



Digitale sporen ontcijferen met taalmodellen

Hans Henseler

Lector Digital Forensics & E-Discovery HS Leiden
Senior Wetenschappelijk Medewerker NFI

E-Discovery Symposium, 2 April 2024

Vorig jaar op het symposium:

Eerste resultaten van ons (empirisch) onderzoek

Vier experimenten:

- Zoekvragen formulere
- Chat samenvatten en
- Analyse van sporen
- Reverse engineering

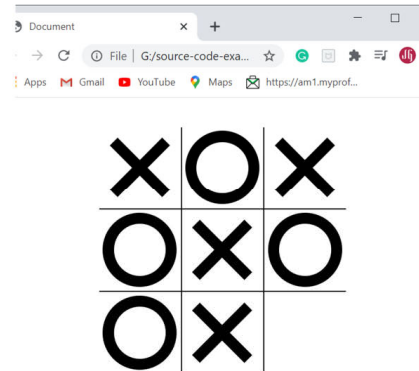
Zoekvragen formuleren

Rapporteren 4: visualisatie

Analyse van sporen

Reverse engineering experiment

- Boter kaas en eieren (Tic-Tac-Toe):
 - script.js
 - index.html
 - style.css
- Dubbele obfuscation van script.js:
 - Handmatig hernoemen van namen naar function1, function2, variable1, variable2 etc.
 - Vervolgens met <https://obfuscator.io/> obfuscaten
- Als prompt aan ChatGPT met vragen:
 1. *What does the following javascript code do?*
 2. *Can you convert this to javascript?*
 3. *Fix the code so that it works index.html, style.css; add comments and use meaningful names*



Wat is er sindsdien gebeurd?

- 22-2-2022: GPT3.5, 170 miljard parameters
- 14-3-2023: GPT4 1.760 miljard
- 10-5-2023: Palm 2 van Google, 530 miljard
- 1-7-2023: LLaMA 2 van Meta AI (Facebook), 7, 13, 33 en 70 miljard
- 11-7-2023: Claude 2 van Anthropic, 130 miljard
- 8-7-2023: Falcon, 7, 40 en 180 miljard
- 27-9-2023: Mistral 7B van Mistral, 7 miljard
- 4-11-2023: Grok-1 van xAI (Elon Musk), 63 miljard
- 6-11-2023: GPT4-turbo van OpenAI (GTP4 basis)
- 20-11-2023: Orca-2 van Microsoft, 7 en 13 miljard (Llama basis)
- 6-12-2023: Gemini van Google, nano, pro en ultra
- 12-12-2023: Phi2 van Microsoft, 2.7 miljard
- 22-2-2024: Gemma van Google, 2 en 7 miljard



Het vervolg: copiloot voor digitaal forensisch onderzoek

- Maak gebruik van Azure OpenAI API (GPT4) in trusted cloud van Ministerie van J&V
- Geen finetuning, vooral prompting. Verschillende *agents*:
 - Schrijven van HQL
 - Samenvatten resultaten
 - Dialoog voor het onderzoek
- Eerste prototype werkt maar dit is pas het begin...



Voorbeeld van een onderzoek naar een telefoon

- Welke accountnamen zijn er van Whatsapp, Snapchat, enz.
- Van wie is de telefoon?
- Welke contacten kun je vinden?
- Welke chat conversaties zijn er en waar gaan die over?
- Welke zoekopdrachten zijn er uitgevoerd?
- Hoe ziet de internetgeschiedenis eruit?
- Welke locaties kun je vinden?
- Met welke draadloze netwerken is er verbinding gemaakt?



Menu search

- Home
- Case alerts
- Search
- Timeline (Beta)
- Tags and Notes
- Overview
- Multimedia
- File explorer
- Locations
- Accounts
- Communications
- Browser artifacts
- E-commerce
- Reports
- Entities
- System

Tags (15)

- watch (543)
- Not-important (55)
- Delivery (21)
- Preserve (14)
- M2309 0716 (8)
- gjr (7)
- Firearm (5)
- New-tag (5)
- OPS Arie (5)
- Shipping (5)
- Carly (1)
- My new tag (1)
- Picture3 (1)
- Picture4 (1)
- reg (1) Opgeslagen in \\nfi\shares

Evidence Container (4)

- Subjects (0)
- Suspects (18)
- Crimes (3)
- Victims (0)

