal aan de orde komen bij de door het CBV uitgevoerde evaluatie.

Zoals we zagen word er binnen de CBV-ontologie gebruik gemaakt van 7987 conceptcriteria. Dit betekent dat er evenzoveel linktypes⁵ zijn gebruikt. Het soort linktype, 43 verschillende, en het bijbehorende aantal keren dat deze binnen de CBV- CvV-modellering is gebruikt wordt in onderstaande tabel vermeld.

⁵ Voor de betekenis van de verschillende linktypes wordt verwezen naar het uitgebreide projectverslag, te downloaden vanaf www.cbv.nl.

Soort Linktype	Aantal	Soort Linktype	Aantal
DISJOINT	1	HAS-EXCLUSIVE	24
HAS-ACTOR	1	HAS-PURPOSE	24
HAS-CEN-OCCURRENCE-SINCE	1	HAS_ASSOC	26
HAS-CEN-OCCURRENCE-UNTIL	1	HAS-OVERLAPPING-REGION	26
HAS-CONSEQUENCE	1	HAS-TO-DO-WE-STATE	27
HAS-ENABLER	1	HAS-HAPPENING-LATER-THAN	35
IS-CONNECTOR-OF	1	HAS-INSTRUMENT	35
IS-PROPER-SPATIAL-PART-OF	1	IS-PARTICIPATING-PROCESS-OF	37
IS-SPATIAL-PART-OF	1	HAS-GENERALISED-MEANS	38
HAS-CEN-OCCURRENCE-FOLLOWS	2	HAS-TARGET	39
HAS-CEN-OCCURRENCE-INCLUDES	2	HAS-SPATIAL-POINT-REFERENCE	45
HAS-FINISH-LATER-THAN	2	HAS-INCLUSIVE	53
HAS-HAPPENING-EARLIER-THAN	2	HAS-PROCEDURAL-APPROACH	58
HAS-INDICATION	2	HAS-REASON	75
HAS-CEN-OCCURRENCE-AFTER	3	HAS-SOURCE	127
HAS-PATH-OF-THEME	3	HAS-THEME	160
IS-INCLUSIVE-OF	5	HAS-PARTICIPATING-PROCESS	164
HAS-PATH	9	HAS-ACTEE	328
HAS-CEN-OCCURRENCE-DURING	11	IS_A	2597
HAS-CREATIVE-RESULT	14	IS-INCLUSION-OF	23
DEFINES-SAME-SITUATION-AS	16	HAS-CCC	3944
HAS-EPONYMIC-ASSOC	22	TOTAAL	7987

Tabel 3: Soorten en aantallen linktypes

Tijdens het toekennen van conceptcriteria wordt een rechtstreekse relatie gelegd tussen de CBV- en CvV-metaconcepten en de domeinconcepten. De domeinconcepten kennen op hun beurt weer rechtstreekse relaties met metaconcepten uit andere classificatiesystemen. Hierdoor wordt het mogelijk vast te stellen of er CBV- en CvV-concepten zijn die wat betreft betekenis overeenkomen met concepten uit andere classificatiesystemen. In onderstaande tabel is per classificatiesysteem het aantal overeenkomstige metaconcepten aangegeven:

HOMEY2	1
FDB-USE	2
ICPC-2	3
FDB-US	5
FDB DX	6
MEL2002	9
ICD-10	20
MESH 2001	27
MEDDRA	80
MEDCIN	101
OPCS4	199
SNOMED-RT	224
ICD-9-CM	265
SNOMED-CT	323

Tabel 4: Aantallen overeenkomstige metaconcepten per classificatiesysteem

3.3.2.2 Doublures en synoniemen

Door L&C zijn een aantal bestanden aangeleverd waarmee inzicht kan worden verkregen in hoeverre de omschrijvingen zoals die gehanteerd worden binnen het CBV- en het CvV-bestand eenzelfde semantische betekenis hebben. Er is sprake van doublures indien omschrijvingen binnen het CBV- óf het CvV-bestand eenzelfde betekenis hebben. Wanneer omschrijvingen eenzelfde betekenis hebben en gebruikt worden binnen het CBV-bestand én het CvV-bestand wordt dit als synonymie aangeduid.

Uit de resultaten blijkt dat 17 van de in totaal 1111 aangeleverde CBV-codes (1,5%) een omschrijving hebben die minimaal één doublure kent binnen het CBV-bestand zélf.

Een voorbeeld van een doublure:

CBV-code 336095: "nier – percutane lithotripsie" en CBV-code 336097 "nier – percutane litholapaxie".

