



**KWALITEITSZORG EN AI**  
*BELEID SCHRIJVEN MET CHATGPT*

THEORIE

ChatGPT

Veiligheid en ethiek

PROMPT  
ENGINEERING

4 prompt engineering stappen

Structured Prompts

OPDRACHT

Digitale geletterdheid in het onderwijs

REFLECTIE

Wanneer is dit bruikbaar?

Wanneer absoluut niet?

PLANNING

# INTRODUCING



Bram Wiggers

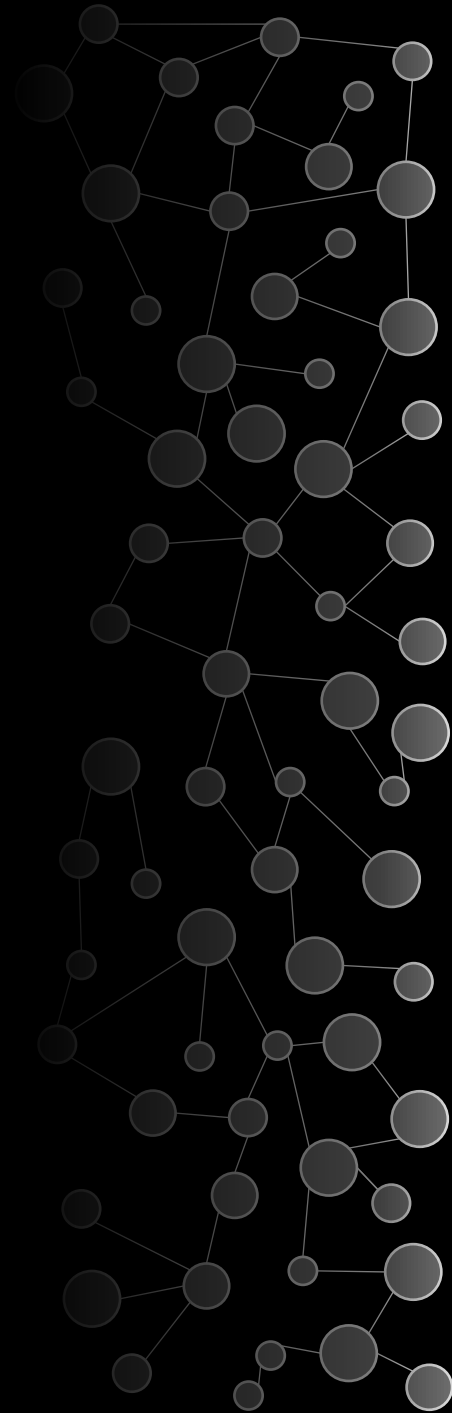


BSc, MSc Artificial Intelligence

Track32

OrangeNodes

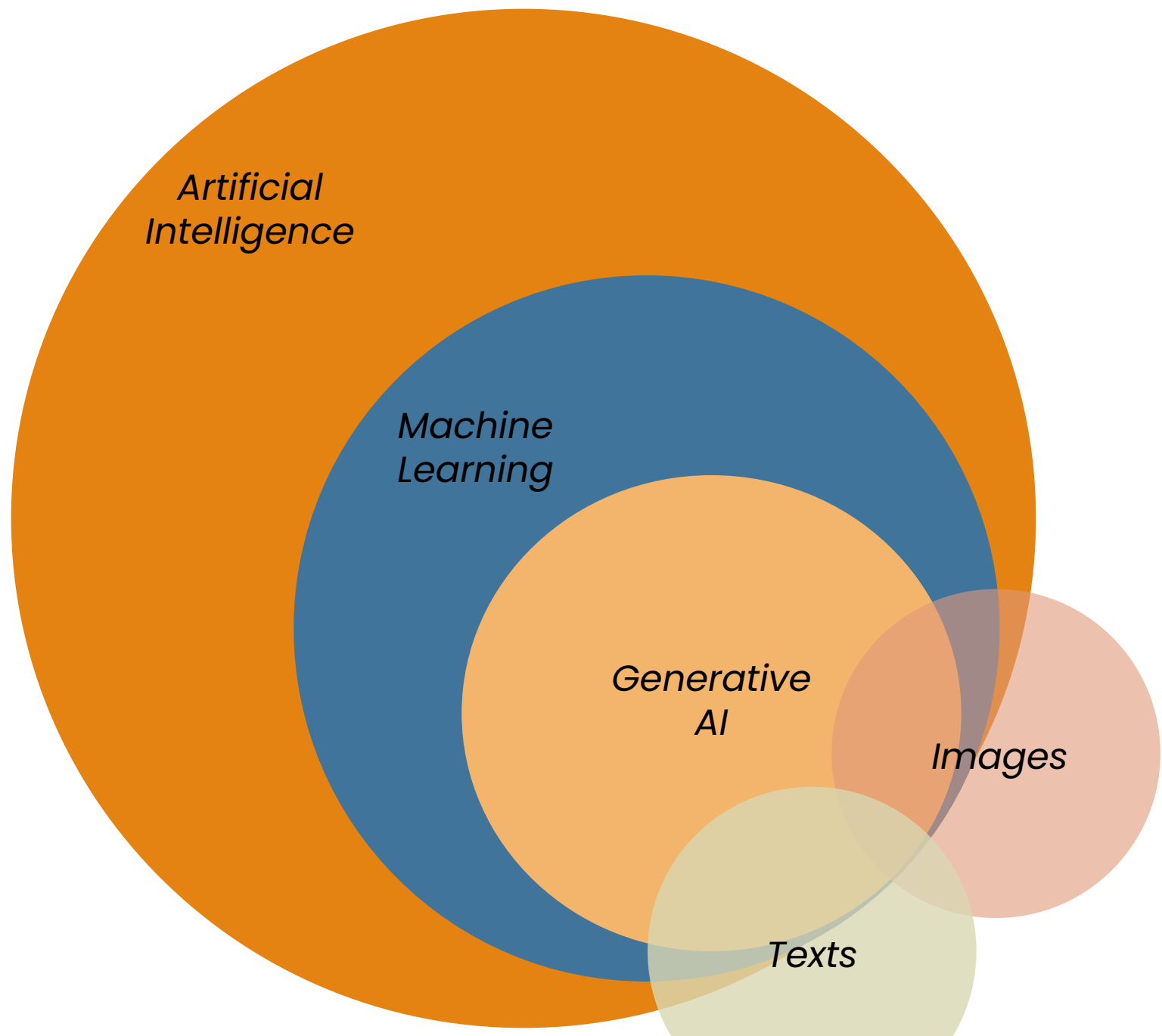
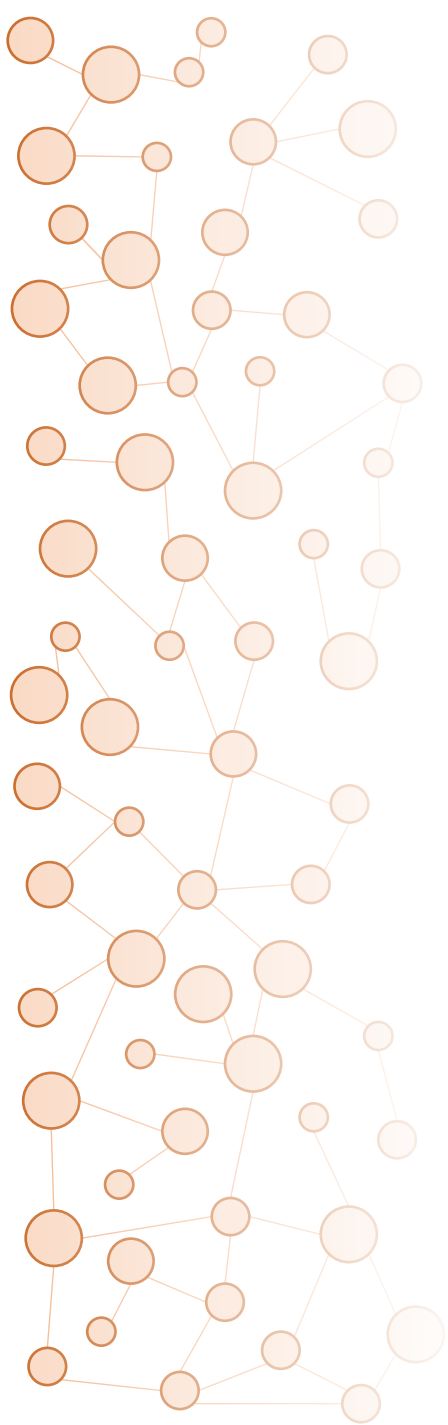
Gebruik je AI?

