Netherlands Forensic Institute
Ministry of Justice and Security

Large Language Models for Digital Forensics

Hans Henseler (NFI, UoSL)
Gaëtan Michelet (UNIL)

DFRWS EU 2024, Workshop,
March 19, 2024, Zaragoza, Spain

UNIL | Université de Lausanne
hogeschool Leiden
Netherlands Forensic Institute
Ministry of Justice and Security

Agenda

Time	Title
11:00	Introduction Large Language Models
11:30	Part I: Hands-on prompt engineering for digital forensics
12:15	Break
12:30	Part II: Hands-on with Llama2
12:50	Wrap up

UNIL | Université de Lausanne

hogeschool Leiden

Netherlands Forensic Institute
Ministry of Justice and Security

Introduction Large Language Models

Part I: Hans Henseler

Microsoft Copilot

- 2021 GitHub Copilot
- February 1: Bing Chat
- September 26 : Windows 11 Copilot
- November 1: Microsoft Office 365 Copilot:

[Introducing Microsoft 365 Copilot | Your Copilot for Work - YouTube]

The rise of deep learning 2012-2022

2012: AlexNet wins the ImageNet Large Scale Visual Recognition Challenge

2014: Introduction of Generative Adversarial Networks (GAN's)

2015: AlphaGo defeats world champion Go, Lee Sedol

2017: Google introduces BERT improving ML translations

2018-2021: Introduction of GPT-2, DALL-E, CLIP, GPT-3, …

2022: DALL-E2, Midjourney, Stable Diffusion, ChatGPT, …

2023: GPT4, Llama2 (Meta), Claude 2 (Anthropic), Mistral, Grok (X)

2024: Gemini (Google), Mistral Large, Claude 3 …

What is ChatGPT?

- **ChatGPT is a large language model (LLM)**
  - Essentially a machine learning model that learns *an algorithm* to predict the next word based on many text examples
- **Based on GPT3.5/GPT4 (Generative Pre-trained Transformer)**
  - Improved version of GPT-3 that "understands" text and program code
  - Different models for performance, chat, text and code completion
  - GPT3.5 was trained on 570 GB data from the internet (articles, posts, web pages and books)
- **Available as**
  - Free version
  - ChatGPT-plus €23 per month
  - OpenAI playground (API access):
    - GPT3.5-turbo API 0,002 dollar per 1.000 tokens, ~700 words
    - GPT4 API 0,03 dollar per 1.000 tokens, ~700 words

What can ChatGPT do?

Chat. Like a chatbot that…

- Assists with writing and brainstorming
- Tells riddles, jokes, stories
- Plays games
- Gives compliments and advise
- Helps with filling in forms

But it that can also:

- Summarise
- Translate
- Analyse and structure (unstructured) information
- Answer questions (but the answer may not be right)
- Assist with software writing and debugging
- Generate (anonymous) testdata

What can ChatGPT not do?

- It hallucinates facts
- It gives wrong answers
- Replies can be biased
- Can not act spontaneously (needs to be prompted)
- Is not good at making calculations (e.g. 4213x8242)
- Is limited to generating text output
- Can accidentally reveal sensitive training data
- …?

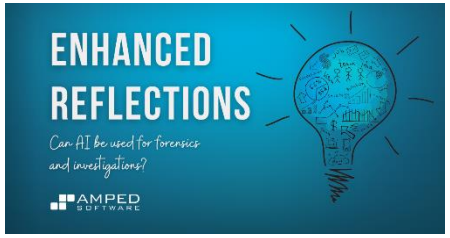
REMINDER

use LLMs as co-pilot

UNIL | Université de Lausanne
hogeschool Leiden
Netherlands Forensic Institute
Ministry of Justice and Security

Thoughts on using AI for forensic purposes

- Hansken is:
  - used for investigations, bit
  - designed for evidentiary use
- Evidentiary use is more strict:
  - Accurate
  - Repeatable
  - Reproducible
- So algorithms must be:
  - Explainable
  - Validated
  - Deterministic
  - Not depend on external data



- Artificial Intelligence:
  - Use external data: Training sets
  - Use external data: Cause bios
  - Lacks explainability
- Use AI for investigative purposes,with
  - Disclaimer
  - Education
- By the way:
  - Not all currently used algorithms are good
  - Data under investigation can results from AI itself

https://blog.ampedsoftware.com/2021/10/05/can-ai-be-used-for-forensics-and-investigations

Hallucinations, data privacy and explainability

- Preventing hallucinations:
  - Provide clear prompts to ChatGPT to base its response on digital traces
  - ChatGPT should not hallucinate but inform that there are no relevant traces
  - Retrieval-Augmented-Generation (RAG) comes to the rescue
  - Explicit references to the source on which a response is based
- Maintaining data privacy:
  - Digital traces and case specific details can not be send to the public cloud (e.g., ChatGPT in the OpenAI cloud)
  - Powerfull Large Language Models can already be deployed on premise (e.g., Meta's Llama 2)
  - Assumption: Open source LLMs with RAG do not need the extensive factual knowledge as ChatGPT/GPT-4
- Explaining responses:
  - Identify the sources that were retrieved as part of the RAG method to explain the response
  - Reproducability over creativity (experiment with "temperature" of the LLM)

Topics for future work

- Can we do this off-line with the same quality?
- Build a co-pilot in Hansken leveraging Retrieval Augmented Generation (RAG)
- Evaluate with (real) users
- Advanced topics:
  - Multi-modal generative transformers (Visual ChatGPT)
  - Augmented language models
  - Planning an investigation