Het merendeel van de doublures wordt echter verklaard doordat binnen de CBV-omschrijvingen gebruik wordt gemaakt van woorden als "overige", "niet nader omschreven" en "overig gespecificeerd". Woorden die, zoals eerder aangegeven, geen eigenschappen aan een concept toevoegen en daarom binnen de ontologie als betekenisloos worden gekenmerkt. Zelfs vanuit een classificatietechnisch standpunt is het gebruik van deze woorden binnen het CBV-bestand, een platte codelijst met vele tienduizenden verrichtingen, discutabel.

158 CBV-codes hebben een omschrijving die synoniem is met de omschrijving van tenminste één CvV-code. Uitgaande van deze codes met overeenkomstige omschrijvingen binnen zowel het CBV- als het CvV-bestand kan meer inzicht worden verkregen in de door L&C gelegde relatie tussen de CBV- en CvV-codes. Deze relatie tussen het CBV- en het CvV-bestand zal in paragraaf 3.3.2.3 verder worden uitgewerkt.

Daarnaast hebben van de 710 aangeleverde CvV-codes er 103 (15%) een omschrijving die minimaal één doublure kent binnen het CvV-bestand zélf. Wat hierbij opvalt is dat het merendeel van de doublures is toe te rekenen aan het gebruik van woorden als "overig gespecificeerd" en "niet gespecificeerd" (resp. NEC en NNO). Ook hier geldt dat deze woorden op zichzelf geen eigenschappen toevoegen aan een bepaald concept. Omdat het CvV-bestand echter een hiërarchisch gestructureerd bestand is, is het gebruik van deze woorden vanuit classificatietechnisch standpunt wél zinvol.

160 CvV-codes hebben een omschrijving die synoniem is met de omschrijving van tenminste één CBV-code⁶.

3.3.2.3 CBV-CvV-relatie

Door het bureau CBV wordt in het CBV-bestand, waar mogelijk, aan de verrichtingencodes op basis van medisch inhoudelijke overeenkomsten een CvV-code gekoppeld. Om een beter inzicht te krijgen in de registratieve waarde van de CBV-code in het algemeen en de relatie tot de CvV-code in het bijzonder wordt aan die CBV-codes waarbij een relatie met een CvV-code wordt gelegd een zogenaamde CvV-categorie toegekend. Eén van deze categorieën is "Categorie 1" welke aangeeft dat de betreffende verrichting een één op één relatie heeft met een CvV-code (en vice versa).

Van een CBV-code die volgens L&C semantisch exact overeenkomt met een CvV-code mag verwacht worden dat door het bureau CBV aan deze CBV-code een CvV-code is gekoppeld die gekenmerkt is als "Categorie 1". Tevens zou men het omgekeerde verwachten: alle aan L&C aangeleverde CBV-codes met een "Categorie 1"-aanduiding zouden terug moeten komen in de door L&C aangeleverde bestanden met synoniemen en doublures.

Zoals eerder aangegeven blijkt uit de resultaten dat de omschrijvingen van 158 CBV-codes overeenkomen met de omschrijving van één of meerdere CvV-codes. Bij 154 van de 158 CBV-codes is de gekoppelde CvV-code inderdaad gekenmerkt als "Categorie 1". Bij de vier resterende codes

⁶ N.b. het verschil in aantal (158 CBV-codes / 160 CvV-codes) wordt o.a. verklaard doordat in de bronbestanden de omschrijvingen van meerdere CBV-codes overeenkomen met de omschrijving van één CvV-code en de omschrijvingen van meerdere CvV-codes overeenkomen met de omschrijving van één CBV-code. Dit wordt veroorzaakt door de aanwezigheid van doublures.

werden door het bureau CBV in het verleden dezelfde CvV-codes als passend gezien, echter werd er besloten een andere categorie toe te kennen. Bij drie van deze vier CBV-codes blijkt er sprake te zijn van doublures. Aangezien het bureau CBV slechts bij één koppeling tussen een CBV- en een CvV-code het "Categorie 1"-kenmerk toekent en in andere gevallen een categorie 2 of 3 is het verschil in categorie-indeling hierdoor verklaarbaar. In één geval zouden L&C's resultaten leiden tot een betere categorie toekenning. L&C geeft aan dat CBV-code 336421 "URETHRA – OPERATIE URETHRAFISTELS" en CvV-code 5-583.3 "OPHEFFEN VAN URETHRAFISTEL" semantisch overeenkomen. Door het bureau CBV is echter de koppeling tussen CBV-code 336443 "URETHRA – PLASTISCHE OPERATIE – SLUITEN FISTEL" met CvV-code 5-583.3 "OPHEFFEN VAN URETHRAFISTEL" gekenmerkt als "CATEGORIE 1". Bij CBV-code 336443 is sprake van een plastische operatie, terwijl hierover bij CBV-code 336421 en CvV-code 5-583.3 geen uitspraak wordt gedaan (m.a.w. de omschrijving van CBV-code 336443 kent een extra eigenschap). L&C beschouwt CBV-code 336443 dan ook terecht als een kind van CBV-336421 en CvV-code 5-583.3 waardoor deze koppeling als "Categorie 1" aangeduid zou moeten worden i.p.v. de koppeling tussen CBV-code 336443 en CvV-code 5-583.3.