Media	Encrypted	Communication	Documents	Database	Virtual																																																																																		
<table border="1"> <tr><td>Pictures</td><td>293,459</td></tr> <tr><td>Videos</td><td>1,597</td></tr> <tr><td>Audio</td><td>4,283</td></tr> </table>	Pictures	293,459	Videos	1,597	Audio	4,283	<table border="1"> <tr><td>Archive</td><td>2</td></tr> <tr><td>Configuration</td><td>1</td></tr> <tr><td>Container</td><td>65</td></tr> <tr><td>Driver</td><td>0</td></tr> <tr><td>Executable</td><td>2</td></tr> <tr><td>File</td><td>6</td></tr> <tr><td>Installer</td><td>0</td></tr> <tr><td>Key</td><td>0</td></tr> <tr><td>Library</td><td>0</td></tr> </table>	Archive	2	Configuration	1	Container	65	Driver	0	Executable	2	File	6	Installer	0	Key	0	Library	0	<table border="1"> <tr><td>Email</td><td>91,618</td></tr> <tr><td>Chat conversations</td><td>4,091</td></tr> <tr><td>Chat log</td><td>60</td></tr> <tr><td>Chat messages</td><td>89,810</td></tr> <tr><td>Contacts</td><td>15,242</td></tr> <tr><td>Email Archive</td><td>116</td></tr> </table>	Email	91,618	Chat conversations	4,091	Chat log	60	Chat messages	89,810	Contacts	15,242	Email Archive	116	<table border="1"> <tr><td>DOC</td><td>10,100</td></tr> <tr><td>PDF</td><td>1,345</td></tr> <tr><td>RAR</td><td>1</td></tr> <tr><td>ODT</td><td>1</td></tr> <tr><td>RTF</td><td>3,100</td></tr> <tr><td>XLS</td><td>4,188</td></tr> <tr><td>PPT</td><td>1,285</td></tr> </table>	DOC	10,100	PDF	1,345	RAR	1	ODT	1	RTF	3,100	XLS	4,188	PPT	1,285	<table border="1"> <tr><td>sqlite</td><td>475</td></tr> <tr><td>db</td><td>3,954</td></tr> <tr><td>edb</td><td>7</td></tr> <tr><td>mdf</td><td>0</td></tr> <tr><td>ipd</td><td>0</td></tr> <tr><td>mysql</td><td>0</td></tr> <tr><td>crypt*</td><td>0</td></tr> <tr><td>wal</td><td>0</td></tr> </table>	sqlite	475	db	3,954	edb	7	mdf	0	ipd	0	mysql	0	crypt*	0	wal	0	<table border="1"> <tr><td>vmdk</td><td>0</td></tr> <tr><td>vmx</td><td>0</td></tr> <tr><td>iso</td><td>2</td></tr> <tr><td>img</td><td>19</td></tr> <tr><td>dd</td><td>0</td></tr> <tr><td>e01</td><td>0</td></tr> <tr><td>vhd</td><td>0</td></tr> <tr><td>hdd</td><td>0</td></tr> </table>	vmdk	0	vmx	0	iso	2	img	19	dd	0	e01	0	vhd	0	hdd	0
Pictures	293,459																																																																																						
Videos	1,597																																																																																						
Audio	4,283																																																																																						
Archive	2																																																																																						
Configuration	1																																																																																						
Container	65																																																																																						
Driver	0																																																																																						
Executable	2																																																																																						
File	6																																																																																						
Installer	0																																																																																						
Key	0																																																																																						
Library	0																																																																																						
Email	91,618																																																																																						
Chat conversations	4,091																																																																																						
Chat log	60																																																																																						
Chat messages	89,810																																																																																						
Contacts	15,242																																																																																						
Email Archive	116																																																																																						
DOC	10,100																																																																																						
PDF	1,345																																																																																						
RAR	1																																																																																						
ODT	1																																																																																						
RTF	3,100																																																																																						
XLS	4,188																																																																																						
PPT	1,285																																																																																						
sqlite	475																																																																																						
db	3,954																																																																																						
edb	7																																																																																						
mdf	0																																																																																						
ipd	0																																																																																						
mysql	0																																																																																						
crypt*	0																																																																																						
wal	0																																																																																						
vmdk	0																																																																																						
vmx	0																																																																																						
iso	2																																																																																						
img	19																																																																																						
dd	0																																																																																						
e01	0																																																																																						
vhd	0																																																																																						
hdd	0																																																																																						

Evidence items (10) Select all/Deselect all

Search evidence items

Search name, description, file type, or size

Menu search

- Home
- Case alerts
- Search
- Timeline (Beta)
- Tags and Notes
- Overview
- Multimedia
- File explorer
- Locations
- Accounts
- Communications
- Browser artifacts
- E-commerce
- Reports
- Entities
- System

Tags (15)

- watch (543)
- Not-important (55)
- Delivery (21)
- Preserve (14)
- M2309 0716 (8)
- gjr (7)
- Firearm (5)
- New-tag (5)
- OPS Arie (5)
- Shipping (5)
- Carly (1)
- My new tag (1)
- Picture3 (1)
- Picture4 (1)
- reg (1)

Evidence Container (4)

- Subjects (0)
- Suspects (18)
- Crimes (3)
- Victims (0)

Media	Encrypted	Communication	Documents	Database	Virtual																																																																																				
<table border="1"> <tr><td>Pictures</td><td>293,459</td></tr> <tr><td>Videos</td><td>1,597</td></tr> <tr><td>Audio</td><td>4,283</td></tr> </table>	Pictures	293,459	Videos	1,597	Audio	4,283	<table border="1"> <tr><td>Archive</td><td>2</td></tr> <tr><td>Configuration</td><td>1</td></tr> <tr><td>Container</td><td>65</td></tr> <tr><td>Driver</td><td>0</td></tr> <tr><td>Executable</td><td>2</td></tr> <tr><td>File</td><td>6</td></tr> <tr><td>Installer</td><td>0</td></tr> <tr><td>Key</td><td>0</td></tr> <tr><td>Library</td><td>0</td></tr> </table>	Archive	2	Configuration	1	Container	65	Driver	0	Executable	2	File	6	Installer	0	Key	0	Library	0	<table border="1"> <tr><td>Email</td><td>91,618</td></tr> <tr><td>Chat conversations</td><td>4,091</td></tr> <tr><td>Chat log</td><td>60</td></tr> <tr><td>Chat messages</td><td>89,810</td></tr> <tr><td>Contacts</td><td>15,242</td></tr> <tr><td>Email Archive</td><td>116</td></tr> </table>	Email	91,618	Chat conversations	4,091	Chat log	60	Chat messages	89,810	Contacts	15,242	Email Archive	116	<table border="1"> <tr><td>DOC</td><td>10,100</td></tr> <tr><td>PDF</td><td>1,345</td></tr> <tr><td>RAR</td><td>1</td></tr> <tr><td>ODT</td><td>1</td></tr> <tr><td>RTF</td><td>3,100</td></tr> <tr><td>XLS</td><td>4,188</td></tr> <tr><td>PPT</td><td>1,285</td></tr> </table>	DOC	10,100	PDF	1,345	RAR	1	ODT	1	RTF	3,100	XLS	4,188	PPT	1,285	<table border="1"> <tr><td>sqlite</td><td>475</td></tr> <tr><td>db</td><td>3,954</td></tr> <tr><td>edb</td><td>7</td></tr> <tr><td>mdf</td><td>0</td></tr> <tr><td>ipd</td><td>0</td></tr> <tr><td>mysql</td><td>0</td></tr> <tr><td>crypt*</td><td>0</td></tr> <tr><td>wal</td><td>0</td></tr> </table>	sqlite	475	db	3,954	edb	7	mdf	0	ipd	0	mysql	0	crypt*	0	wal	0	<table border="1"> <tr><td>vmdk</td><td>0</td></tr> <tr><td>vmx</td><td>0</td></tr> <tr><td>iso</td><td>0</td></tr> <tr><td>ova</td><td>0</td></tr> <tr><td>img</td><td>0</td></tr> <tr><td>dd</td><td>0</td></tr> <tr><td>e01</td><td>0</td></tr> <tr><td>vhd</td><td>0</td></tr> <tr><td>hdd</td><td>0</td></tr> </table>	vmdk	0	vmx	0	iso	0	ova	0	img	0	dd	0	e01	0	vhd	0	hdd	0
Pictures	293,459																																																																																								
Videos	1,597																																																																																								
Audio	4,283																																																																																								
Archive	2																																																																																								
Configuration	1																																																																																								
Container	65																																																																																								
Driver	0																																																																																								
Executable	2																																																																																								
File	6																																																																																								
Installer	0																																																																																								
Key	0																																																																																								
Library	0																																																																																								
Email	91,618																																																																																								
Chat conversations	4,091																																																																																								
Chat log	60																																																																																								
Chat messages	89,810																																																																																								
Contacts	15,242																																																																																								
Email Archive	116																																																																																								
DOC	10,100																																																																																								
PDF	1,345																																																																																								
RAR	1																																																																																								
ODT	1																																																																																								
RTF	3,100																																																																																								
XLS	4,188																																																																																								
PPT	1,285																																																																																								
sqlite	475																																																																																								
db	3,954																																																																																								
edb	7																																																																																								
mdf	0																																																																																								
ipd	0																																																																																								
mysql	0																																																																																								
crypt*	0																																																																																								
wal	0																																																																																								
vmdk	0																																																																																								
vmx	0																																																																																								
iso	0																																																																																								
ova	0																																																																																								
img	0																																																																																								
dd	0																																																																																								
e01	0																																																																																								
vhd	0																																																																																								
hdd	0																																																																																								

