



Digitaler Spürhund

**11. Internationales Symposium
Neue Technologien**

**Michael Fila, BMI, Wien
Katja Kraus, K-Analytica, Klagenfurt**

6. November 2024

Michael Fila



- Seit 2024 Kriminal-Analytiker am Bundesamt zur Korruptionsprävention und -bekämpfung
- > 20 Jahre Erfahrung als Kriminalanalytiker im Bundeskriminalamt Österreich
- > 10 Jahre Durchführung von verdeckten Ermittlungen, (technischen) Observationen
- Tätigkeitsfelder: Operative Kriminalanalyse, Textanalyse, grafische Auswertungen , ...

Überblick



Large Language Models - LLMs

06. November 2024 Digitaler Spürhund Folie 4 von 24

Digitaler Spürhund - Methodik

06. November 2024 Digitaler Spürhund Folie 10 von 24

Digitaler Spürhund - Kernfragen

06. November 2024 Digitaler Spürhund Folie 13 von 24

Fazit

06. November 2024 Digitaler Spürhund Folie 21 von 24



Large Language Models - LLMs

LLMs - Allgemein



- Basieren auf Verfahren der Artificial Intelligence (AI) bzw. Künstliche Intelligenz (KI)
- Nutzen ANNs (Artificial Neural Networks)
- 1970 - erste Ideen
- Praktische Umsetzung durch die rasante Entwicklung der Hardware (Firma Nvidia)
- Konzept aus dem Bereich der Generativen AI (Generierung von Text, Bild, Video, Audio, Musik, Design und Kunst)

LLMs - Aktuelle Entwicklungen



- OpenAI: DALL-E 3 (Oktober 2023), ChatGPT (November 2022), Sora (in Arbeit)
- Meta (Facebook): LLaMa 3.2 (Large Language Model Meta AI)
- Amazon: Alexa 2.0
- Kostenlose Werkzeuge und Anbieter: GPT4All, OpenGPT-X, Google Gemini, Anthropic Claude, DeutschlandGPT; Mistral, Google, Samsung
- Anwendungsbereiche: Medizin, Datenanalyse, Statistik, Juristische Textverarbeitung, Chemie...

LLMs - Grenzen



- Statistische Verfahren im Kern
- Im Regelfall kein Zugriff auf aktuelle Ereignisse und Echtzeitinformationen
- Kein Beweis für die Richtigkeit und Qualität der Ergebnisse
- Können moralische und ethische Aspekte nicht berücksichtigen
- Falschaussagen und Halluzinieren

LLMs - Grenzen



ChatGPT 4o >



welche zahl ist größer, 5.9 oder 5.11



Die Zahl 5.11 ist größer als 5.9. Das liegt daran, dass die Dezimalstellen nach dem Punkt bei 5.11 (also 11) größer sind als bei 5.9 (also 9).

und bei 5.90 und 5.11



In diesem Fall ist die Zahl 5.90 größer als 5.11. Auch wenn die Nullen hinter der 9 in 5.90 keine Auswirkung auf den Wert haben und es effektiv 5.9 ist, bleibt sie dennoch größer als 5.11.

LLMs - Optimieren



Warum die vorhandenen Daten nicht einfach in ein LLM einspielen?