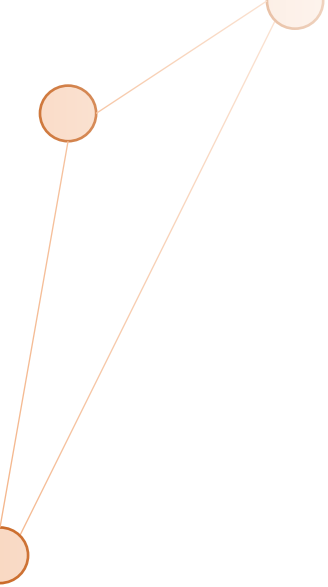




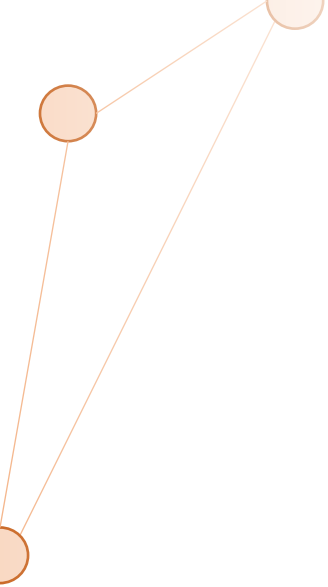
**BACKGROUND**

**ORANGENODES**





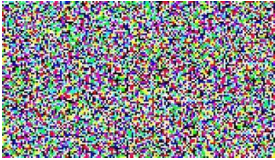
CHATGPT



GPT

CHAT

GENERATIVE



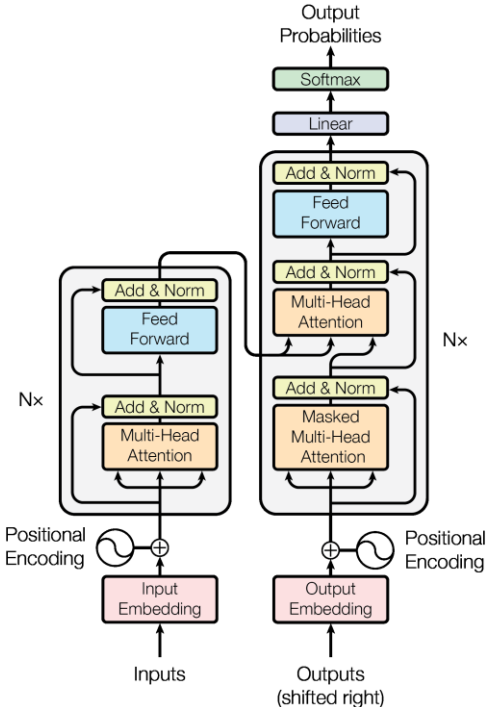
“ONCE UPON A TIME...”



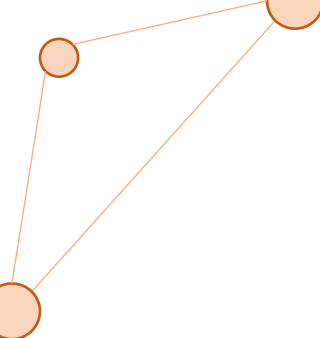
PRE-TRAINED



TRANSFORMER







# WORD PREDICTION

DE ZON ...

ZWEEFT [60%]

APPEL [0.001%]

SCHIJNT [99.9%]

SCHIJNEN [10%]

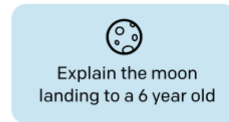
# CHAT

# GPT

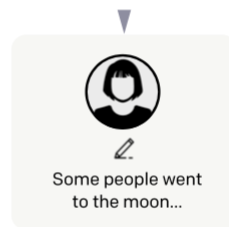
Step 1

Collect demonstration data, and train a supervised policy.

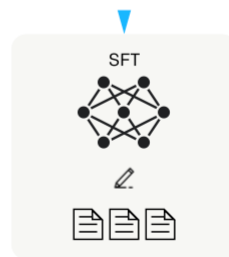
A prompt is sampled from our prompt dataset.



A labeler demonstrates the desired output behavior.



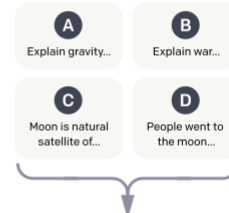
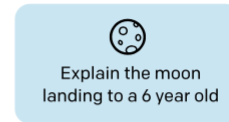
This data is used to fine-tune GPT-3 with supervised learning.



