
# KomMKonLLM

Zwischenbericht | Call 19 | Projekt ID 7409

# Inhalt



# 1 Einleitung

Mit der zunehmenden Integration von Large Language Models (LLMs) in Produktentwicklung und Dienstleistungen wird deren Konsistenz und Verlässlichkeit ein zentraler Erfolgsfaktor. Gerade in unternehmenskritischen Anwendungen müssen LLMs zuverlässig auf ähnliche Eingaben konsistente Antworten liefern, auch wenn sich die exakte Formulierung ändert. Das Projekt **KomMKonLLM** widmet sich dieser Herausforderung, indem es systematische Methoden für das Konsistenztesten von LLMs betrachtet. In **KomMKonLLM** wenden wir kombinatorische Testmethoden aus dem Softwaretesten an, um Konsistenztests für LLMs zu generieren.
Das Projektziel ist der Aufbau einer automatisierten Plattform, die mithilfe kombinatorischer Testmethoden LLMs auf Konsistenz testet. Zu den Kernfunktionen zählen die Erstellung, Durchführung und Auswertung von Tests sowie die Bereitstellung von Ergebnissen über eine benutzerfreundliche Schnittstelle. Zielgruppen sind Wissenschaft, Forschung sowie Unternehmen, die LLMs sicher und effizient einsetzen möchten.

In diesem Zwischenbericht präsentieren wir Fortschritte in den Arbeitspaketen: den Abschluss der ersten zwei Arbeitspakete (AP1 and AP2), der methodischen Entwicklung (AP3) und die ersten Schritte bezüglich der Implementierung der technischen Komponenten (AP4).

# 2 Status der Arbeitspakete

## 2.1 Arbeitspaket 1 - Detailplanung und Formales am Projektstart

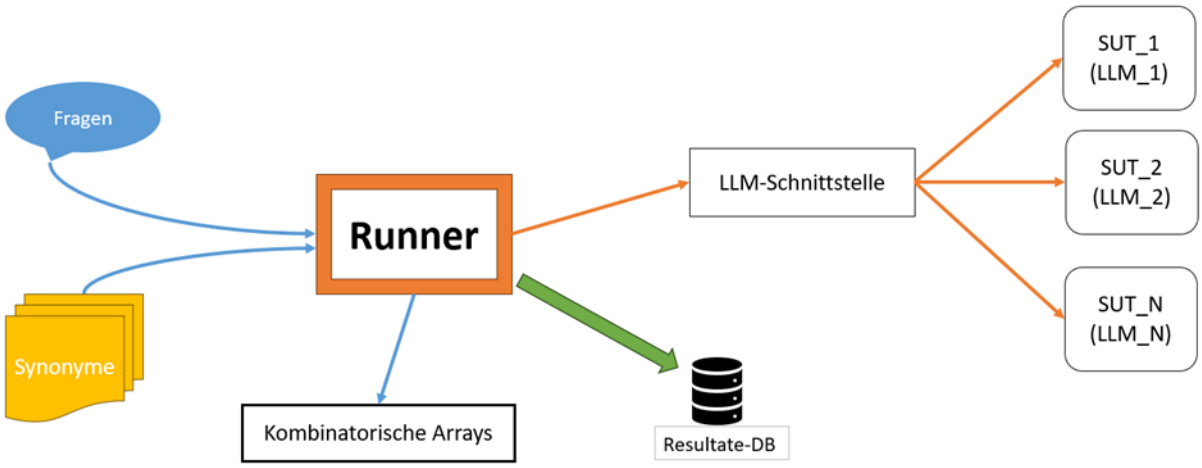
Das **erste Arbeitspaket** wurde im Dezember 2024 erfolgreich abgeschlossen. Insbesondere sind folgende Ergebnisse entstanden:

- Beidseitig unterschriebener Fördervertrag;
- Detailprojektplan (inklusive Plan der einzelnen Arbeitspakete) erstellt und abgenommen;
- Liste der Projektergebnisse mit zugehöriger Lizenz erstellt;
- Projektwebseite von **KomMKonLLM** innerhalb von Netidee in Betrieb genommen und ersten Blogeintrag erstellt;
- Beantragung von erster Förderrate und deren Genehmigung nach Feedback.

## 2.2 Arbeitspaket 2 - Technische Planung

Das **zweite Arbeitspaket** beschäftigte sich mit einer Anforderungsanalyse von **KomMKonLLM**, der technischen Gesamtarchitektur und der Auswahl technischer Komponenten (wie – unter anderem – Frameworks, Programmiersprachen und Schnittstellen) und wurde Anfang Jänner 2025 abgeschlossen. Die erhaltenen Erkenntnisse haben wir in unserem zweiten Blogpost zugänglich gemacht, den wir im Folgenden kurz zusammenfassen.

