

Hochschule Luzern
Fachbereich Informatik
Modul Wissenschaftliches Schreiben & Forschungsmethodik
Leitung: Benjamin Haymond

HSLU Mobile Apps - Android

Jetpack Compose - AI First

Forschungsskizze (Leistungsnachweis WSFM)

Arbeit von:

Raphael Eiholzer
raphael.eiholzer@stud.hslu.ch
Studiengang Informatik

Inhalt

1	Ausgangslage.....	3
2	Forschungsfragen	4
3	Persönliche Reflexion.....	5
4	Quellenverzeichnis	6
5	Anhang	7
5.1	Einordnung der Literaturquellen	7
5.2	Zugrundeliegende Aufgabenstellung (WIPRO).....	9

1 Ausgangslage

Die Ausgangslage ist gegeben durch ein Wirtschaftsprojekt der Hochschule Luzern. Eine auf iOS bestehende Codebasis soll mithilfe von KI-Tools für Android nachgerüstet werden. Dabei sollen vorgängig geeignete KI-Tools evaluiert werden. Mithilfe dieser Tools soll dann die Umsetzung des Projektes erfolgen.

KI-Werkzeuge haben das Potenzial, Entwickler bei der Arbeit zu unterstützen. Insbesondere bei repetitiven Aufgaben wie Codegenerierung, Dokumentation oder automatisiertem Testing können KI-gestützte Tools Entwickler*innen entlasten und Entwicklungsprozesse beschleunigen (Noy & Zhang, 2023).

Trotz dieses Potenzials ist der praktische Einsatz von KI-Werkzeugen nicht unproblematisch. Viele Tools liefern nur eingeschränkt brauchbare Ergebnisse, insbesondere wenn der Kontext komplex ist oder spezifische Anforderungen bestehen. Häufig entstehen unvollständige, fehleranfällige oder schwer wartbare Codevorschläge, was im schlimmsten Fall zu einem höheren Mehraufwand führt, als wenn der Code manuell geschrieben worden wäre. Laut Yetiştireن et al., 2023 erzeugen KI-gestützte Codegenerierungstools durchschnittlich technische Schulden von bis zu 9 Minuten pro Codebeispiel, die im Nachgang manuell behoben werden müssen. Dies relativiert den vermeintlichen Effizienzgewinn solcher Tools und unterstreicht die Notwendigkeit einer kritischen Evaluation vor ihrem produktiven Einsatz.

Im Rahmen dieses Wirtschaftsprojekts der Hochschule Luzern soll genau dieser Spannungsbogen untersucht werden: Ziel ist es, geeignete KI-Tools für die Nachrüstung einer bestehenden iOS-App in einer Android-Codebasis zu evaluieren und deren Praxistauglichkeit unter realen Bedingungen zu testen. Dabei stehen neben der technischen Umsetzung auch dokumentationsrelevante und qualitätssichernde Aspekte wie Testing und CI/CD-Prozesse im Fokus.

2 Forschungsfragen

Mit welchem KI-Tools erzielen Softwareentwickler*innen die grösste Zeitsparnis bei der Implementierung standardisierter Codeaufgaben?

Diese Forschungsfrage hat zum Zweck, dass ein objektiver Vergleich der Effizienz verschiedener KI-Codegenerierungstools möglich wird. Dabei soll untersucht werden, welches Tool Entwickler*innen am besten dabei unterstützt, repetitive oder standardisierte Aufgaben im Entwicklungsaltag schneller zu lösen.

Zur Beantwortung der Frage werden standardisierte Programmieraufgaben definiert und von mehreren Entwickler*innen mit Hilfe unterschiedlicher Tools (z. B. *GitHub Copilot*, *ChatGPT*) bearbeitet. Die Bearbeitungszeit wird dabei jeweils gemessen und vergleichend ausgewertet. Ergänzend kann ein kurzer Fragebogen zur subjektiven Einschätzung der wahrgenommenen Unterstützung durch das Tool erhoben werden. Es ist dabei wichtig, dass die Entwickler*innen ungefähr die gleichen Fähigkeiten mitbringen müssen und dass mehrere Tests durchgeführt werden, damit das Ergebnis reproduzierbar bleibt.