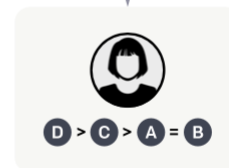
Step 2

Collect comparison data, and train a reward model.

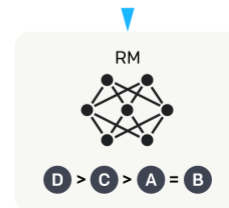
A prompt and several model outputs are sampled.



A labeler ranks the outputs from best to worst.



This data is used to train our reward model.



Step 3

Optimize a policy against the reward model using reinforcement learning.

A new prompt is sampled from the dataset.

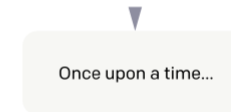


The policy generates an output.

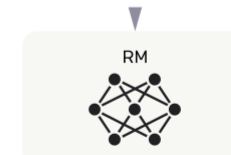


Once upon a time...

The reward model calculates a reward for the output.

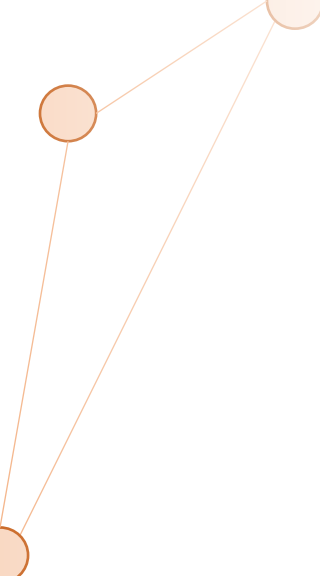


The reward is used to update the policy using PPO.



$r_k$





# VEILIGHEID

## CHATGPT

Stop er geen privé data in. OpenAI traint door op wat jij erin stopt. Dus als je iemands naam erin stopt, kan die bij iemand anders eruit komen.

Vraag altijd je manager.

## WEL PRIVE GEBRUIKEN?

Premium account

GPT API

Privé LLM

## HALLUCINATIES

ChatGPT is generatief. Het is dus absoluut niet geoptimaliseerd om te fact-checken.

Het zal blijven genereren, ook als het het fout heeft.

# ETHIEK

## CREATIEF EIGENDOM

Wie is de eigenaar van kunst?  
Mag je iemand zijn of haar stijl kopiëren?

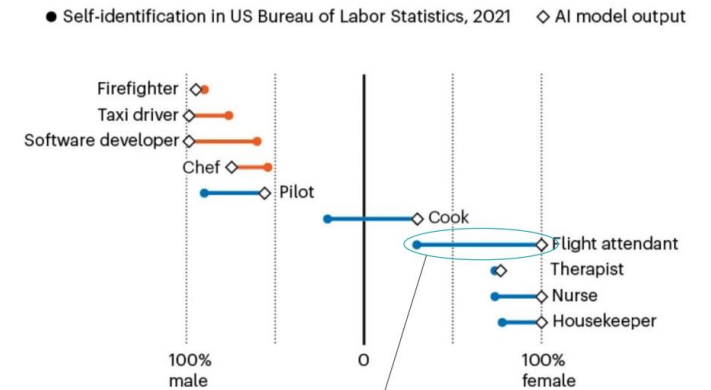
## MILLIEU

Het trainen van een GPT model kost hetzelfde als 1000 gezinnen hun jaarlijkse stroom.

Met 4 prompts kan je je telefoon weer helemaal opladen.