Het blijkt niet zo te zijn dat alle aan L&C aangeleverde CBV-codes met een "Categorie 1"-aanduiding terugkomen in de door L&C aangeleverde bestanden met synoniemen en doublures. Van de 1111 aan L&C aangeleverde CBV-codes hebben er 361 een door het bureau CBV gekoppelde "Categorie 1" CvV-code. Voor 203 CBV-codes (361-(158-4)) geldt echter dat door L&C niet een overeenkomstige semantisch gelijke omschrijving binnen de aangeleverde CvV-subset is gevonden. De verklaring hiervoor is tweeledig. De CBV- en CvV-omschrijvingen kunnen ten opzichte van elkaar globaler dan wel gedetailleerder zijn (minder of juist meer eigenschappen bevatten) waardoor ze niet exact overeenkomen, terwijl het bureau CBV aan de betreffende CBV-CvV-koppeling, als best passende relatie, wel een "Categorie 1" heeft toegekend. Daarnaast kunnen soms meerdere CvV-codes gekoppeld worden. Steekproefsgewijze controle m.b.v. LinkFactory Lite (Fastcode) bevestigt deze verklaring.

4 Conclusie

4.1 Conclusie pilotproject en indicatoren

Resultaten van de indicatoren

Om vanuit het CBV te kunnen beoordelen in welke mate het formaliseren bijdraagt aan een kwalitatieve verbetering van het bestand zijn bij aanvang van het project een vijftal indicatoren bepaald (zie ook het hoofdstuk werkwijze). In onderstaande tabel zijn de resultaten van deze vijf indicatoren kort weergegeven.

<i>Indicator</i>	<i>Resultaat</i>
Het aantal verrichtingen waarbij sprake is van doublures, spellingsfouten en niet reguliere afkortingen.	<ul style="list-style-type: none">- Een belangrijk deel van het modellerwerk is handmatig verricht vanwege het hoge detailniveau van de CBV-codes.- Daarnaast zorgde het niet volgen van grammaticale regels, spellingsfouten, niet reguliere afkortingen en vage omschrijvingen voor handmatige modellering;- 1,5% van de CBV-codes hebben een doublure;- 15% van de CvV-codes heeft een doublure, dit wordt veroorzaakt door het (gewenste) gebruik van de woorden 'overige gespecificeerd' en 'niet gespecificeerd'.
Het aantal verrichtingen waarvan de betekenis niet rechtstreeks is af te leiden uit de omschrijving en die dus voor meerdere uitleg vatbaar zijn.	<ul style="list-style-type: none">- Tijdens het modelleren zijn er over 44 van de 1821 verrichtingencodes aanvullende vragen gesteld. Deze antwoorden gaven veelal antwoorden op latere vragen.- Onduidelijke omschrijvingen zoals 'en', 'of', 'en/of' en '-' roepen veelal inhoudelijke vragen op.
Het aantal verrichtingen dat ontbreekt in het bestand wanneer gelet wordt op de uiteindelijke semantische structuur. Dit kan worden gezien als een maat voor de volledigheid van het CBV- en CvV-bestand.	<ul style="list-style-type: none">- Het blijkt dat de verrichtingen van vooral de CBV-codes een hoog detailniveau bevatten. De overeenkomsten met de metaconcepten van andere al aanwezige classificaties in de terminologieserver tonen het hoge detailniveau aan (tabel 4).- De huidige CvV heeft een hiërarchische opbouw met een nadruk op de eindklassen. Dit wordt veroorzaakt door de registratievoorschriften. Door de modellering van de hogere klassen kunnen deze ontbrekende klassen (de zgn. 'gaten') worden toegevoegd.
Het aantal niet correct gelegde relaties tussen het CBV- en het CvV-bestand wanneer de door het CBV gelegde relatie wordt vergeleken met de relaties gelegd bij het formeel modelleren.	<ul style="list-style-type: none">- Door L&C zijn er 158 koppelingen gelegd. Van deze 158 CBV-CvV relaties zijn er 154 'één op één' relaties, drie relaties blijken een doublure te betreffen en één relatie kon worden verbeterd.- Door het CBV zijn er 361 gekoppelde codes met een één op één relatie ("categorie 1") aangeleverd.- Door L&C werden 203 van deze relaties als niet overeenkomstig bevonden. In veel gevallen gaat het hierbij om meer globale of meer gedetailleerde omschrijvingen.

