

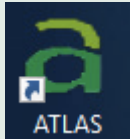
1^e MIM Snomed Hackathon 2018

Elzenveld Antwerpen, 23 november 2018

MIM-BMIA (Belgian Medical Informatics Association)
AZ Sint-Lucas Gent

1. Context AZ Sint-Lucas Gent

Vandaag - één ziekenhuisbreed kerndossier (medisch)



- met één centrale medische voorgeschiedenis
- gedeeld en samen beheerd
- synthese, voorgeschiedenis, lijn/lijn, **in vrije tekst**
- mogelijkheid tot gespecialiseerde aanvulling in detailveld

Gepland – evolutie naar Centraal EPD (medisch-zorg-spoed-OK-IZ-...)

- met één centrale medische probleemlijst
- gedeeld en samen beheerd
- synthese, probleemlijst, lijn/lijn, **gestructureerd met Snomed CT**
- **niet met lege voorgeschiedenis beginnen**
- **voorbereiden migratie**

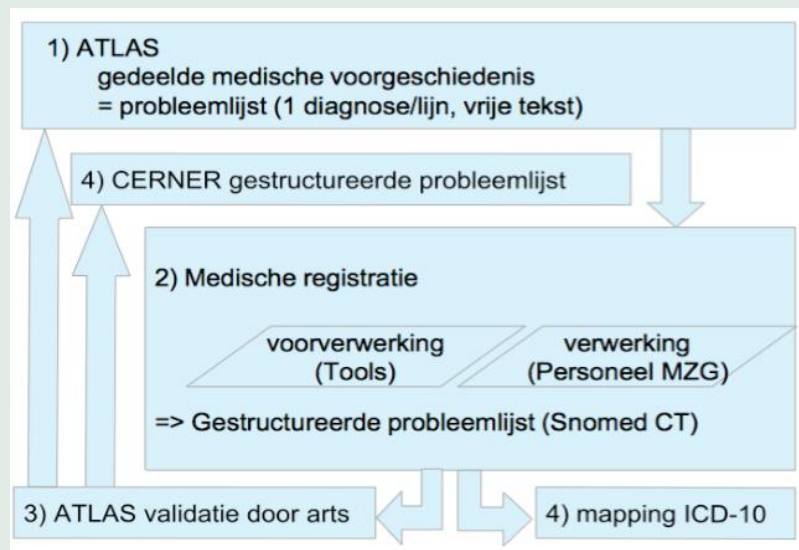


2. Hoe 400.000+ voorgeschiedenislijnen structureren en migreren?

Rosetta project



- 2 talen (Nederlands en Engels)
- 3 geschriften (vrije tekst, Snomed CT en ICD-10)
- hulp bij omzetten naar Snomed CT



voorverwerking

acceptatie

validatie

3. Hulp van NLP (natural language processing) en AI (artificiële intelligentie)?

Vereisten: snel, schaalbaar, betrouwbaar, begrijpbaar, bruikbaar

Idee: Snomed Hackathon

- Met z'n allen samenwerken om een gezamenlijk probleem oplossen (400.000 voorgeschiedenislijnen structureren naar Snomed CT)
- Naar analogie met IHE-Connectathon, focus op semantische interoperabiliteit



3. Snomed Hackathon – evaluatie commissie

MIM (Filip Veldeman, Michel Walravens)

AZ Sint-Lucas Gent (Thomas Caesens, Günther Kostka)

FOD Volksgezondheid / Belgisch Snomed NRC
(Ingrid Mertens)



LT3, Universiteit Gent
(Els Lefever, Veronique Hoste)



4. Eerste MIM Snomed Hackathon 2018, in de marge van MIC2018

Contacten met geïnteresseerde bedrijven en onderzoeksgroepen

Kick-off meeting: BT Clinical Computing, Earlytracks, IBIS.ai, Lynxcare

Prescreening, individuele gesprekken

Bevestigen van deelname

Conference calls

Voortschrijdend inzicht

- Praktisch – wat is haalbaarbaar in beschikbare tijd, hoe evalueren?
- Medico-legaal – afspraken ivm privacy, rapportering, IP
- Commercieel – geen wedstrijd, wel evaluatie van stand van de technologie
- Win/win: voor deelnemende bedrijven, MIM, AZ Sint-Lucas, NRC, LT3