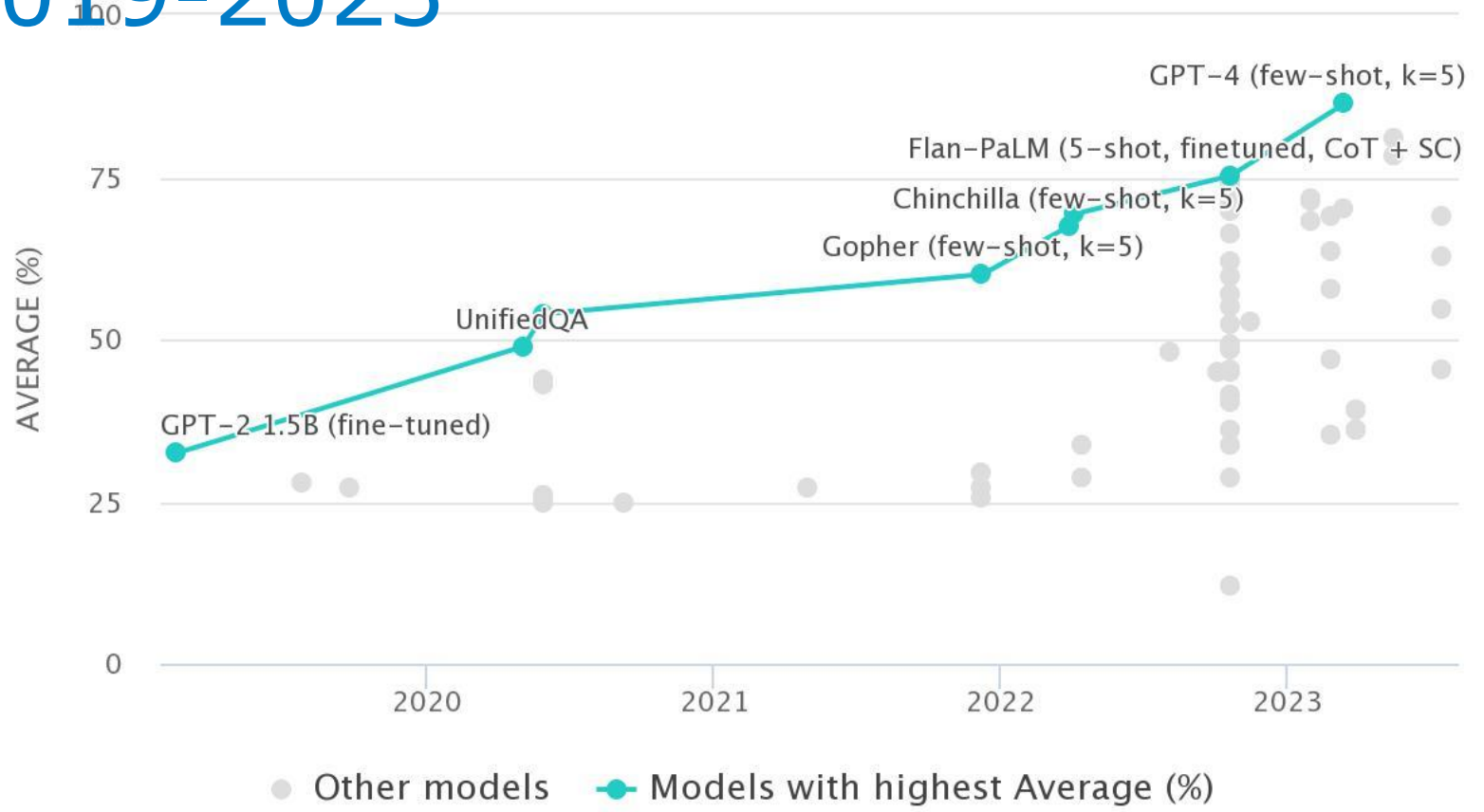
Midjourney prompt: Looking in a crystal ball seeing the future of artificial intelligence, ultra HD, super realistic, cinematic lighting. (fast)

How smart are LLMs?

- Hugging Face **open** LLM leaderboard:
  - https://huggingface.co/spaces/HuggingFaceH4/open_llm_leaderboard
- Score gebaseerd op:
  - **ARC**: Abstraction en Reasoning Challenge
  - **HellaSwag**: een benchmark die zich richt op gezond verstand redeneren
  - **MMLU**: Massive Multitask Language Understanding
  - **TruthfulQA**: een benchmark die beoordeelt of een taalmodel waarheidsgetrouwe antwoorden genereert

Model	Score
garage-bAInd/Platypus2-70B-instruct	73.13
upstage/Llama-2-70b-instruct-v2	72.95
fangloveskari/Platypus_QLoRA_LLaMA_70b	72.94
yeontaek/llama-2-70B-ensemble-v5	72.86
TheBloke/Genz-70b-GPTQ	72.82
TheBloke/Platypus2-70B-Instruct-GPTQ	72.81
psmathur/model_007	72.72
yeontaek/llama-2-70B-ensemble-v4	72.64
psmathur/orca_mini_v3_70b	72.64
ehartford/Samantha-1.11-70b	72.61
MayaPH/GodziLLa2-70B	72.59
psmathur/model_007_v2	72.49
chargoddard/MelangeA-70b	72.43
ehartford/Samantha-1.1-70b	72.42
psmathur/model_009	72.36