## BIAS

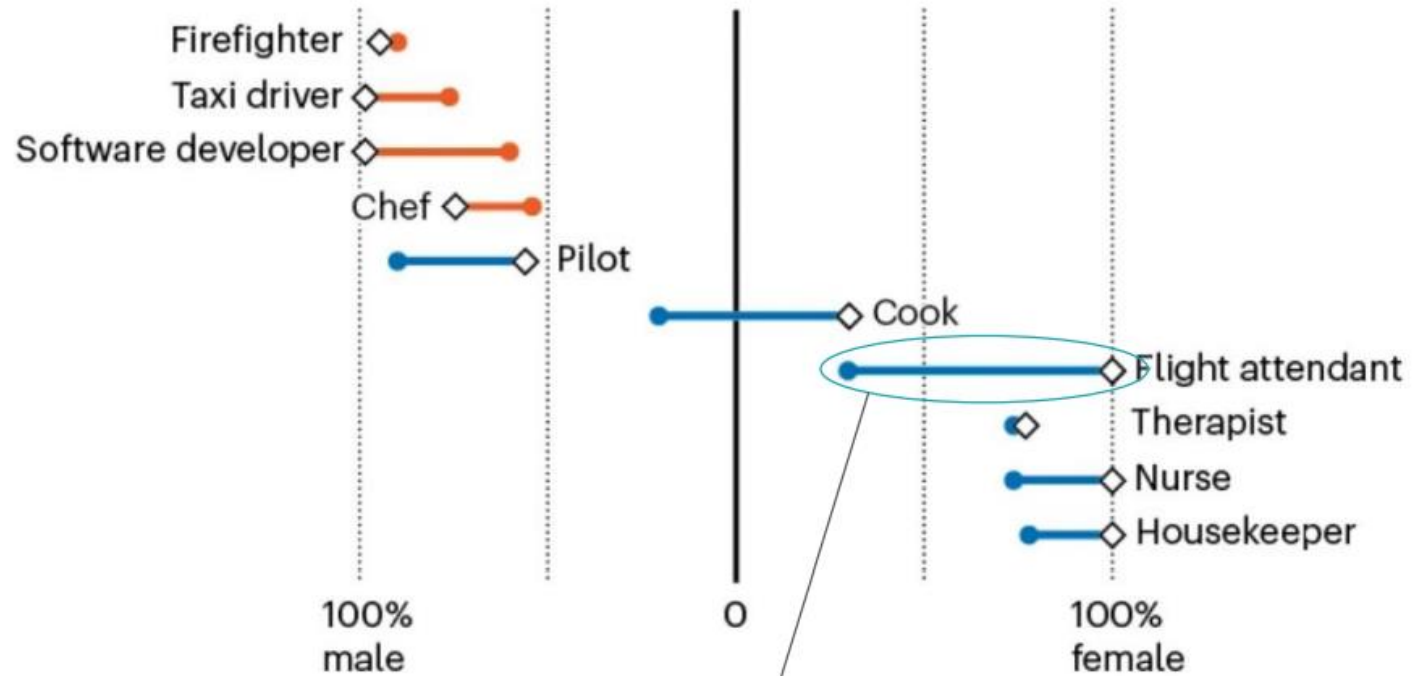
### Gender Gaps



For example,  
All AI-generated flight attendant images portrayed women.  
But, in reality, only about 65% of US flight attendants identify as women.

# Gender Gaps

● Self-identification in US Bureau of Labor Statistics, 2021    ◇ AI model output



For example,

All AI-generated flight attendant images portrayed women.  
But, in reality, only about 65% of US flight attendants identify as women.