Einen Überblick über die technische Architektur zeigt das folgende Diagramm:




Eine detaillierte Beschreibung der Aufgaben bzw. Anforderungen an diese Komponenten ist online zu finden.

Weiters wählten wir zu verwendende Technologien in **KomMKonLLM** unter Berücksichtigung folgender Kriterien aus:

- **Popularität bei Entwicklern**: Um die Verbreitung und Weiterentwicklung von **KomMKonLLM** zu vereinfachen, sollten Technologien weit genutzt werden.
- **Aktive Entwicklung**: Um das Risiko von Softwareobsoleszenz von verwendeten Komponenten zu mitigieren, sollten eingebundene Technologien nicht nur zum jetzigen Zeitpunkt unter aktiver Entwicklung sein, sondern es sollte auch eine aktive Weiterentwicklung in der Zukunft anzunehmen sein. Für diese Entscheidungen werden wir auch die Einschätzungen von CrOSSD[1] verwenden.
- **Lizenzkompatibilität**: Ausgewählte Technologien dürfen keine Einschränkungen auf die Lizenzierung von **KomMKonLLM**[2] haben.

Die anhand dieser Kriterien ausgewählten Technologien und Bibliotheken sind online zusammengefasst.

## 2.3 Arbeitspaket 3 - Methodische Entwicklung LLM Testing Framework

Das dritte Arbeitspaket wurde im Jänner 2025 weitgehend fertiggestellt und beschäftigt sich mit der Methodik zur Erzeugung kombinatorischer Konsistenztests für LLMs. Wir haben unseren gewählten methodischen Ansatz von theoretischer Seite her in zwei Blog-Posts dargelegt (Teil 1 und Teil 2) und in einem Weiteren ein konkretes Beispiel durchgeführt. Nachfolgend geben wir eine Zusammenfassung.

Die in **KomMKonLLM** implementierte Methode zur Erzeugung von Konsistenztests für LLMs kann auf jede binäre Entscheidungsfrage (vereinfacht als Ja/Nein- bzw richtig/falsch-Frage beschreibbar) in natürlicher Sprache angewendet werden. Wir verwenden Wortersetzungen durch Synonyme und kombinatorische Methoden, um strukturiert und in (kombinatorisch-) messbarer Diversität abgewandelte Fragen zur Konsistenzevaluierung zu erstellen. Ausgehend von einer gegebenen binären Entscheidungsfrage, gliedert sich der Gesamtprozess zur Erzeugung von kombinatorischen Konsistenztests in folgende Schritte:


- **„Linguistische Satzanalyse“**: In diesem Schritt wenden wir NLP-Techniken an, um aus der gegebenen binären Entscheidungsfrage ein diskretisiertes Satzmodell zu erzeugen.
- „**Kombinatorisches Satzmodell**“: In diesem Schritt verwenden wir Synonym-Datenbanken, um für die in der Frage vorkommenden Wörter der ausgewählten lexikalischen Klassen eine bestimmte Anzahl an Synonymen auszuwählen, wobei jedes Wort auch selbst als erstes Synonym in die entsprechende Liste hinzugefügt wird.

[1] https://crossd.tech/ (Netidee gefördertes Projekt 2022: https://www.netidee.at/crossd; bzw Nachfolge-Projekt **CrOSSD2 | netidee** Netidee Call #19 (2024): **https://www.netidee.at/crossd2**).
[2] Die Lizenz von Softwareergebnissen von **KomMKonLLM** ist *MIT license*.

- „**Kombinatorische Frage-Erzeugung**“: Basierend auf dem abgeleiteten IPM[3] verwenden wir nun *abdeckende Array-Strukturen* (engl.: covering arrays), welche in minimalisierter Zeilenanzahl das Auftreten von allen t-fachen Parameter-Wertekombinationen garantieren.
- **„t-fache Konsistenzfragen“**: Jede Zeile der im vorherigen Schritt erzeugten Array-Struktur kann nun wieder in eine Frage zurückgeführt werden, indem die Wörter durch ihre in der Zeile angegebenen Synonyme ersetzt werden.
- „**Anfrage an LLM**“: Jede der so erzeugten Fragen wird nun mit einem speziellen Prompt, der eine binäre Antwort des LLMs forcieren soll, an ein LLM gesendet und das Frage-Antwort Paar in einer Datenbank gespeichert.
- **„Auswertung der Antworten**“: Alle Frage/Resultat-Paare, welche basierend auf der abdeckenden Array-Struktur erstellt wurden, können nun ausgewertet werden.

