Netherlands Forensic Institute
Ministry of Justice and Security

Deep Reasoning and Large Context Windows:
Next-generation AI in Digital Forensic Investigations

Dr. Hans Henseler

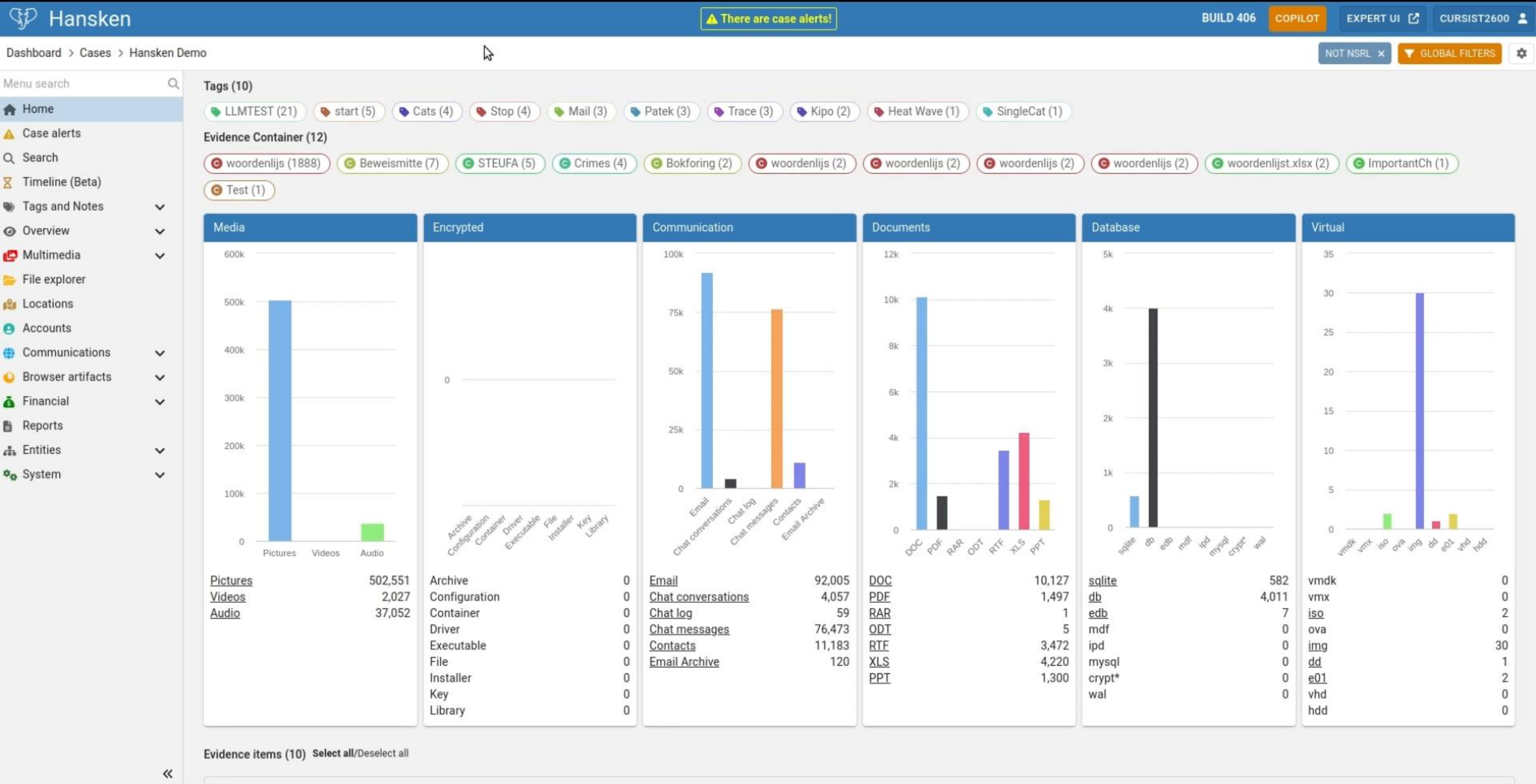
Netherlands Forensics Institute &
University of Applied Sciences Leiden

DFRWS EU 2025

April 3rd, 2025

The Hansken Copilot

- Building on experiments from 2023 and a prototype in 2024
- Currently working with an open source LLM on the local computer
  - DeepSeek R1 Distill (Llama 8B)
- Will be replaced by a LLM server in the Hansken backend
- PhD researcher **Timo Meconi** NFI Hansken togetger with the University of Groningen
- University of Applied Sciences Leiden Bachelor Forensic ICT thesis **Nickey Beekers**



Media

Pictures	502,551
Videos	2,027
Audio	37,052

Encrypted

Archive	0
Configuration	0
Container	0
Driver	0
Executable	0
File	0
Installer	0
Key	0
Library	0

Communication

Email	92,005
Chat conversations	4,057
Chat log	59
Chat messages	76,473
Contacts	11,183
Email Archive	120

Documents

DOC	10,127
PDF	1,497
RAR	1
ODT	5
RTF	3,472
XLS	4,220
PPT	1,300

Database

sqlite	582
db	4,011
edb	7
mdf	0
ipd	0
mysql	0
crypt*	0
wal	0

Virtual

vmdk	0
vmx	0
iso	2
ova	0
img	30
dd	1
e01	2
vhd	0
hdd	0

Recent developments

September 2024
- GPT-o1-preview (OpenAI)
- Llama 3.2 (Meta)
- Gemini Pro 1.5 & Notebook LM
- Qwen 2.5 (Alibaba)

October 2024
- Claude 3.5 Sonnet (Anthropic)

November 2024
- QwQ 32B Preview & Qwen 2.5 Coder (Alibaba)
- DeepSeek r1 preview (DeepSeek)
- Claude 3.5 Haiku (Anthropic)
- SmolLM 2 & SmolVLM (HuggingFace)

December 2024
- GPT-o1 & o1 Pro (OpenAI, and announces o3 & o3-mini)
- Gemini 2.0, Flash & Flash Thinking
- Phi-4 (Microsoft)
- DeepSeek V3 (DeepSeek)

January 2025
- GPT-o3-mini en o3-mini-high (OpenAI)
- Mistrall 3 Small 24B (Mistral)
- DeepSeek-r1 (DeepSeek)
- DeepSeek-r1-distill-Qwen |Llama-1.5|3|7|8|14|32|70 (DeepSeek)
- Sky-T1 32B (NovaSky)

February 2025
- Grok-3 (Xai)
- Claude 3.7 (Anthropic)
- ChatGPT-4.5 (OpenAI)
- Phi-4-mini en Phi-4-multimodal (Microsoft)

March 2025
- Gemma 3, 1/4/12/27B (Google)
- Bernie 4.5 and x1 (Baidu)

See https://huggingface.co/spaces/reach-vb/2024-ai-timeline

2025: Reasoning LLMs with large contexts

Can LLMs with extended context and reasoning capabilities transform digital investigations?

Long-term vs short-term memory in LLMs

- Pre-training:
  - Long-term memory
  - Information is stored in parametric memory, i.e., encoded in the weights of the model
  - Only stores information that the model has seen during pre-training
  - Pre-training takes a long time for large models
- Context window:
  - Short-term memory
  - Relevant information is stored in the context, i.e., the prompt
  - Can be provided at prompt time, including recent and private information
  - Less chance of hallucinations
  - Long contexts also take more time but much less time and resources than pre-training.

LLMs with large contexts

- ChatGPT-3.5 in November 2023 started with only 2,048 tokens
- GPT increases to 8,192 – 32,768 – 128K tokens over the last 2 years
- New Google Gemini has over 2M tokens

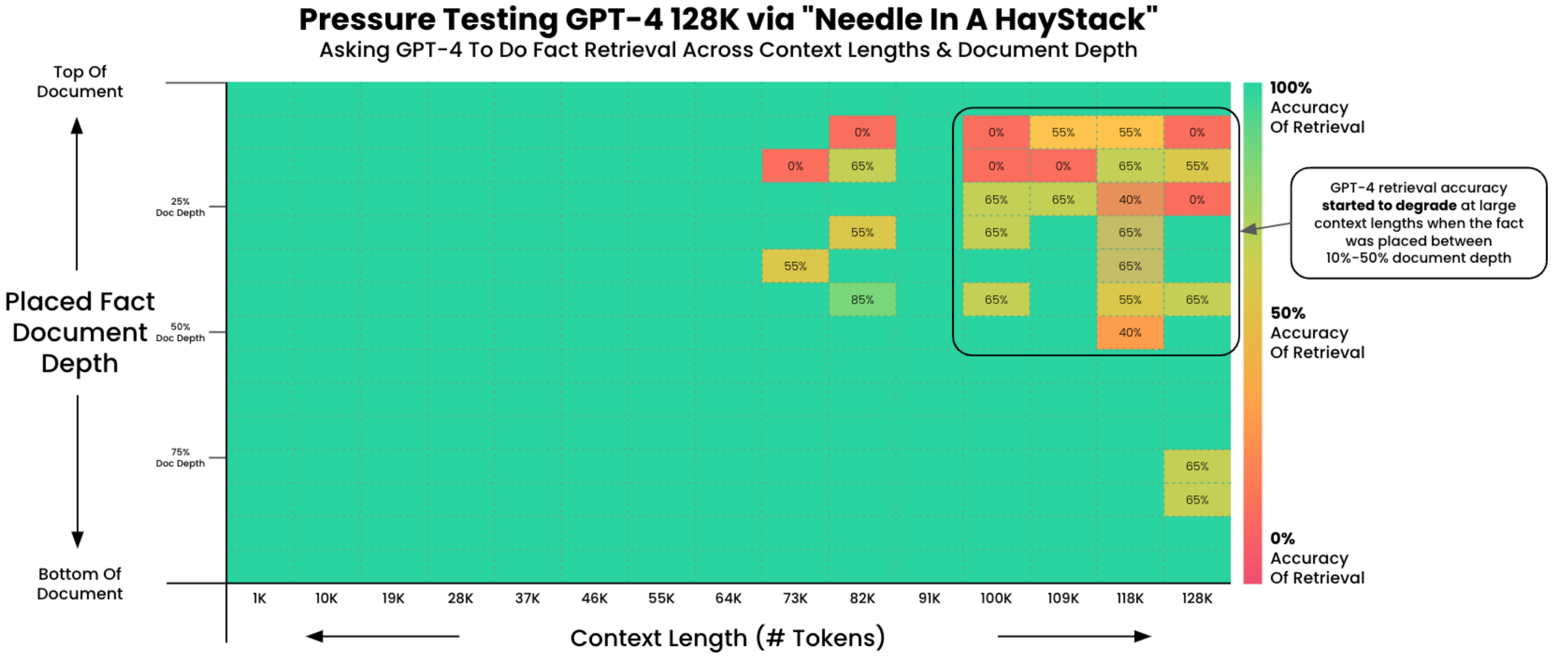
How well do LLMs remember details?

