

**Health
Campus**

**Den
Haag**

Natuurlijke Taalverwerking in de Zorg

Natural Language Processing (NLP)

Via fundamentele taalmodellen naar slimme zorg

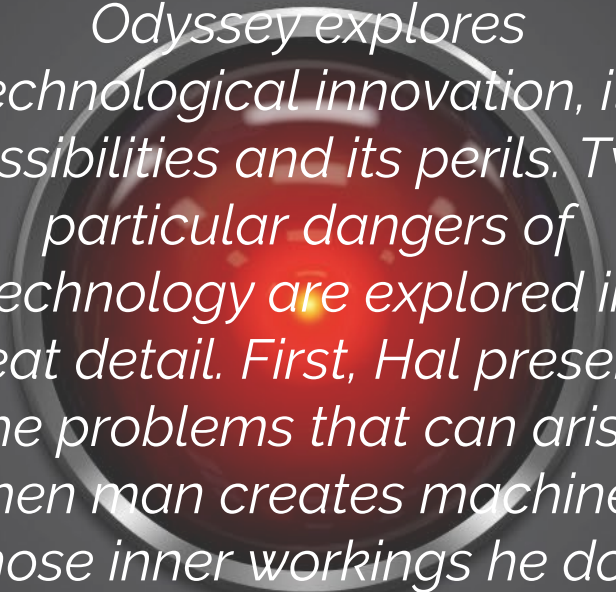
AI & Technologiweek Geneeskunde jaar 1, 2 juli 2024

Prof.dr. Marco Spruit





Dave Bowman: Open the pod bay doors, HAL.
HAL: I'm sorry Dave. I'm afraid I can't do that.



"2001: A Space Odyssey explores technological innovation, its possibilities and its perils. Two particular dangers of technology are explored in great detail. First, Hal presents the problems that can arise when man creates machines, whose inner workings he does not fully understand [...]"

"2001: A Space Odyssey"

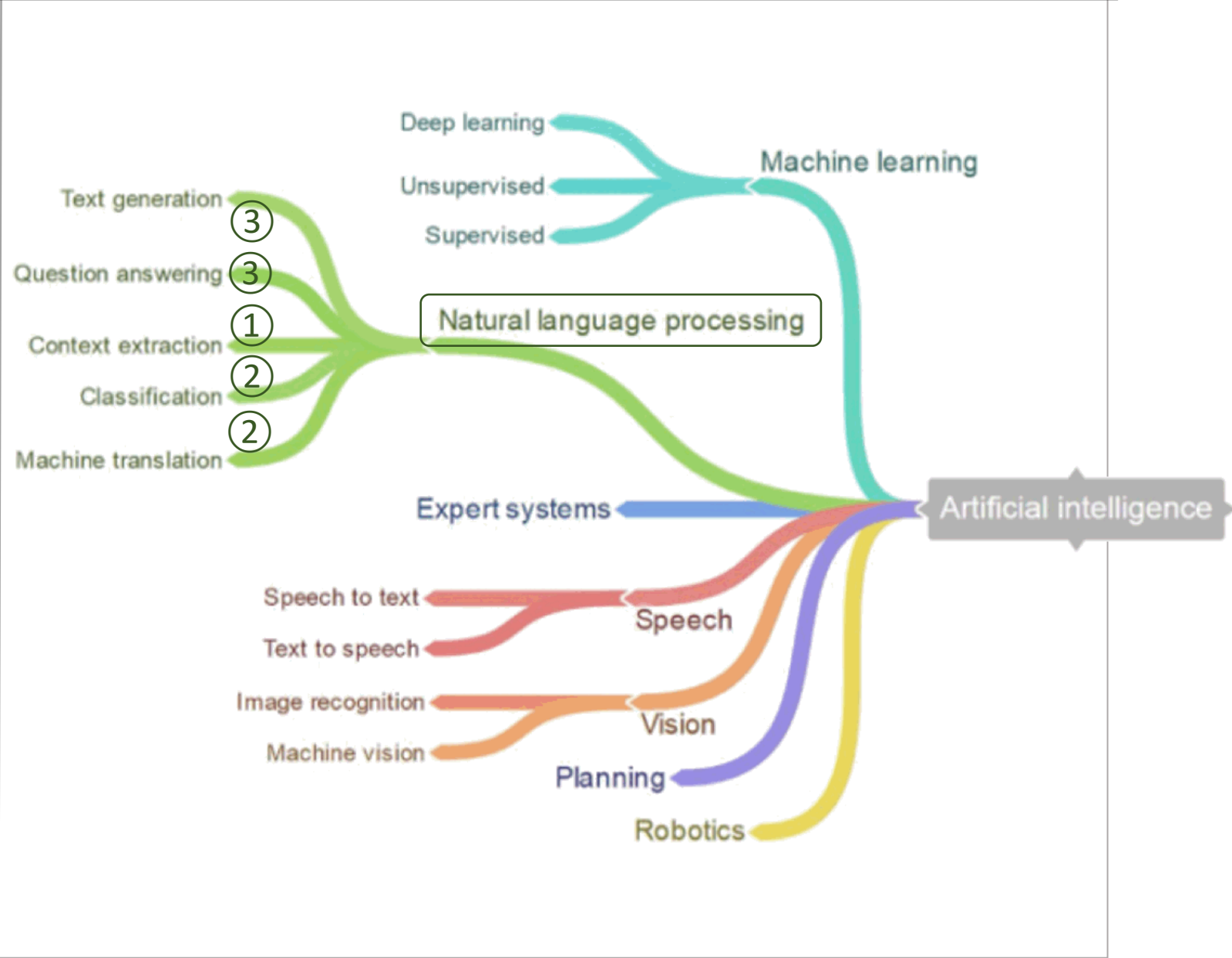
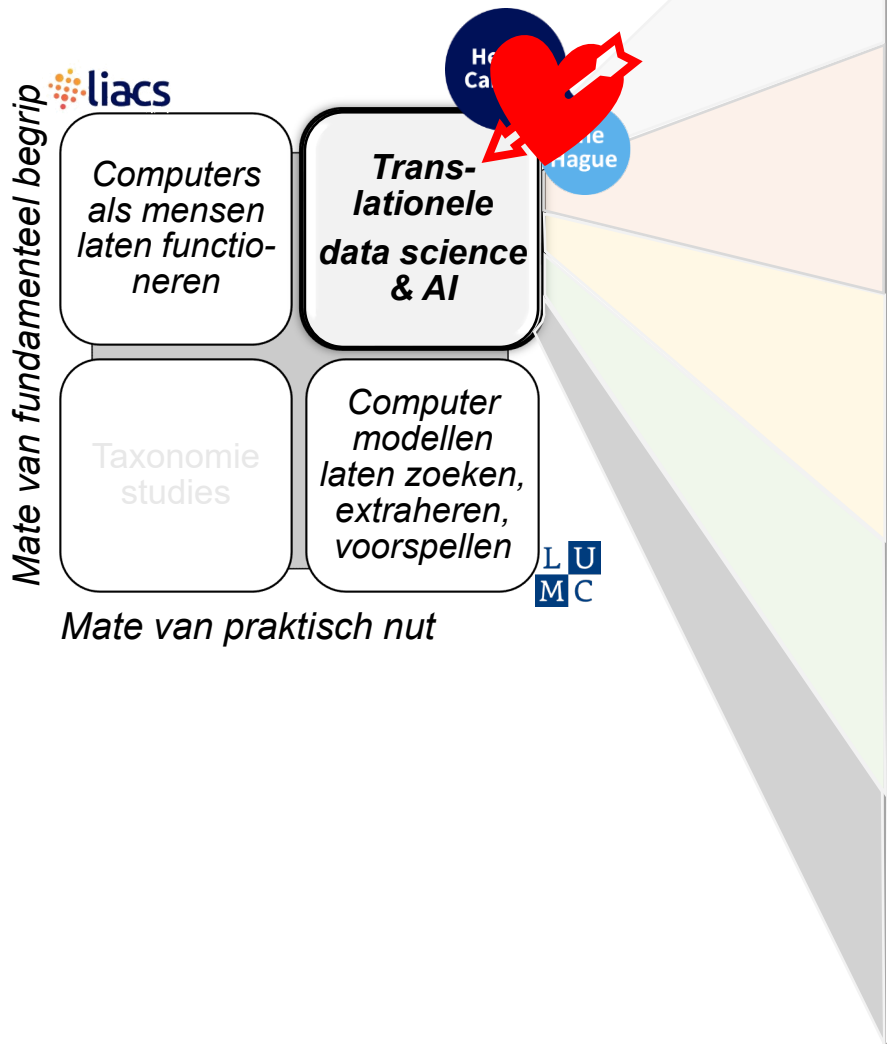
Een film uit 1968 van Stanley Kubrick, gebaseerd op werk van Arthur C. Clarke.