MMLU 2019-2023



Llama 2 70B
Llama 2 35B
Llama 2 13B
Llama 2 7B

https://paperswithcode.com/sota/multi-task-language-understanding-on-mmlu

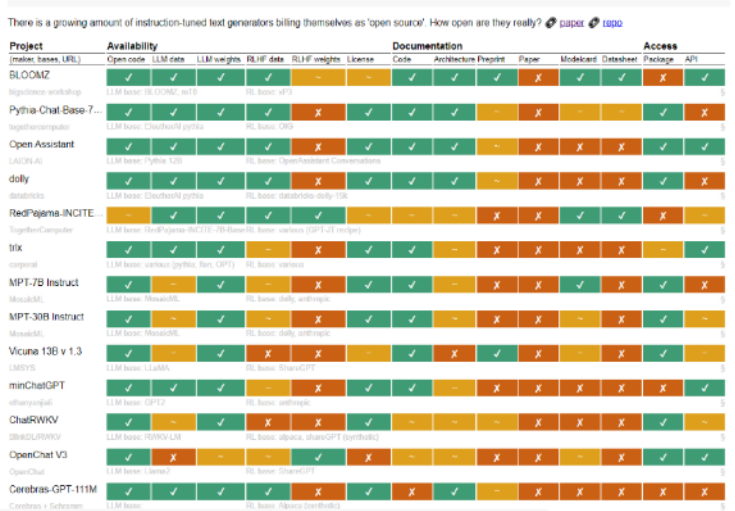
UNIL | Université de Lausanne

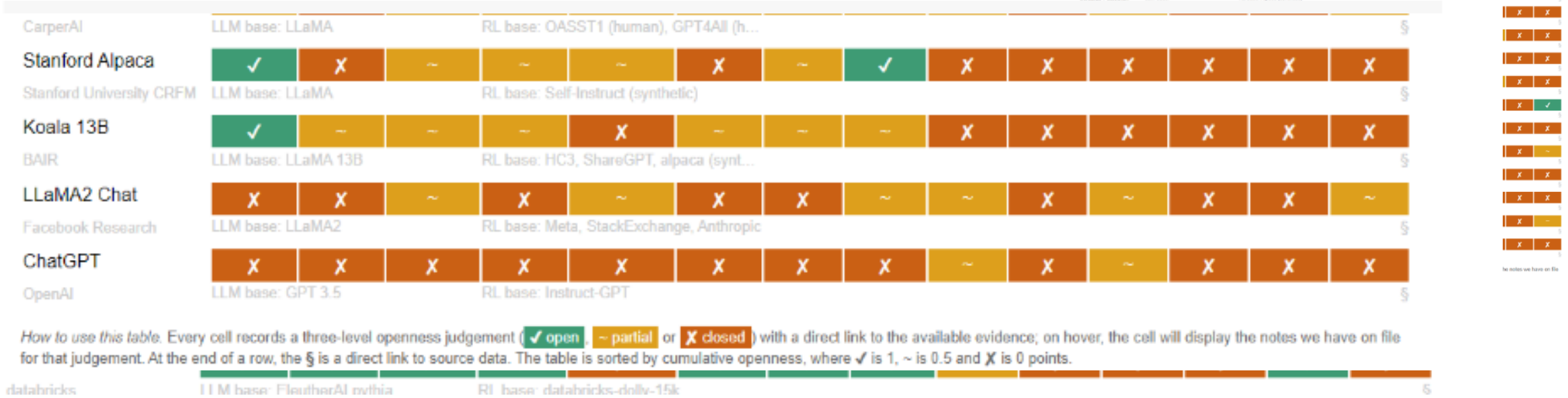
hogeschool Leiden

Netherlands Forensic Institute
Ministry of Justice and Security

How open are LLMs?

- Researchers from Nijmegen University:
  - Opening up ChatGPT: tracking "open source" LLM + RLH
  - https://opening-up-chatgpt.github.io/





UNIL | Université de Lausanne
hogeschool Leiden
Netherlands Forensic Institute
Ministry of Justice and Security

Agenda

Time	Title
11:00	Introduction Large Language Models
11:30	Part I: Hands-on prompt engineering for digital forensics
12:15	Break
12:30	Part II: Hands-on with Llama2
12:50	Wrap up

UNIL | Université de Lausanne

hogeschool Leiden

Netherlands Forensic Institute
Ministry of Justice and Security

Part I

Hands-on prompt engineering for digital forensics

Github & Google CoLab

Link:

- https://github.com/NetherlandsForensicInstitute/DFRWS-EU-2024-LLM4DF-Workshop
- Notebooks:
- Part I: Prompt engineering with ChatGPT for Digital Forensics
- Part II: Hands-on with Llama2

Requirements:

- Google CoLab is free but you need a Gmail account!
- Account for accessing free version of OpenAI ChatGPT
- Make sure to select a T4 GPU

Reductive operations

- **Summarization** — Say the same thing with fewer words. Can use list, notes, executive summary.
- **Distillation** — Purify the underlying principles or facts. Remove all the noise, extract axioms, foundations, etc.
- **Extraction** — Retrieve specific kinds of information. Question answering, listing names, extracting dates, etc.
- **Characterizing** — Describe the content of the text. Describe either the text as a whole, or within the subject.
- **Analyzing** — Find patterns or evaluate against a framework. Structural analysis, rhetorical analysis, etc
- **Evaluation** — Measuring, grading, or judging the content. Grading papers, evaluating against morals
- **Critiquing** — Provide feedback within the context of the text. Provide recommendations for improvement