Needle in the haystack benchmark: measure how well a hidden “needle” can be retrieved from the models context. See:

- https://cloud.google.com/blog/products/ai-machine-learning/the-needle-in-the-haystack-test-and-how-gemini-pro-solves-it
- https://github.com/gkamradt/LLMTest_NeedleInAHaystack

RULER benchmark is more reliable

- https://github.com/NVIDIA/RULER



Pressure Testing GPT-4 128K via "Needle In A HayStack"

Asking GPT-4 To Do Fact Retrieval Across Context Lengths & Document Depth

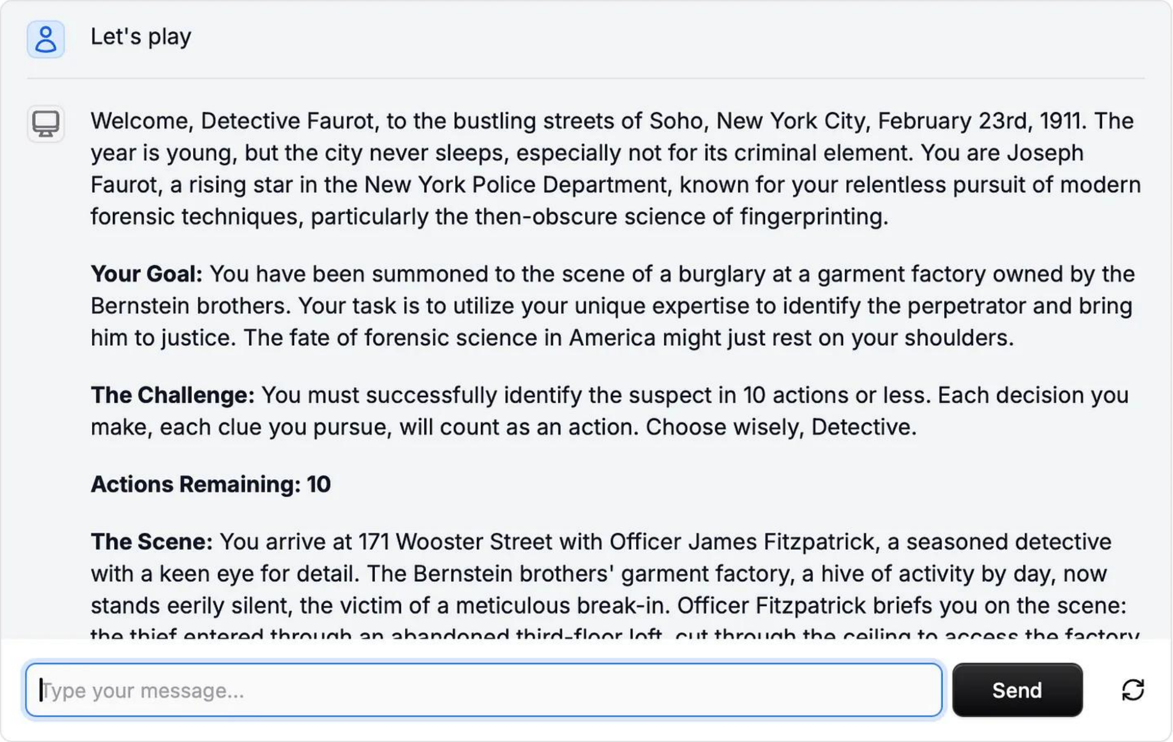
GPT-4 retrieval accuracy **started to degrade** at large context lengths when the fact was placed between 10%-50% document depth

Goal: Test GPT-4 Ability To Retrieve Information From Large Context Windows

A fact was placed within a document. GPT-4 (1106-preview) was then asked to retrieve it. The output was evaluated for accuracy. This test was run at 15 different document depths (top > bottom) and 15 different context lengths (1K >128K tokens). 2x tests were run for larger contexts for a larger sample size.

From story to adventure game

- Essay published by Steven Johnson in November 2024:
  - https://thelongcontext.com
- Steven identifies *"the ability to transform linear narratives into immersive adventures"* where the language model is acting as a game host



Let's play

Welcome, Detective Faurot, to the bustling streets of Soho, New York City, February 23rd, 1911. The year is young, but the city never sleeps, especially not for its criminal element. You are Joseph Faurot, a rising star in the New York Police Department, known for your relentless pursuit of modern forensic techniques, particularly the then-obscure science of fingerprinting.

Your Goal: You have been summoned to the scene of a burglary at a garment factory owned by the Bernstein brothers. Your task is to utilize your unique expertise to identify the perpetrator and bring him to justice. The fate of forensic science in America might just rest on your shoulders.

The Challenge: You must successfully identify the suspect in 10 actions or less. Each decision you make, each clue you pursue, will count as an action. Choose wisely, Detective.

Actions Remaining: 10

The Scene: You arrive at 171 Wooster Street with Officer James Fitzpatrick, a seasoned detective with a keen eye for detail. The Bernstein brothers' garment factory, a hive of activity by day, now stands eerily silent, the victim of a meticulous break-in. Officer Fitzpatrick briefs you on the scene: the thief entered through an abandoned third-floor loft, cut through the ceiling to access the factory

LLMs that can reason

OpenAI's o1 and o3 models represent significant advancements in AI reasoning capabilities:

- o1: September 2024, was designed to enhance reasoning abilities in AI. It achieved high score on various benchmarks:
  - https://openai.com/index/introducing-openai-o1-preview
- o3: Building upon o1, OpenAI introduced o3 in December 2024. This model showcased remarkable improvements, achieving a score of 87.5% on the ARC-AGI benchmark in high-compute mode, surpassing the human-level threshold of 85%.
  - https://en.wikipedia.org/wiki/OpenAI_o3

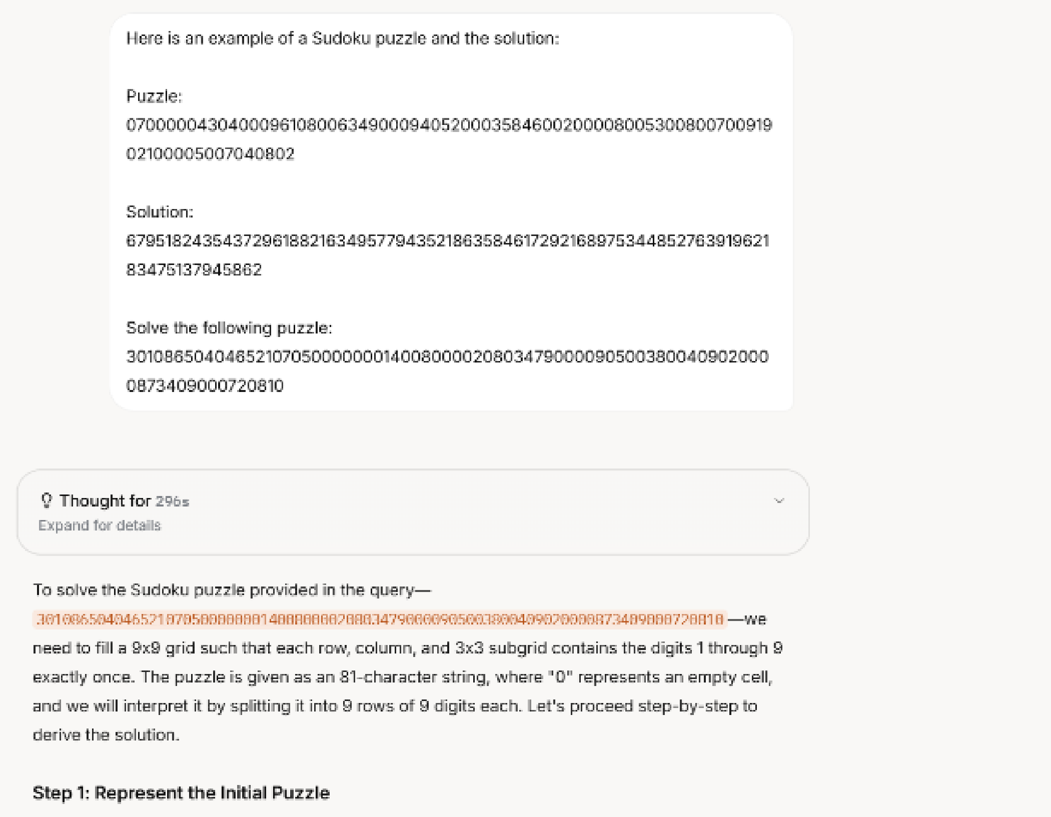
Grok-3 reasoning example: Sudoku

Prompt to Grok-3

- **`Here is an example of a Sudoku puzzle and the solution:`**
  `Puzzle:`
  `070000043040009610800634900094052000358460020000800530080070091902100005007040802`

  `Solution:`
  `679518243543729618821634957794352186358461729216897534485276391962183475137945862`

  `Solve the following puzzle:`
  `301086504046521070500000001400800002080347900009050038004090200008734090007208103`



Here is an example of a Sudoku puzzle and the solution:

Puzzle:
070000043040009610800634900094052000358460020000800530080070091902100005007040802

Solution:
679518243543729618821634957794352186358461729216897534485276391962183475137945862