Evidence items (10) Select all/Deselect all

Search evidence items

Search name, description, file type, or size

20:41:03

@hansken zoek naar whatsapp accounts

HQL: account.application.whatsapp
 Query: account.application.whatsapp
 Show 7 results

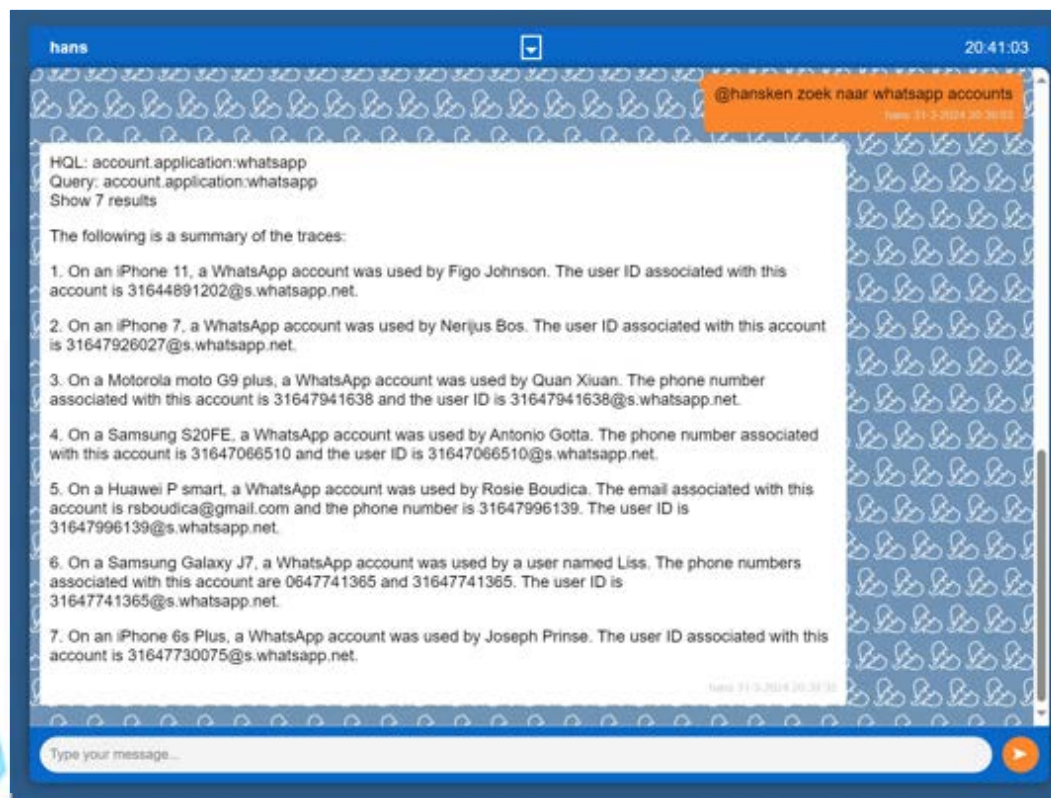
The following is a summary of the traces:

1. On an iPhone 11, a WhatsApp account was used by Figo Johnson. The user ID associated with this account is 31644891202@s.whatsapp.net.
2. On an iPhone 7, a WhatsApp account was used by Nenjus Bos. The user ID associated with this account is 31647928027@s.whatsapp.net.
3. On a Motorola moto G9 plus, a WhatsApp account was used by Quan Xuan. The phone number associated with this account is 31647941638 and the user ID is 31647941638@s.whatsapp.net.
4. On a Samsung S20FE, a WhatsApp account was used by Antonio Gotta. The phone number associated with this account is 31647066510 and the user ID is 31647066510@s.whatsapp.net.
5. On a Huawei P smart, a WhatsApp account was used by Rosie Boudica. The email associated with this account is rsboudica@gmail.com and the phone number is 31647996139. The user ID is 31647996139@s.whatsapp.net.
6. On a Samsung Galaxy J7, a WhatsApp account was used by a user named Liss. The phone numbers associated with this account are 0647741365 and 31647741365. The user ID is 31647741365@s.whatsapp.net.
7. On an iPhone 6s Plus, a WhatsApp account was used by Joseph Prinse. The user ID associated with this account is 31647730075@s.whatsapp.net.