Transformative Operations

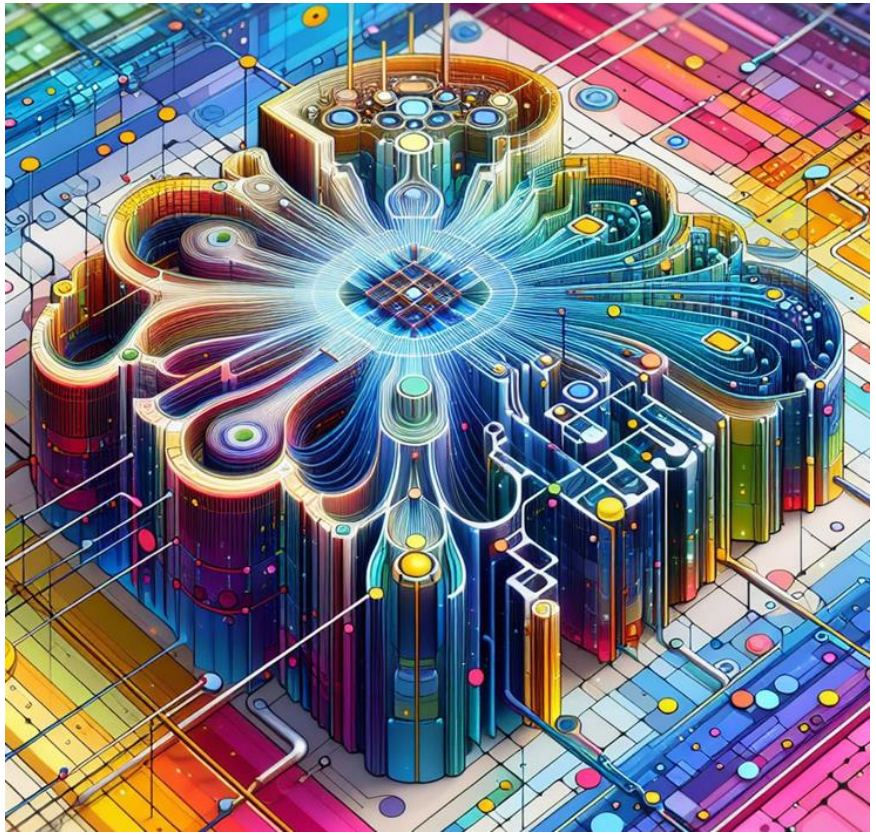
- **Reformatting** — Change the presentation only. Prose to screenplay, XML to JSON.
- **Refactoring** — Achieve same results with more efficiency. Say the same exact thing, but differently.
- **Language Change** — Translate between languages. English to Russian, C++ to Python.
- **Restructuring** — Optimize structure for logical flow, etc. Change order, add or remove structure.
- **Modification** — Rewrite copy to achieve different intention. Change tone, formality, diplomacy, style, etc.
- **Clarification** — Make something more comprehensible. Embellish or more clearly articulate.

Generative (Expansion) Operations

- **Drafting** — Generate a draft of some kind of document. Code, fiction, legal copy, KB, science, storytelling.
- **Planning** — Given parameters, come up with plans. Actions, projects, objectives, missions, constraints, context.
- **Brainstorming** — Use imagine to list out possibilities. Ideation, exploration of possibilities, problem solving, hypothesizing.
- **Amplification** — Articulate and explicate something further. Expanding and expounding, riffing on stuff.

Maximum prompt size

- The maximum size of the prompt in a LLM is called *context size*
- Prompts are converted into tokens
  - English: 1 word ≈ 1.3 tokens
- GPT3 from 22-11-2022:
  - 2.048 tokens ≈ 5 pages
- GPT4 from 14-3-2023:
  - 8.096 tokens ≈ 20 pagina's



What to do if your prompt doesn't fit?

- Cut the information into smaller pieces and present them one by one
  - After the last part you ask the question
- Or, search the information for relevant paragraphs with a regular search engine
  - Create a prompt with the found paragraphs and the question to the user and offer it to ChatGPT
- The latter can be automated:
  - Retrieval Augmented Generation (RAG)

UNIL | Université de Lausanne
hogeschool Leiden
Netherlands Forensic Institute
Ministry of Justice and Security

Prompt engineering with ChatGPT for DF

Our 4 case study experiments:

1. Writing search queries
2. Summarising chat conversations
3. Analysing search results
4. Reverse engineering

In part I Colab we will focusses on #1, #3 and #4

Midjourney prompt: photorealistic picture of a digital sleuth in the style of Sherlock Holmes as a robot investigating a crime scene with digital traces in smartphones and computers (fast)

More on prompt engineering

- Videos and articles by David Shapiro:
  - https://medium.com/@dave-shap/become-a-gpt-prompt-maestro-943986a93b81
  - On YouTube: https://www.youtube.com/watch?v=aq7fnqzeaPc
  - About System Prompts: https://www.youtube.com/watch?v=oILYjtbmLgc&t=760s
- Video and notebook by AssemblyAI:
  - Prompt Engineering 101
    - https://www.youtube.com/watch?v=aOm75o2Z5-o
  - Prompt_Engineering_101.ipynb
    - https://colab.research.google.com/drive/1lHd9b8C4ccAGpkK06dzcFB0asjXWGZi0