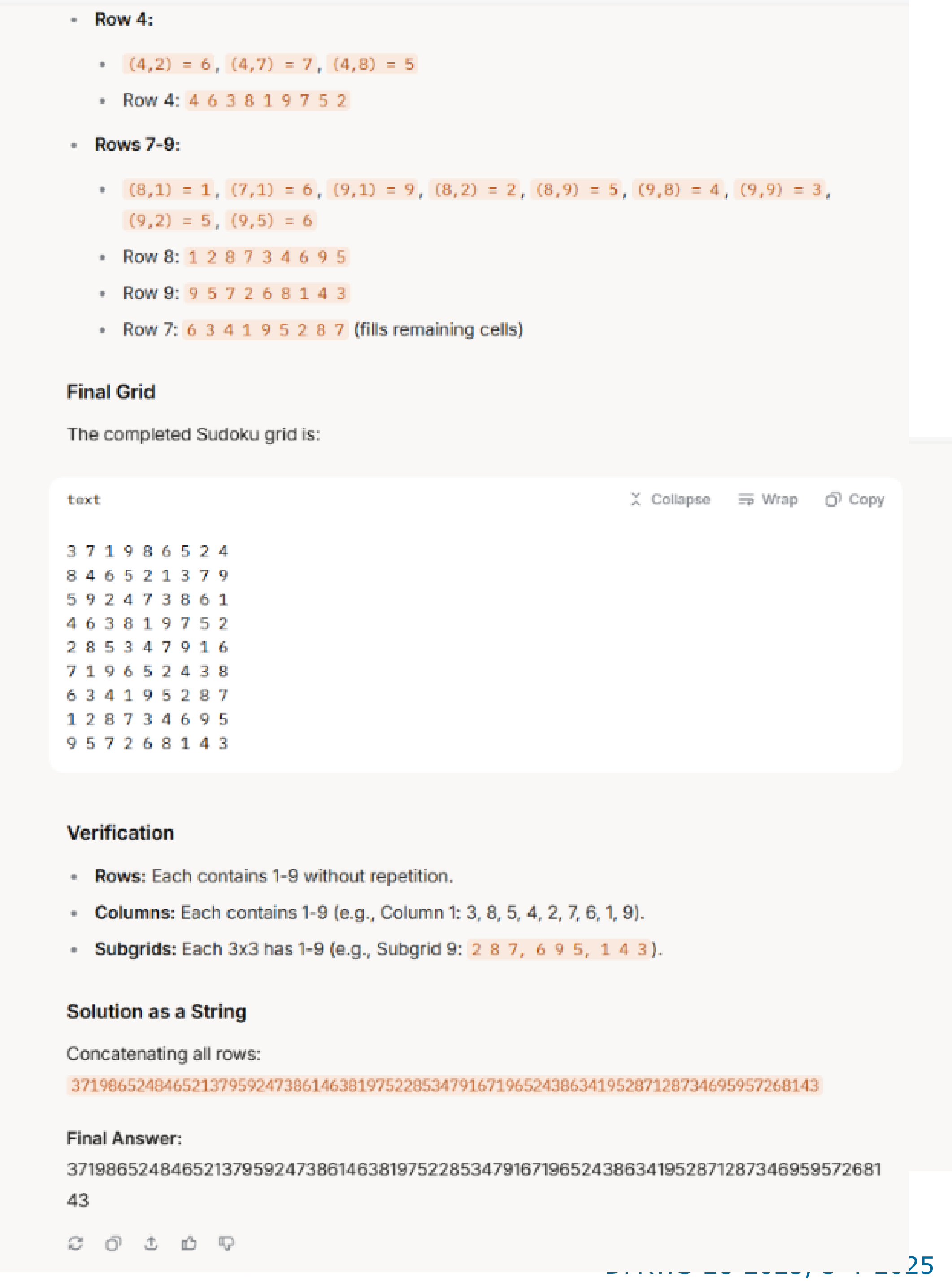
Solve the following puzzle:
301086504046521070500000001400800002080347900009050038004090200008734090000720810

To solve the Sudoku puzzle provided in the query—[illegible]—we need to fill a 9x9 grid such that each row, column, and 3x3 subgrid contains the digits 1 through 9 exactly once. The puzzle is given as an 81-character string, where "0" represents an empty cell, and we will interpret it by splitting it into 9 rows of 9 digits each. Let's proceed step-by-step to derive the solution.

Step 1: Represent the Initial Puzzle

Sudoku solution

- By Grok-3 in ≈ 5m
- The others can also do this:
  - DeepSeek-r1 (12m15s)
  - GPT-o1 (3m53s)
  - GPT-o3-mini (1m50s)
  - Gemini 2.0 Flash Thinking Experimental 01-21 (thinks for 4s writes a working python script in 6s)
  - Gemini 2.0 Pro Experimental 02-05 (in 6.5s without thinking)



- **Row 4:**
  - `(4,2) = 6`, `(4,7) = 7`, `(4,8) = 5`
  - Row 4: `4 6 3 8 1 9 7 5 2`
- **Rows 7-9:**
  - `(8,1) = 1`, `(7,1) = 6`, `(9,1) = 9`, `(8,2) = 2`, `(8,9) = 5`, `(9,8) = 4`, `(9,9) = 3`, `(9,2) = 5`, `(9,5) = 6`
  - Row 8: `1 2 8 7 3 4 6 9 5`
  - Row 9: `9 5 7 2 6 8 1 4 3`
  - Row 7: `6 3 4 1 9 5 2 8 7` (fills remaining cells)

Final Grid

The completed Sudoku grid is:

```text
3 7 1 9 8 6 5 2 4
8 4 6 5 2 1 3 7 9
5 9 2 4 7 3 8 6 1
4 6 3 8 1 9 7 5 2
2 8 5 3 4 7 9 1 6
7 1 9 6 5 2 4 3 8
6 3 4 1 9 5 2 8 7
1 2 8 7 3 4 6 9 5
9 5 7 2 6 8 1 4 3
```

Verification

- **Rows:** Each contains 1-9 without repetition.
- **Columns:** Each contains 1-9 (e.g., Column 1: 3, 8, 5, 4, 2, 7, 6, 1, 9).
- **Subgrids:** Each 3x3 has 1-9 (e.g., Subgrid 9: `2 8 7, 6 9 5, 1 4 3`).

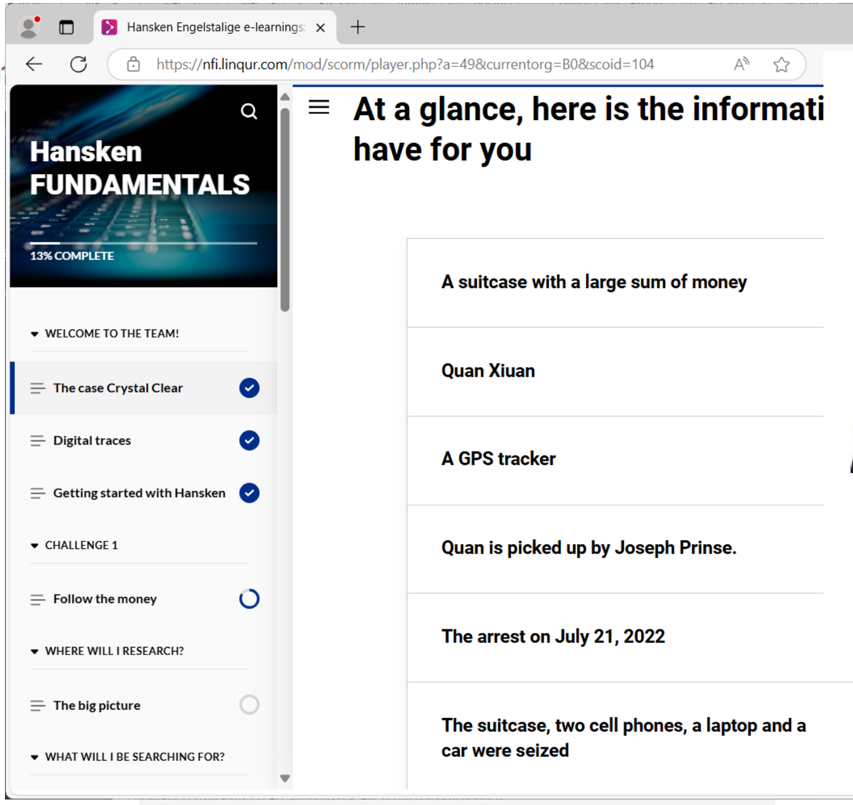
Solution as a String

Concatenating all rows:
`371986524846521379592473861463819752285347916719652438634195287128734695957268143`

Final Answer:
371986524846521379592473861463819752285347916719652438634195287128734695957268143

Changing roles: from game host to detective

ChatGPT o


Hansken Engelstalige e-learnings
https://nfi.linqur.com/mod/scorm/player.php?a=49&currentorg=B0&scoid=104

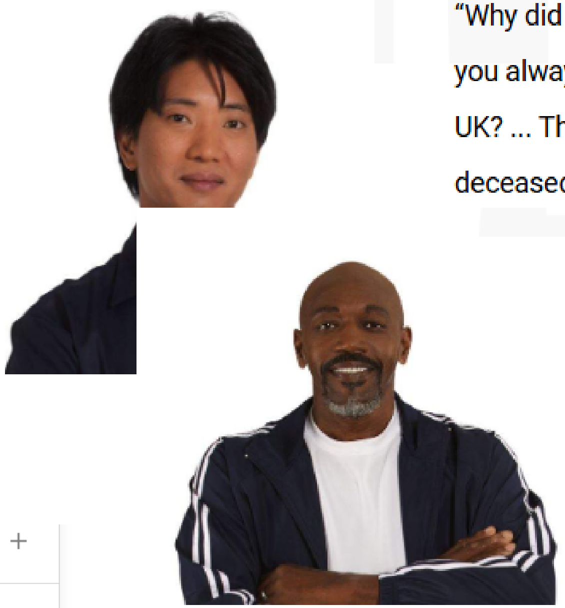
Hansken FUNDAMENTALS

13% COMPLETE

- WELCOME TO THE TEAM!
  - The case Crystal Clear ✓
  - Digital traces ✓
  - Getting started with Hansken ✓
- CHALLENGE 1
  - Follow the money
- WHERE WILL I RESEARCH?
  - The big picture ○
- WHAT WILL I BE SEARCHING FOR?

