

WordGraph2Vec: using language constructs to create sentence embeddings

ONS Machine Learning Group

29-9-2021

Marc Ponsen



What we will discuss today

- The context how WordGraph2Vec is used (why we developed the algorithm)
- How WordGraph2Vec works
- Future plans with WordGraph2Vec



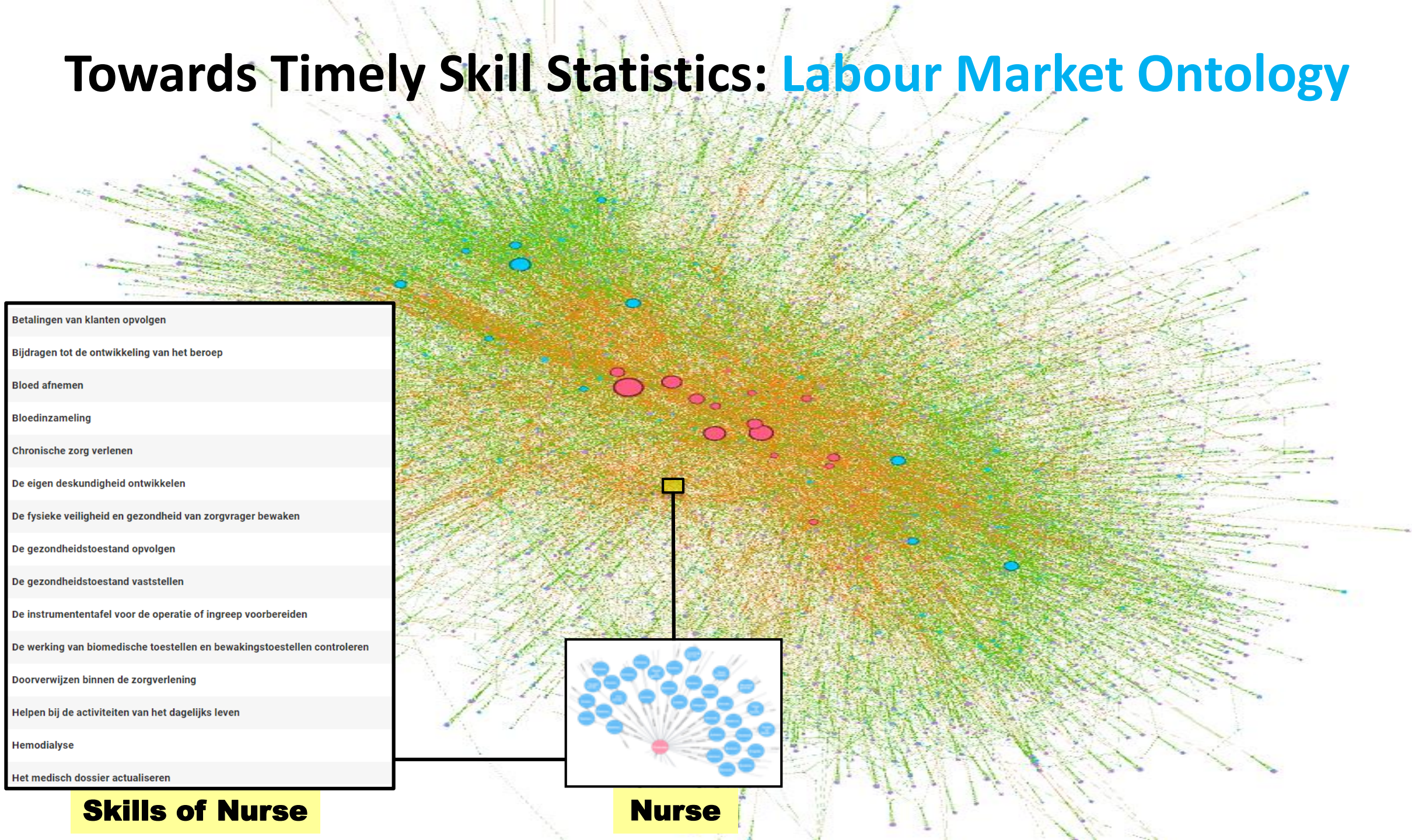
Towards Timely Skill Statistics: Labour Market Ontology

Betalingen van klanten opvolgen
Bijdragen tot de ontwikkeling van het beroep
Bloed afnemen
Bloedinzameling
Chronische zorg verlenen
De eigen deskundigheid ontwikkelen
De fysieke veiligheid en gezondheid van zorgvrager bewaken
De gezondheidstoestand opvolgen
De gezondheidstoestand vaststellen
De instrumententafel voor de operatie of ingreep voorbereiden
De werking van biomedische toestellen en bewakingstoestellen controleren
Doorverwijzen binnen de zorgverlening
Helpen bij de activiteiten van het dagelijks leven
Hemodialyse
Het medisch dossier actualiseren

Skills of Nurse



Nurse



Towards Timely Skill Statistics: Vacancies

Als projectmanager wegen onderhoud je intensieve klantrelaties waaruit je zelfstandig opdrachten verwerft en aanstuurt. Naast je projectmatige werkzaamheden houdt je je bezig met de aansturing van medewerkers en het implementeren van processen uit onder andere het managementsysteem. Je bent het aanspreekpunt voor de opdrachtgevers, projectpartners en andere betrokkenen en je zorgt voor de coördinatie van alle betrokkenen. Op basis van technisch vakmanschap geef je inhoud aan het project en neem je besluiten op het gebied van tijd, geld, kwaliteit, informatie en organisatie.