Exercise II: prompt engineering & ChatGPT 3.5

Link:

- https://github.com/HansHenseler/DFRWS-APAC-LLM-Workshop

Notebooks:

- Part II: Prompt engineering with ChatGPT for Digital Forensics
- Part III: Handson with Llama2
- Part IV: Retrieval Augmented Generation with Llama2

Requirements:

- You need to have an account to chat with ChatGPT 3.5 (free)
- You can open the notebook in Google CoLab for better navigation

Break

Time	Title
11:00	Introduction Large Language Models
11:30	Part I: Hands-on prompt engineering for digital forensics
12:15	Break
12:30	Part II: Hands-on with Llama2
12:50	Wrap up

UNIL | Université de Lausanne
hogeschool Leiden
Netherlands Forensic Institute
Ministry of Justice and Security

Agenda

Time	Title
11:00	Introduction Large Language Models
11:30	Part I: Hands-on prompt engineering for digital forensics
12:15	Break
12:30	Part II: Hands-on with Llama2
12:50	Wrap up

UNIL | Université de Lausanne

hogeschool Leiden

Netherlands Forensic Institute
Ministry of Justice and Security

Part II

Hands-on with a local LLM in a Google Colab notebook

How to get LLMs

Subscribe to OpenAI GPT4, Google PaLM

- Models are generally more powerful (Higher scores in various assessments)
- No need to setup and maintain the models and hardware
- Need to pay
- Privacy problems

Setup a local in-house LLM

- Many models are free
- No privacy issue
- Mid range hardware required
- Self maintenance (very limited support from publishers)

Local LLMs

What
Hardware
is required?

What
Models
to be used?

Background Info for Model Selection

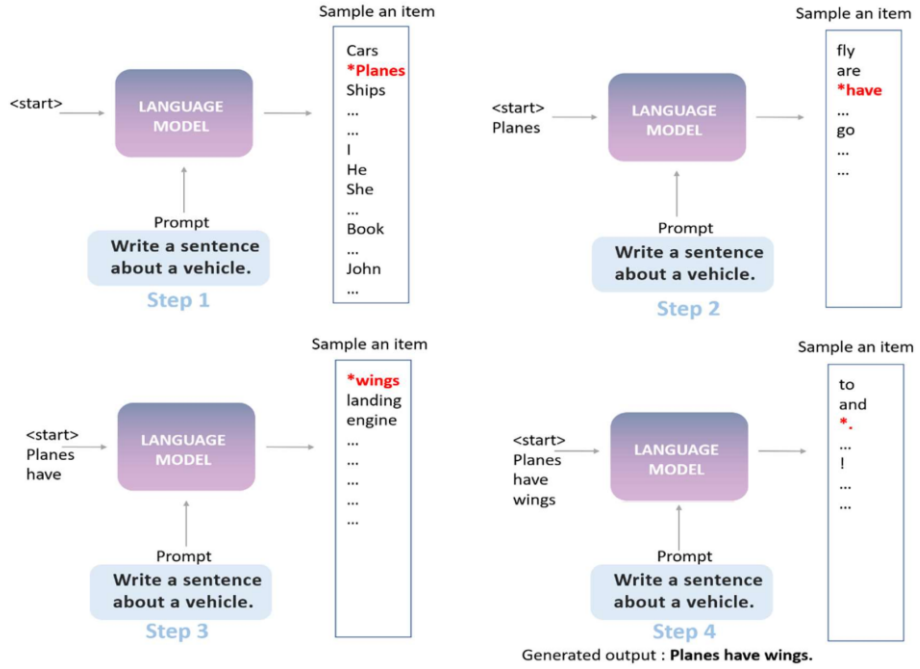
- What do LLMs perform?
  - Generate coherent text (semantically related text) which can delivery meaningful contents:

$$P(w_n | w_{n-1}, w_{n-2}, \ldots, prompt)$$

  - Words are probabilistically generated one by one: depending on the previous generated words and the user given “prompt”. Different LLMs have different probability distribution functions!!!

An example illustrating how LLMs generate words

- using the prompt: “Write a sentence about a vehicle.”



Important Parameters for Local LLMs

Top_k	Only consider the top *k* words
Top_p	Only consider the top words having total probabilities ≤ *Top_p*
Temperature	Higher value → more diverse and creative content, but content may not be coherent or even irrelevant
n_ctx, max_length	Max. context length
Max_new_tokens	Max. number of tokens to be generated
Repeat_penalty	Discourage repetitive or redundant output

Local LLM Selection

- Features of local LLMs to be considered:

Size of the models (num. of parameters/weights)	• 7B, 13B, 30B, etc. • Larger size models usually give better performance but require better hardware and slower
Nature of the models	Use instruct model or chat models for Q&A and Retrieval Augmented Generation (RAG)
Weight Quantization	Usually map floating point values (16bits/32bits) to integer values (int8, int4, etc)
Model Data Format/Structure	Hugging Face, GGUF, GGML (now replaced by GGUF), GPTQ, AWQ
Context length (tokens)	• 2K, 4K, 8K, 32K…. • ChatGPT : 8K, GPT4: 32K • Number of context words ~ (0.6 or 0.7) * number of tokens

Parameter Quantization

- Models are **too big**!
- High VRAM GPU cards are **too expensive**! Almost no competitor !!!
- **Limited supply** of high VRAM GPUs
- Model computation is **slow**!