Tabel 5: Overzicht indicatoren

Door het formaliseren van de aangeleverde CBV- en CvV-omschrijvingen de medisch inhoudelijke betekenis geëxpliciteerd. Uit het pilot project blijkt dat de betrouwbaarheid en de juistheid van de medische inhoud van de diverse codeomschrijvingen wordt verbeterd.

Tevens is gebleken dat door het formaliseren van de verrichtingenomschrijvingen het mogelijk wordt om op betekenisniveau op geautomatiseerde wijze relaties tussen identieke verrichtingen uit verschillende classificaties te leggen.

De kwaliteit van de door L&C gelegde relaties tussen het CBV-bestand en het CvV-bestand komt grotendeels overeen met de door het bureau CBV handmatig gelegde relaties. Het handmatig leggen van relaties door het bureau CBV is van een goed niveau.

De omschrijvingen van de registratiecodes zijn zonder het formele karakter door de gebruiker ook goed te begrijpen voor wat er staat en wat er bedoeld wordt.

Op basis van de ervaringen van het pilot project is gebleken dat door het formaliseren van het CBV- en CvV-bestand bestaande classificaties onderzocht kunnen worden op lacunes, ze kunnen automatisch uitgebreid of geordend worden en opnieuw gegenereerd worden. M.a.w het gebruik van een formele terminologie kan leiden tot een meer efficiënt beheer van classificaties.

4.2 Meerwaarde eindresultaat

Meerwaarde voor de zorgpraktijk

De meerwaarde van de L&C-systematiek moet vooral gezocht worden in de meerwaarde bij het dagelijkse gebruik van zorginhoudelijke gegevens (i.e. zorgvraag, klacht, diagnose, behandeling) in de praktijk. Hierbij gaat het om het uitvoeren van analyses op (geaggregeerde) gegevens, de administratieve afhandeling (bijvoorbeeld voor facturatie- en rapportagedoeleinden), het leggen van relaties naar verschillende (inter)nationale classificaties en terminologieën en het gebruik van applicaties die aansluiten op het vakjargon van de medische professional lijken door het gebruik van L&C's ontologie voor de dagelijkse praktijk beschikbaar te komen.

Ontsluiting en mapping naar (inter)nationale classificaties en terminologieën, bijv. de ICD9 en ICD10, de ICPC en SNOMED CT, is mogelijk. De flexibiliteit en het adaptievermogen van de verschillende classificaties en onderlinge relaties nemen toe door de opname in LinkBase®. De verwachting is dat het hierdoor mogelijk wordt om de registratielast binnen het primaire zorgproces te verminderen en de betrouwbaarheid van de informatievoorziening te vergroten.

Meerwaarde voor de Stichting CBV

Op basis van de resultaten van het pilot project is als andere belangrijke meerwaarde voor de Stichting CBV gebleken dat door de formele L&C-aanpak de werkwijze geobjectiveerd wordt waardoor de verrichtingenomschrijvingen worden 'opgeschoond'.

Doublures, spellingsfouten en niet reguliere afkortingen komen tijdens het modelleerwerk nadrukkelijk naar voren waardoor na correctie de aanwezige ambiguïteit afneemt. Daarnaast neemt door het expliciet formaliseren van de betekenis van de verschillende verrichtingenomschrijvingen de eenduidigheid binnen het CBV-bestand toe. De meerwaarde in het gebruik van de terminologieserver m.b.t. beheer en onderhoud van codestelsels en classificaties kan gezien worden als een belangrijke extra meerwaarde naast de meerwaarde voor de zorgpraktijk.

Vanuit bedrijfseconomisch oogpunt liggen de kansen voor de terminologieserver en de afgeleide toepassingen vooral op het terrein van de registratie informatievoorziening van het primaire zorgproces.