ChatGPT

Voer je vraag in...

NLP: Deelgebied binnen Artificiële Intelligentie



Logisch regel
systeem
ontwerpen
dat entiteiten
in context
destilleert

**Trans-
lationele
data science
& AI**

Taxonomie
studies

Computer
modellen
laten
zoeken,
extraheren,
voorspellen

L U
M C

Mate van praktisch nut

Voorbeeld #1: Zoeken naar SNPs in de literatuur

SNP Curator

Results for **ibd**:

- A total of **1535** articles were fetched by PubMed.
- The extracted list of abstracts was shortened to **280** via selecting those comprised of SNP mentions.
- **150** PubMed article(s) had statistical results reported withing the abstract text with a total of **524** SNP pairs.

[Go back to home page](#) [Export Data to CSV File](#)

SNP	PMID	Title	Date	Pvalue	ORvalue	Ethnicity	P g S
rs61750370	29788244	Nonsynonymous Polymorphism in Guanine Monophosphate Synthetase Is a Risk Factor for Unfavorable Thiopurine Metabolite Ratios in Patients With Inflammatory Bowel Disease.		0.031		Caucasian	2
The SNP rs61750370 was significantly associated with 6-MMP:6-TGN ratios ≥ 100 odds ratio, 5.64; 95% confidence interval, Caucasian IBD patients. The GMPS SNP rs61750370 may be a reliable risk factor for extreme 6MMP preferential metabolism							
rs61750370	29788244	Nonsynonymous Polymorphism in Guanine Monophosphate Synthetase Is a Risk Factor for Unfavorable Thiopurine Metabolite Ratios in Patients With Inflammatory Bowel Disease.		0.031		Caucasian	2
rs16969968	29688464	Smoking Interacts With CHRNA5, a Nicotinic Acetylcholine Receptor Subunit Gene, to Influence the Risk of IBD-Related Surgery.		0.05			

Voorbeeld #2: Deïdentificatie van persoonlijke info

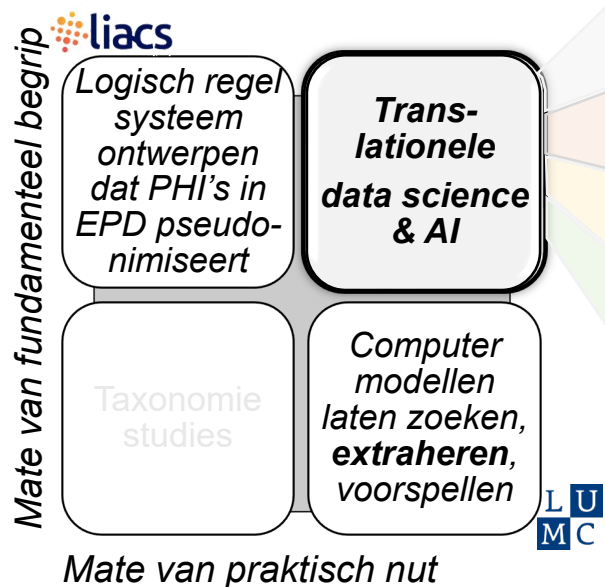
[Legend: Patient Persoon Locatie Instelling Datum Leeftijd Patientnummer
Telefoonnummer Url]

Annotated text

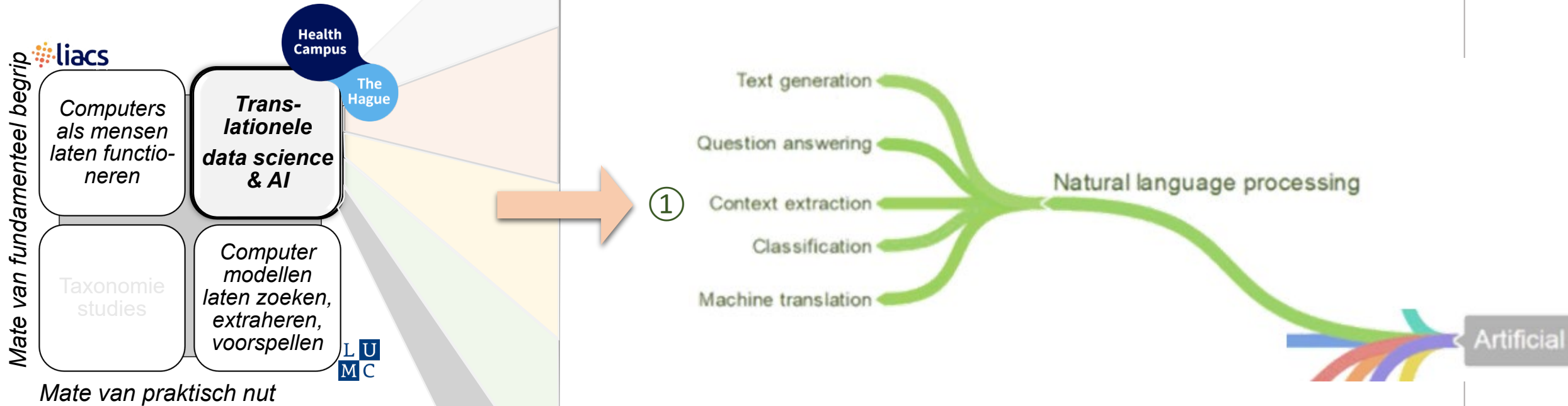
Intakegesprek met Jan Jansen (e:j.g.jsnen_1966@email.com, t:0612345678, patnr:1243567). Het betreft een 51-jarige man die van 14 maart t/m 31 juli op de polikliniek van het umcu zal worden behandeld i.v.m. somberheidsklachten. Patient is woonachtig aan de Voorstraat 45b in Utrecht en zal hier onder behandeling komen te staan van Peter de Visser.