How to make models smaller, while preserving the number of parameters/weights, or minimizing the degradation of performance?

Use **smaller number of bits** to store the parameters/weights

Float32, float16 ---→ int8 (8-bit integer), int4, …

Faster computation

Frameworks

Hugging Face: Traditional framework

GGUF/GGML: Optimized for CPU and (CPU + GPU)

GPTQ: Optimized for GPU and (GPU + CPU)

AWQ: Recent efficient quantization method (size, speed)

meta-llama/Llama-2-7b-chat-hf
Text Generation • Updated Aug 9 • ⬇ 1.09M • ♡ 1.32k

meta-llama/Llama-2-7b
Text Generation • Updated Jul 20 • ♡ 2.65k

meta-llama/Llama-2-70b-chat-hf
Text Generation • Updated Aug 9 • ⬇ 141k • ♡ 1.39k

meta-llama/Llama-2-7b-hf
Text Generation • Updated Aug 9 • ⬇ 563k • ♡ 627

meta-llama/Llama-2-13b-chat-hf
Text Generation • Updated Aug 9 • ⬇ 240k • ♡ 585

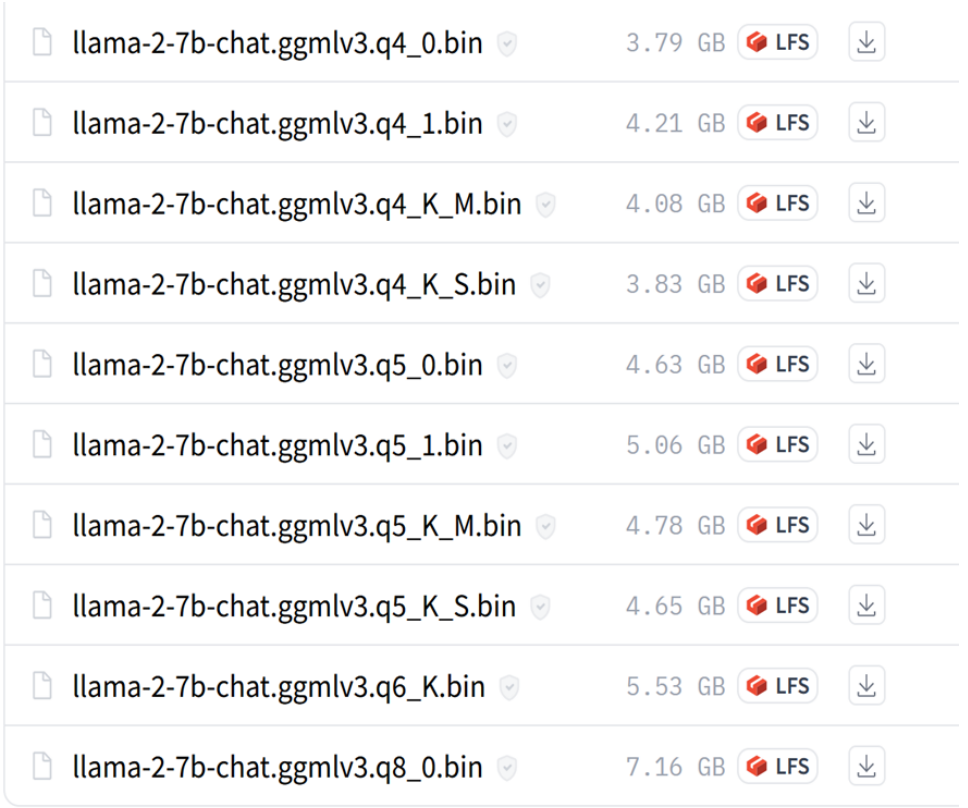
TheBloke/Llama-2-7B-Chat-GGML
Text Generation • Updated 6 days ago • ⬇ 7.04k • ♡ 570

Llama 2 7B chat model

No. of parameters: 7B (float 16)

Memory required: ~ 14GB

- Q4_0: 4bit quantization
  - 7B parameters ~ 3.5GB
- Q5_0: 5bit quantization
  - 7B parameters ~ 4.4GB
- Q6_K_S: 6bit K-quantization
  - 7B parameters ~ 5.3GB
- Q8_0: 8bit quantization
  - 7B parameters~ 7.0GB



Exercise II: Hands on with Llama2

Link:

- https://github.com/NetherlandsForensicInstitute/DFRWS-EU-2024-LLM4DF-Workshop

Notebooks:

- Part I: Prompt engineering with ChatGPT for Digital Forensics
- Part II: Hands-on with Llama2

Requirements:

- Google CoLab is free but you need a Gmail account!
- Make sure to select a T4 GPU

UNIL | Université de Lausanne
hogeschool Leiden
Netherlands Forensic Institute
Ministry of Justice and Security

Agenda

Time	Title
11:00	Introduction Large Language Models
11:30	Part I: Hands-on prompt engineering for digital forensics
12:15	Break
12:30	Part II: Hands-on with Llama2
13:00	Wrap up

UNIL | Université de Lausanne

hogeschool Leiden

Netherlands Forensic Institute
Ministry of Justice and Security

Wrap up

What’s new and what’s coming?