Documenteren van afspraken (praktisch, technisch, inhoudelijk)

4. Snomed Hackathon kader - proces

Op voorhand 25.000 lijnen om systemen te trainen

Op moment van Hackathon 25.000 lijnen verwerken

Beoordeling door evaluatiecommissie van 2500 testlijnen

Automatische vergelijking met 'gouden standaard' en 'base line'

Keuze om breed te gaan, volumes (eerder dan 'in de diepte', meer detail)

Kwalitatieve vergelijking van aantal patronen door evaluatiecommissie

Geen wedstrijd, geen ranking

Update over stand van de technologie

Resultaten niet toewijsbaar aan individuele leveranciers

4. Snomed Hackathon kader - inhoudelijk

- Werken met vrije tekstlijnen uit de voorgeschiedenis – gegroepeerd per patiënt, verder géén context over specialisme of verdere opvolging
- Enkel concepten met semantic tags 'disorder', 'finding' en 'procedure'
- Zo veel mogelijk informatie bundelen in één concept
 - Emergency appendectomy (procedure), is beter dan appendectomy met vermelding resttekst 'urgent', is beter dan appendectomy
- Automatische vergelijking met 'gouden standaard':
 - 1 punt voor gewenste concept, 0,3 punten voor minder specifieke parent
- Kwalitatieve evaluatie (steekproeven, patronen)

5. Gouden standaard

2500 voorgeschiedenislijnen

Toekenning van Snomed CT concept door 2 senior MZG-codeurs

Interne Snomed opleiding om het potentieel van Snomed CT te kaderen
Ervaring opgedaan met coderen van AZ Sint-Lucas top 500 en top 5000 (samen met 4 medische jobstudenten tijdens de zomer).

Documenteren van keuzes/afspraken/richtlijnen in 'codeerfiches'

Codes opzoeken met gebruik van de Snomed browser

4 categorieën: 1 = testlijnen met één concept uit top 500

2 = testlijnen met één concept uit top 5000

3 = testlijnen met meerdere concepten, incl. een top 5000

4 = overige testlijnen

6. Baseline

Doel baseline:

Welke performantie kunnen we verwachten met bestaande off-the-shelf technologie?

2 automatische stappen:

1. Automatische vertaling (Machine Translation, MT)
2. cTakes (open source software om Snomed codes op te halen voor het Engels: clinical Text analysis knowledge extraction system)

6. Baseline stap 1: Machine Translation Nederlands-Engels

Technische setup:

- Preprocessing testlijnen: omzetten in UTF-8 (Unicode) codering
- Aanmaken van een automatische vertaling voor de testlijnen met de standaard Google Translate API (via installatie SDK)
- Postprocessing: vertaling omzetten in Ansi codering + vervangen speciale karakters (bijv. ' > ')

6. Baseline stap 1: Machine Translation Nederlands-Engels

Kwaliteit:

- kwaliteit van het Google Translate MT-systeem is **verrassend goed** (algemeen systeem, niet getraind op het medische domein)
- sanity check van de MT-kwaliteit op een sample van 100 entries: slechts een 10-tal entries verkeerd vertaald
- Diverse soorten vertaalproblemen

6. Baseline stap 1: Machine Translation Nederlands-Engels

Vertaalproblemen: **Samenstellingen**

- vaak woord voor woord vertaald, bijv.
 - liesbreukherstel > Lose fracture repair (ipv. Inguinal hernia), liesbreukheelkunde > lies fracture surgery, liesbreuk > “groin rupture”
 - maar “liesbreukoperatie” wordt wel correct vertaald als “inguinal hernia operation”
- Soms wordt maar een deel van de samenstelling vertaald, bijv: *valneiging* > *fall*
- Soms klopt de letterlijke vertaling wel, maar is het niet de meest gangbare term in de medisch Engelse woordenschat, bijv.
 - chronische nierinsufficiëntie > chronic renal insufficiency (ipv. Kidney failure)
 - hartritmestoornis > heart rythm disorder (ipv. Heart arrhythmia)
- Anderzijds zijn er heel wat voorbeelden van medische termen die wel correct vertaald worden, bijv. *dunne darm* > *small intestine*, *bundeltakblok* > *bundle branch block*, *schedelbreuk* > *skull fracture*, *wiegendood* > *cot death*