De-identified text

Intakegesprek met <PATIENT> (e:<URL-1>, t:<TELEFOONNUMMER-1>, patnr:<PATIENTNUMMER-1>). Het betreft een <LEEFTIJD-1>-jarige man die van <DATUM-1> t/m <DATUM-2> op de polikliniek van het <INSTELLING-1> zal worden behandeld i.v.m. somberheidsklachten. Patient is woonachtig aan de <LOCATIE-1> in <LOCATIE-2> en zal hier onder behandeling komen te staan van <PERSOON-1>.



Agenda [1/3]



Welk LLM is
het meest
geschikt voor
klinische
tekst
classificatie?

**Trans-
lationele
data science
& AI**

Taxonomie
studies

Leefstijl
kenmerken
extractie uit
EPD voor
persoonlijke
prognoses

① Informatie Extractie in NLP: Probleem Begrip



Voorspellingsmodel onderschat risico op hart- en vaatziekten

Redactioneel



Michel van Dijk

1 februari 2023

Het model dat door huisartsen gebruikt wordt om te voorspellen hoe groot het risico is dat iemand binnen tien jaar een hart- of vaatziekte krijgt, onderschat het aantal mensen dat dit ook daadwerkelijk krijgt. Het risico is anderhalf keer zo groot dan het model verwacht, blijkt uit een studie van het Radboudumc. Marion Biermans, epidemioloog en onderzoeker bij de afdeling Eerstelijns geneeskunde, over de uitkomsten van de studie.

Om het risico te voorspellen dat iemand een hart- en vaatziekte krijgt, bijvoorbeeld bij mensen die roken of met obesitas, maken huisartsen gebruik van meerdere voorspellingsmodellen. De multidisciplinaire richtlijn Cardiovasculair

Classificatie, NLP taken

Vertalen, samenvatten

Conversatie, Creatie, ...

Mate van fundamenteel begrip

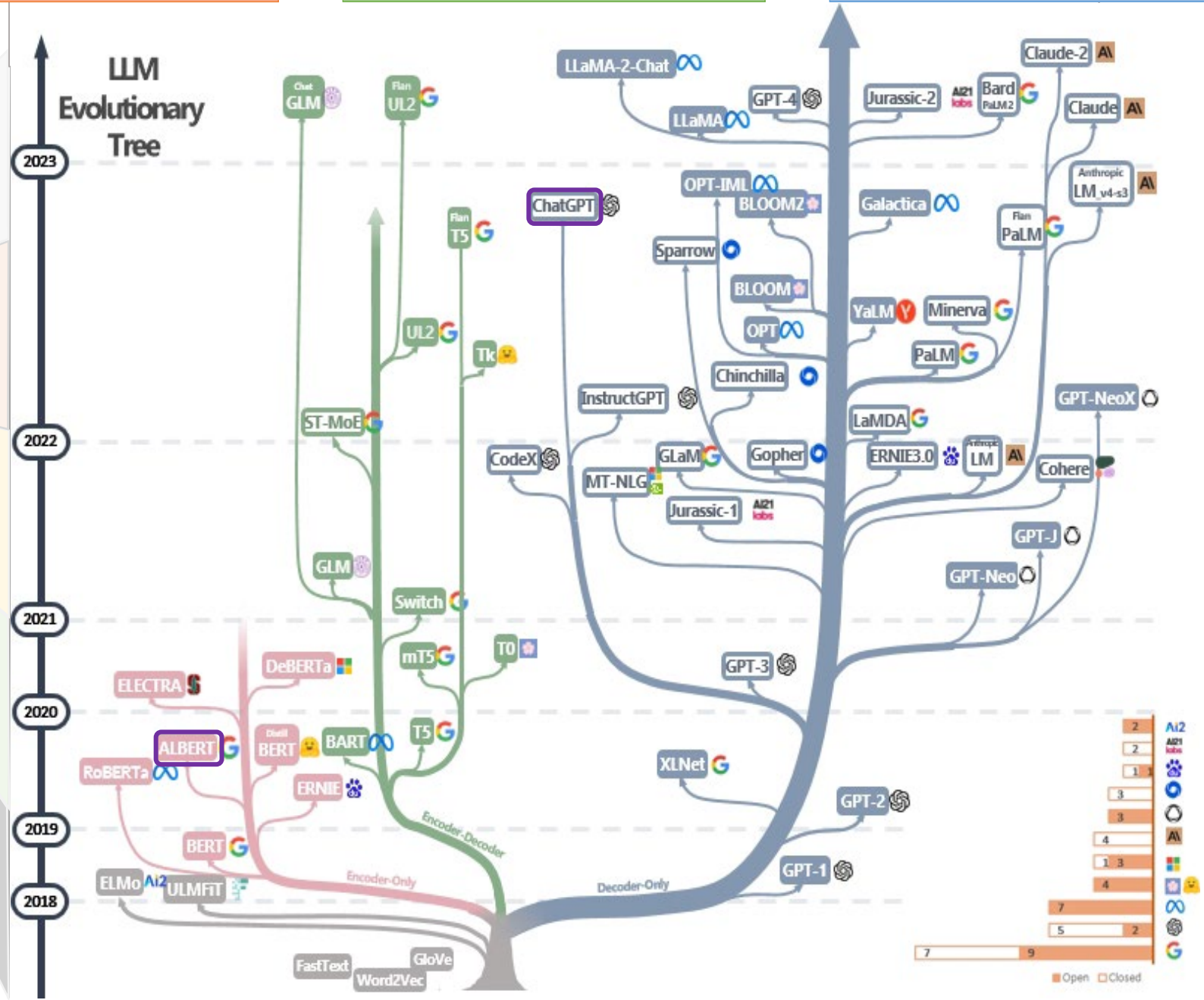
Welk LLM is het meest geschikt voor klinische tekst classificatie?



Taxonomic studies

Leefstijl kenmerken extractie uit EPD voor persoonlijke prognoses

Mate van praktisch nut



① Informatie Extractie in NLP: Data Omschrijving

Mate van fundamenteel begrip

Welk LLM is het meest geschikt voor klinische tekst classificatie?