At a glance, here is the informati have for you

- A suitcase with a large sum of money
- Quan Xiuan
- A GPS tracker
- Quan is picked up by Joseph Prinse.
- The arrest on July 21, 2022 +
- The suitcase, two cell phones, a laptop and a car were seized +



"Why did you arrest me? I haven't done anything wrong. Do you always arrest people returning from a family visit to the UK? ... That is, indeed, my suitcase. A gift from my recently deceased grandfather. All I want to do with this money is

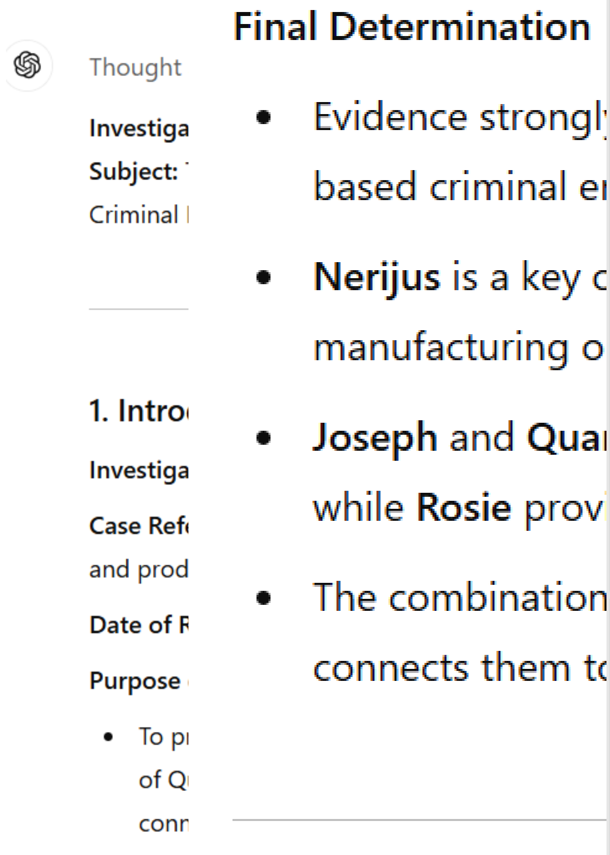
hat the mentioned
es were found.
t individuals, their

messages that
nd carefully crafted
nstrate how the

"I received a message asking if I could pick up someone at the airport. This is something I do on a daily basis. Prinse rental is run by me. My company provides car rentals. I have a long list of satisfied customers who recommend my services. We occasionally pick up people who have rented cars. I'm not sure why he (Quan) had so much money on him. He most likely won the lottery or something ..."

- Joseph Prinse

The report



Thought

Investiga

Subject: '

Criminal |

1. Intro

Investiga

Case Ref

and prod

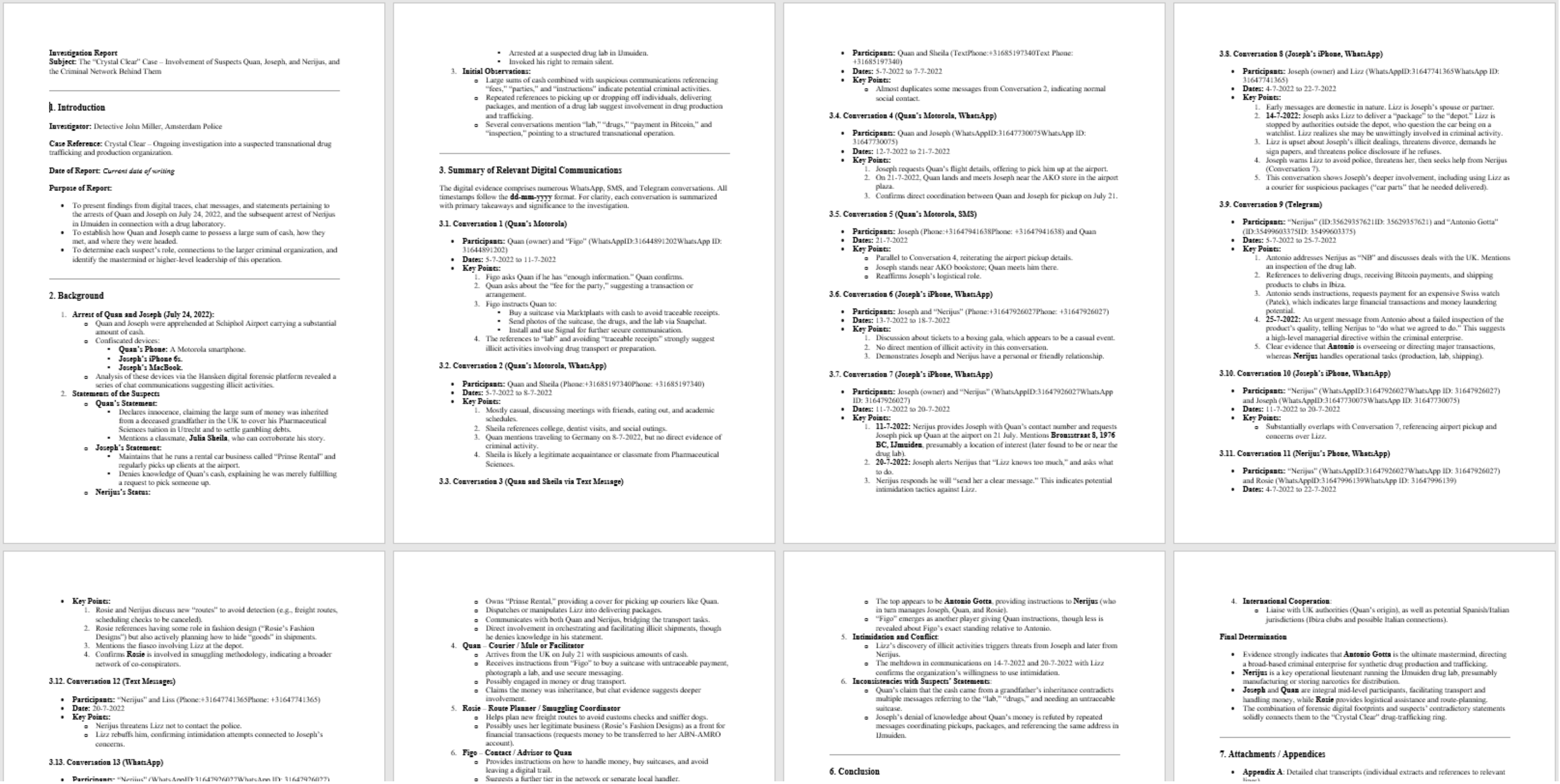
Date of R

Purpose

- To pr
  of Q
  conn

Final Determination

- Evidence strongl
  based criminal e
- **Nerijus** is a key c
  manufacturing o
- **Joseph** and **Quan**
  while **Rosie** prov
- The combination
  connects them to



Investigation Report
Subject: The "Crystal Clear" Case – Involvement of Suspects Quan, Joseph, and Nerijus, and the Criminal Network Behind Them

---

1. Introduction

Investigator: Detective John Miller, Amsterdam Police

Case Reference: Crystal Clear – Ongoing investigation into a suspected transnational drug trafficking and production organization.

Date of Report: *Current date of writing*

Purpose of Report:

- To present findings from digital traces, chat messages, and statements pertaining to the arrests of Quan and Joseph on July 24, 2022, and the subsequent arrest of Nerijus in IJmuiden in connection with a drug laboratory.
- To establish how Quan and Joseph came to possess a large sum of cash, how they met, and where they were headed.
- To determine each suspect's role, connections to the larger criminal organization, and identify the mastermind or higher-level leadership of this operation.

---

2. Background

1. **Arrest of Quan and Joseph (July 24, 2022):**
   - Quan and Joseph were apprehended at Schiphol Airport carrying a substantial amount of cash.
   - Confiscated devices:
     - **Quan's Phone:** A Motorola smartphone.
     - **Joseph's iPhone 6s.**
     - **Joseph's MacBook.**
   - Analysis of these devices via the Hansken digital forensic platform revealed a series of chat communications suggesting illicit activities.
2. **Statements of the Suspects**
   - **Quan's Statement:**
     - Declares innocence, claiming the large sum of money was inherited from a deceased grandfather in the UK to cover his Pharmaceutical Sciences tuition in Utrecht and to settle gambling debts.
     - Mentions a classmate, **Julia Sheila**, who can corroborate his story.
   - **Joseph's Statement:**
     - Maintains that he runs a rental car business called "Prinse Rental" and regularly picks up clients at the airport.
     - Denies knowledge of Quan's cash, explaining he was merely fulfilling a request to pick someone up.
   - **Nerijus's Status:**
     - Arrested at a suspected drug lab in IJmuiden.
     - Invoked his right to remain silent.
3. **Initial Observations:**
   - Large sums of cash combined with suspicious communications referencing "fees," "parties," and "instructions" indicate potential criminal activities.
   - Repeated references to picking up or dropping off individuals, delivering packages, and mention of a drug lab suggest involvement in drug production and trafficking.
   - Several conversations mention "lab," "drugs," "payment in Bitcoin," and "inspection," pointing to a structured transnational operation.

---

3. Summary of Relevant Digital Communications

The digital evidence comprises numerous WhatsApp, SMS, and Telegram conversations. All timestamps follow the **dd-mm-yyyy** format. For clarity, each conversation is summarized with primary takeaways and significance to the investigation.

3.1. Conversation 1 (Quan's Motorola)

- **Participants:** Quan (owner) and "Figo" (WhatsAppID:31644891202WhatsApp ID: 31644891202)
- **Dates:** 5-7-2022 to 11-7-2022
- **Key Points:**
  1. Figo asks Quan if he has "enough information." Quan confirms.
  2. Quan asks about the "fee for the party," suggesting a transaction or arrangement.
  3. Figo instructs Quan to:
     - Buy a suitcase via Marktplaats with cash to avoid traceable receipts.
     - Send photos of the suitcase, the drugs, and the lab via Snapchat.
     - Install and use Signal for further secure communication.
  4. The references to "lab" and avoiding "traceable receipts" strongly suggest illicit activities involving drug transport or preparation.

3.2. Conversation 2 (Quan's Motorola, WhatsApp)

- **Participants:** Quan and Sheila (Phone:+31685197340Phone: +31685197340)
- **Dates:** 5-7-2022 to 8-7-2022
- **Key Points:**
  1. Mostly casual, discussing meetings with friends, eating out, and academic schedules.
  2. Sheila references college, dentist visits, and social outings.
  3. Quan mentions traveling to Germany on 8-7-2022, but no direct evidence of criminal activity.
  4. Sheila is likely a legitimate acquaintance or classmate from Pharmaceutical Sciences.