Samenvattend:

De meerwaarde van het gebruik van de terminologieserver in de praktijk ligt:

1. bij degene die registreert aan de bron;
2. bij degene die op basis van de vastgelegde gegevens rapportages moet maken en daarvoor verschillende gezichtspunten wil hanteren (daarvoor zijn de specifieke kenmerken van belang);
3. bij degene die op basis van kenmerken patiënten wil selecteren voor diverse doeleinden (bijvoorbeeld onderwijs, onderzoek, opleiding, vergelijking van casuïstiek, etc.);
4. in de toekomst bij de verdere ontwikkeling van transmurale communicatie, het Elektronisch Patiënten Dossier, ordercommunicatie, etc.;
5. bij het beheer en onderhoud van classificaties en codestelsels en de onderlinge relaties.

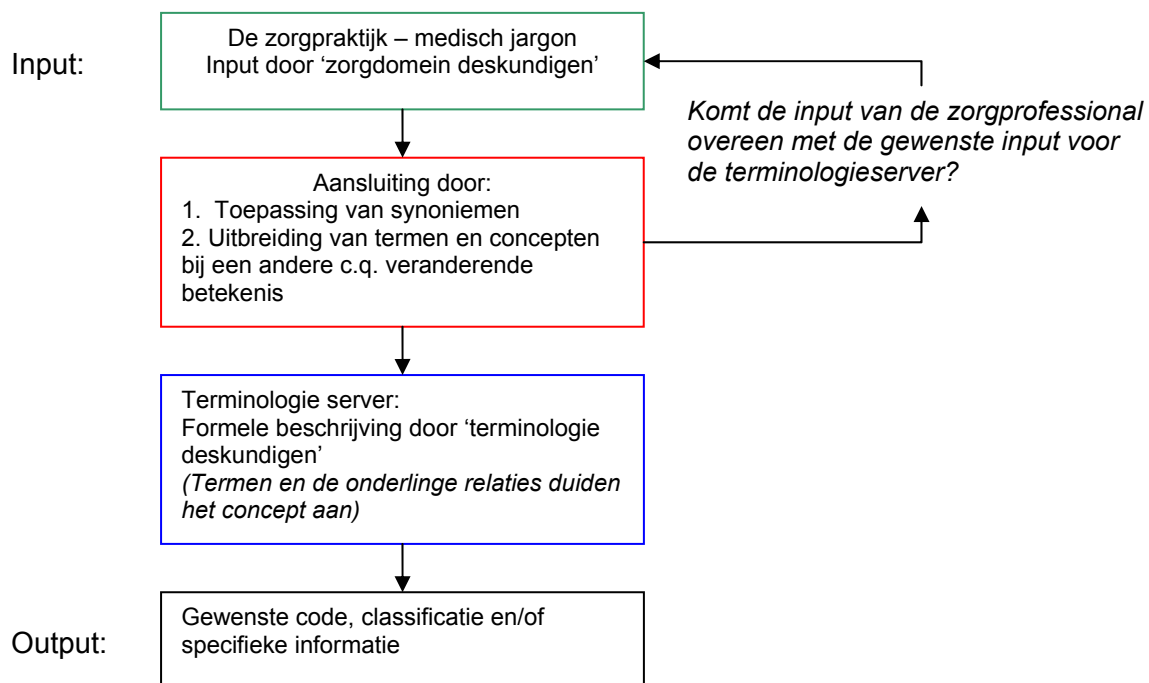
Het succes van een verdere implementatie en de toepasbaarheid van de L&C systematiek is (mede) afhankelijk van:

- het gebruiksgemak voor en acceptatie door de zorgprofessional;
- bewezen werkzaamheid in de praktijk;
- de diversiteit en omvang van de opgenomen codestelsels en classificaties (ICPC, ICD, CvV, SNOMED);
- verdere opbouw van de vakinhoudelijke expertise.

4.3 Aanbevelingen

Het succes bij de implementatie van de terminologieserver in de zorgpraktijk is afhankelijk van de aansluiting van het jargon van de medisch specialist (de input) aan de uitkomst van de terminologieserver in de vorm van codes, klassen, etc. (de output). Het uiteindelijke resultaat wordt bepaald door enerzijds de kwaliteit van de bestaande omschrijving (CBV/CvV) en anderzijds de kwaliteit van de formele beschrijving door de terminologiespecialisten.

Gezien de gekozen projectbenadering, waarbij door L&C terminologiespecialisten de formalisatie is uitgevoerd, is het wenselijk meer inzicht te verkrijgen in de mate waarin de interpretatie van de individuele verrichtingen door L&C overeenkomt met die van 'zorgdomein deskundigen'. Omdat de CBV-codelijst niet expliciet de betekenis van een beschrijving (code) volledig beschrijft is het dus mogelijk dat de betekenis van een omschrijving zoals die in de praktijk van de verrichtingenregistratie wordt gehanteerd anders is dan door L&C is begrepen op basis van de aangeleverde beschrijving. De context waarin gegevens voor registratie wordt gebruikt bepaald immers mede de betekenis. Vooral de aansluiting van het medische jargon op de termen en de strikte taalkundige regels bepalen de kwaliteit en de herkenbaarheid van de uiteindelijke uitgevoerde formalisatie (figuur 7).