**Trans-
lationele
data science
& AI**

Taxonomie
studies

Leefstijl
kenmerken
extractie uit
EPD voor
persoonlijke
prognoses

Mate van praktisch nut

Voorbeeld tekstfragmenten		Roken	Alcohol	Drugs	
Patient rookt drinkt niet gebruikt geen drugs		Huidige gebruiker	Niet gebruiker	Niet gebruiker	
roken (+ (20 packyears, is gestopt); drinkt 1 bier per dag		Voormalig gebruiker	Huidige gebruiker	Onbekend	
Roken: - (20 jaar geleden na 50 py), Cardiovasculaire risicofactoren: roken:+; gebruikt dagelijks marihuana		Voormalig gebruiker	Onbekend	Huidige gebruiker	
Type label	# gelabelde teksten	Huidige Gebruikers	Voormalige Gebruikers	Niet Gebruikers	Onbekend
Roken	148.768	7.015 (4.72%)	32.230 (21.66%)	44.677 (30.03%)	64.846 (43.59%)
Alcohol	143.166	16.017 (11.25%)		39.119 (27.32%)	87.940 (61.43%)
Drugs	147.999	1.443 (0.98%)		53.005 (35.81%)	93.551 (63.21%)

① Informatie Extractie in NLP: Data Voorbeelden

Text with highlighted words

Beloop: Beloop: TC na opname Af en toe nog wat buikpijn, maar niet meer zo hevig. Geen RBV meer. Is niet gestopt met roken, moeder erg ziek, stress, moeder heeft ongeneeselijke longca. Conclusie: Conclusie: 19-11 tot 20-11 opname MDL ivm een ischaemische colitis waarvoor conservatief beleid. Beleid: Beleid: Nogmaals besproken staken roken voor eigen gezondheid Geinstrueerd afspraak te maken bij HA voor CRM + stoppen met roken begeleiding EB geen brief

Text with highlighted words

Reden van komst / Verwijzing: Reden verwijzing: Oogleden Anamnese: Voorgeschiedenis Blanco Medicatie Geen Anamnese Sinds 0,5jr progressieve hoofdpijn en vermoeidheid na ontstaan afhangen bovenooglid li. Kan hierbij ook iets minder zien, moet mn bij autorijden de wenkbrauwen helemaal optillen. Kan 's avonds niet meer goed lezen vanwege de vermoeidheid. Merkt zelf geen verschil tussen li en re. Intoxicaties Roken 20jr terug gestopt, alcohol-, drugs- Allergie's n Geen bekend Lichamelijk onderzoek: Wenkbrauwen bdz op niveau van de orbitarand. Dermatochalasis re mild, li tot op de wimpers. Levatorfunctie bdz intact, goede functie. Geen ptosis. Conclusie: Conclusie: Milde dermatochalasis. Beleid: Beleid: Helaas voldoet de ernst van de dematochalasis niet aan de voorwaarden die de zorgverzekeraar stelt aan vergoeding voor een blepharoplastiek. Patiënt werd op de mogelijkheden gewezen van zelfbetaling. De patiënt gaat hier over nadenken.

Text with highlighted words

Beloop: Beloop: Patient gesproken: ja Verslag gesprek met patient: gb RAST: HSM 1.5, rest negatief. Conclusie: intolerant voor rook van vuur en graspollen. Gemaakte afspraken: uitslag besproken, stop immunotherapie HSM (3,5 jaar gehad). Geen indicatie voor opstarten andere immunotherapie. EB Vervangt dit gesprek een herhaal polikliniekbezoek: ja

① Informatie Extractie in NLP: Data Preparatie

Mate van fundamenteel begrip

Welk LLM is het meest geschikt voor klinische tekst classificatie?

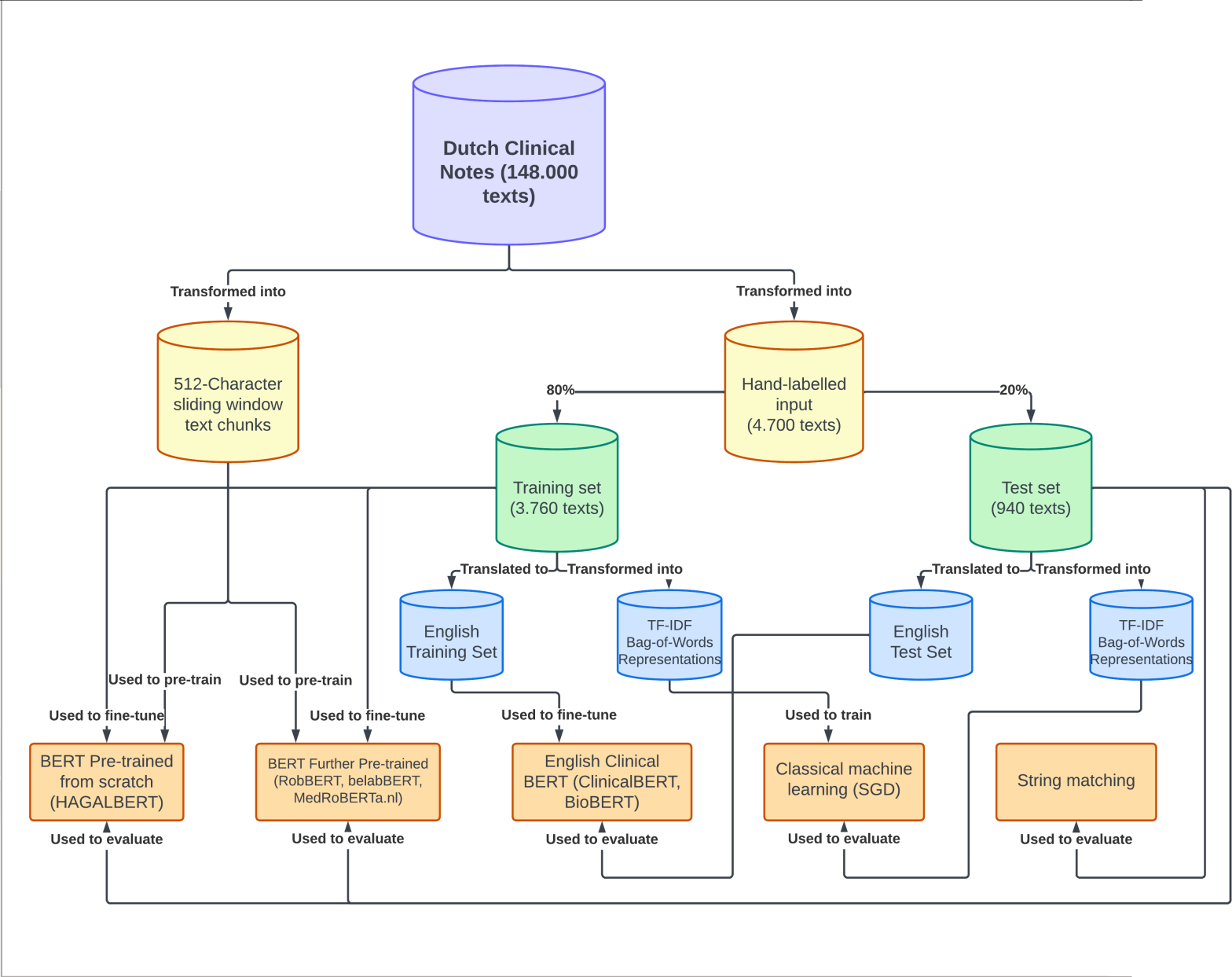
**Trans-
lationele
data science
& AI**

Taxonomic studies

Leefstijl kenmerken extractie uit EPD voor persoonlijke prognoses

Mate van praktisch nut

Computationele experimenten



Mate van fundamenteel begrip

Welk LLM is
het meest
geschikt voor
klinische
tekst
classificatie?