Als projectmanager wegen ben je ondernemend, klantgericht, commercieel en communicatief vaardig. Je hebt een groot probleemoplossend vermogen en neemt verantwoordelijkheid. Je bent beheersmatig sterk en een teamspeler richting de eigen organisatie. Je hebt een opleiding HTS of TU Civiele Techniek en aantoonbare ervaring met het managen en begeleiden van complexe wegenbouwprojecten.

Inr is Personeelsdiensten is een Werving & Selectie bureau dat gespecialiseerd is in de Civiele Techniek & Infrastructuur en al onze vacatures zijn voor een VAST DIENSTVERBAND bij onze opdrachtgevers.

Onze opdrachtgever is een internationale onderneming die management-, advies- en ingenieursdiensten levert op maatschappelijke relevante vraagstukken, die te maken hebben met mobiliteit, milieu & ruimte, gebouwen en water.

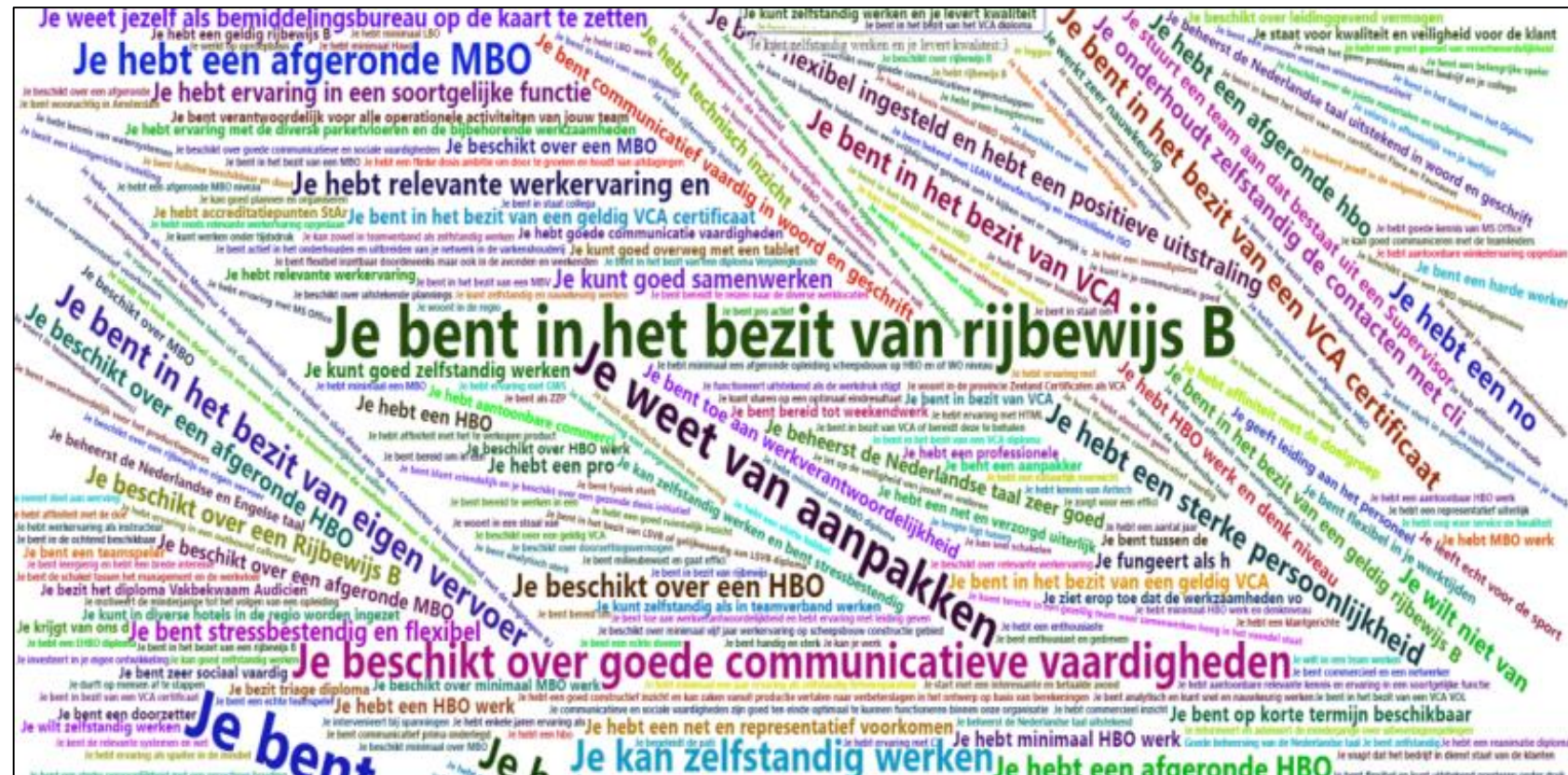
De divisie Mobiliteit realiseert omvangrijke en complexe projecten op het gebied van rail- en weginfrastructuur en luchtavens, (geo)grafische data-inwinning, -verwerking en toepassingen.