# PROMPT ENGINEERING

# CHATGPT

PROMPT/INPUT

MODEL

ANSWER/OUTPUT

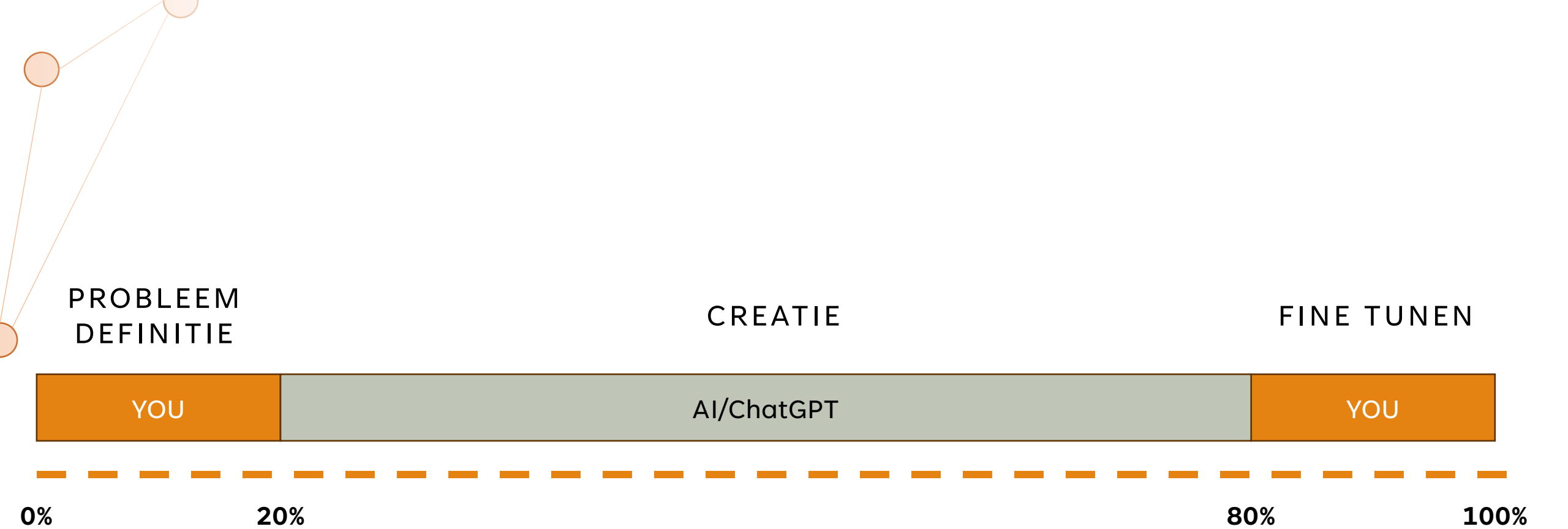
WRITE THE BEST  
POEM IN THE  
WORLD



ChatGPT

PROMPT'S REFINED,  
PATHWAYS ALIGNED,  
CHATGPT,  
CONVERSATIONS FOR FREE.

PROMPT  
ENGINEERING





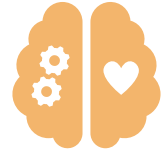
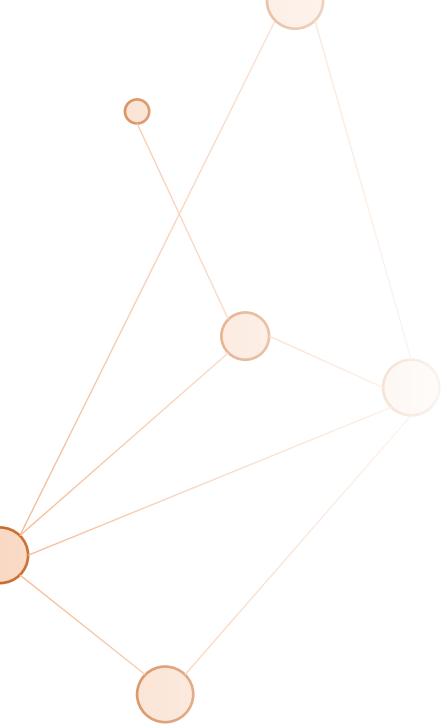


## PROBLEEM

*“In 2021 werd in het kader veel ingegaan op burgerschap. In de laatste communicatie van OCW ligt er meer focus op digitale vaardigheden. Er zijn ‘potjes’ beschikbaar, maar er moet een plan komen om aanspraak te maken op deze potjes.”*