**Trans-
lationele
data science
& AI**

Taxonomic
studies

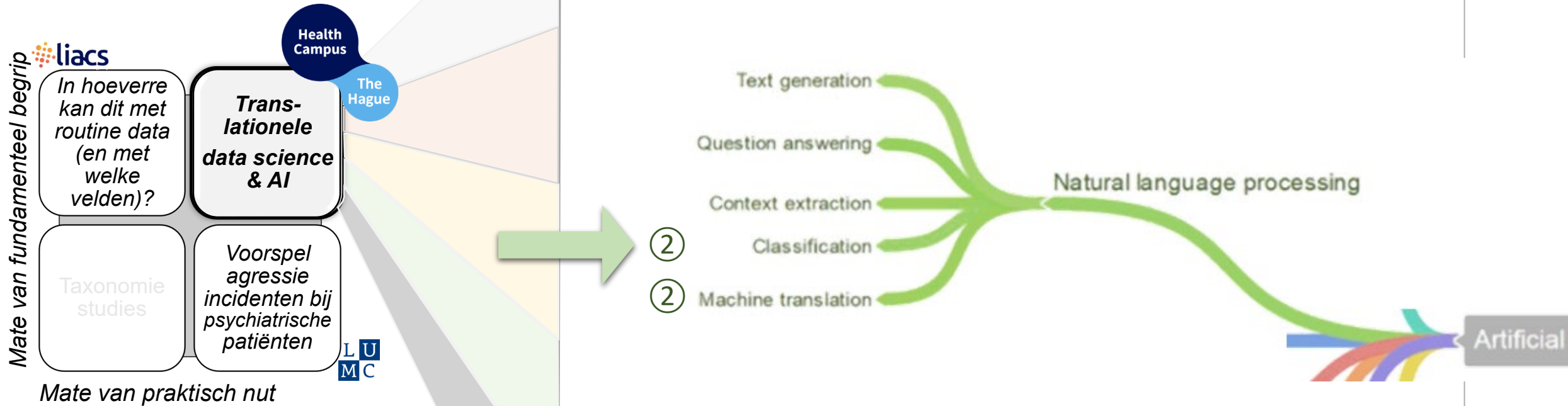
Leefstijl
kenmerken
extractie uit
EPD voor
persoonlijke
prognoses

Mate van praktisch nut

① Informatie Extractie in NLP: Data Modelleren

Model	Smoking	Alcohol	Drugs
String Matching	0.84	0.74	0.68
Machine Learning (SGD)	0.85	0.71	0.60
HAGALBERT	0.66	0.54	0.43
RobBERT-HAGA	0.87	0.71	0.63
belabBERT-HAGA	0.48	0.64	0.57
MedRoBERTa.nl-HAGA	0.93	0.79	0.77
BioBERT (translated)	0.91	0.72	0.52
ClinicalBERT (translated)	0.92	0.80	0.61

Agenda [2/3]

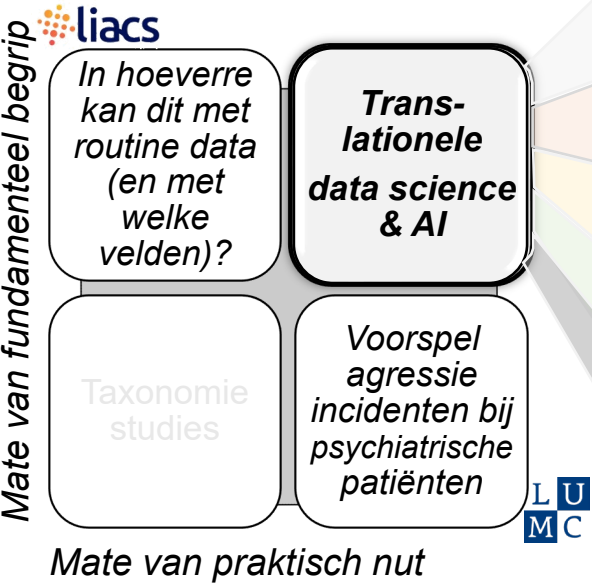


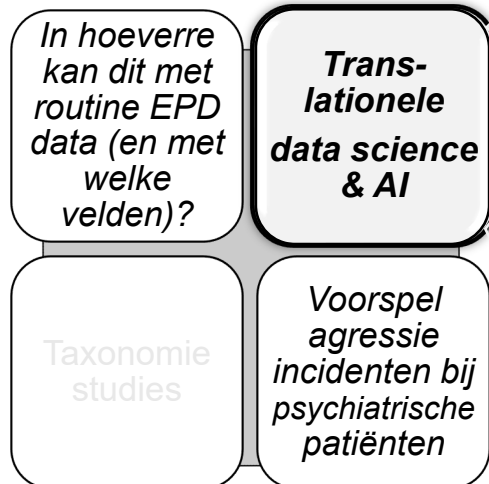
② Classificatie in NLP: Probleem Begrip



NU'91: 'Agressie tegen zorgprofessionals blijft groot probleem'

29 december 2023 – Meer dan de helft van de zorgprofessionals kreeg het afgelopen jaar maandelijks of vaker te maken met een vorm van agressie. Dat varieerde van schelden en schreeuwen tot aan spugen, doodsbedreigingen of fysiek geweld. Dit blijkt uit een peiling van NU'91 in samenwerking met WNL. Femke Merel van Kooten, voorzitter van NU'91, noemt de resultaten schokkend.