3.3. Conversation 3 (Quan and Sheila via Text Message)

- **Participants:** Quan and Sheila (TextPhone:+31685197340Text Phone: +31685197340)
- **Dates:** 5-7-2022 to 7-7-2022
- **Key Points:**
  - Almost duplicates some messages from Conversation 2, indicating normal social contact.

3.4. Conversation 4 (Quan's Motorola, WhatsApp)

- **Participants:** Quan and Joseph (WhatsAppID:31647730075WhatsApp ID: 31647730075)
- **Dates:** 12-7-2022 to 21-7-2022
- **Key Points:**
  1. Joseph requests Quan's flight details, offering to pick him up at the airport.
  2. On 21-7-2022, Quan lands and meets Joseph near the AKO store in the airport plaza.
  3. Confirms direct coordination between Quan and Joseph for pickup on July 21.

3.5. Conversation 5 (Quan's Motorola, SMS)

- **Participants:** Joseph (Phone:+31647941638Phone: +31647941638) and Quan
- **Dates:** 21-7-2022
- **Key Points:**
  - Parallel to Conversation 4, reiterating the airport pickup details.
  - Joseph stands near AKO bookstore; Quan meets him there.
  - Reaffirms Joseph's logistical role.

3.6. Conversation 6 (Joseph's iPhone, WhatsApp)

- **Participants:** Joseph and "Nerijus" (Phone:+31647926027Phone: +31647926027)
- **Dates:** 13-7-2022 to 18-7-2022
- **Key Points:**
  1. Discussion about tickets to a boxing gala, which appears to be a casual event.
  2. No direct mention of illicit activity in this conversation.
  3. Demonstrates Joseph and Nerijus have a personal or friendly relationship.

3.7. Conversation 7 (Joseph's iPhone, WhatsApp)

- **Participants:** Joseph (owner) and "Nerijus" (WhatsAppID:31647926027WhatsApp ID: 31647926027)
- **Dates:** 11-7-2022 to 20-7-2022
- **Key Points:**
  1. **11-7-2022:** Nerijus provides Joseph with Quan's contact number and requests Joseph pick up Quan at the airport on 21 July. Mentions **Bronsstraat 8, 1976 BC, IJmuiden**, presumably a location of interest (later found to be or near the drug lab).
  2. **20-7-2022:** Joseph alerts Nerijus that "Lizz knows too much," and asks what to do.
  3. Nerijus responds he will "send her a clear message." This indicates potential intimidation tactics against Lizz.

3.8. Conversation 8 (Joseph's iPhone, WhatsApp)

- **Participants:** Joseph (owner) and Lizz (WhatsAppID:31647741365WhatsApp ID: 31647741365)
- **Dates:** 4-7-2022 to 22-7-2022
- **Key Points:**
  1. Early messages are domestic in nature. Lizz is Joseph's spouse or partner.
  2. **14-7-2022:** Joseph asks Lizz to deliver a "package" to the "depot." Lizz is stopped by authorities outside the depot, who question the car being on a watchlist. Lizz realizes she may be unwittingly involved in criminal activity.
  3. Lizz is upset about Joseph's illicit dealings, threatens divorce, demands he sign papers, and threatens police disclosure if he refuses.
  4. Joseph warns Lizz to avoid police, threatens her, then seeks help from Nerijus (Conversation 7).
  5. This conversation shows Joseph's deeper involvement, including using Lizz as a courier for suspicious packages ("car parts" that he needed delivered).

3.9. Conversation 9 (Telegram)

- **Participants:** "Nerijus" (ID:35629357621ID: 35629357621) and "Antonio Gotta" (ID:35499603375ID: 35499603375)
- **Dates:** 5-7-2022 to 25-7-2022
- **Key Points:**
  1. Antonio addresses Nerijus as "NB" and discusses deals with the UK. Mentions an inspection of the drug lab.
  2. References to delivering drugs, receiving Bitcoin payments, and shipping products to clubs in Ibiza.
  3. Antonio sends instructions, requests payment for an expensive Swiss watch (Patek), which indicates large financial transactions and money laundering potential.
  4. **25-7-2022:** An urgent message from Antonio about a failed inspection of the product's quality, telling Nerijus to "do what we agreed to do." This suggests a high-level managerial directive within the criminal enterprise.
  5. Clear evidence that **Antonio** is overseeing or directing major transactions, whereas **Nerijus** handles operational tasks (production, lab, shipping).

3.10. Conversation 10 (Joseph's iPhone, WhatsApp)

- **Participants:** "Nerijus" (WhatsAppID:31647926027WhatsApp ID: 31647926027) and Joseph (WhatsAppID:31647730075WhatsApp ID: 31647730075)
- **Dates:** 11-7-2022 to 20-7-2022
- **Key Points:**
  - Substantially overlaps with Conversation 7, referencing airport pickup and concerns over Lizz.

3.11. Conversation 11 (Nerijus's Phone, WhatsApp)

- **Participants:** "Nerijus" (WhatsAppID:31647926027WhatsApp ID: 31647926027) and Rosie (WhatsAppID:31647996139WhatsApp ID: 31647996139)
- **Dates:** 4-7-2022 to 22-7-2022
- **Key Points:**
  1. Rosie and Nerijus discuss new "routes" to avoid detection (e.g., freight routes, scheduling checks to be canceled).
  2. Rosie references having some role in fashion design ("Rosie's Fashion Designs") but also actively planning how to hide "goods" in shipments.
  3. Mentions the fiasco involving Lizz at the depot.
  4. Confirms **Rosie** is involved in smuggling methodology, indicating a broader network of co-conspirators.

3.12. Conversation 12 (Text Messages)

- **Participants:** "Nerijus" and Liss (Phone:+31647741365Phone: +31647741365)
- **Date:** 20-7-2022
- **Key Points:**
  - Nerijus threatens Lizz not to contact the police.
  - Lizz rebuffs him, confirming intimidation attempts connected to Joseph's concerns.

3.13. Conversation 13 (WhatsApp)

- **Participants:** "Nerijus" (WhatsAppID:31647926027WhatsApp ID: 31647926027)

  - Owns "Prinse Rental," providing a cover for picking up couriers like Quan.
  - Dispatches or manipulates Lizz into delivering packages.
  - Communicates with both Quan and Nerijus, bridging the transport tasks.
  - Direct involvement in orchestrating and facilitating illicit shipments, though he denies knowledge in his statement.
4. **Quan** – **Courier / Mule or Facilitator**
   - Arrives from the UK on July 21 with suspicious amounts of cash.
   - Receives instructions from "Figo" to buy a suitcase with untraceable payment, photograph a lab, and use secure messaging.
   - Possibly engaged in money or drug transport.
   - Claims the money was inheritance, but chat evidence suggests deeper involvement.
5. **Rosie** – **Route Planner / Smuggling Coordinator**
   - Helps plan new freight routes to avoid customs checks and sniffer dogs.
   - Possibly uses her legitimate business (Rosie's Fashion Designs) as a front for financial transactions (requests money to be transferred to her ABN-AMRO account).
6. **Figo** – **Contact / Advisor to Quan**
   - Provides instructions on how to handle money, buy suitcases, and avoid leaving a digital trail.
   - Suggests a further tier in the network or separate local handler.

  - The top appears to be **Antonio Gotta**, providing instructions to **Nerijus** (who in turn manages Joseph, Quan, and Rosie).
  - "Figo" emerges as another player giving Quan instructions, though less is revealed about Figo's exact standing relative to Antonio.
5. **Intimidation and Conflict**:
   - Lizz's discovery of illicit activities triggers threats from Joseph and later from Nerijus.
   - The meltdown in communications on 14-7-2022 and 20-7-2022 with Lizz confirms the organization's willingness to use intimidation.
6. **Inconsistencies with Suspects' Statements**:
   - Quan's claim that the cash came from a grandfather's inheritance contradicts multiple messages referring to the "lab," "drugs," and needing an untraceable suitcase.
   - Joseph's denial of knowledge about Quan's money is refuted by repeated messages coordinating pickups, packages, and referencing the same address in IJmuiden.

---

6. Conclusion

4. **International Cooperation**:
   - Liaise with UK authorities (Quan's origin), as well as potential Spanish/Italian jurisdictions (Ibiza clubs and possible Italian connections).

Final Determination

- Evidence strongly indicates that **Antonio Gotta** is the ultimate mastermind, directing a broad-based criminal enterprise for synthetic drug production and trafficking.
- **Nerijus** is a key operational lieutenant running the IJmuiden drug lab, presumably manufacturing or storing narcotics for distribution.
- **Joseph** and **Quan** are integral mid-level participants, facilitating transport and handling money, while **Rosie** provides logistical assistance and route-planning.
- The combination of forensic digital footprints and suspects' contradictory statements solidly connects them to the "Crystal Clear" drug-trafficking ring.