*“Ohja, de deadline is **vanavond**”*





BRAINSTORM



CREATIE

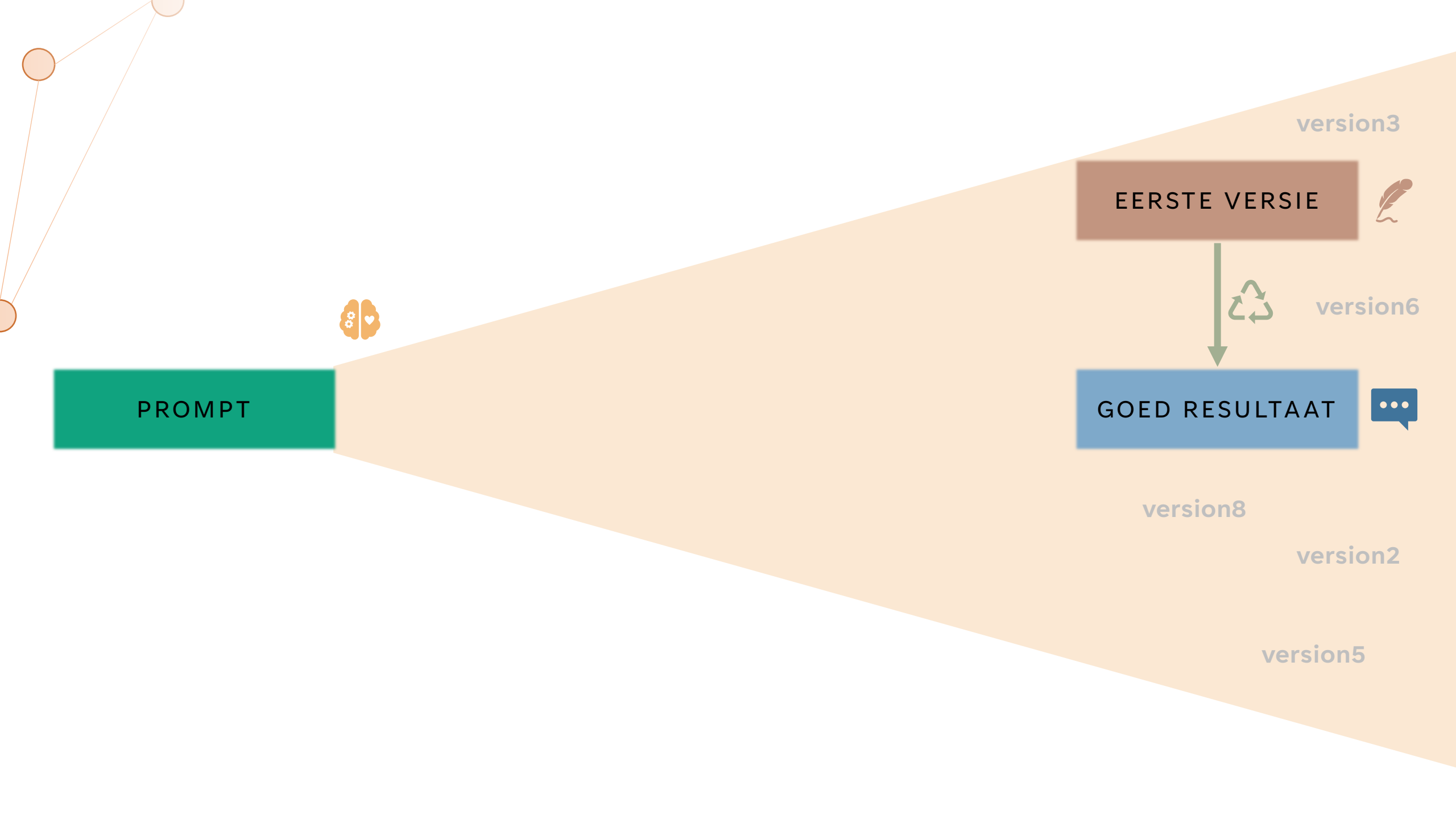


ITERATIE



OEFENEN

TIPS AND TRICKS



PROMPT



EERSTE VERSIE



GOED RESULTAAT



version3

version6

version8

version2

version5



## TIPS AND TRICKS *GEEF EEN LIJST MET IDEEËN*

1. *Goed start punt*
2. *ChatGPT heeft altijd meer ideeën*

*GEEF ME EEN LIJST VAN 5 MOGELIJKHEDEN OM DIGITALE  
GELETERDHEID TE VERBETEREN OP MIJN SCHOOL*



## TIPS AND TRICKS CREATIVITEIT

1. *Stuur de hoeveelheid creativiteit van ChatGPT*
2. *Kies een waarde tussen 0 en 1*  
*0: Basis, niet overdreven, niet enthousiast, basis*  
*1: Creatief, uitbundig, hyped*

... *GEBRUIK CREATIVITEIT 0.2*



## TIPS AND TRICKS ACHTERGROND

1. *Voeg achtergrond informatie toe*
2. *Vraag om feedback als er dingen niet duidelijk zijn*

**HET DOEL IS XXX. HET PUBLIEK IS XXX. IK WIL HET KRIJGEN ALS XXX.**

*STEL ME VRAGEN ALS ER DINGEN NIET DUIDELIJK ZIJN*



## TIPS AND TRICKS DOCUMENTEN

1. *Upload een document waaruit informatie gehaald kan worden.*
2. *Vraag ChatGPT om het document te gebruiken*
3. *Vraag om specifieke referenties, en **controleer***

*GEBRUIK VOORBEELDEN UIT HET DOCUMENT EN GEEF REFERENTIES*



## TIPS AND TRICKS *VERDEEL EN HEERS*

1. Er is een maximum aan de hoeveelheid output van GPT.
2. Vraag eerst om een globaal overzicht, en ga daarna in op de details.