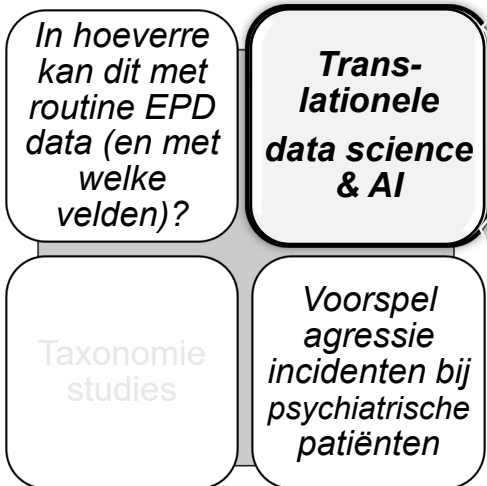
② Classificatie in NLP: Data Omschrijving

(2012-07-29)

“Mw heeft matig geslapen, sliep van 1.00 uur tot 4.00 uur. Kwam toen uit bed, at koekjes en dronk thee. Nog geadviseerd medicatie te nemen en mijn zorgen geuit over evt. doorschieten in een manie. Mw was er niet gevoelig voor en reageerde geagiteerd. Mw had spreekdrang maar gaf aan dat wanneer zij zich goed voelt ook veel praat. Mw gaat vandaag naar <PERSOON-1> met haar zoon, ziet daar nu niet meer tegenop omdat de klachten die zij gisteren aan haar voeten ervaren verdwenen zijn. Mw ging na 4.00 uur weer naar bed en kwam niet meer uit haar kamer tot de ochtend.”

... Verwacht je, op basis van klinische notities zoals deze, dat er een agressie incident zou kunnen gaan plaatsvinden in de komende 30 dagen, **JA** of **NEE**?





② Classificatie in NLP: Data Omschrijving

(2012-07-29)

“Mw heeft matig geslapen, sliep van 1.00 uur tot 4.00 uur. Kwam toen uit bed, at koekjes en dronk thee. Nog geadviseerd medicatie te nemen en mijn zorgen geuit over evt. doorschieten in een manie. Mw was er niet gevoelig voor en reageerde geagiteerd. Mw had spreekdrang maar gaf aan dat wanneer zij zich goed voelt ook veel praat. Mw gaat vandaag naar <PERSOON-1> met haar zoon, ziet daar nu niet meer tegenop omdat de klachten die zij gisteren aan haar voeten ervaren verdwenen zijn. Mw ging na 4.00 uur weer naar bed en kwam niet meer uit haar kamer tot de ochtend.”

(2012-08-05)

[...] “Uit verzet trapte mw. mij tegen de zijkant van mijn been” [...]

Mate van fundamenteel begrip

In hoeverre
kan dit met
routine data
(en met
welke
velden)?

**Trans-
lationele
data science
& AI**

Taxonomie
studies

Voorspel
agressie
incidenten bij
psychiatrische
patiënten

Mate van praktisch nut

② Classificatie in NLP: Data Preparatie

(2012-07-29)

“Mw heeft matig geslapen, sliep van 1.00 uur tot 4.00 uur. Kwam toen uit bed, at koekjes en dronk thee. Nog geadviseerd medicatie te nemen en mijn zorgen geuit over evt. doorschieten in een manie. Mw was er niet gevoelig voor en reageerde geagiteerd. Mw had spreekdrang maar gaf aan dat wanneer zij zich goed voelt ook veel praat. Mw gaat vandaag naar <PERSOON-1> met haar zoon, ziet daar nu niet meer tegenop omdat

[0.341, -0.359, 0.7, 0.926, -0.004, ..., -0.129]

ervaarde verdwenen zijn. Mw ging na 4.00 uur weer naar bed en kwam niet meer uit haar kamer tot de ochtend.



[Positief, Negatief]

van mijn been” [...]

② Classificatie in NLP: Data Modelleren

Mate van fundamenteel begrip

In hoeverre kan dit met routine data (en met welke velden)?

**Trans-
lationele
data science
& AI**

Taxonomie
studies

Voorspel
agressie
incidenten bij
psychiatrische
patiënten

Mate van praktisch nut



Model	Bag-of-words binary	Bag-of-words tf-idf	Word embeddings	Document embeddings
RNN ¹	.771±.018 ^b	.753±.031	.654±.043	.788±.018 ^{ab}
CNN ²	.729±.030	.716±.038	.684±.038	.763±.024 ^a
NN ³	.727±.033	.717±.038	.751±.036 ^a	.745±.022
NB ⁴	.686±.026	.704±.034 ^a	.700±.051	.692±.046
SVM ⁵	.759±.040	.756±.036 ^b	.764±.024 ^b	.770±.029 ^a
DT ⁶	.727±.018 ^a	.719±.041	.685±.041	.665±.035

¹Recurrent Neural Network; ²Convolutional Neural Network; ³Neural Network;

⁴Naive Bayes; ⁵Support Vector Machine; ⁶Decision Tree

- Beste resultaat?
- Combi van *Recurrent Neural Network* voorspelmodel met Document Embeddings representatie van de bronteksten

Maar... het *Accuracy vs Explainability* dilemma...

In hoeverre
kan dit met
routine data
(en met
welke
velden)?