Voor haar vestiging in Amersfoort zijn wij op zoek naar een Projectmanager Wegen.

Je bent woonachtig in de regio van Midden Nederland of je bent bereid te verhuizen.

De arbeidsvoorwaarden bij onze opdrachtgever zijn zeer goed te noemen.

Vacancy text



Potential skills extracted from vacancies

WordGraph2Vec

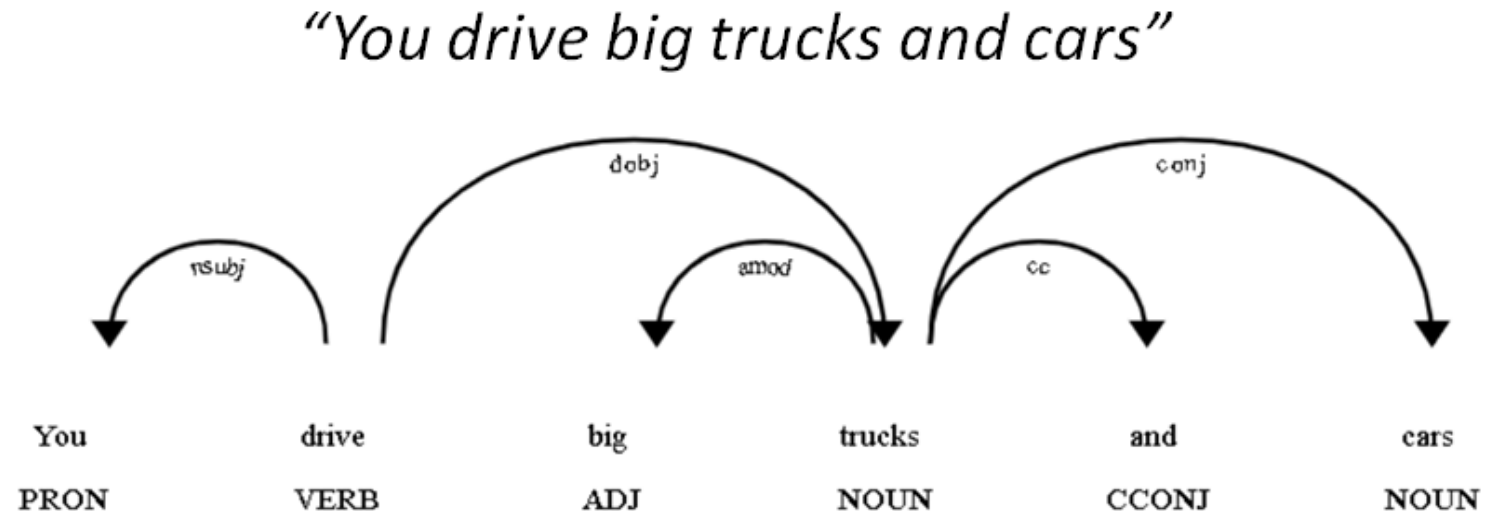
using language constructs to create sentence embeddings

- **WordGraph2Vec** combines two state-of-the-art Natural Language Processing (NLP) techniques:
 - 1) *Language models*: syntactic understanding of language
 - 2) *Vector models*: semantic understanding of language
- **Language models**: focus on relevant parts of a sentence
- **Vector models**: represent parts of sentence as numbers (i.e., a sentence embedding)

WordGraph2Vec

Language models: syntactic understanding of the sentence

- Extract so-called *WordGraphs* from sentences with language models
- *WordGraphs* can be tailored to the problem domain, e.g., finding skills



Simplified

“You drive big trucks”
“You drive big cars”