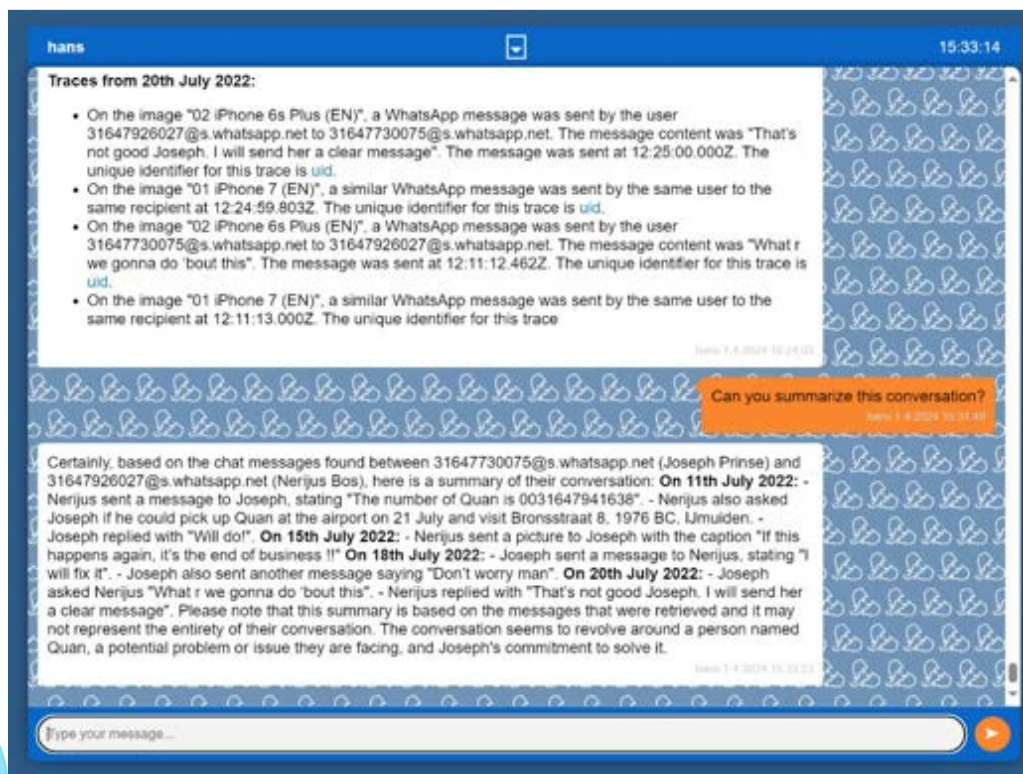
Mar 31 9:20:41 AM '24

Type your message...

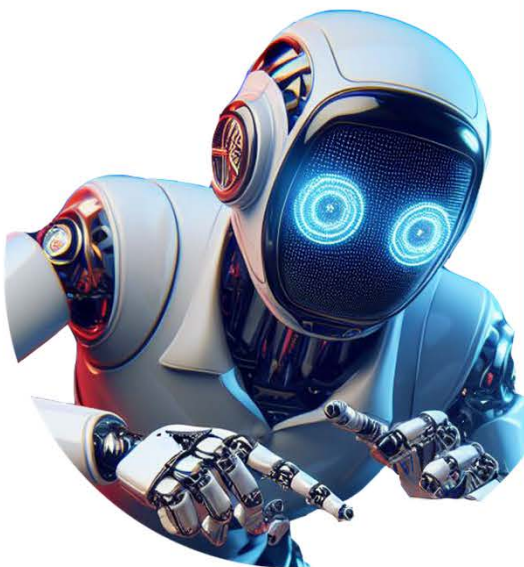
Dialog met de Hansken copiloot: accounts



Dialog met de Hansken copiloot: chats



Dialog met de Hansken copiloot: zoekgeschiedenis



hansken 21:45:30

- search.user: Empty
- search.source: chrome (486), safari (189), knowldegec (110), play store (31), instagram (26), google maps (13), gmail (11), google home page (11), googlequicksearchbox (6), firefox (4), gmail webcache - chrome (4), apple maps (1)
- search.accessedOn: 2021 (188), 2022 (653)

Too many results. Refine your query by adding or excluding a facet, date range or keywords.

hansken 1-4-2024 21:41:38

Do you see anything interesting in the search history?

hansken 1-4-2024 21:44:59

Based on the search history, several interesting patterns and themes emerge: 1. **Location Searches:** There are numerous queries related to specific locations, including 'liverpool stadium', 'schiphol airport', 'de baanderij, sisalbaan, leiderdorp', and 'london city airport'. This suggests that the user may have been planning travel or researching these locations. 2. **Social Media and Communication:** There are also several searches related to social media and communication platforms, such as 'what to do when someone ignores you on social media', 'message', 'facebook', 'whatsapp group fitness', and 'telegram'. This could indicate the user's active engagement with various social media and messaging platforms. 3. **Legal and Law Enforcement:** Some searches are related to legal matters and law enforcement, for example, 'border patrol car uk', 'divorce settlements', 'testify against husband', and 'drug trafficking punishment uk'. This could suggest that the user might be involved in legal issues or has an interest in law enforcement activities. 4. **Interest in Cryptocurrency:** There are several searches related to cryptocurrency, including 'actual rate ripple', 'course ripple', 'bitcoin price usd', which could signal interest or involvement in cryptocurrency trading. 5. **Personal Interests:** The search history reveals some personal interests of the user. For example, 'chocolate cake', 'group app cooking', 'patek philippe', 'boxing rotterdam 14 july' suggest hobbies and interests in cooking, luxury watches, and boxing. Please note that these are just observations and should be considered in the broader context of the investigation.

hansken 1-4-2024 21:45:36

Type your message...