**Trans-
lationele
data science
& AI**

Taxonomie
studies

**Voorspel
agressie
incidenten bij
psychiatrische
patiënten**

Mate van praktisch nut

② Classificatie in NLP: Model Evaluatie

Table 3. Results of Exploratory Analysis

Rank ^a	Site 1				Site 2			
	Term (English Translation) ^b	Ratio	MCC (95% CI) ^c	P Value ^d	Term (English Translation) ^b	Ratio	MCC (95% CI) ^c	P Value ^d
1	Agressief (aggressive)	1.00	0.17 (0.13 to 0.21)	<.001	Verbaal (verbal)	1.00	0.14 (0.10 to 0.18)	<.001
2	Reageert (reacts)	1.00	0.15 (0.11 to 0.19)	<.001	Dreigend (threatening)	1.00	0.13 (0.08 to 0.16)	<.001
3	Aangeboden (offered)	1.00	0.14 (0.11 to 0.18)	<.001	Agressie (aggression)	1.00	0.15 (0.11 to 0.17)	<.001
4	Boos (angry)	1.00	0.16 (0.12 to 0.19)	<.001	Hierop ([up]on this)	1.00	0.13 (0.09 to 0.16)	<.001
5	Deur (door)	1.00	0.14 (0.10 to 0.18)	<.001	Kantoor (office)	1.00	0.12 (0.08 to 0.16)	<.001
6	Loopt (walks)	1.00	0.15 (0.11 to 0.18)	<.001	Personeel (staff)	1.00	0.12 (0.07 to 0.16)	<.001
7	Ibs (arrest)	1.00	0.14 (0.10 to 0.17)	<.001	Aangesproken (spoke to)	1.00	0.11 (0.08 to 0.15)	<.001
8	Aanbieden (offer)	1.00	0.12 (0.08 to 0.15)	<.001	Agressief (aggressive)	0.99	0.11 (0.08 to 0.15)	<.001
9	Noodmedicatie (emergency medication)	0.99	0.14 (0.10 to 0.17)	<.001	Gevaar agressie (danger aggression)	0.99	0.11 (0.07 to 0.15)	<.001
10	Liep (walked)	0.99	0.12 (0.08 to 0.16)	<.001	Agitatie (agitation)	0.99	0.11 (0.07 to 0.14)	<.001
11	Agressie (aggression)	0.99	0.13 (0.09 to 0.18)	<.001	Geirriteerd (irritated)	0.99	0.10 (0.06 to 0.14)	.001
12	Vraagt (asks)	0.99	0.13 (0.10 to 0.17)	<.001	Separeer (seclusion room)	0.99	0.10 (0.06 to 0.15)	<.001
13	Status vrijwillig (status voluntary)	0.99	-0.12 (-0.14 to -0.09)	<.001	Loopt (walks)	0.99	0.11 (0.08 to 0.14)	.02
14	Psychotisch (psychotic)	0.98	0.12 (0.09 to 0.16)	<.001	Grond (ground)	0.98	0.10 (0.06 to 0.14)	<.001
15	Collega (colleague)	0.98	0.11 (0.07 to 0.15)	<.001	Aanvang (commencement)	0.98	0.11 (0.08 to 0.14)	.01
16	Spreekt (speaks)	0.97	0.12 (0.08 to 0.15)	<.001	Mede (also)	0.98	0.10 (0.07 to 0.14)	.001
17	Gehouden (obliged)	0.97	0.11 (0.07 to 0.15)	<.001	Dhr wilde (Mr wanted)	0.98	0.10 (0.06 to 0.14)	.001
18	Beoordelen (judge), verb	0.96	0.11 (0.07 to 0.15)	<.001	Liep (walked)	0.98	0.10 (0.06 to 0.14)	.006
19	Momenten (moments)	0.96	0.12 (0.08 to 0.15)	<.001	Geagiteerd (agitated)	0.96	0.10 (0.06 to 0.14)	.01
20	Somber (dejected)	0.95	-0.14 (-0.17 to -0.11)	<.001	cvd (not available)	0.96	0.10 (0.06 to 0.14)	.004

Abbreviation: MCC, Matthews correlation coefficient.

^a The top 20 terms with highest within-data set generalizability (ratio) are included.

^b The *Van Dale Dutch-English Dictionary*, 3rd edition,³⁰ was used for translations.

^c Matthews correlation coefficient is computed to assess the direction of association between the term and outcome.

^d P values derived from χ^2 test, and a Holm-Bonferroni correction was applied to obtain corrected P values.

Mate van fundamenteel begrip

In hoeverre kan dit met routine data (en met welke velden)?

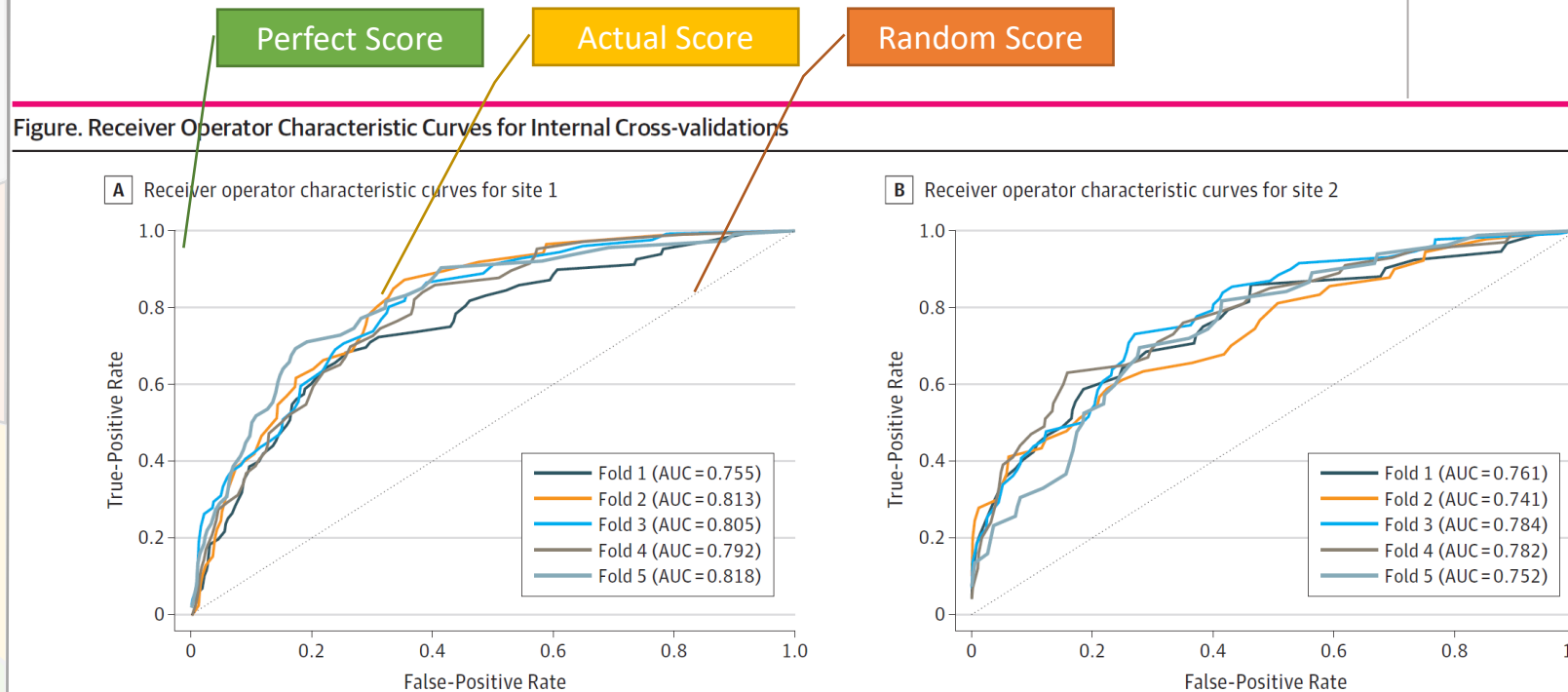
**Trans-
lationele
data science
& AI**

Taxonomie studies

Voorspel agressie incidenten bij psychiatrische patiënten

Mate van praktisch nut

② Classificatie in NLP: Model Evaluatie

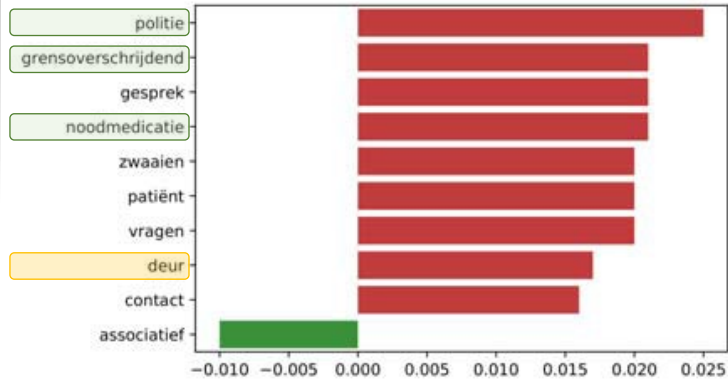


Receiver operator characteristic curves are shown for each fold, according to internal cross-validation in site 1 (A) and site 2 (B). Dashed diagonal lines denote an area under the curve (AUC) of 0.5, ie, predictive validity equivalent to chance. AUC indicates area under the curve.