### 2.4 Arbeitspaket 4 - Implementierung LLM Testing Framework

Das vierte Arbeitspaket wurde geändert schon mit Ende Dezember 2024 gestartet. Dies ermöglichte eine frühe Koordination zwischen dem Arbeitspaket 3 und Arbeitspaket 4.

Aktuell sind sowohl der zentrale Runner sowie die Datenbank-Anbindung in einer Entwurfs-Version implementiert und für interne Tests bereit. Es sind allerdings höchstwahrscheinlich noch kleinere Anpassungen basierend auf den Anforderungen der LLM-Anbindungen sowie der Komponenten zur Testfallgenerierung notwendig.
Unsere nächsten Schritte bestehen aus der Implementierung der Testfallgenerierung, der Erstellung einer technischen Interface-Dokumentation sowie dem Setup von LLM-Instanzen, um die Schnittstellen daran zu implementieren.

### 2.5 Arbeitspaket 5 - Technische Dokumentation und Release-Vorbereitung

*Dieses Arbeitspaket wurde zum Zeitpunkt dieses Zwischenberichts noch nicht begonnen.*

### 2.6 Arbeitspaket 6 - Dokumentation und Formales am Projektende

*Dieses Arbeitspaket wurde zum Zeitpunkt dieses Zwischenberichts noch nicht begonnen.*

# 3 Umsetzung Förderauflagen

*Im unterzeichneten Fördervertrag wurden keine Förderauflagen festgelegt.*

[3] Die Bezeichnung *IPM* steht für *„input parameter model“* und ergibt sich aus dem vorherigen Schritt.

# 4 Zusammenfassung Planaktualisierung

Aufgrund von Urlauben and Krankenständen über Weihnachten und Silvester 2024/2025 haben sich folgende Verschiebungen ergeben:

- AP2 endet statt Weihnachten 2024 erst Anfang 2025;
- AP3 wurde bis Anfang Februar 2025 verlängert;
- Die Implementierung im Rahmen von AP4 wurde bereits Ende Dezember 2024 begonnen.

Es wurden in AP1-3 einige Planstunden nicht verbraucht; diese werden auf AP4 verschoben, wo sich ein leicht erhöhter Implementierungsaufwand abzeichnet.

# 5 Öffentlichkeitsarbeit/ Vernetzung

Die KomMKonLLM Projekt-Homepage auf der Netidee-Homepage ist die zentrale Anlaufstelle für Informationen über **KomMKonLLM**, wobei es aber auch eine Ankündigung von **KomMKonLLM** auf der Homepage des Forschungszentrums SBA Research[4] sowie einen Eintrag in der Projektliste der MATRIS-Forschungsgruppe gibt.
Zum Zeitpunkt dieses Zwischenberichtes haben wir fünf Blog-Beiträge innerhalb der Netidee-**KomMKonLLM** Homepage erstellt (Überblick; Architektur & Technologien; Methodik von KomMKonLLM (Teil 1 von 2); Methodik von KomMKonLLM (Teil 2 von 2); Beispiel für das Erzeugen von kombinatorischen Konsistenzfragen). Wir benutzen den BlueSky Account @kommkonllm.bsky.social, um über **KomMKonLLM** auf dieser Social-Media Plattform zu informieren und haben dort fünf Beiträge erstellt. Im Jänner 2025 erfolgte eine Online-Präsentation des Projektes **KomMKonLLM** an das Management des Forschungszentrums SBA Research.

Wir werden unser Projekt **KomMKonLLM** am 25. Februar 2025 im Rahmen der Veranstaltung „*SBA Security Meetup hosted by Dynatrace!*“ präsentieren und im März 2025 der Belegschaft – mit besonderem Augenmerk auf die Forscherinnen und Forscher - von SBA Research in einer Online-Präsentation vorstellen. Weitere Aktivitäten der *Öffentlichkeitsarbeit* befinden sich derzeit in der Planung.

# 6 Eigene Projektwebsite

*Es wird aktuell keine weitere Projektseite für **KomMKonLLM** betrieben. Weitere veröffentlichungsfähige Materialien werden unter https://github.com/KomMKonLLM/KomMKonLLM zur Verfügung gestellt.*

[4] Das COMET-Zentrum SBA Research (SBA-K1) wird im Rahmen von COMET – Competence Centers for Excellent Technologies durch BMK, BMAW und das Land Wien gefördert. COMET wird durch die FFG abgewickelt.