



netidee

PROJEKTE

IndeRun

Zwischenbericht | Call 20 | Projekt ID 7913

Lizenz: CC BY-SA

Inhalt

1 Einleitung.....	3
2 Status der Arbeitspakete.....	3
2.1 Arbeitspaket 1 - Detailplanung und Formales am Projektstart.....	3
2.2 Arbeitspaket 2 - Architektur und Prototypen.....	3
2.3 Arbeitspaket 3 - Runtime-Kern mit Policy-Engine.....	5
2.4 Arbeitspaket 4 - LLM-Provider mit Fallback-Logik.....	5
2.5 Arbeitspaket 5 - Konsistente SDKs.....	5
2.6 Arbeitspaket 6 - Demo Apps und Auswertung.....	5
2.7 Arbeitspaket 7 - Dokumentation und Formales am Projektende.....	6
3 Zusammenfassung Planaktualisierung.....	6
4 Öffentlichkeitsarbeit/ Vernetzung.....	7
5 Eigene Projektwebsite.....	7

1 Einleitung

IndeRun entwickelt eine offene Softwarebibliothek, die Anwendungen den Einsatz von KI-Modellen auf verschiedenen Plattformen erleichtert. Das Projekt trainiert keine eigenen KI-Modelle. Stattdessen stellt IndeRun eine gemeinsame Schnittstelle bereit, über die eine Anwendung KI-Funktionen nutzen kann, egal ob die Ausführung direkt am Gerät oder über einen Cloud-Dienst erfolgt. Die Bibliothek soll automatisch einen passenden Weg wählen, zum Beispiel mit Blick auf Datenschutz, Verfügbarkeit, Kosten und Geschwindigkeit.

Der aktuelle Projektstand zeigt, dass die Grundfunktionen für einfache Textanfragen bereits integriert sind. Im GitHub-Repository sind gemeinsame Datenformate, die Auswahllogik für Anbieter, erste Anbieter-Anbindungen, Softwarepakete für Web, iOS und Android, automatische Prüf- und Veröffentlichungsprozesse sowie erste Dokumentation und Demo-Bausteine sichtbar. Die nächsten Arbeiten betreffen vor allem zusätzliche Anbieter, anschauliche Demos, Tests, eine gut nutzbare Entwickler:innen-Erfahrung und die öffentliche Dokumentation.

2 Status der Arbeitspakete

2.1 **Arbeitspaket 1 - *Detailplanung und Formales am Projektstart***

Die Haupttätigkeiten in Arbeitspaket 1 umfassten die Vertragsprüfung und -unterzeichnung. Weiter wurde die Projektwebseite überprüft, mit Details zum Team ergänzt und das Abrufformular für die erste Förderrate abgeschickt.

Nachdem die Webseite angepasst wurde, begann die Detailplanung des Projekts IndeRun, es wurden Arbeitspakete für die Core-Library, Mobile- und Backend-Entwicklung geschnürt.

Bei der Detailplanung und dem Projektstart traten keine nennenswerten Probleme oder Veränderungen auf.

2.2 **Arbeitspaket 2 - *Architektur und Prototypen***

Das Ziel dieses Arbeitspakets ist es, eine einheitliche Architektur und ein erstes Proof-of-Concept für den Provider-Flow zu entwickeln.

AP2 ist inhaltlich abgeschlossen. In diesem Arbeitspaket wurde geklärt, wie IndeRun grundsätzlich funktionieren soll:

App-Entwickler:innen sollten nicht für jeden KI-Anbieter eigene Integrationen schreiben müssen. Stattdessen nimmt IndeRun eine Anfrage entgegen, prüft die verfügbaren Möglichkeiten und leitet die Anfrage an einen passenden Anbieter weiter.

Der aktuelle Stand zeigt, dass aus dem Architekturkonzept bereits erste lauffähige Bausteine für Web, iOS und Android entstanden sind. Die Workflow-Grafik gehört inhaltlich zu AP2: Sie erklärt den geplanten Ablauf von der App-Anfrage bis zur einheitlichen Rückgabe. Links stellt die App eine KI-Aufgabe; danach prüft IndeRun unter anderem vorhandene Hardware, Internetverbindung, Zugangsdaten und Datenschutzvorgaben. Anschließend entscheidet IndeRun, ob die Aufgabe lokal am Gerät, über einen Cloud-Anbieter oder gar nicht ausgeführt werden soll. Rechts zeigt die Grafik, dass die App am Ende immer dieselbe Art von Ergebnis, Fehlermeldung und Metadaten erhält.



2.3 **Arbeitspaket 3 - *Runtime-Kern mit Policy-Engine***

Die Haupttätigkeit dieses Arbeitspakets besteht in der Implementierung der zentralen Runtime mit Provider-Registry, Policy-/Routing-Regeln, Fehlerbehandlung, Task-Schemas und Capability-Checks.