Figuur 7: aansluiting van het medische jargon op de formalisatie van de zorgactiviteiten

Aanbevolen wordt een evaluatie waarbij, uitgaande van een multidisciplinaire aanpak, de aansluiting van de (weerbarstige) zorgpraktijk (input) wordt gecorreleerd met de taalkundige correcte input voor de terminologieserver.

Het verdient daarnaast aanbeveling om de meerwaarde van de L&C systematiek in een breder samenwerkingsverband in de praktijk te toetsen. Aandachtspunt hierbij is de vereenvoudiging van de vastlegging van medisch inhoudelijke gegevens aan de bron door de zorgprofessional zoals de verrichtingen en diagnoses. Dit kan in de toekomst van belang zijn bij de verdere ontwikkeling van het Elektronisch Patiënten Dossier, uitwisseling van relevante informatie bij (transmurale)communicatie, ordercommunicatie, gestandaardiseerde zorgpaden, ontslagbrieven, etc.. Deze aanvullende pilot studie in de zorgpraktijk moet duidelijkheid verschaffen over de mogelijke meerwaarde in het gebruik van een formeel terminologiesysteem en de diverse applicaties in de zorgpraktijk. De resultaten van deze studie moeten een bestuurlijke verantwoording mogelijk maken door het leggen van een relatie tussen de financiële investering en de meerwaarde in het gebruik in de dagelijkse zorgpraktijk.

Aanbevolen wordt om de meerwaarde van de L&C systematiek voor de zorgpraktijk in brede zin t.o.v. platte lijsten te bepalen, daarbij rekening houdend met de bedrijfseconomische haalbaarheid .

5 Bijlagen

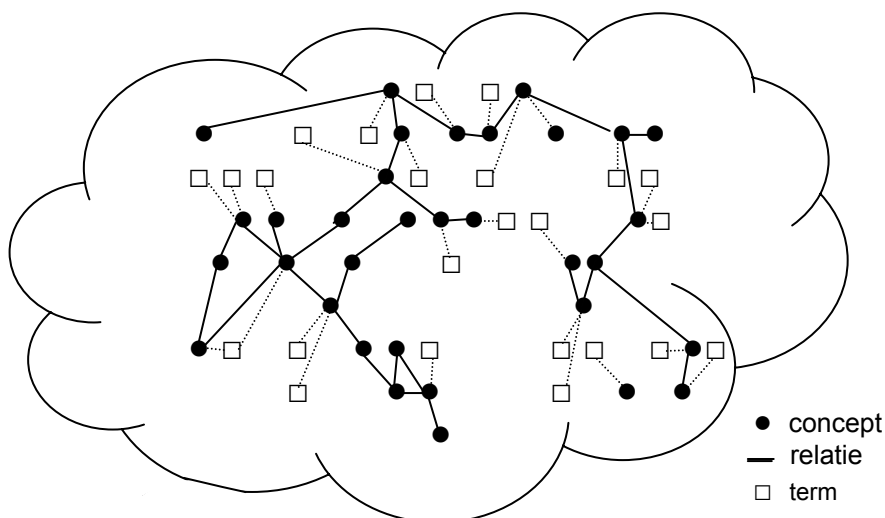
5.1 Een formele medische ontologie: theoretische achtergrond

Het is een feit dat computers alleen overweg kunnen met zorggegevens indien deze op een gestructureerde manier beschikbaar zijn. De conclusie die hieruit jarenlang volgde was dat er dus ook een noodzaak was om op een gestructureerde manier zorggegevens vast te leggen en aan elkaar te relateren. Voorbeelden hiervan te over: het gebruik van codestelsels, diagnose- en verrichtingenclassificaties binnen (elektronische)aankruis- en invulformulieren voor het vastleggen van de gezondheidstoestand en het behandelproces van de patiënt zijn gemeengoed binnen de Nederlands gezondheidszorg. Het gestructureerd vastleggen is weinig flexibel, interpretatie is afhankelijk van het doel waarvoor wordt vastgelegd en kost veel tijd. Het grootste bezwaar is echter dat dergelijke oplossingen niet aansluiten op de manier hoe artsen, verpleegkundigen en paramedici kennis opbouwen, denken en handelen.