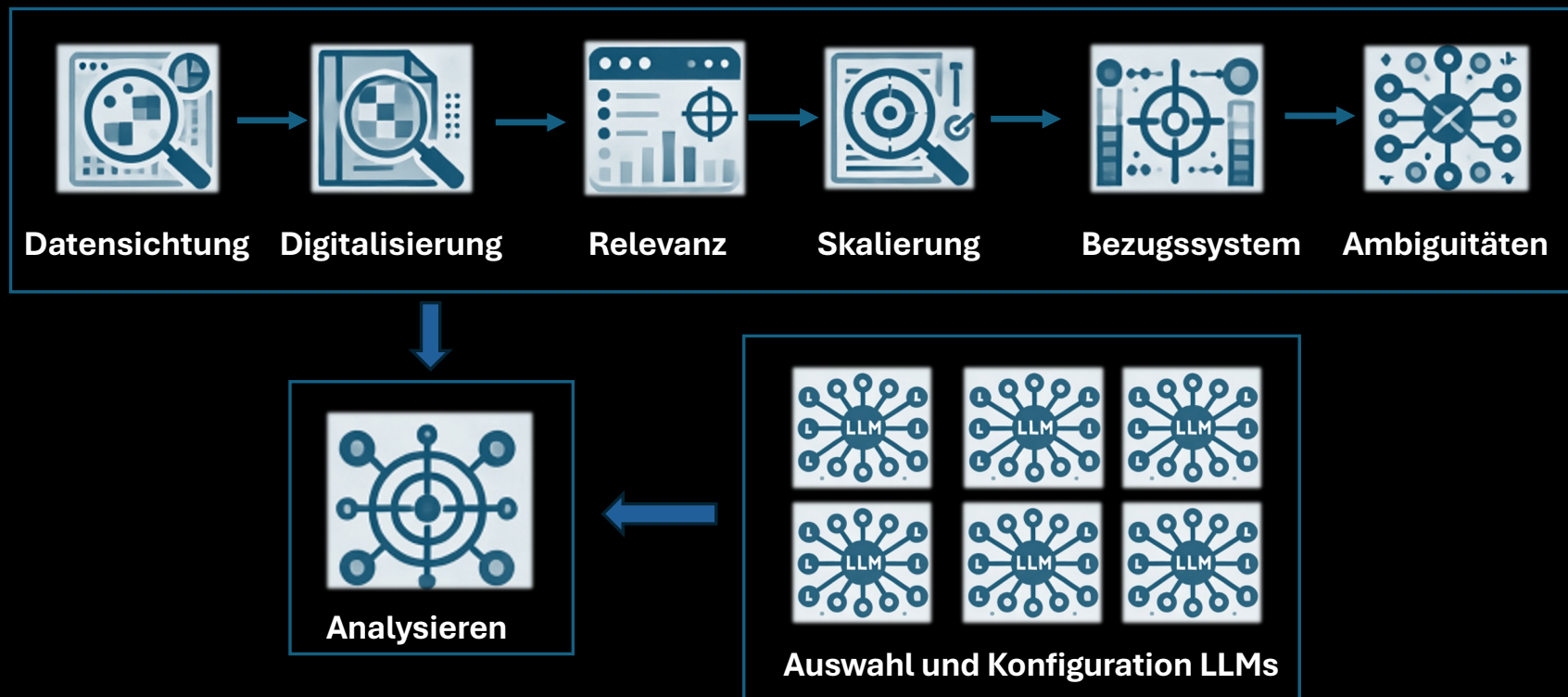
Qualität der Ergebnisse nimmt drastisch ab – Falschaussagen und Halluzinieren nehmen zu

Digitaler Spürhund basiert auf der Verbesserung der Datenqualität
durch optimierte LLMs und Datenverarbeitungsprozesse



Digitaler Spürhund - Methodik

Digitaler Spürhund - Methodik





Digitaler Spürhund - Kernfragen

Digitaler Spürhund – Kernfragen 1/5



Wie wird der rasante technische Fortschritt berücksichtigt?

- Modularer Aufbau
- Verwendung der aktuellsten LLMs (z.B. aktuelle Sprach- oder Bilderkennungs-LLMs)
- Kontinuierliche Integration neuer LLMs – keine Auswirkung auf die Anwendung

Digitaler Spürhund – Kernfragen 2/5



Wie werden das Halluzinieren-Problem bzw. Falschaussagen behandelt?

- Aufzeigen von Zusammenhängen aus unüberschaubaren Datenpool
- Aufbereitung qualitativ hochwertiger Daten
- **Hinweisgeber** - kein eigenständiger Solver



Hier ist die Analyse des Fotos (Indiz 004):

Fotoanalyse: Indiz 004

Beschreibung des Fotos:

- Das Foto zeigt einen Küchenbereich mit einem Herd.
- Auf der Arbeitsfläche neben dem Herd befinden sich:
 - Eine Schüssel mit braunen Stücken (vermutlich Zutaten).
 - Ein Karton mit Eiern.
 - Eine Packung Mehl.
 - Eine Packung Zucker.
 - Eine leere Schüssel.
- Die Herdplatten sind sauber, aber es gibt Anzeichen dafür, dass sie kürzlich benutzt wurden (leichte Verfärbungen).
- An der Wand über der Arbeitsfläche befinden sich zwei Steckdosen und zwei Schalter, von denen einer eingeschaltet ist (Licht an).

Relevanz für den Fall:

1. Zustand des Tatorts:

- Das Foto zeigt einen ordentlichen Küchenbereich, was im Gegensatz zur Beschreibung des chaotischen Tatorts steht. Dies deutet darauf hin, dass der Tatort möglicherweise manipuliert wurde oder dass das Chaos nur in bestimmten Bereichen vorhanden war.
- Die Vorbereitungen für das Backen könnten darauf hinweisen, dass Harold Green kurz vor seiner Ermordung etwas backen wollte oder dass jemand anderes diese Vorbereitungen getroffen hat.