*GEEF ME EEN OVERZICHT VAN DE TITEL VAN ELKE ALINEA*

*SCHRIJF NU ALINEA 1*





## TIPS AND TRICKS *CHAIN OF THOUGHT*

1. Vraag ChatGPT om ieder antwoord te eindigen met een wedervraag
2. Geen ongemakkelijke stiltes!

*EINDIG IEDER ANTWOORD MET EEN NIEUWE VRAAG*



## TIPS AND TRICKS *CHAT 'OPSLAAN'*

1. Puinhoop-chat
2. Vraag om een samenvatting
3. Verder met een schone lei

*VAT ALLES SAMEN ZODAT IK DAT IN EEN NIEUWE CHAT KAN PLAKKEN.*

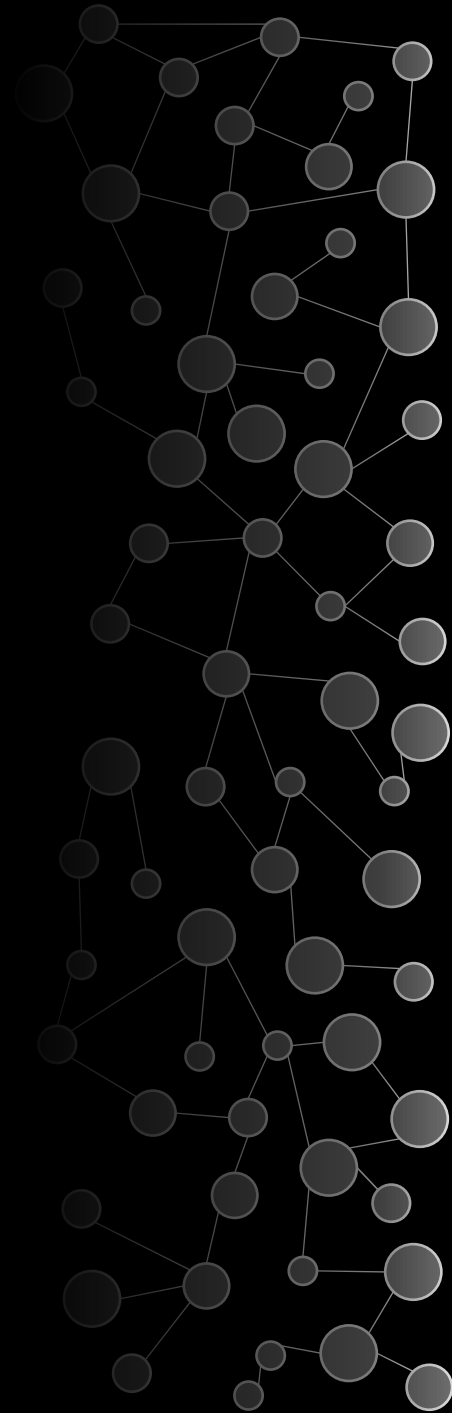


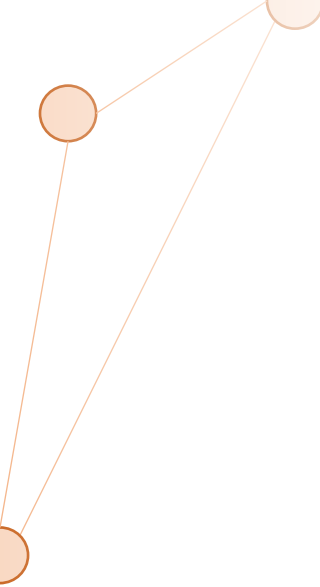
## TIPS AND TRICKS *OEFENEN*

1. Test jezelf door ChatGPT de inspectie te laten spelen.
2. Achtergrond maakt sterker.
3. Codewoordt om aan- en uit te zetten

*GEDRAAG JE ALS XXX TOT IK 'KOEKENPAN' ZEG, DAN GEDRAAG JE JE  
WEER ALS JEZELF.*

TRY IT YOURSELF





# REFLECTION

WAT IS MAKKELIJK?

WAT IS MOEILIJK?



## VOLGENDE STAP?

1. Dataanalyse met GPT
2. Custom GPT's
3. Custom instructions en geheugen
4. Premium account
5. Vragen?

[bram@orangenodes.com](mailto:bram@orangenodes.com)

QUESTIONS