---

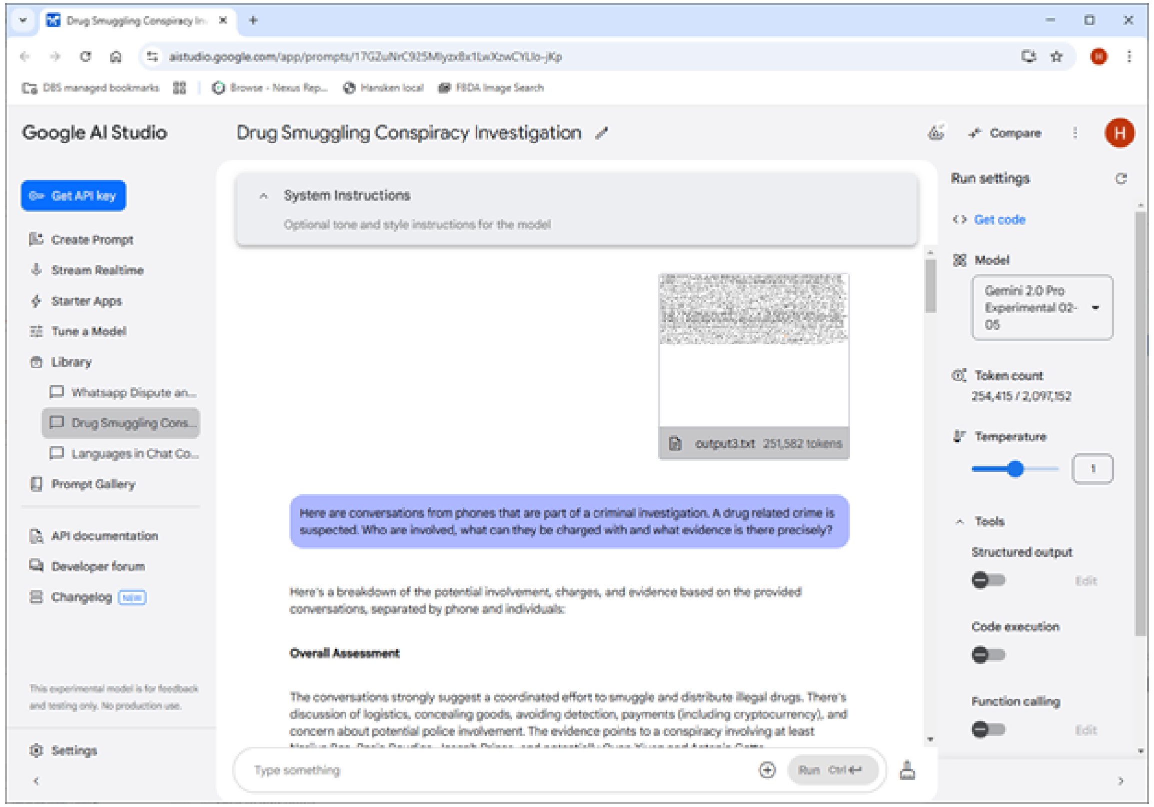
7. Attachments / Appendices

- **Appendix A**: Detailed chat transcripts (individual extracts and references to relevant lines)

Netherlands Forensic Institute
Ministry of Justice and Security

Gemini 2.0 Pro Experimental

- All conversations with max 3 participants (252K tokens!)
- **`Here are conversations from phones that are part of a criminal investigation. A drug related crime is suspected. Who are involved, what can they be charged with and what evidence is there precisely?`**
- Provides an overview



Drug Smuggling Conspiracy Investigation

output3.txt 251,582 tokens

Model: Gemini 2.0 Pro Experimental 02-05

Token count: 254,415 / 2,097,152

Temperature: 1

Here are conversations from phones that are part of a criminal investigation. A drug related crime is suspected. Who are involved, what can they be charged with and what evidence is there precisely?

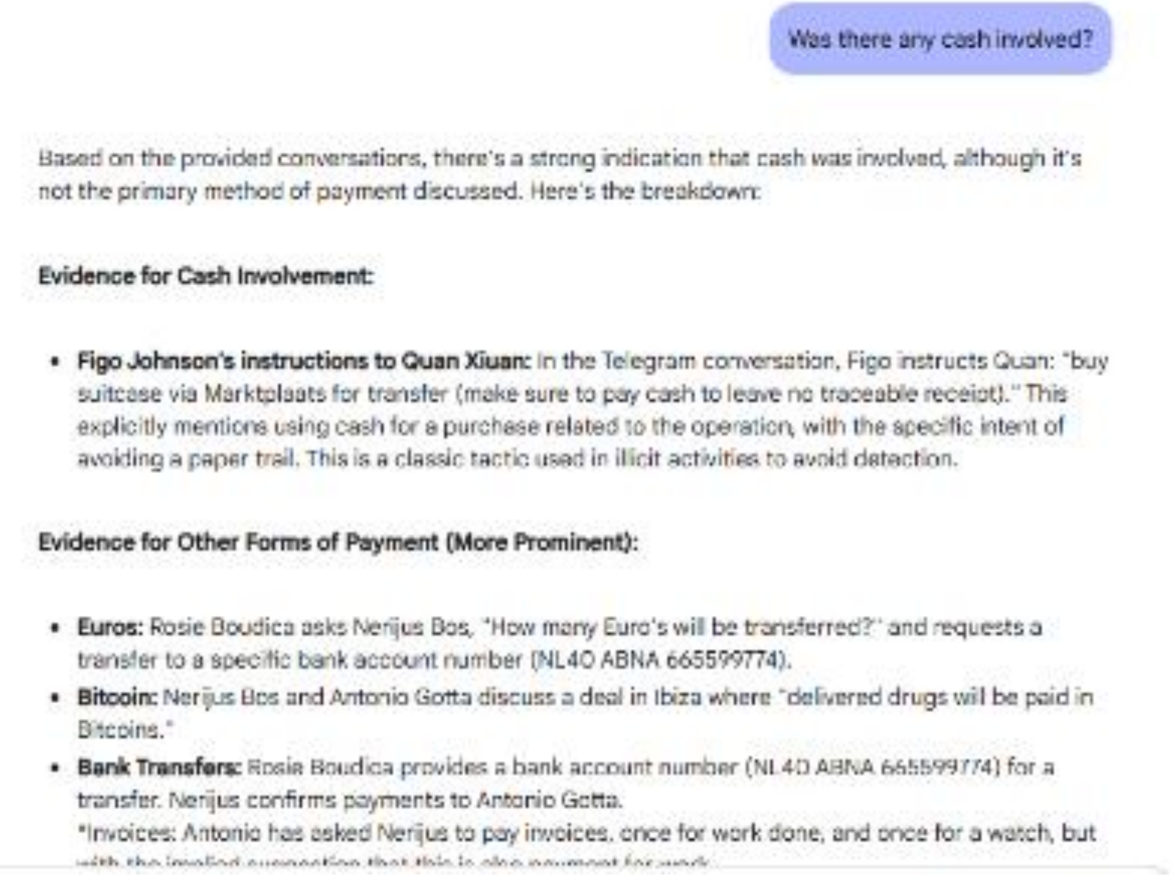
Here's a breakdown of the potential involvement, charges, and evidence based on the provided conversations, separated by phone and individuals:

Overall Assessment

The conversations strongly suggest a coordinated effort to smuggle and distribute illegal drugs. There's discussion of logistics, concealing goods, avoiding detection, payments (including cryptocurrency), and concern about potential police involvement. The evidence points to a conspiracy involving at least

Gemini 2.0 Pro Experimental

- All conversations with max 3 participants (252K tokens!)
- **`Here are conversations from phones that are part of a criminal investigation. A drug related crime is suspected. Who are involved, what can they be charged with and what evidence is there precisely?`**
- Provides an overview
- Ask follow-up questions!
  - Is there cash money involved



Was there any cash involved?

Based on the provided conversations, there's a strong indication that cash was involved, although it's not the primary method of payment discussed. Here's the breakdown:

Evidence for Cash Involvement:

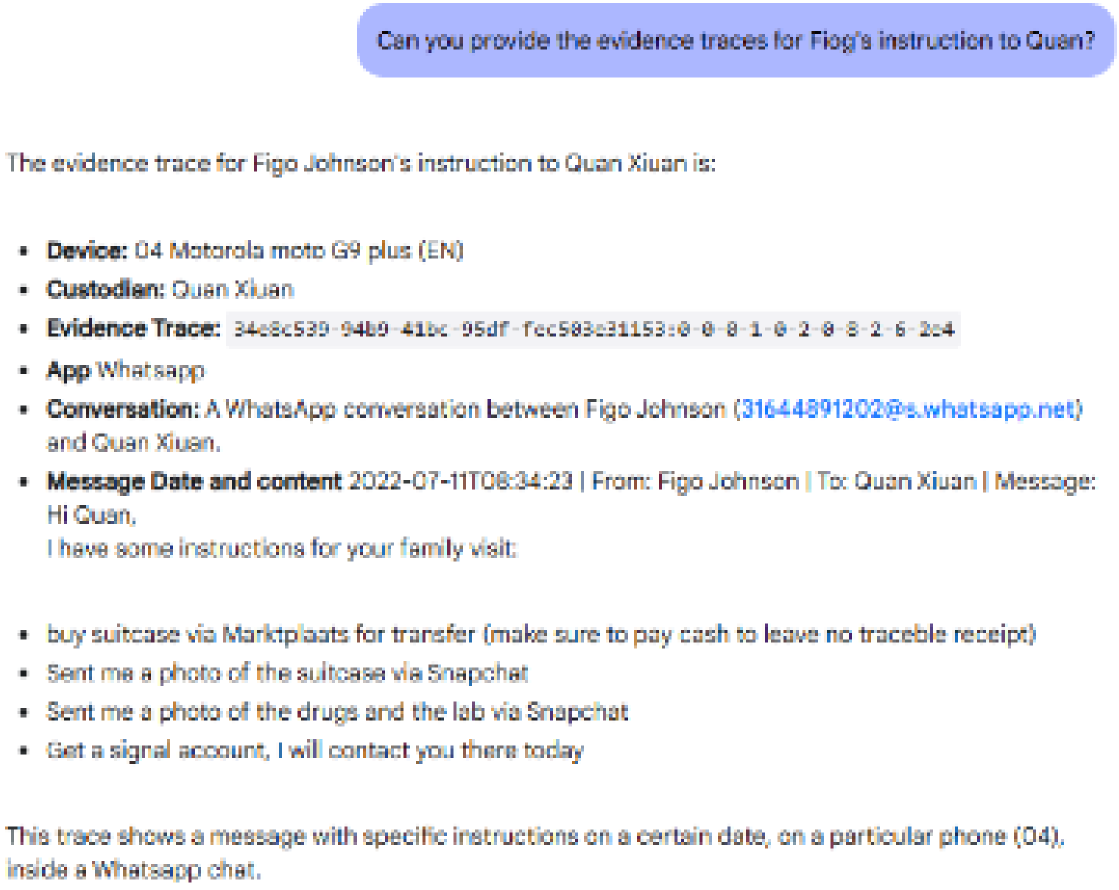
- **Figo Johnson's instructions to Quan Xiuan:** In the Telegram conversation, Figo instructs Quan: "buy suitcase via Marktplaats for transfer (make sure to pay cash to leave no traceable receipt)." This explicitly mentions using cash for a purchase related to the operation, with the specific intent of avoiding a paper trail. This is a classic tactic used in illicit activities to avoid detection.