AP3 ist im Kern umgesetzt. Die zentrale Bibliothek kann einfache Anfragen entgegennehmen, Anbieter nach Regeln auswählen und Ergebnisse oder Fehler in einer einheitlichen Form zurückgeben. Wichtig ist dabei, dass Anwendungen nicht selbst wissen müssen, welcher Anbieter im Hintergrund genutzt wird. Die Regeln für Datenschutz, Cloud-Nutzung und Verfügbarkeit sind bereits in den gemeinsamen Datenformaten und in der Dokumentation sichtbar.

2.4 **Arbeitspaket 4 - *LLM-Provider mit Fallback-Logik***

Im Arbeitspaket 4 wird besonderes Augenmerk auf die Integration der wichtigsten LLM-Provider (Anbieter für große Sprachmodelle), insbesondere OpenAI-kompatible Cloud-Ausführung, Apple Foundation Models und Android ML Kit GenAI, inklusive Fallback-Verhalten und erster Benchmarks gelegt.

Der aktuelle Stand zeigt Anbindungen an cloudbasierte und lokale KI-Anbieter: OpenAI-kompatible Anbieter können plattformübergreifend genutzt werden, während Apple- und Android-Geräte lokale KI-Funktionen verwenden können, sofern das jeweilige Gerät diese unterstützt. Noch offen sind eine übersichtliche Anbieter-Matrix, weitere lokale Anbieter-Familien sowie ergänzende Tests und Dokumentation.

2.5 **Arbeitspaket 5 - *Konsistente SDKs***

Im Arbeitspaket 5 wird ein einheitliches API-Design für Swift, Kotlin und TypeScript entwickelt sowie Packaging, Versionierung und Release-Pipelines.

Für Entwickler:innen gibt es bereits klare Hinweise, wie IndeRun in Web-, iOS- und Android-Projekten eingebunden werden kann. Außerdem sind Veröffentlichungsprozesse integriert, damit neue Versionen nachvollziehbar erstellt und verteilt werden können. Der Capacitor-Teil wurde in ein eigenes Repository ausgelagert, damit die Verteilung für hybride Apps sauberer funktioniert.

2.6 **Arbeitspaket 6 - *Demo Apps und Auswertung***

Das Ziel dieses Arbeitspakets ist die Entwicklung von Demo-Apps für iOS, Android und Web, Pilot-Integrationen und die Auswertung von Feedback.

AP6 ist gestartet, aber noch in einem sehr frühen Stadium. Es gibt eine erste Web-Demo, mit der einfache Anfragen ausgeführt werden können. Die Demo zeigt entweder ein Ergebnis oder eine einheitliche Fehlermeldung und gibt Informationen darüber aus, welcher Anbieter genutzt wurde. Außerdem gibt es erste Demo-Ansätze für Android und iOS.

Für den Abschluss dieses Arbeitspakets fehlen noch umfassendere iOS- und Android-Demos, Pilot-Integrationen, gesammeltes Feedback und eine Auswertung der Erkenntnisse.

2.7 Arbeitspaket 7 - *Dokumentation und Formales am Projektende*

Dieses Arbeitspaket wurde noch nicht gestartet.

3 Zusammenfassung Planaktualisierung

IndeRun kommt gut voran und hat eine solide Grundlage erreicht. Der formale Projektstart und die Architekturarbeit sind abgeschlossen, und das Kernsystem für einfache textbasierte KI-Anfragen ist bereits vorhanden. Anwendungen können einen Text-Prompt senden und erhalten entweder ein einheitliches Textergebnis mit Metadaten oder eine standardisierte Fehlermeldung.

Im GitHub-Repository sind aktuell gemeinsame Anfrage- und Fehlerformate, die Auswahllogik für Anbieter, erste Anbieterintegrationen, SDKs für Web, iOS und Android, Release-Prozesse sowie erste Dokumentation sichtbar. OpenAI-kompatible Cloud-Nutzung und erste lokale Anbieterpfade für Apple und Android sind dokumentiert.

Die wichtigsten verbleibenden Arbeiten liegen in der Fertigstellung und Dokumentation der Anbieterlandschaft, im Ausbau von Demos und Tests, in der Sammlung und Auswertung von Feedback sowie in der finalen Entwickler:innen- und Anwender:innen-Dokumentation.

Insgesamt liegt das Projekt im geplanten Zwischenstand: AP1 bis AP3 sind abgeschlossen, AP4 ist weit fortgeschritten, AP5 ist teilweise umgesetzt, und AP6/AP7 bilden den Schwerpunkt der nächsten Projektphase.

4 Öffentlichkeitsarbeit/ Vernetzung

Der Projektleiter hat am NetIdee Summer Talk teilgenommen und sich mit den Teilnehmer*innen der anderen, im Call 20 geförderten Projekte ausgetauscht.

Auf der Projektwebseite wurden 2 Berichte zum Projektfortschritt veröffentlicht.

5 Eigene Projektwebsite

Für das Projekt wurde ein GitHub-Repository angelegt:

- [IndeRun Github Repo](#)