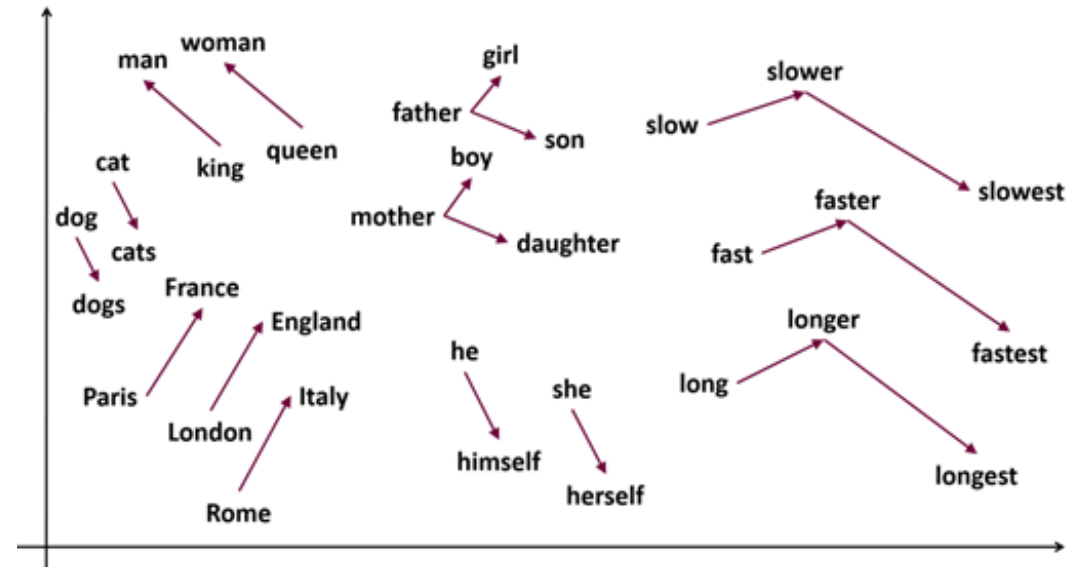
Word graphs

Activity	drive trucks	drive cars
Leading	trucks	cars
Specific	big trucks	big cars
Noun	trucks	cars
Verb	drive	

WordGraph2Vec

Vector models: semantic understanding of the sentence

- Represent WordGraphs as vectors using baseline vector model, such as Word2Vec
- We use a Word2Vec model trained on vacancy data



Word2Vec

$$w_1 = \{a_1, a_2, a_3, \dots, a_{298}, a_{299}, a_{300}\}$$

$$w_2 = \{b_1, b_2, b_3, \dots, b_{298}, b_{299}, b_{300}\}$$

WordGraph2Vec

$$np_1 = \{a_1, a_2, a_3, \dots, a_{298}, a_{299}, a_{300}, b_1, b_2, b_3, \dots, b_{298}, b_{299}, b_{300}\}$$

WordGraph2Vec:

Compute similarity between ontology skill and potential skill

Some ontology skill

Bijdragen tot de ontwikkeling van het beroep

OS₁ = {
wordgraph₁{a₁, a₂, ..., a₂₀₀, b₁, b₂, ..., b₂₀₀}
wordgraph₂{c₁, c₂, ..., c₂₀₀}
}

De eigen deskundigheid ontwikkelen

De fysieke veiligheid en gezondheid van zorgvrager bewaken

De gezondheidstoestand opvolgen

De gezondheidstoestand vaststellen

De instrumententafel voor de operatie of ingreep voorbereiden

De werking van biomedische toestellen en bewakingstoestellen controleren

Doorverwijzen binnen de zorgverlening

Helpen bij de activiteiten van het dagelijks leven

Hemodialyse

Het medisch dossier actualiseren

Ontology skills

Some potential skills

PS₁ = {
wordgraph₁{a₁, a₂, ..., a₂₀₀, b₁, b₂, ..., b₂₀₀}
wordgraph₂{c₁, c₂, ..., c₂₀₀}
}

PS₂ = {
wordgraph₂{c₁, c₂, ..., c₂₀₀}
}

...

PS_n = {
wordgraph₁{a₁, a₂, ..., a₂₀₀, b₁, b₂, ..., b₂₀₀}
wordgraph₂{c₁, c₂, ..., c₂₀₀}
wordgraph₃{d₁, d₂, ..., d₂₀₀, e₁, e₂, ..., e₂₀₀}
}