Evidence for Other Forms of Payment (More Prominent):

- **Euros:** Rosie Boudica asks Nerijus Bos, "How many Euro's will be transferred?" and requests a transfer to a specific bank account number (NL40 ABNA 665599774).
- **Bitcoin:** Nerijus Bos and Antonio Gotta discuss a deal in Ibiza where "delivered drugs will be paid in Bitcoins."
- **Bank Transfers:** Rosie Boudica provides a bank account number (NL40 ABNA 665599774) for a transfer. Nerijus confirms payments to Antonio Gotta.
  *Invoices: Antonio has asked Nerijus to pay invoices, once for work done, and once for a watch, but with the implied suggestion that this is also payment for work.

Gemini 2.0 Pro Experimental

- All conversations with max 3 participants (252K tokens!)
- **`Here are conversations from phones that are part of a criminal investigation. A drug related crime is suspected. Who are involved, what can they be charged with and what evidence is there precisely?`**
- Provides an overview
- Ask follow-up questions!
  - Is there cash money involved
  - Ask for the trace id



Can you provide the evidence traces for Fiog's instruction to Quan?

The evidence trace for Figo Johnson's instruction to Quan Xiuan is:

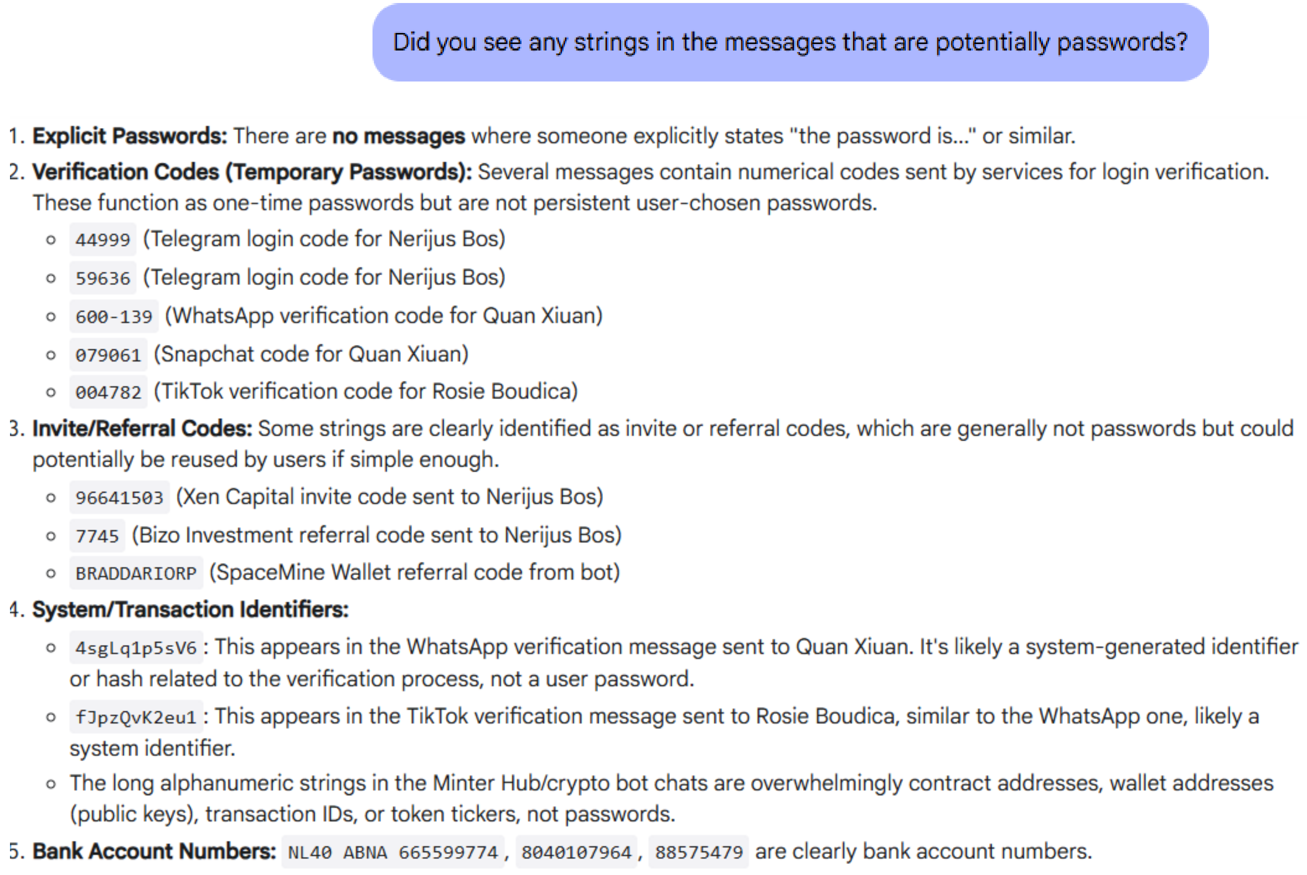
- **Device:** 04 Motorola moto G9 plus (EN)
- **Custodian:** Quan Xiuan
- **Evidence Trace:** `34e8c539-94b9-41bc-95df-fec583e31153:0-0-8-1-0-2-0-8-2-6-2e4`
- **App** Whatsapp
- **Conversation:** A WhatsApp conversation between Figo Johnson (31644891202@s.whatsapp.net) and Quan Xiuan.
- **Message Date and content** 2022-07-11T08:34:23 | From: Figo Johnson | To: Quan Xiuan | Message:
  Hi Quan,
  I have some instructions for your family visit:

- buy suitcase via Marktplaats for transfer (make sure to pay cash to leave no traceble receipt)
- Sent me a photo of the suitcase via Snapchat
- Sent me a photo of the drugs and the lab via Snapchat
- Get a signal account, I will contact you there today

This trace shows a message with specific instructions on a certain date, on a particular phone (04), inside a Whatsapp chat.

Gemini 2.0 Pro Experimental

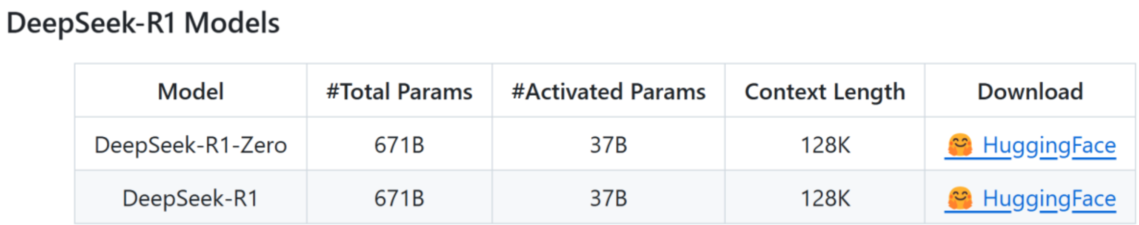
- All conversations with max 3 participants (252K tokens!)
- **Here are conversations from phones that are part of a criminal investigation. A drug related crime is suspected. Who are involved, what can they be charged with and what evidence is there precisely?**
- Provides an overview
- Ask follow-up questions!
  - Is there cash money involved
  - Ask for the trace id
  - Ask for passwords…



Did you see any strings in the messages that are potentially passwords?

1. **Explicit Passwords:** There are **no messages** where someone explicitly states "the password is..." or similar.
2. **Verification Codes (Temporary Passwords):** Several messages contain numerical codes sent by services for login verification. These function as one-time passwords but are not persistent user-chosen passwords.
   - `44999` (Telegram login code for Nerijus Bos)
   - `59636` (Telegram login code for Nerijus Bos)
   - `600-139` (WhatsApp verification code for Quan Xiuan)
   - `079061` (Snapchat code for Quan Xiuan)
   - `004782` (TikTok verification code for Rosie Boudica)
3. **Invite/Referral Codes:** Some strings are clearly identified as invite or referral codes, which are generally not passwords but could potentially be reused by users if simple enough.
   - `96641503` (Xen Capital invite code sent to Nerijus Bos)
   - `7745` (Bizo Investment referral code sent to Nerijus Bos)
   - `BRADDARIORP` (SpaceMine Wallet referral code from bot)
4. **System/Transaction Identifiers:**
   - `4sgLq1p5sV6`: This appears in the WhatsApp verification message sent to Quan Xiuan. It's likely a system-generated identifier or hash related to the verification process, not a user password.
   - `fJpzQvK2eu1`: This appears in the TikTok verification message sent to Rosie Boudica, similar to the WhatsApp one, likely a system identifier.
   - The long alphanumeric strings in the Minter Hub/crypto bot chats are overwhelmingly contract addresses, wallet addresses (public keys), transaction IDs, or token tickers, not passwords.
5. **Bank Account Numbers:** `NL40 ABNA 665599774`, `8040107964`, `88575479` are clearly bank account numbers.