A



Hoe nu verder?



- Uitleggen van het onderzoek
- Doornemen van verklaringen en ten laste legging (of verdenking)
- Copiloot stelt vragen en verwerkt antwoorden tot een knowledge graph

- Copilot heeft een plan van aanpak. Samen met de onderzoeker wordt op basis daarvan een onderzoeksplan opgesteld
- Het onderzoeksplan is een lijst met onderzoeksvragen die met hulp van digitaal bewijs dmv Hansken beantwoord moeten worden.

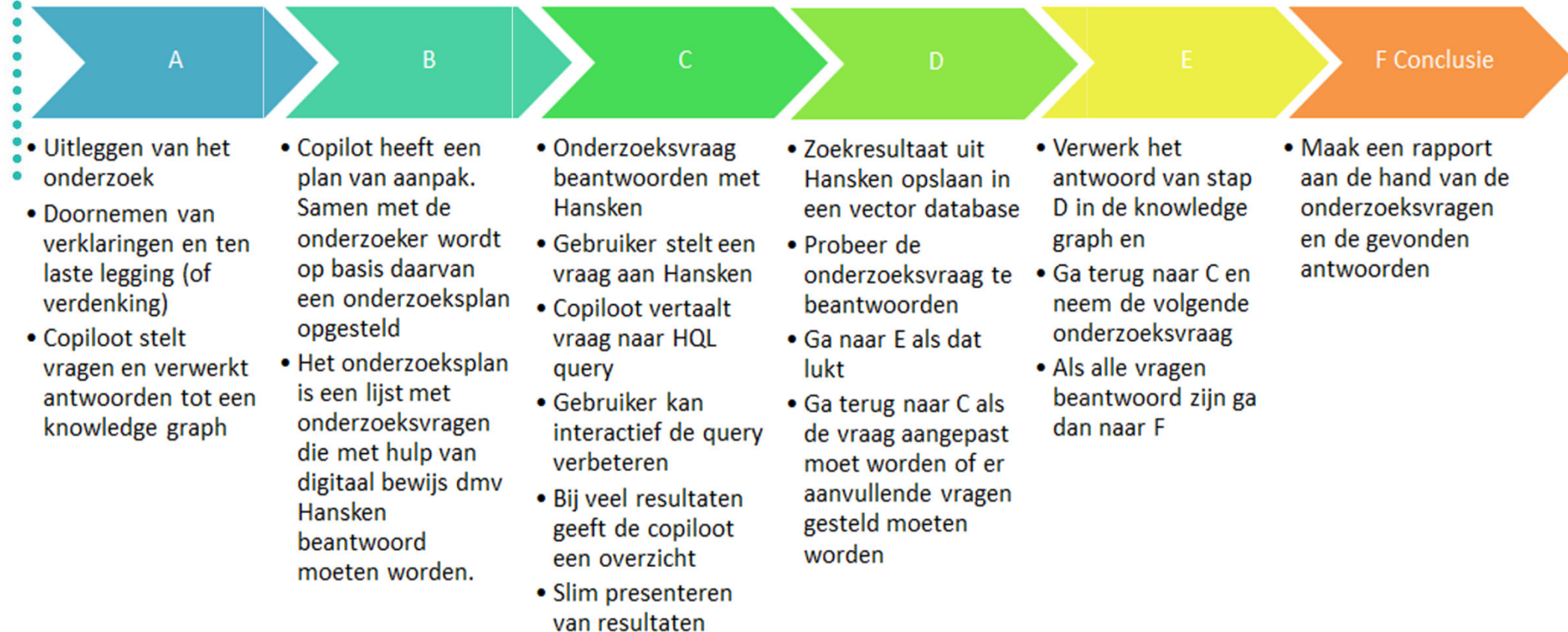
- Onderzoeksvraag beantwoorden met Hansken
- Gebruiker stelt een vraag aan Hansken
- Copiloot vertaalt vraag naar HQL query
- Gebruiker kan interactief de query verbeteren
- Bij veel resultaten geeft de copiloot een overzicht
- Slim presenteren van resultaten

- Zoekresultaat uit Hansken opslaan in een vector database
- Probeer de onderzoeksvraag te beantwoorden
- Ga naar E als dat lukt
- Ga terug naar C als de vraag aangepast moet worden of er aanvullende vragen gesteld moeten worden

- Verwerk het antwoord van stap D in de knowledge graph en
- Ga terug naar C en neem de volgende onderzoeksvraag
- Als alle vragen beantwoord zijn ga dan naar F

- Maak een rapport aan de hand van de onderzoeksvragen en de gevonden antwoorden

Vervolg experimenten en onderzoek



Zelf aan de slag

- ChatGPT van OpenAI
 - ChatGPT 3.5 turbo is nog steeds gratis
 - ChatGPT 4 en custom GPTs €23 per maand inclusief afbeeldingen
- Gemini van Google
 - Gemini pro is gratis
 - Gemini advanced is €22 per maand (geen afbeeldingen)
- Microsoft Copilot
 - Windows Copilot (Bing Chat) inclusief afbeeldingen
 - Copilot pro €20 per maand
- Of open source op eigen computer of in de cloud.



Stoomcursus prompt schrijven: Wat is het?

- Wat is prompt engineering eigenlijk?
 - Om ChatGPT tekst te laten generen, moet je een vraag stellen of een instructie geven. Dat heet een **prompt**. De manier waarop je een prompt schrijft bepaalt in sterke mate wat voor soort antwoord je krijgt.
- Er zijn drie fundamentele manieren om met ChatGPT te werken:
 - Verkleinen: van groot naar klein (bijv. samenvatten)
 - Veranderen: grote en betekenis veranderen niet (bijv. stijl of vertalen)
 - Genereren: van klein naar groot (bijv. verhaal schrijven)

Zie David Shapiro op Medium voor meer details:
<https://medium.com/@dave-shap/become-a-gpt-prompt-maestro-943986a93b81>



Prompt schrijven: Hoe doe je dat?