Potential skills extracted from vacancies

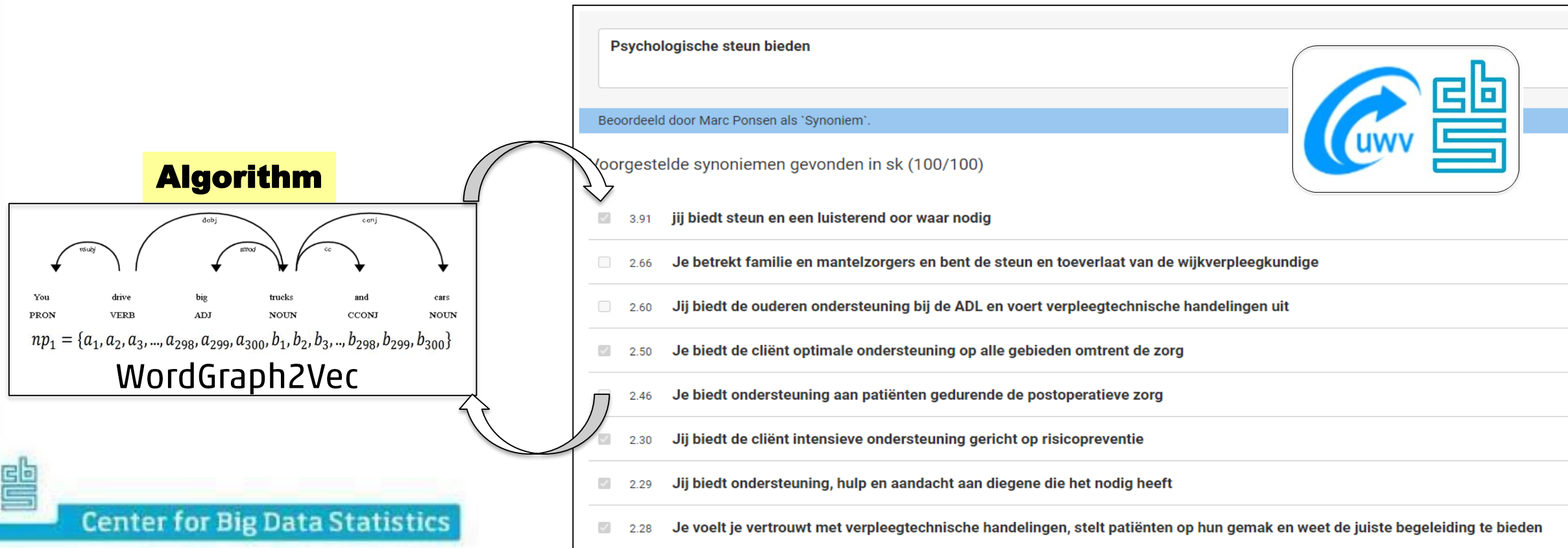
Algorithm 1 One suggested similarity score used by WordGraph2Vec.

Require: Word2vec W , Language model L , Word Graphs N , Word Graph Kernel K , Vector Kernel V

```
1: WG2V_SIMILARITY(Sentence  $S_1$ , Sentence  $S_2$ ):  
2:  $final\_score = 0$   
3: for Each word_graph_type  $n \in N$  do  
4:    $A = S_1[word\_graph\_type]$   
5:    $B = S_2[word\_graph\_type]$   
6:    $intersection = A \cap B$   
7:    $overlap\_score = |intersection|$   
8:    $symmetric\_difference\_A = (A - B)$   
9:    $symmetric\_difference\_B = (B - A)$   
10:   $difference\_score = 0$   
11:  for Each word_graph_type  $i \in symmetric\_difference\_A$  do  
12:    for Each word_graph_type  $j \in symmetric\_difference\_B$  do  
13:       $V_i = Wordgraph2vec(i)$   
14:       $V_j = Wordgraph2vec(j)$   
15:       $difference\_score += cosine\_similarity(V_i, V_j) \cdot K$   
16:    end for  
17:  end for  
18:   $final\_score += (overlap\_score + difference\_score) \cdot V$   
19: end for  
20: return  $final\_score$ 
```