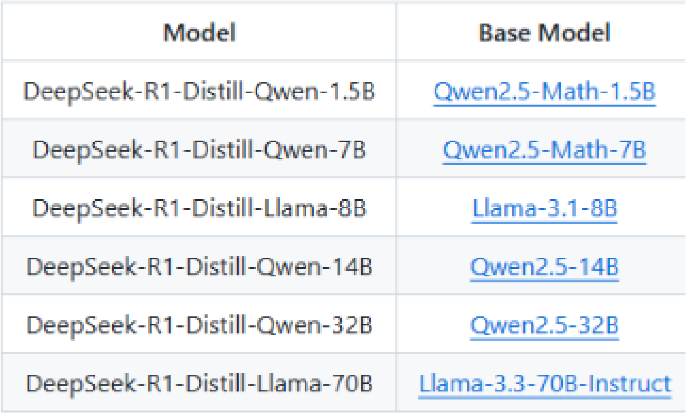
Can we do this with open source? DeepSeek R1



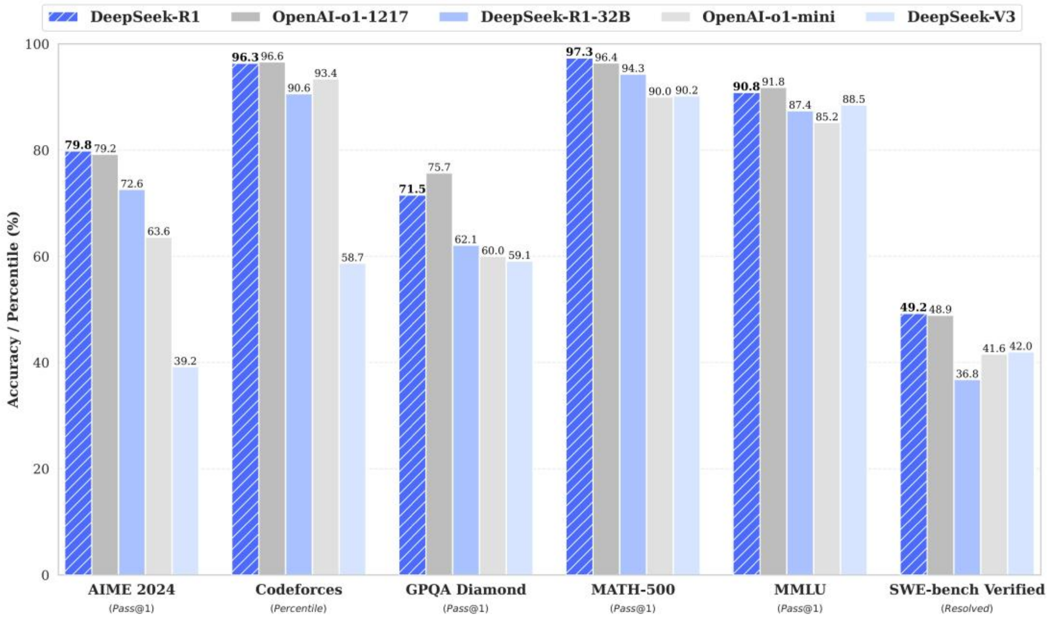
DeepSeek-R1 Models

Model	#Total Params	#Activated Params	Context Length	Download
DeepSeek-R1-Zero	671B	37B	128K	🤗 HuggingFace
DeepSeek-R1	671B	37B	128K	🤗 HuggingFace

- 671B is a bit too large to do this ourselves
- But DeepSeek has taught small open source models how to reason:



Model	Base Model
DeepSeek-R1-Distill-Qwen-1.5B	Qwen2.5-Math-1.5B
DeepSeek-R1-Distill-Qwen-7B	Qwen2.5-Math-7B
DeepSeek-R1-Distill-Llama-8B	Llama-3.1-8B
DeepSeek-R1-Distill-Qwen-14B	Qwen2.5-14B
DeepSeek-R1-Distill-Qwen-32B	Qwen2.5-32B
DeepSeek-R1-Distill-Llama-70B	Llama-3.3-70B-Instruct



https://github.com/deepseek-ai/DeepSeek-R1

Experiment: Entity Identification

- This is the prompt:

```
Based on the following information from the extracted details
and the background provided in Parts 1 and 2, identify and list
all individuals involved in the case. For each person, include:

- Their full name (and any known aliases),

- Their claimed role as provided in the statements,

- Relevant hints or evidence regarding their involvement (with
references such as conversation dates or markers).

Extracted Details:

{step1_output}

Background (Parts 1 and 2):

{part1}

{part2}

Please output the list in a structured, numbered format or as a
well-structured table with clear columns.
```

Using DeepSeek-R1-Distill-Llama-8B-Q8

```
Thinking Output:

Alright, so I'm trying to figure out all the individuals
involved in this case based on the information provided. Let
me start by going through each section step by step.

First, from the extracted details, there are nine names
listed: Quan, Joseph, Nerijus, Figo, Sheila, Antonio Gotta,
Rosie, Liss, and Lizz. I need to make sure I include all of
them with their roles and relevant evidence.

Starting with Quan. He's the owner of a Motorola phone. His
role is claimed as someone who recently returned from a
family visit in the UK, saying the cash is a gift from his
deceased grandfather for his studies and gambling debts. The
evidence includes WhatsApp chats where he talks about
meeting Sheila and Figo, and SMS messages confirming an
airport pickup on July 21, 2022.

Next, Joseph owns an iPhone 6s and a MacBook. His role is as
the owner of "Prinse Rental," a car rental business. He
claims to have been picking up someone from the airport but
denies knowledge about the cash. Evidence includes Telegram
chats with Nerijus about drug deals, SMS messages about the
pickup, and his car being spotted at a depot.
……
```

Identified entities

Comments:

- Response time 28s
- Liss and Lizz are the same persons but the LLM doesn’t get that
- If we summarize information between prompts the LLM may not see all the details

Name	**Aliases**	**Claimed Role**	**Relevant Evidence/References**
Quan	Motorola phone owner	Recently returned from a family visit, claims cash is a gift for studies and debts.	- WhatsApp chats with Sheila and Figo about meetings and payments. - SMS messages confirming airport pickup on July 21, 2022.
Joseph	iPhone 6s and MacBook owner	Car rental business ("Prinse Rental") owner, claims to have picked up someone from the airport.	- Telegram chats with Nerijus about drug deals. - SMS messages about the pickup. - Car spotted at a depot.
Nerijus	Arrested in IJmuiden	Key player coordinating drug deals and Bitcoin transfers.	- Telegram conversations with Antonio Gotta discussing drug deals and payments.
Figo	WhatsApp ID: 31644891202	Involved in transactions or meetings.	- WhatsApp chats with Quan about meeting again and payment for a party.
Sheila	WhatsApp ID: +31685197340	Regular meetups with Quan at McDonald's and Club Thouma.	- Quan mentions her as a potential witness to his claims.
Antonio Gotta	Telegram ID: 35499603375	Likely part of the criminal organization involved in drug deals.	- Telegram conversations with Nerijus about drug deals and Bitcoin payments.
Rosie	WhatsApp ID: 31647741365	Involved in logistics or delivery discussions.	- WhatsApp chats with Nerijus about new routes for transporting goods.
Liss	Hospital worker	Spotted with Joseph's car at a depot, leading to police scrutiny.	- Police visited her home regarding Joseph's car.
Lizz	Mentioned in conversation	Potential co-conspirator who knows too much about the operation.	- Joseph expresses concern about her knowledge.

What’s next?

New open source models are arriving

New small open source LLMs

- Microsoft Phi-4-mini (3.8B) and Phi-4-multimodal (5.6B) with 128K context with 24 languages
- Google Gemma (1/4/12/27B) with 128K context and also multimodal and 140+ languages.
- We can expect Llama 4 to be released this year (probably 1H2025)
- Use cases:
  - Translation
  - Besides text also image recognition, OCR, speech-to-text
  - Useful for chat conversations with pictures,videos and voice messages

Thank you!

Netherlands Forensic Institute
Ministry of Justice and Security

Dr. Hans Henseler
h.henseler@nfi.nl
www.linkedin.com/in/henseler

Related abstracts, papers and articles:

Deep Reasoning and Large Context Windows:
Next-Generation AI in Digital Forensic Investigations

Hans Henseler[a,b,*], Timo Meconi[b]

[a] *University of Applied Sciences Leiden, The Netherlands*
[b] *Netherlands Forensic Institute, The Hague, The Netherlands*

Abstract

Recent advances in large language models (LLMs) allow the processing of entire case dossiers "in one breath" while employing deeper, step-by-step reasoning, heralding a new era in digital forensics. Where existing retrieval-augmented approaches can miss crucial clues by splitting long documents into fragments, larger context windows enable a more holistic view of all text, significantly reducing the risk of overlooking pivotal evidence. Coupled with explicit multi-step reasoning, these models can propose hypotheses, link scattered evidence, and illuminate investigative scenarios otherwise hidden from manual review.

This short paper reports on experiences gained from a proof-of-concept experiment involving the so-called "Crystal Clear-case," a fictitious yet realistic digital forensics scenario. We also draw parallels to Steven Johnson's "Long Context" experiment, illustrating how entire books or voluminous dossiers can be read and reasoned about cohesively. We discuss the potential for AI to automate tasks such as privilege filtering and data relevance assessments, thus helping protect fundamental rights. Finally, we present the Hansken Copilot Prototype—currently under development—a local LLM-based assistant designed to streamline investigative workflows in real forensic environments.

This document is intended as a long abstract for a presentation at DFRWS EU 2025 in which we will elaborate on the technical and procedural aspects of harnessing extended context windows, stepwise AI reasoning, and local deployments in forensic labs.

Keywords: Digital forensics, Large Language Models, Extended context windows, Stepwise reasoning, Evidence analysis, Hansken Copilot, Forensic AI

1. Introduction

Digital forensic practitioners routinel
varied collections of electronic data—for
backups, chat logs, or emails. As the E
phasizes in its e-evidence publication, "
is redefining every aspect of society, and
(European Council, 2024). In many inv
lenge is akin to searching for a "needle
based on Retrieval-Augmented Generati
mitigate this by splitting documents into
cross-referenced clues may be missed if
do not appear together (Henseler, 2023).

Emerging large language models (LL
text windows promise a new solution: f
"in one breath," thus reducing reliance o
methods. In tandem, explicit multi-step
referred to as "slow reasoning") provides
tail, enabling investigators to trace the l
ences (OpenAI, 2024b; DeepMind, 2024

---

[*] Corresponding author.
Email addresses: `henseler.h@hsleiden.nl`
`timo.meconi@nfi.nl` (Timo Meconi)

Preprint submitted to Presentation proposal submit

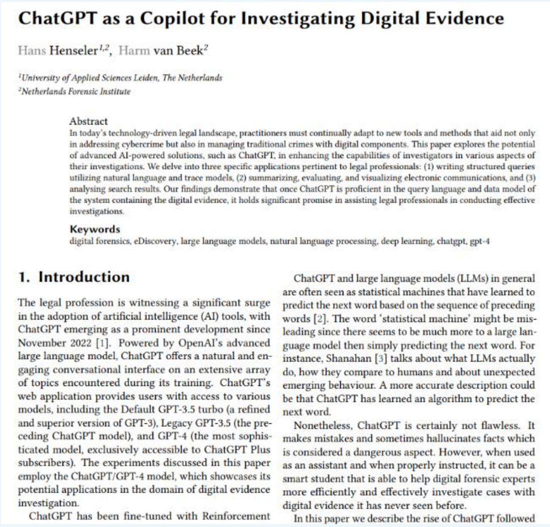
eForensics magazine & courses
LOGIN / SIGN UP

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....

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

ChatGPT: A Digital

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