2. Zeitliche Hinweise:

- Der eingeschaltete Schalter und die sauberen Herdplatten könnten darauf hindeuten, dass der Herd kürzlich benutzt wurde. Dies könnte die Zeit des Verbrechens weiter eingrenzen.

3. Fingerabdrücke und DNA:

- Die Gegenstände auf der Arbeitsfläche könnten Fingerabdrücke oder DNA-Spuren von der Person enthalten, die sich zuletzt in der Küche aufgehalten hat. Dies könnte helfen, die Anwesenheit von Personen zu bestätigen.

Mögliche Hypothesen und nächste Schritte:

1. Ermittlung der letzten Aktivitäten des Opfers:

- Feststellen, ob Harold Green tatsächlich die Person war, die die Backzutaten vorbereitet hat, oder ob jemand anderes diese Aktivitäten durchgeführt hat.

2. Fingerabdruckanalyse:

- Untersuchung der Gegenstände auf der Arbeitsfläche auf Fingerabdrücke und Abgleich mit den bereits genommenen Fingerabdrücken (Harold Green, Barbara Fuller, James Foster und Jason Green).

3. Befragung zu Backaktivitäten:

- Befragung von Personen, die Harold Green gut kannten, um herauszufinden, ob er regelmäßig backte oder ob dies ungewöhnlich war. Dies könnte Hinweise darauf geben, ob jemand anderes in der Küche war.

4. Überprüfung der Schalter und Steckdosen:

- Überprüfen, ob der eingeschaltete Schalter mit den Tatbeschreibungen übereinstimmt und ob dies möglicherweise eine Spur hinterlassen hat.

Falls weitere spezifische Analysen oder Untersuchungen erforderlich sind, lassen Sie es mich wissen.

Digitaler Spürhund – Kernfragen 3/5



Wie kann die Nachvollziehbarkeit der Ergebnisse gewährleistet werden (Gerichtsverwertbarkeit)?

- Protokollierung von jedem einzelnen Schritt - Transparenz
- Digitales Signieren bei Bedarf
- Berücksichtigung der Datenschutzaspekte
- Berücksichtigung des AI Act der EU

Digitaler Spürhund – Kernfragen 4/5



Welche Daten können verarbeitet bzw. zusammengeführt werden?

- Text: Textanalyse, Übersetzung, Textgenerierung, Zusammenfassung, Frage-Antwort-System, Übersetzung
- Bild: Bilderkennung, -beschreibung, Bildanalyse (z.B. für medizinische Zwecke), Text-Erkennung (OCR)
- Audio: In Entwicklung: Spracherkennung, Transkription, Audibearbeitung und -analyse, Dialogsystem
- Video: In Entwicklung: Videozusammenfassungen, Erkennung von Szenen und Ereignissen

Digitaler Spürhund – Kernfragen 5/5



Für welche Ermittlungen kann die Methode angewandt werden?

Generell für jede Art von Ermittlungen, beispielsweise:

- Offene Ermittlungen
- Gerichtsanhängige Fälle
- Cold Cases



Fazit

Fazit – LLMs Allgemein



LLMs – Stärken:

Daten- und Informationsanalyse, Bildanalyse, Audio-Transkription, Zusammenfassungen, Übersetzungen

LLMs – Herausforderungen und Grenzen:

Qualität der Ergebnisse (Falschaussagen, Halluzinieren), Datenschutz, Präzision und Verständnis (echter Kontext fehlt)

LLMs – Forschungsfelder:

Integration von Text, Bild, Video und Audio, Verbesserung des Verständnisses und des Kontexts, Ethik und Regulierung

Fazit – Digitaler Spürhund



Potenzial für die Ermittlungsarbeit:

Analyse großer Datenmengen, Steigerung der Effizienz und Qualität der Analyse

Innovativer methodischer Ansatz:

Informationszusammenfassung aus verschiedenen Quellen und Verknüpfung, Hinweis auf Widersprüche oder fehlende Informationen, modularer Aufbau erlaubt Integration neuester Technologien

Forschung und Weiterentwicklung:

Anpassung an behördliche Bedürfnisse, maßgeschneiderte Weiterentwicklung für spezifische Anforderungen einzelner Ermittlungsbehörden



Digitaler Spürhund



Michael Fila
BMI

+43 664 4020107
michael.fila@bak.gv.at

Dipl.-Ing. Katja Kraus
K-Analytica

+43 677 636 879 62
katja.kraus@k-analytica.com