- Geef expliciete instructies. Dat helpt om de context te schetsen en de taak duidelijk te definiëren.
- Je kunt de kwaliteit verbeteren door vervolgvragen te stellen of om aanvullingen te vragen. Je kunt ook aangeven dat je iets zo niet had bedoeld.
- Je kunt aangeven in wat voor vorm je het antwoord wil hebben. Bijvoorbeeld in de vorm van een tabel.
- Je kunt ook aangeven of je iets in een bepaalde stijl wilt hebben. Bijvoorbeeld schrijf de uitleg voor een student op de universiteit of voor een leerling in groep 2.



Prompt schrijven: Om te beginnen

- Zorg dat je duidelijk bent.
 - Als je niet de juiste woorden kent dan beperkt dat ChatGPT in het antwoorden.
- Het is ook niet goed om teveel woorden te gebruiken.
 - Het is best lastig om de juiste balans te vinden en woorden te kiezen om het beste resultaat te krijgen.
- Rollen en doelen. Je kunt ChatGPT een rol toebedelen.
 - Bijvoorbeeld:
"je bent een makelaar met 10 jaar ervaring in het verkopen van huizen. Je doel is om een samenvatting te geven van 3 wijken in Delft voor een publiek dat bestaat uit kopers die weinig ervaring hebben met het kopen van een huis."
- Positief en negatief.
 - Je kunt aangeven dat het taalmodel bepaalde dingen juist wel of niet moet doen.



Prompt schrijven: Voor gevorderden

- Invoer-uitvoer opdracht:
 - beschrijf aan ChatGPT de invoer en welke uitvoer je verwacht
- Zero/one/few-shot prompting:
 - Zero-shot: vraag aan ChatGPT om een antwoord te geven op een vraag zonder dat je een voorbeeld geeft
 - One-shot: vraag aan ChatGPT om antwoord te geven op een vraag en geef zelf een voorbeeld
 - Few-shot: vraag aan ChatGPT om antwoord te geven en geef zelf een paar voorbeelden
- Denk stap voor stap:
 - Zgn. chain-of-thought thinking of think step-by-step. GPT3.5 en 4 doen dat nu al standaard.



Prompt schrijven: Samenwerken met ChatGPT

- Laat ChatGPT zichzelf controleren.
 - Vraag aan ChatGPT of zijn antwoord klopt
- Stapsgewijs
 - Bijv. schrijf een stuk, geef het onderwerp en vraag om een titel, vraag dan om inhoudsopgave, vraag dan om hoofdstuk 1 enz,
- Beschouw ChatGPT als je assistent of copiloot.
 - Als je niet weet hoe je een prompt moet schrijven kun je het ook vragen.
- Vraag aan ChatGPT om vragen te stellen als informatie nodig is die niet bekend is.
 - Raad met ja/nee vragen aan welk dier ik in gedachten heb



Over de maximale prompt lengte

- De maximale grootte van de prompt heet *context size*
- Prompts worden omgezet in tokens:
 - Engels: 1 woord \approx 1.3 tokens
- GPT3 van 22-11-2022:
 - 2.048 tokens \approx 5 pagina's
- GPT4 van 14-3-2023:
 - 8.096 tokens \approx 20 pagina's



Prompt schrijven: als het niet meer past

- Knip de informatie op in kleinere stukken en biedt die één voor één aan
 - Na het laatste stuk stel je de vraag
- Of, zoek in de informatie naar relevante alinea's met een gewone zoekmachine
 - Maak een prompt met de gevonden alinea's en met de vraag aan de gebruiker en biedt die aan ChatGPT
- Dit laatste kan geautomatiseerd worden:
 - Retrieval Augmented Generation (RAG)



CustomGPTs van OpenAI

- Met een ChatGPT plus abonnement kun je een eigen GPT met eigen informatie
 - Een speciale instructie
 - Eigen documenten
 - Koppeling met online databases
- OpenAI heeft een GPT store gelanceerd begin dit jaar
 - Daarin staan nu 159.000 GPTs

<https://chat.openai.com/gpts>



OpenAI is zeker niet de enige

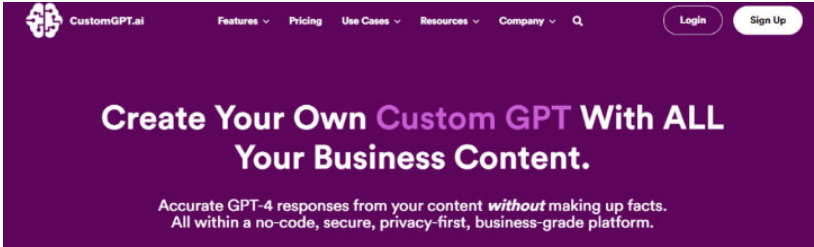
<https://customgpt.ai>

<https://www.perplexity.ai>

<https://auth.askyourpdf.com>

<https://pdf.ai>

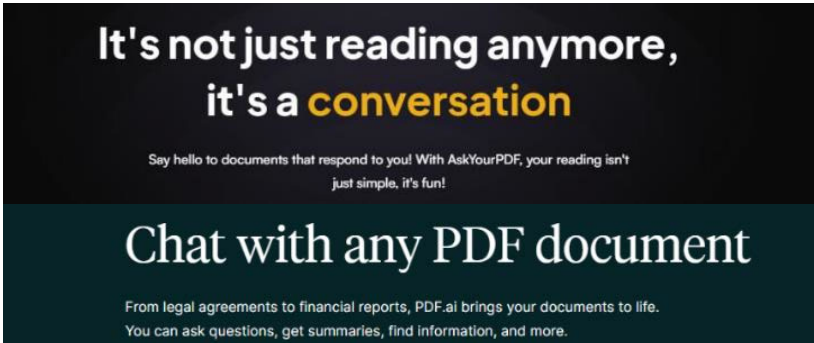
<https://chatdoc.com>



CustomGPT.ai Features Pricing Use Cases Resources Company Q Login Sign Up

Create Your Own Custom GPT With ALL Your Business Content.