Medische professionals doen tijdens hun studie en later in de praktijk kennis op door zich vertrouwd te maken met de begrippen die zij in de praktijk gebruiken. Neem bijvoorbeeld het begrip “bloeddruk”. Een arts of verpleegkundige heeft dit begrip ooit geleerd door kennis te nemen van de normale bloeddruk, de diverse meetmethoden, de definitie van hoge bloeddruk, de verschillende soorten hoge bloeddruk, de definitie van lage bloeddruk, factoren die de bloeddruk beïnvloeden, enzovoorts. Het gaat voornamelijk om losse feiten; het feitelijke kennisdomein.

Voor diepgaandere kennis over het begrip “bloeddruk” moet de arts of verpleegkundige ook weten hoe al deze losse feiten over de bloeddruk onderling met elkaar samenhangen. De professional moet als het ware, behalve de anatomie, ook de fysiologie van het begrip bloeddruk leren, want dan weet hij pas welke rol de bloeddruk speelt, waarom een normale bloeddruk gewenst is, hoe de bloeddruk gereguleerd wordt, etc. Het verhelderen van begrippen staat centraal en begripvorming vindt plaats; het conceptuele kennisdomein.

Begrijpen is dus een vaardigheid die betrekking heeft op het conceptuele kennisdomein. Dit domein bestaat uit kennis over: (a) concepten, zoals “bloeddruk”, “hoge bloeddruk” en “lage bloeddruk” en (b) de relaties tussen deze concepten. Een concept kan hierbij worden gedefinieerd als het geheel van betekenissen, associaties, ideeën en beelden dat verbonden is aan een begrip. Om over concepten te kunnen praten moeten we gebruik maken van natuurlijke taal, termen, waarmee een concept wordt aangeduid. Meerdere termen kunnen hierbij betrekking hebben op één concept, zoals het geval is bij synoniemen (b.v. “hoge bloeddruk” en “hypertensie”), maar één term kan ook meerdere betekenissen hebben zoals het geval is bij homoniemen (b.v. “tumor” als aanduiding voor “neoplasma” of “zwellings” als klinisch kenmerk bij een ontsteking) Voorgaande wordt schematisch weergegeven in figuur 2.

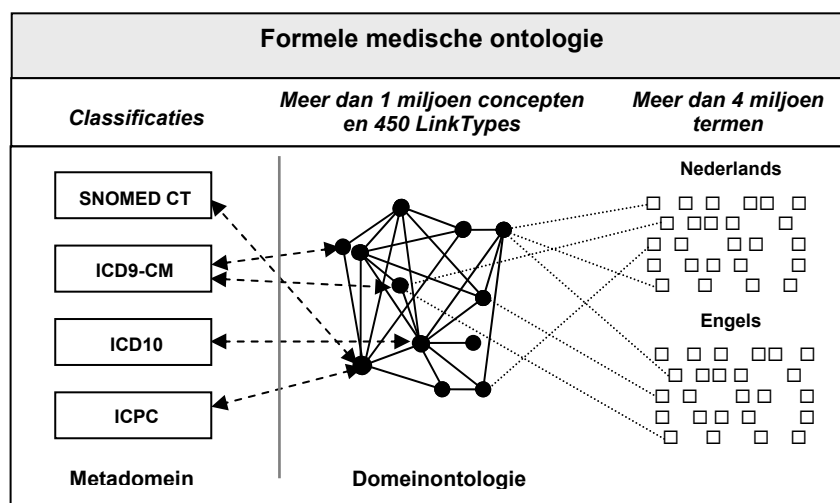


Figuur 2: Het conceptuele kennisdomein

Waarom uitwiden over kennisopbouw en begripsvorming bij medische professionals? Dit om aan te geven dat indien men de nadelen van gestructureerde gegevensinvoer, ontsluiting en presentatie zoals hierboven aangegeven wil voorkomen, men zal moeten aansluiten op de denk- en handelwijze van artsen en verpleegkundigen. Het modelleren van het conceptuele kennisdomein, en daarmee het vastleggen van de betekenis (semantiek) van zorggegevens waardoor medische terminologie interpreteerbaar wordt gemaakt voor de computer, is een eerste stap in die richting.