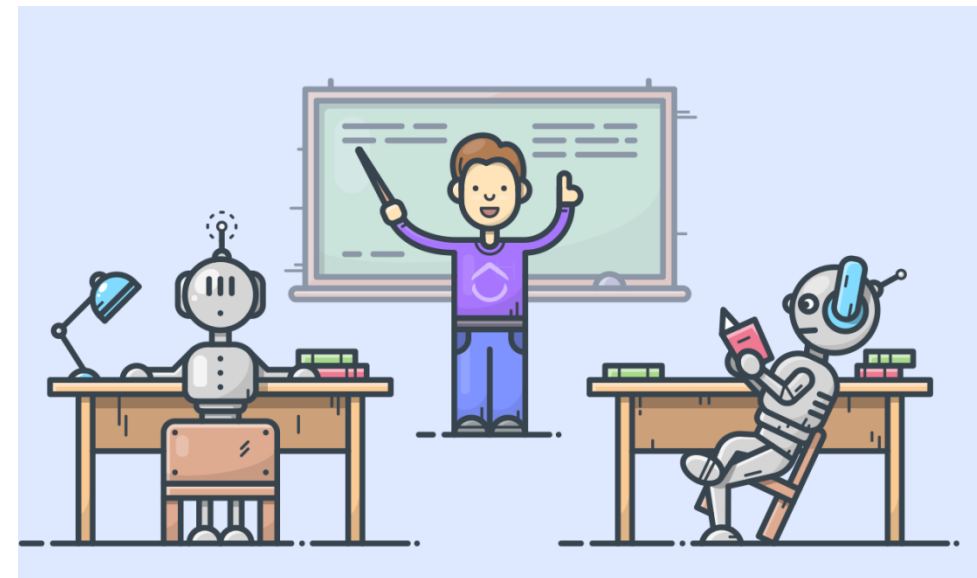
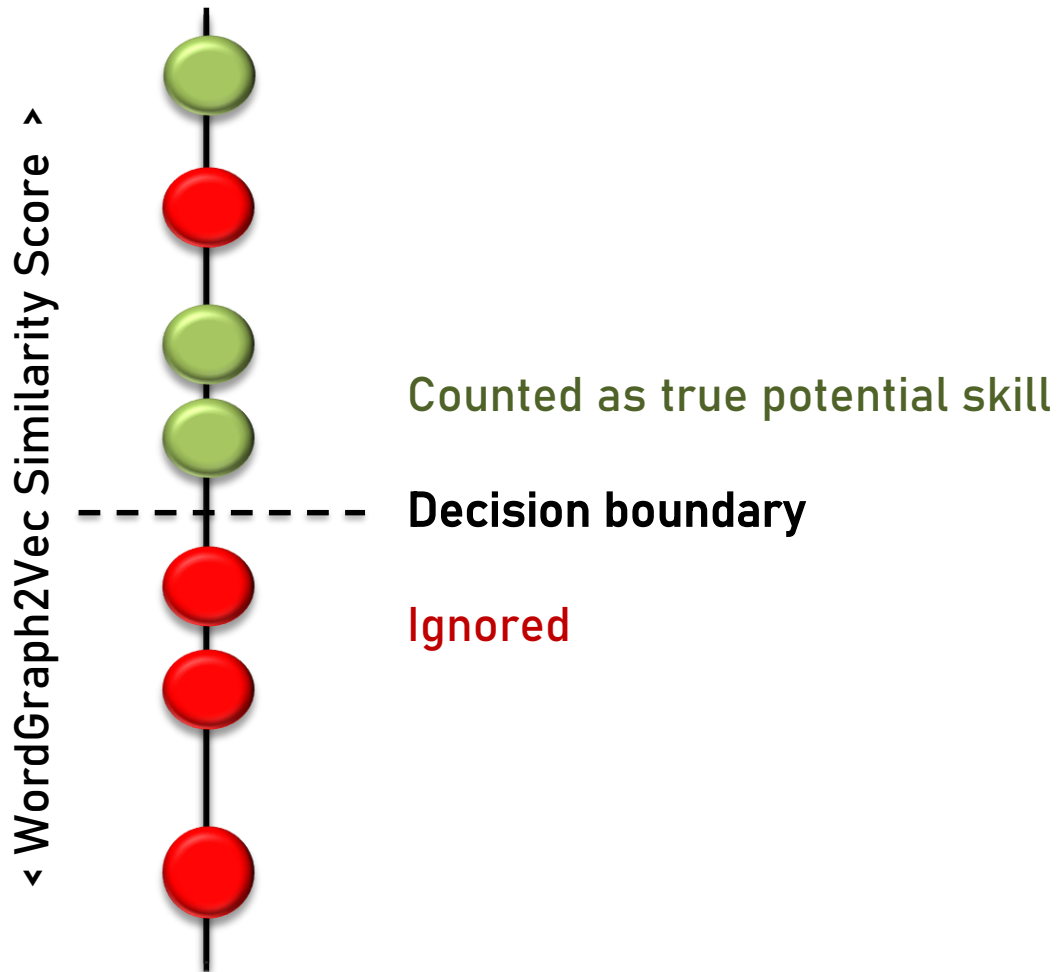
Towards Timely Skill Statistics: The Algorithm-Human Feedback Loop

- The **algorithm** makes suggestions for good connections, the **human** (two different workflows for unemployment agency Netherlands and Statistics Netherlands recognized) verifies these suggestions. The **algorithm** learns from human observations.



Towards Timely Skill Statistics: **Timely Skill Counts**

Train statistical model with features (X) (e.g., the algorithm similarity score) and human expert annotations (Y) to successfully classify “true potential skills”





Towards Timely Skill Statistics: **Timely Skill Counts**

Train statistical model with features (X) (e.g., the algorithm similarity score) and human expert annotations (Y) to successfully classify “true potential skills”