6. Baseline stap 1: Machine Translation Nederlands-Engels

Vertaalproblemen: **Polysemie en ambiguïteit**

- (1) de entries zijn vaak heel kort, waardoor het MT-systeem niet over voldoende context beschikt
- (2) de niet-medische betekenis komt vaker voor, bijv.

Opname > recording

Blaaspoliepen > blower polyps, blaashalsincisie > blow neck incision, blaasingreep > blow stripe, blaasoperatie > blow operation, blaasopnaaiing > blown-up (in andere gevallen is blaas wel als “bladder” vertaald)

Kleplijden > flap

Amandelen > almonds

6. Baseline stap 1: Machine Translation Nederlands-Engels

Andere vertaalproblemen:

- Overgeneratie: soms worden eigennamen (bijv. geneesmiddelen) ten onrechte vertaald, bijv. *Risperdal* > *risper deal*
- Anderzijds worden sommige woorden niet vertaald omdat ze verkeerdelijk als eigennamen beschoud worden, bijv. *Gaatje'* > *Gaatje*
- Hetzelfde probleem voor ongekeerde afkortingen, bijv. *Thv* > *thv*, *Resp* > *resp*
- Problemen met diacritische tekens (niet systematisch), bijv.
 - Flecaïnide > Fleca ï nide
 - Hemorroïdale bloeding > Hemorrhoid dialect
- Door de verkorte manier van noteren vallen er soms werkwoorden, voorzetsels, etc. weg, wat tot vertaalproblemen leidt, bijv.
 - Buisjes oren > ears of tubes

6. Baseline stap 1: Machine Translation Nederlands-Engels

Onverklaarbare vertaalfouten: kenmerkend voor de recente familie van automatische vertaalsystemen die op “deep learning” gebaseerd zijn, bijv.

- Mechanieker chauffage > mechanic driver
- Koffiegruisbraken > coffee break, koffiegruisbraken ulcer > coffee breakers ulcer
- Kleine penisingreep > small pen grip

6. Baseline stap 1: Machine Translation Nederlands-Engels

Openstaande vraag:

In hoeverre kunnen we de kwaliteit van de automatische vertaling verbeteren door:

- Een MT-systeem te trainen op medische data
- Een medisch lexicon te integreren in het proces

7. Baseline stap 2: cTakes – ophalen van bijhorende Snomed codes



clinical Text Analysis Knowledge Extraction System

Zie ctakes.apache.org

Krachtig, snel, schaalbaar, modulair, systeemonafhankelijk, gratis
NLP-systeem voor medische informatie extractie uit engelstalige vrije tekst

Ontwikkeld in 2006 in de Mayo Clinic, sinds 2013 Open Source

Gebruikt UIMA (Unstructured Information Management Architecture), XML
ontwikkeld voor IBM Watson, nu ook Open Source

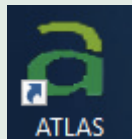
Met interactief en batch programma

Zet vrije tekst om in XML waar je de Snomed codes kan uithalen

Resultaten van de eerste MIM Snomed Hackathon

1. Context AZ Sint-Lucas Gent

Vandaag - één ziekenhuisbreed kerndossier (medisch)



- met één centrale medische voorgeschiedenis
- gedeeld en samen beheerd
- synthese, probleemlijst, lijn/lijn, **in vrije tekst**
- mogelijkheid tot gespecialiseerde aanvulling in detailveld

Gepland – evolutie naar Centraal EPD (medisch-zorg-spoed-OK-IZ-...)