Custom GPTs from OpenAI: RAG & functions

- With a ChatGPT plus subscription you can build a custom GPT
  - Tailored instruction
  - Proprietary documents (RAG)
  - Connection to online API's (Functions)
- OpenAI launched their GPT store beginning of 2024
  - In february it had 159.000 GPTs

https://chat.openai.com/gpts

OpenAI Custom GPTs are not alone

https://customgpt.ai

https://www.perplexity.ai

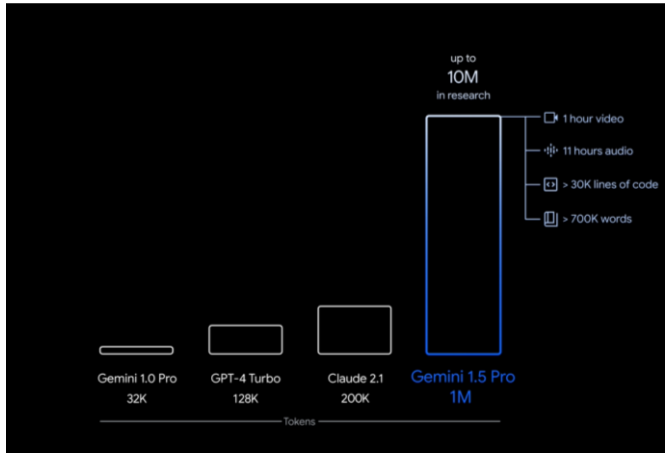
https://auth.askyourpdf.com

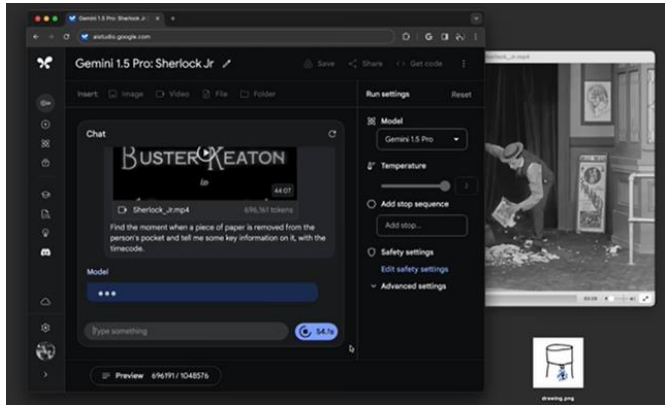
https://pdf.ai

https://chatdoc.com

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 ≈ 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 eedere modellen.
- Volgens OpenAI is SORA kun je met SORA een wereld simuleren

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

UNIL | Université de Lausanne

hogeschool Leiden

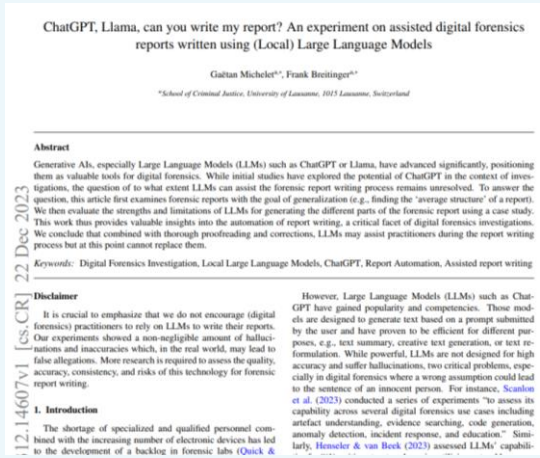
Netherlands Forensic Institute
Ministry of Justice and Security

Thank you!

Hans Henseler
h.henseler@nfi.nl

Gaëtan Michelet
victor.cheng@tauexpress.com

Published papers and articles:



ChatGPT, Llama, can you write my report? An experiment on assisted digital forensics reports written using (Local) Large Language Models

Gaëtan Michelet[a,*], Frank Breitinger[a,*]

[a]School of Criminal Justice, University of Lausanne, 1015 Lausanne, Switzerland

Abstract

Generative AIs, especially Large Language Models (LLMs) such as ChatGPT or Llama, have advanced significantly, positioning them as valuable tools for digital forensics. While initial studies have explored the potential of ChatGPT in the context of investigations, the question of to what extent LLMs can assist the forensic report writing process remains unresolved. To answer the question, this article first examines forensic reports with the goal of generalization (e.g., finding the 'average structure' of a report). We then evaluate the strengths and limitations of LLMs for generating the different parts of the forensic report using a case study. This work thus provides valuable insights into the automation of report writing, a critical facet of digital forensics investigations. We conclude that combined with thorough proofreading and corrections, LLMs may assist practitioners during the report writing process but at this point cannot replace them.

Keywords: Digital Forensics Investigation, Local Large Language Models, ChatGPT, Report Automation, Assisted report writing

Disclaimer

It is crucial to emphasize that we do not encourage (digital forensics) practitioners to rely on LLMs to write their reports. Our experiments showed a non-negligible amount of hallucinations and inaccuracies which, in the real world, may lead to false allegations. More research is required to assess the quality, accuracy, consistency, and risks of this technology for forensic report writing.