>
< WordGraph2Vec Similarity Score

Je krijgt veel tijd en ruimte om bewoners de aandacht en zorg te geven die zij nodig hebben
 Je draagt bij aan de continuïteit en kwaliteit van het welbevinden van de cliënt
 Je wordt ingezet op jouw eigen verpleegkundige deskundigheid
 Je ondersteunt bijvoorbeeld bij de lichamelijke verzorging, het geven van injecties of je doelt medicatie uit
 Jij stelt de bewoners op hun gemak en zorgt naast de lichamelijke verzorging ook voor een sociaal moment
 Je gaat patiënten ondersteunen door hen te leren omgaan met hun psychische stoornis en je streeft samen met de patiënt naar optimaal functioneren
 Je verricht uiteenlopende taken om de verpleegkundige zorgverlening te verbeteren
 Je voert de testen uit met volledige persoonlijke bescherming
 Jij begeleidt en verleent zorg aan cliënten zoals die is afgesproken in het individuele behandelplan
 Je vast welke zorg er nodig is en als het kan, betrek je de familie hierbij
 Je gaat aan de slag in de ouderenzorg en weet de bewoners veel persoonlijke aandacht en zorg te geven
 Je controleert bewoners onder andere op hun gezondheid en je geeft medicatie wanneer dit nodig is
 Je begeleidt de patiënt na de ingreep en monitort de fysieke toestand waarin de patiënt na een operatie in verkeerd
 Je coördineert de zorg op een effectieve en efficiënte manier waarbij je de kernwaarden hartelijk, huiselijk en
 Jij bewaakt de kwaliteit van de zorg en zorgt ervoor dat het zorgdossier altijd up to date is
 Jij haalt voldoening uit de begeleiding van psychiatrisch patiënten en zorgt er samen met de patiënt voor dat hij zijn doelen behaalt
Je bewaakt de kwaliteit van de zorg en waar mogelijk verhoog je deze ook
 Je verleent de benodigde zorg, verricht medische handelingen en maakt de zorgplannen
 Je stelt samen met de bewoners persoonlijke zorgleefplannen op en coördineert de zorg
 Je werkt samen met je collegas, monitort de kwaliteit en levert een positieve bijdrage aan de ontwikkeling van de afdeling
 Je voert gesprekken en analyseert hun lichamelijke en geestelijke functioneren en ondersteunt ze hier in
 Je begeleidt jouw patiënten in de dagelijkse leefomgeving
 Je traint en begeleidt de medewerkers in een methodische aanpak ter verbetering van de kwaliteit van zorg
 Je stuurt een team aan en bewaakt de kwaliteit
 Je levert een bijdrage aan het optimaliseren van verpleegkundige zorg en neemt deel aan afdelingsoverstijgende werkgroepen en commissies ter bevordering van de kwaliteit van zorg
Je draagt bij aan het opstellen en uitvoeren van het zorgplan en denkt graag mee over de zorg die geboden wordt
 Je betreft het hele gezin actief bij de zorg om een goede hechting met de pasgeborene te bevorderen
 Je begeleidt onder andere patiënten met aangeboren afwijkingen, participeert in research en registreert, bewaakt en interpreteert de hemodynamische en elektrocardiografische parameters

Je werkt aan je eigen persoonlijke en professionele ontwikkeling
 Je blijft inzetten in dit mooie vak
 Je enthousiast geworden
 Je voert zorgtaken en bijzondere verpleegtechnische handelingen uit
 Je maakt zelf de afspraken met de cli

Je krijgt uitgebreide begeleiding van je consultant vanuit Randstad Zorg Professionals
 Je komt in dienst bij deze zorgorganisatie wordt onderdeel van een hecht team waar collegialiteit, persoonlijke ontwikkeling en vakinhoudelijke uitdagingen en groei centraal staan

Je werkt aan je persoonlijke ontwikkeling maar ook aan je professionele ontwikkeling
 Je voert verzorgende handelingen uit, signaleert mogelijke gezondheidsproblemen bij cliënten en helpt met de ADL-zorg
 Jij zorgt ervoor
 Jij contact legt met de cliënt
 Je werkt nauw samen met je collega
 Jij doet alles om kwalitatieve zorg te kunnen bieden
 Je wordt ingezet op verschillende projecten
 Jij maakt echt het verschil
 Je geeft je collega-verzorgenden regelmatig advies

Je werkt binnen een vast team waar wordt gewerkt volgens de richtlijnen van het RIVM
 Jij bepaalt met het beste bij jou past en waarvan jij gelukkig wordt
 Jij voelt je als een vis in het water in de zorg
 Je hoeft niet te zoeken tussen 100 vacatures
 Je contact op met Gijs Oosterveld op 06-12631989 of via gijs
 Je verricht verpleegtechnische handelingen en gestrekt medicatie aan de cli

Je geeft zelf aan hoeveel je wil werken
 Jij staat stevig in je schoenen en houdt rustig contact met hen
 Jij geeft de ouderen de nodige medische en sociale ondersteuning
 Je verricht of assisteert bij diagnostisch en therapeutisch onderzoek uitgevoerd door een medisch specialist
 Je ingezet wordt

je gaat doen
 Je werkt straks als verpleegkundige op contract bij Randstad Zorg Professionals diverse uitdagende projecten
 Je werkt op basis van detachering

Je werkt straks als verpleegkundige in dienst bij Randstad Zorg Professionals op diverse uitdagende projecten
 Je gaat werken aan je professionele en persoonlijke ontwikkeling
 Je komt te werken op een groep met Moeilijk Verstaanbaar Gedrag en je krijgt echt de kans om de cliënten goed te leren kennen

Jij maakt iedere dag het verschil voor jouw cliënten of patiënten
 Je onderhoudt contacten met de bewoners en hun relaties, met collegas en met externe behandelaars en hulpverleners