Accurate GPT-4 responses from your content *without* making up facts. All within a no-code, secure, privacy-first, business-grade platform.

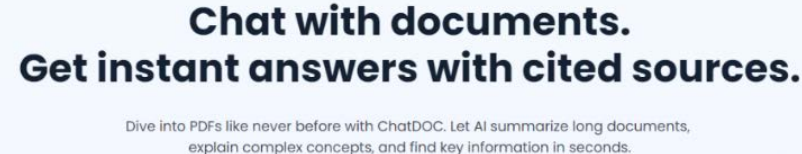


It's not just reading anymore, it's a conversation

Say hello to documents that respond to you! With AskYourPDF, your reading isn't just simple, it's fun!

Chat with any PDF document

From legal agreements to financial reports, PDF.ai brings your documents to life. You can ask questions, get summaries, find information, and more.



Chat with documents. Get instant answers with cited sources.

Dive into PDFs like never before with ChatDOC. Let AI summarize long documents, explain complex concepts, and find key information in seconds.

Multi modale modellen: afbeelding analyseren

- ChatGPT4 en Google Gemini kunnen plaatjes interpreteren

Wat voor eten staat er op deze tafel?

ChatGPT

Op de tafel staat een verscheidenheid aan eten dat eruitziet als een uitgebreide brunch of ontbijt. Hier zijn enkele van de items die ik kan identificeren:

- Vers fruit: Er zijn schalen met citroenen, bessen (waaronder blauwe bessen en aardbeien), druiven en wat lijkt op perziken of nectarines.
- Gebak: Er is een bord met wat lijkt op wafels of pannenkoeken en mogelijk ook enkele croissants of ander bladerdeeggebak.
- Granen: Er zijn kommen met wat lijkt op havermout of een andere soort ontbijtgranen.
- Dranken: Er zijn verschillende glazen die mogelijk sap of water bevatten, en een fles die wijn of sap zou kunnen zijn.
- Overige: Er zijn ook andere items die minder duidelijk zijn, zoals kaas, noten, en mogelijk wat jams of spreads.

Het is een vrij rijke en gevarieerde tafel met veel opties die men zou kunnen genieten bij een speciale gelegenheid of een luxe brunch.



https://nl.freepik.com/premium-photo/een-lange-gedekte-tafel-met-verschillend-eten_47942160.htm

Multi modale modellen: afbeeldingen genereren

- Kun je een prompt maken waarmee dezelfde afbeelding kan worden gegenereerd?

De afbeelding is gegenereerd op basis van uw verzoek. Hierboven kunt u het resultaat zien.

You

Bedankt. Dat ziet er mooi uit. Kun je me ook de tekst van deze prompt geven?

ChatGPT

Natuurlijk, hier is de tekst van de prompt die gebruikt werd om de afbeelding te genereren:

"A sumptuous brunch spread on an elegant dining table in a room with vintage decor. The table is covered with a dark blue tablecloth and is heavily laden with various foods. There are bowls filled with fresh fruits, including lemons, berries, and grapes. Plates with waffles, pancakes, and croissants add to the variety. Several glasses that could contain juice or water, along with a bottle that resembles a wine bottle, are present. Bowls of what appear to be oatmeal or cereal are also on the table. The setting is near a window with natural light streaming in, highlighting the vibrant colors of the food and creating a warm, inviting atmosphere. The room has an antique wooden cabinet and the overall aesthetic suggests a luxurious and leisurely late morning meal."



Wat kunnen we het komende jaar verwachten?

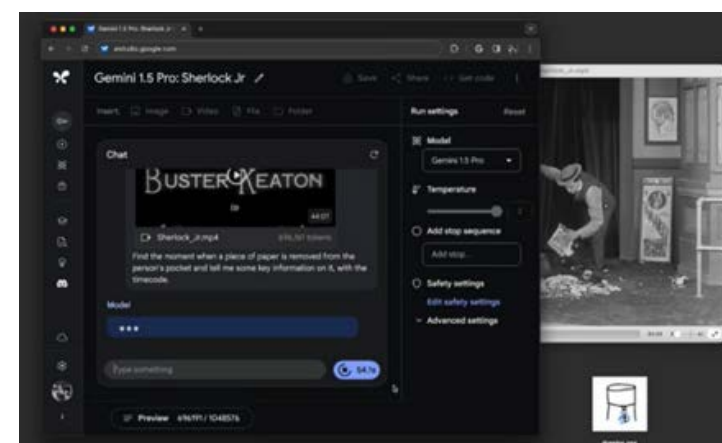
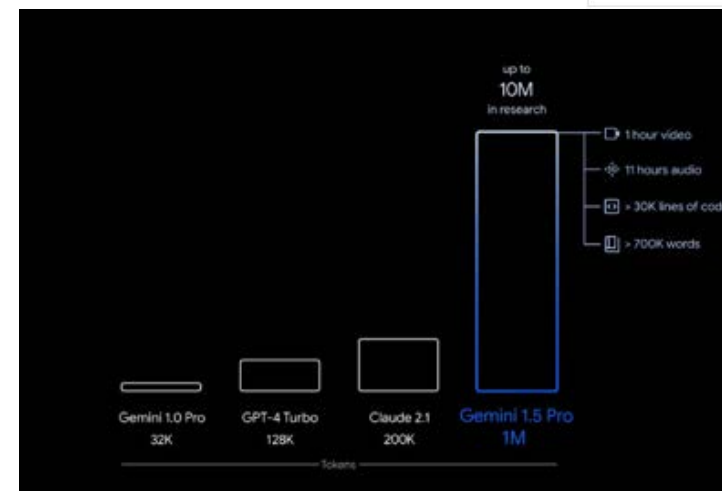
- Nog grotere context windows
- Kleinere modellen die erg goed zijn
- AI agents die samenwerken
- Helpdesk wordt vervangen door taalmodel met RAG
- Tekst naar video
- Alvast een voorproefje ...