1. Introduction

The shortage of specialized and qualified personnel combined with the increasing number of electronic devices has led to the development of a backlog in forensic labs (Quick &

However, Large Language Models (LLMs) such as ChatGPT have gained popularity and competencies. Those models are designed to generate text based on a prompt submitted by the user and have proven to be efficient for different purposes, e.g., text summary, creative text generation, or text reformulation. While powerful, LLMs are not designed for high accuracy and suffer hallucinations, two critical problems, especially in digital forensics where a wrong assumption could lead to the sentence of an innocent person. For instance, Scanlon et al. (2023) conducted a series of experiments "to assess its capability across several digital forensics use cases including artefact understanding, evidence searching, code generation, anomaly detection, incident response, and education." Similarly, Henseler & van Beek (2023) assessed LLMs' capabili-

https://arxiv.org/pdf/2312.14607.pdf

ChatGPT: A Digital Sleuth For Detectives?

21st February 2023 by Forensic Focus

By Hans Henseler, Professor of Digital Forensics & E-Discovery, University of Leiden Applied Sciences, and Senior Digital Forensic Scientist at the Netherlands Forensic Institute.

Helping to formulate search questions

https://www.forensicfocus.com/articles/chatgpt-a-digital-sleuth-for-detectives

eForensics

LOGIN / SIGN UP

HOME / NEW EDITION / UNRAVELING DIGITAL MYSTERIES: HOW AI COPILOTS CAN REVOLUTIONIZE DIGITAL FORENSIC INVESTIGATIONS

Unraveling Digital Mysteries: How AI Copilots can Revolutionize Digital Forensic Investigations

By Hans Henseler, Professor of Digital Forensics & E-Discovery, University of Leiden Applied Sciences, and Senior Digital Forensic Scientist at the Netherlands Forensic Institute. Introduction In hindsight, 2021 was a significant inflection point in the world of artificial intelligence, characterized by remarkable developments in deep learning, manifesting in models such....

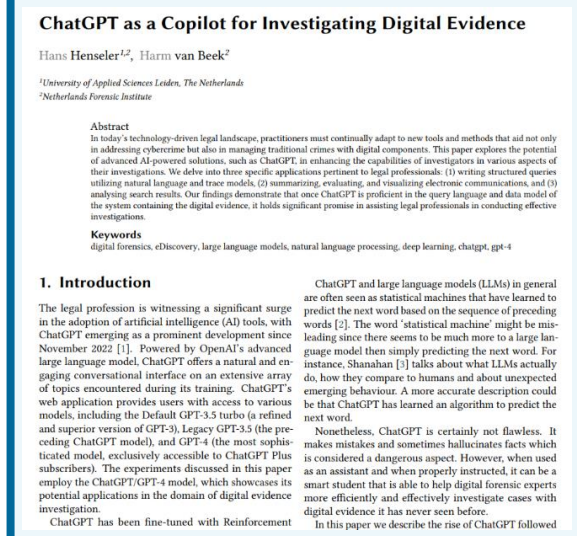
Join Us

Become an Instructor

Become a Reviewer

Writing on eForensics

https://eforensicsmag.com/unraveling-digital-mysteries-how-ai-copilots-can-revolutionize-digital-forensic-investigations/



ChatGPT as a Copilot for Investigating Digital Evidence

Hans Henseler[1,2], Harm van Beek[2]

[1]University of Applied Sciences Leiden, The Netherlands
[2]Netherlands Forensic Institute

Abstract

In today's technology-driven legal landscape, practitioners must continually adapt to new tools and methods that aid not only in addressing cybercrime but also in managing traditional crimes with digital components. This paper explores the potential of advanced AI-powered solutions, such as ChatGPT, in enhancing the capabilities of investigators in various aspects of their investigations. We delve into three specific applications pertinent to legal professionals: (1) writing structured queries utilizing natural language and trace models, (2) summarizing, evaluating, and visualizing electronic communications, and (3) analysing search results. Our findings demonstrate that once ChatGPT is proficient in the query language and data model of the system containing the digital evidence, it holds significant promise in assisting legal professionals in conducting effective investigations.

Keywords

digital forensics, eDiscovery, large language models, natural language processing, deep learning, chatgpt, gpt-4

1. Introduction

The legal profession is witnessing a significant surge in the adoption of artificial intelligence (AI) tools, with ChatGPT emerging as a prominent development since November 2022 [1]. Powered by OpenAI's advanced large language model, ChatGPT offers a natural and engaging conversational interface on an extensive array of topics encountered during its training. ChatGPT's web application provides users with access to various models, including the Default GPT-3.5 turbo (a refined and superior version of GPT-3), Legacy GPT-3.5 (the preceding ChatGPT model), and GPT-4 (the most sophisticated model, exclusively accessible to ChatGPT Plus subscribers). The experiments discussed in this paper employ the ChatGPT/GPT-4 model, which showcases its potential applications in the domain of digital evidence investigation.

ChatGPT has been fine-tuned with Reinforcement

ChatGPT and large language models (LLMs) in general are often seen as statistical machines that have learned to predict the next word based on the sequence of preceding words [2]. The word 'statistical machine' might be misleading since there seems to be much more to a large language model then simply predicting the next word. For instance, Shanahan [3] talks about what LLMs actually do, how they compare to humans and about unexpected emerging behaviour. A more accurate description could be that ChatGPT has learned an algorithm to predict the next word.

Nonetheless, ChatGPT is certainly not flawless. It makes mistakes and sometimes hallucinates facts which is considered a dangerous aspect. However, when used as an assistant and when properly instructed, it can be a smart student that is able to help digital forensic experts more efficiently and effectively investigate cases with digital evidence it has never seen before.

In this paper we describe the rise of ChatGPT followed

https://ceur-ws.org/Vol-3423