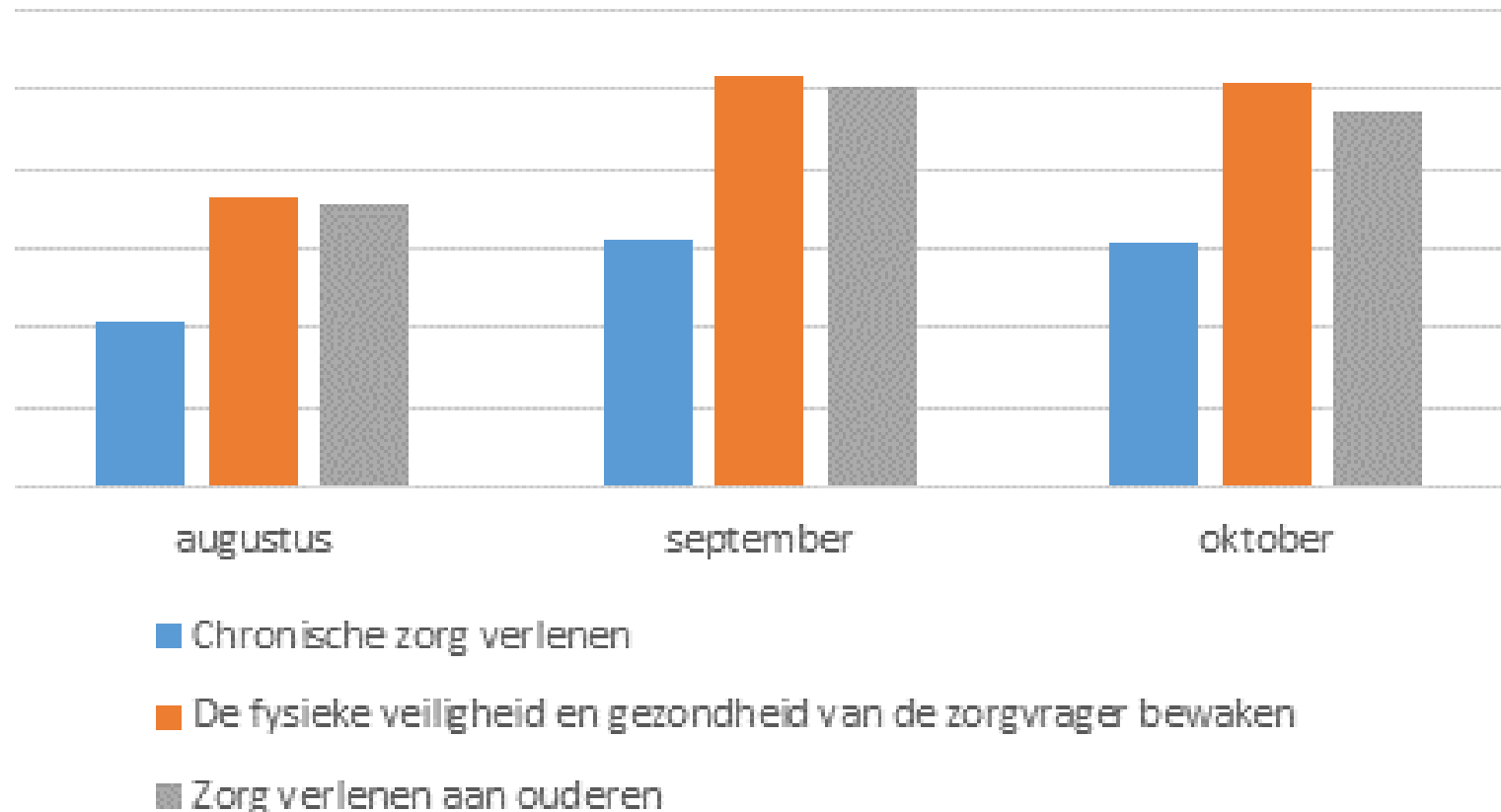
Je werkt tijdens het programma als verpleegkundige diverse uitdagende projecten
 Je staat met regelmaat alleen op de groep waardoor jij de bewoners goed kent en hen goed kan voorzien in hun individuele zorgbehoeften

Je echt zelf waar, wanneer en hoe vaak je werkt



Towards Timely Skill Statistics: **Timely Skill Counts**

Skill counts based on vacancy data



Je krijgt veel tijd en ruimte om b

Je ondersteunt bijvoorbeeld bij de l

Je gaat
Je

Je begeleid

Je bewaakt

Je v

Je stelt samen met de bewoners perso

Je voe

Je draagt bij aa

Je betreft het hele gezin actief bij de zorg

idsnatige taken uit te voeren

deskundigheid

timaal functioneren
bescherming

kan, betrek je de familie hierbij

en je geeft medicatie wanneer dit nodig is

en behaalt

de zorg en
ing

dagelijkse leefomgeving

waliteit van zorg

boden wordt

nische en elektrocardiografische parameters

WordGraph2Vec:

Future plans

- WordGraph2Vec has several hyper parameters that have not been optimized, goal is to use the expert feedback to find optimal parameters and improve similarity scores
- Make a more thorough comparison with existing techniques (e.g., Paragraph2Vec, AverageWord2Vec, ...)
- Apply WordGraph2Vec to different domains