Midjourney prompt: Looking in a crystal ball seeing the future of artificial intelligence, ultra HD, super realistic, cinematic lighting. (fast)

Google Gemini 1.5

- Kort na het vrijgeven van Gemini 1.0 (advanced) kwam Google met de aankondiging van Gemini 1.5 <https://blog.google/technology/ai/google-gemini-next-generation-model-february-2024>
- Gemini 1.5 heeft een context size van 1 miljoen tokens.
 - 1.000.000 tokens \approx 2.500 blz
- Het model is nog niet beschikbaar maar Google heeft wel een aantal indrukwekkende demonstraties als video online gezet.



<https://youtu.be/wa0MT8OwHuk>

OpenAI Sora

- SORA is een tekst naar video model en is op 15-2-2024 gelanceerd
- Is in staat om op basis van een prompt 1 minuut video te genereren.
- Is niet nieuw maar de kwaliteit is veel beter dan eerdere modellen.
- Volgens OpenAI is SORA kun je met SORA een wereld simuleren

<https://openai.com/research/video-generation-models-as-world-simulators>



Stagemogelijkheden bij Hansken met taalmodellen?



Hansken
The open digital forensic platform

Stageopdracht bij Hansken

Onderzoek naar de Integratie van Taalmodellen in Hansken

Projectomschrijving

Met het oog op de toekomst onderzoekt het NFI de mogelijkheden om rechercheurs en digitale experts te ondersteunen met geavanceerde taalmodellen zoals ChatGPT in het gebruik van Hansken. Dit project behelst meerdere deelaspecten, waaronder het beoordelen van de geschiktheid van open source taalmodellen als alternatief voor ChatGPT 4, het ontwikkelen van effectieve prompts voor verschillende taken, en het representeren van zoekresultaten in een knowledge graph (bijvoorbeeld met Neo4j). Daarnaast zal onderzocht worden hoe aanvullende onderzoeksinformatie, zoals aanklachten en verklaringen van verdachten en getuigen, geïntegreerd kan worden in deze knowledge graph zodat taalmodellen deze kunnen benutten bij het beantwoorden van onderzoeksvragen. Een belangrijk aandachtspunt is het voorkomen van hallucinaties of onjuiste generaties door het taalmodel.

Waarom?

Dit project verkent hoe cutting-edge AI en taalmodellen kunnen bijdragen aan forensisch onderzoek. Het onderzoekt hoe deze technologieën kunnen helpen bij het analyseren van complexe data, het bieden van inzichten en het versnellen van onderzoeksprocessen.

Competenties

- Kunstmatige Intelligentie
- NLP
- Data Modelling
- Python
- Neo4j
- Kritisch Denken
- Analyse van Taalmodellen

Wil je meer weten over stage lopen bij Hansken? Ga naar hansken.org



... en nog 7 andere soorten stageopdrachten bij de Hansken tafel op de informatiemarkt.

AG Connect online 20-2-2024



AG CONNECT

ppij Tech & Toekomst Carrière

Expert • Artificial Intelligence +
oegisteren om 08:59 • Seestijd 6 minuten • 0 reacties

ChatGPT als digitale speurneus is er bijna

Is ChatGPT een digitale speurneus voor onderzoekers? Nog steeds niet maar we zijn er bijna, ziet Hans Henseler. Vorig jaar concludeerde hij al wel dat het een slimme assistent voor onderzoekers kon worden. Dat is gedeeltelijk al gelukt maar de uitdaging is nog om dit on-premises te doen en niet afhankelijk te zijn van een groot taalmodel in de BigTech-cloud.

 Dr. ir. Hans Henseler
[Meer van deze auteur](#)



<https://www.agconnect.nl/tech-en-toekomst/artificial-intelligence/taalmodellen-als-copiloot-voor-zoeken-naar-digitaal-bewijs>

Meer lezen, horen en kijken over LLMs?

- POKI: Podcast over Kunstmatige Intelligentie van Alexander Klöpping en Wietse Hage
<https://linktr.ee/poki.show>
- Wes Roth op YouTube: beschouwingen over actuele ontwikkelingen op het gebied van artificial intelligence
<https://www.youtube.com/@WesRoth>
- Two Minute Papers: bespreken en demonstrenen van de nieuwste wetenschappelijke artikelen
<https://www.youtube.com/@TwoMinutePapers>
- David Shapiro: expert in prompt engineering. Veel nieuwe ideeën en filosofeert veel over Artificial General Intelligence AGI
<https://www.youtube.com/@DaveShap>
- Voor als je tijd hebt: Lex Fridman Podcast. Interviews van 1 uur en langer. O.a. met Yann Le Cun (Meta) en Sam Altman (OpenAI)
<https://lexfridman.com/podcast/>

Bedankt voor jullie aandacht



Create a prompt to recreate this picture in the style of painter Rembrandt. On the table a sumptuous still life scene with a variety of cheeses and a bottle of wine, arranged to create an inviting gastronomic display. In the foreground, there's a wooden cutting board adorned with several types of cheese: a tall, white rinded cheese stack, a wedge of soft cheese with a bloomy rind, pieces of hard yellow cheese, and a slice of cheese with an orange rind, possibly indicating a washed rind variety.

Accompaniments include a bowl of mixed grapes, figs halved to show their pink interiors, and assorted nuts like walnuts and almonds. There are also some crackers and possibly dried fruits, adding texture and variety to the board. In the background, two wine glasses, one partially obscured, contain red wine, suggesting a pairing with the cheese. To the left, the wine bottle stands with its label facing away, and behind the glasses, a lush green potted plant adds a fresh, organic touch to the scene. The lighting is warm and soft, indicating either indoor lighting or a mellow natural light, which highlights the rich colors and textures of the food and creates a cozy, inviting atmosphere.