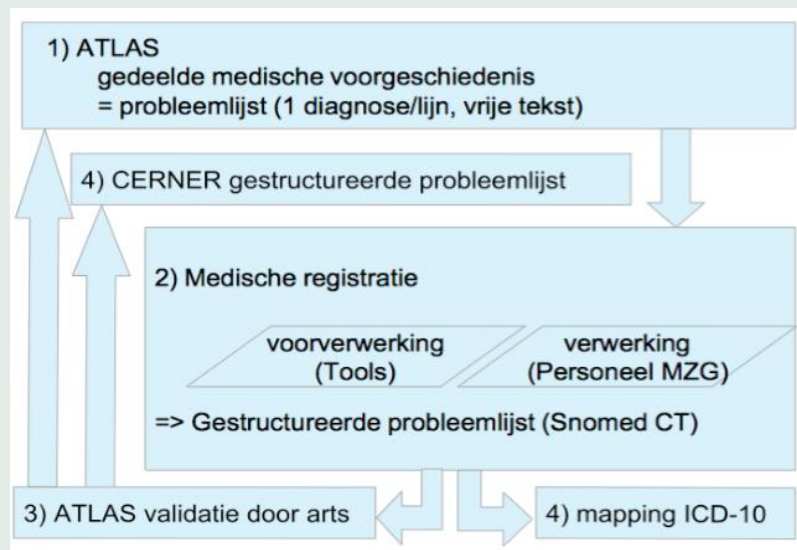
- met één centrale medische voorgeschiedenis
- gedeeld en samen beheerd
- synthese, probleemlijst, lijn/lijn, **gestructureerd met Snomed CT**
- **niet met lege voorgeschiedenis beginnen**
- **voorbereiden migratie**

2. Hoe 400.000+ voorgeschiedenislijnen structureren en migreren?

Rosetta project



- 2 talen (Nederlands en Engels)
- 3 geschriften (vrije tekst, Snomed CT en ICD-10)
- hulp bij structureren en migreren



voorverwerking

acceptatie

validatie

3. Hulp van NLP (natural language processing) en AI (artificiële intelligentie)?

Vereisten: snel, schaalbaar, betrouwbaar, begrijpbaar, bruikbaar

Idee: Snomed Hackathon

- Met z'n allen samenwerken om een gezamenlijk probleem oplossen (400.000 voorgeschiedenislijnen structureren naar Snomed CT)
- Naar analogie met IHE-Connectathon, focus op semantische interoperabiliteit



3. Snomed Hackathon – evaluatie commissie

MIM (Filip Veldeman, Michel Walravens)

AZ Sint-Lucas Gent (Thomas Caesens, Günther Kostka)

FOD Volksgezondheid / Belgisch Snomed NRC
(Ingrid Mertens)



LT3, Universiteit Gent
(Els Lefever, Veronique Hoste)



4. Vier deelnemers – géén wedstrijd – wél update over stand vd technologie

BT Clinical Computing

EarlyTracks

IBIS.ai

Lynxcare

Gouden standaard – gecodeerd door MZG-medewerkers van AZ Sint-Lucas

Baseline – tot stand gekomen via Machine vertaling NL-EN en cTakes tool

4 categorieën: 1 = testlijnen met één concept uit top 500

2 = testlijnen met één concept uit top 5000

3 = testlijnen met meerdere concepten, incl. een top 5000

4 = overige testlijnen

5. Resultaten van automatische evaluatie

	Baseline	Gemiddelde vd 4	Beste vd 4
Categorie 1	30,1 %	73,4 %	87,8%
Categorie 2	36,0 %	58,2 %	72,7%
Categorie 3	29,1 %	40,7 %	47,2%
Categorie 4	23,0 %	26,6 %	37%