Door het formeel specificeren van medische concepten én hun onderlinge relaties begeven we ons op het terrein van de medische ontologie⁷. Binnen een ontologie representeren de concepten, de relaties tussen de concepten en de synonieme en homonieme termen, het gezondheidszorgdomein op een manier die voor een computer middels algoritmen interpreteerbaar is. Dit vereist wel dat de betekenis van de relaties tussen de concepten absoluut consistent is. Een voorbeeld: hiërarchische relaties tussen concepten kunnen alleen voorkomen indien alle eigenschappen van het ouderconcept van toepassing zijn op het kindconcept. Bijvoorbeeld, de eigenschappen van het ouderconcept "nier" zijn ook geldig voor het kindconcept "hoefijzernier" (een hoefijzernier is een nier, beide komen voor in de buikholte en maken deel uit van de urinewegen etc.). Die groep van relaties en bijbehorende concepten waarvoor geldt dat deze absoluut consistent zijn, wordt de "domeinontologie" genoemd. Daar waar sprake is van inconsistenties in de relaties tussen de concepten wordt gesproken van het "metadomein". Binnen het metadomein bevinden zich classificatiesystemen als de UMLS, ICD9-CM, ICD10 en SNOMED CT. Ter verduidelijking een voorbeeld uit de ICD10: niet alle eigenschappen van het ouderconcept "Gastritis en duodenitis" gelden voor het kindconcept "Alkoholische gastritis" (een alcoholische gastritis is geen duodenitis). De domeinontologie ondersteunt het interpreteren door de computer. Omdat de concepten binnen het metadomein formele relaties hebben naar concepten binnen de domeinontologie wordt het tevens mogelijk concepten te presenteren volgens de bestaande classificatiesystemen en codestelsels.

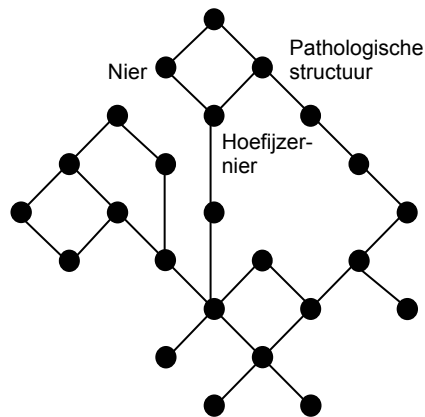
Zoals reeds in hoofdstuk 1 aangegeven heeft het bedrijf L&C een formele medische ontologie (LinKBase®) ontwikkeld, samen met een systeem om deze ontologie te beheren (LinkFactory Workbench®). L&C hanteert hierbij het filosofisch uitgangspunt dat een ontologie moet overeenkomen met de werkelijkheid ("een op realisme gestoelde ontologie"). Hiermee wordt bedoeld dat medische concepten bestaan ongeacht de manier waarop ze taalkundig tot uitdrukking worden gebracht. LinKBase® bevat dan ook meer dan 1 miljoen taalafhankelijke medische concepten, die onderling aan elkaar gerelateerd zijn door meer dan 450 verschillende soorten relaties (in L&C-termen "linktypes" genoemd). De concepten worden hierbij door meer dan 4 miljoen termen, in verschillende talen waaronder het Nederlands, aangeduid. Daarnaast zijn de concepten, uitgaande van hun betekenis, gekoppeld aan bestaande (internationale)classificatiesystemen. Dit laatste óók voor Nederlandse versies van bijvoorbeeld de ICD9-CM en de ICD10. Het voorgaande wordt schematisch in figuur 3 weergegeven.



Figuur 3: LinKBase®, een formele medische ontologie

⁷ "An ontology is a description (like a formal specification of a program) of the concepts and relationships that can exist for an agent or a community of agents." (Tom Gruber)
 "Ontology is the science of what is, of the kinds and structures of objects, properties, events, processes and relations in every area of reality." (Barry Smith)

Om de formele relaties tussen de concepten onderling tot uitdrukking te brengen wordt door L&C gebruik gemaakt van 450 verschillende linktypes. Deze linktypes samen vormen een hiërarchische structuur, waarbij een kindconcept meerdere ouderconcepten kan hebben. De “is-een”-relatie (subsumptie) vormt hierbij een onmisbaar linktype. Dit linktype geeft aan dat alle eigenschappen, in L&C-termen “conceptcriteria” genoemd, van het ouderconcept ook gelden voor het kindconcept én dat het kindconcept daarnaast beschikt over minimaal één conceptcriterium méér dan het ouderconcept (zie figuur 4).



Figuur 4: Hiërarchische netwerkstructuur

Doordat naast het subsumptie-linktype gebruik wordt gemaakt van de vele andere beschikbare linktypes ontstaat er een rijke semantische netwerkstructuur. Doordat alle meertalige termen en alle directe en indirecte hiërarchische en niet hiërarchische relaties die een concept heeft middels een formalisme (“description logics”) worden vastgelegd wordt ieder concept formeel gedefinieerd en wordt het uitwisselen van informatie op betekenisniveau gefaciliteerd.