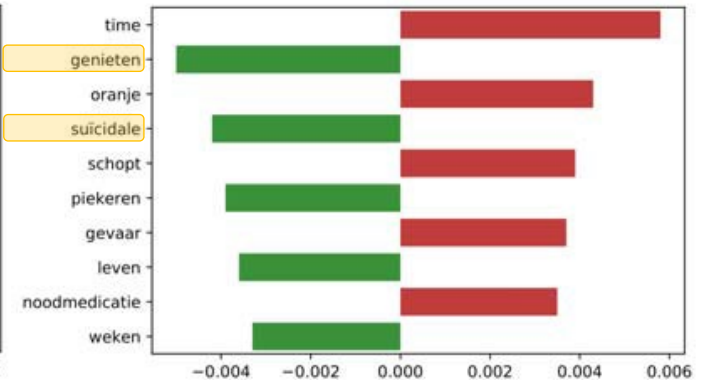
Maar... Hoe verhoudt een AUC van 0.8 zich tot zorgprofessionals die agressie voorspellen voor hun patiënten? *Wie is beter... AI of Professional?*

?

② Classificatie in NLP: Model Implementatie



Waarom **hoog** Risico?



Waarom **laag** risico?

Mate van fundamenteel begrip

In hoeverre kan dit met routine data (en met welke velden)?

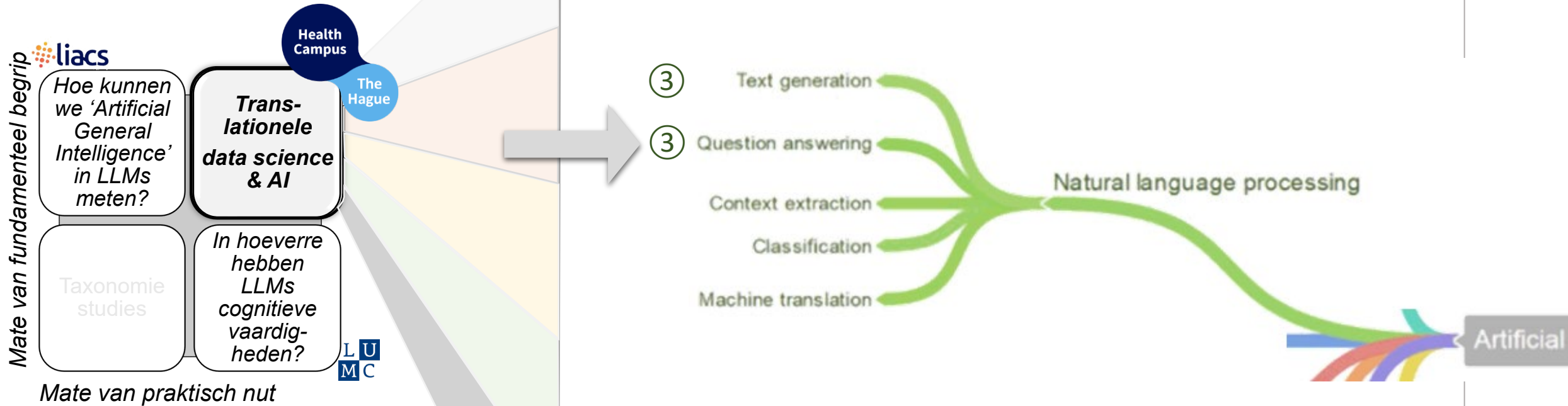
Trans-lationele data science & AI

Taxonomie studies

Voorspel agressie incidenten bij psychiatrische patiënten

Mate van praktisch nut

Agenda [3/3]



③ Vraag-Antwoorden in NLP: “Theory of Mind”

This is Sally

This is Anne

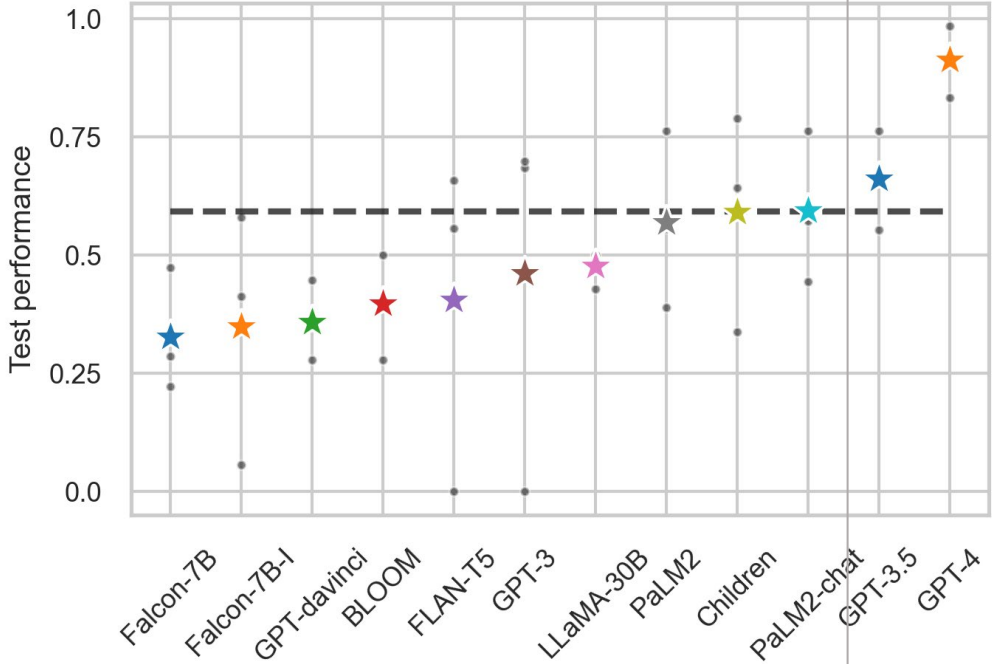
Sally puts her ball in the basket.

Sally goes away.

Anne moves the ball to her box.

Where will Sally look for her ball?

Kinderen (n=36, 9-10j)



Mate van fundamenteel begrip

Hoe kunnen we ‘Artificial General Intelligence’ in LLMs meten?

Trans-lationele data science & AI

Taxonomic studies

In hoeverre hebben LLMs cognitieve vaardigheden?

Mate van praktisch nut

Health
Campus

Den
Haag



Tot ziens!

AI & Technologieweek Geneeskunde jaar 1, 2 juli 2024: ✓

m.r.spruit@lumc.nl