Geen wedstrijd, geen ranking, geen winnaars of verliezers

Eén oefening, geen cyclische verbetercyclus feedback en fine tuning

We zien duidelijke trends en aanknopingspunten voor verdere verbetering

We bespreken individuele resultaten met de leveranciers afzonderlijk

8. Anders kan ook betekenen 'beter dan gouden standaard'

- Deels vooraf opgevangen door meerdere goeie antwoorden te voorzien
specifieke code = 1, minder specifieke 'parent' code = 0,3
varianten wanneer Snomed ruimte laat voor interpretatie
- Na de hackathon werd nog één update gedaan van de gouden
standaard, met zaken die bij de verbetering aan bod kwamen
(occasioniële fout of vaststelling van een 'allen tegen één' resultaat)
 - Vb. laatste Snomed release heeft nog extra termen waar lateraliteit ineens
is mee opgenomen, en die nog niet mee in de gouden standaard zaten.
- Individuele betere resultaten niet mee opgenomen (niet mogelijk in
beperkt tijdsbestek)
- Zo zorgvuldig mogelijk, verder 'bluts met de buil'

9. Fouten tegen technische en inhoudelijke afspraken van de Hackathon

Technische fouten:

mogelijke fouten omwille van conversieprobleem met utf-8
manier van rapporteren wanneer geen concept code gevonden is
manier van rapporteren bij negatie

=> zo goed mogelijk opgevangen om maximaal resultaat te bekomen

=> voor iedereen gelijk

Inhoudelijke fouten:

andere semantische tags (situation, event, body structure, physical object)
onvolledig samengenomen concepten

10. Fouten tegen regels van Snomed

- Niet altijd de laatste Snomed CT versie?
 - Keuze van 'oude' voorkeuromschrijving voor FSN
 - Keuze van een geïnactiveerd concept
- Foute semantic tags: concept krijgt andere tag dan in Snomed release
 - Bron?

11. NLP-gerelateerde fouten, of combinatie medisch inhoudelijk en NLP

Problemen met correcte omzetting door:

- Spelfouten
- Ongekende afkortingen
- Fonetische zoekacties die de mist ingaan
- Herkennen van samengestelde termen tot één concept
- Integreren van medische kennis
- Problematiek van vertaling NL-EN
- Integreren van specifieke richtlijnen bij keuze van concepten

=> op te vangen door feedback en training

12. Lessons learned - algemeen

Voorgeschiedenis omzetten naar Snomed CT is een leerproces

- Een gestructureerde probleemlijst is anders dan een voorgeschiedenis: 'dossierhygiëne', leren werken met structuur, leren kiezen uit lijsten
- Kennisopbouw Snomed CT bij medewerkers:
top 500, top 5000, vastleggen afspraken in codeerfiches, ...
- Nood aan verdere uitbouw van de Belgische Snomed CT release
- Nood aan ondersteunende tools:
Opzetten van een terminologie server, zorgen voor release updates
Integratie van NLP tools kunnen het manuele werk versnellen
- Nood aan integratie in bestaande en nieuwe EPD omgeving (ter klinische validatie)

12. Lessons learned - NLP

	Baseline	Gemiddelde vtd 4	Beste vtd 4
Categorie 1	30,1 %	73,4 %	87,8%
Categorie 2	36,0 %	58,2 %	72,7%
Categorie 3	29,1 %	40,7 %	47,2%
Categorie 4	23,0 %	26,6 %	37%

- We verwachten de grootste meerwaarde van NLP op categorie 3 en 4
- NLP van baseline en Hackathon deelnemers biedt een belangrijke tijds winst in het toekennen en accepteren van Snomed concepten
- Globaal scoren de Hackathon deelnemers in categorie 3 en 4 beter dan baseline
- Er is nog veel ruimte voor verbetering, in onderlinge samenwerking tussen leveranciers en gebruikers
- Nood aan initiatieven om kennis te delen en expertise uit te wisselen

13. Tot slot

Een hackathon organiseren is een leerproces

- Veel overleg en afstemming
- Veel onduidelijkheden en misverstanden wegwerken
- Veel energie stoppen in inhoud (gouden standaard, baseline)

Kan bijdragen tot kennisdeling én oplossen van een probleem waar iedereen mee te maken krijgt in een nieuwe EPD-omgeving

Dank aan alle medewerkers in MIM, AZ Sint-Lucas, Evaluatiecommissie

Dank aan alle medewerkers bij de leveranciers

Dank aan alle aanwezigen op het MIC2018

Voldoende ideeën voor 2^e editie vd Snomed Hackathon