Welche typischen Fehler und Limitationen treten beim Einsatz von KI-Tools zur Code- und Dokumentationsgenerierung auf, und wie kann ein Entwickler*in diese erkennen und im Entwicklungsprozess gezielt darauf reagieren?

Diese Forschungsfrage hat zum Zweck, die Schwächen und Grenzen von KI-basierten Code- und Dokumentationsgeneratoren systematisch zu identifizieren. Gleichzeitig soll untersucht werden, wie Entwickler*innen diese Fehlerquellen erkennen und gezielt mit ihnen umgehen können. Zur Beantwortung der Frage wird zunächst eine qualitative Analyse von KI-generiertem Code und Dokumentation durchgeführt. Dabei werden typische Fehlerarten (z. B. fehlerhafte Logik, inkonsistente Benennungen) kategorisiert. Ziel ist es, Entwickler*innen auf die Grenzen der KI-Tools aufmerksam zu machen. Sie sollen die Codequalität mit menschlichen Korrekturen steigern zu können, um so die technischen Schulden zu reduzieren.

3 Persönliche Reflexion

Die Recherche begann mit einer Schneeballmethode, bei der erste relevante Quellen identifiziert wurden. Mithilfe der Referenzen der Arbeiten und mit Tools wie *Connected Graphs* konnten weitere hilfreiche, teilweise sogar bessere Quellen gefunden werden. Besonders hilfreich war Google Scholar in Kombination mit dem Scanning von Abstracts, um Inhalte effizient zu filtern. Die Bewertung der Quellen erfolgte mithilfe des CRAAP-Tests sowie ergänzend über Kennzahlen wie den h-Index. Die Formulierung der Forschungsfragen begann mit dem Finden einer Wissenslücke. Anschliessend wurden die Fragen überarbeitet mithilfe des Vorgehens wie es Booth et al., 2016 empfehlen.

Rückblickend war das Vorgehen insgesamt erfolgreich. Das Finden geeigneter Quellen zu einem so aktuellen Thema stellte sich aber als Herausforderung dar. Fachliteratur, die bereits ein oder zwei Jahre alt ist, kann in einem so schnelllebigen Bereich bereits als überholt gelten. Die Eingrenzung auf mobile App-Entwicklung erwies sich ebenfalls als herausfordernd, da hierzu weniger spezialisierte Literatur verfügbar war.

Die Beantwortung der wissenschaftlichen Fragen gestaltet sich mit menschlichen Beteiligten schnell als anspruchsvoll, insbesondere wenn reproduzierbare Ergebnisse erzielt werden sollen. Bei der praktischen Umsetzung wäre daher besonders auf gute Konsistenz und Messgenauigkeit zu achten.

Trotz der verschiedenen Herausforderungen bin ich zufrieden mit meiner Arbeit und denke, eine gute Forschungsskizze erstellt zu haben. Ich bin froh, diese Arbeit so gemacht zu haben, denn ich fühle mich jetzt gut vorbereitet für meine WIPRO.

4 Quellenverzeichnis

Booth, W. C., Colomb, G. G., Williams, J. M., Bizup, J., & Fitzgerald, W. T. (2016).

The Craft of Research (4. Aufl.). University of Chicago Press.

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654), 187–192.

<https://doi.org/10.1126/science.adh2586>

Yetişturen, B., Özsoy, I., Ayerdem, M., & Tüzün, E. (2023). *Evaluating the Code*

Quality of AI-Assisted Code Generation Tools: An Empirical Study on

Github Copilot, Amazon CodeWhisperer, and ChatGPT (No.

arXiv:2304.10778). arXiv. <https://doi.org/10.48550/arXiv.2304.10778>

5 Anhang

5.1 Einordnung der Literaturquellen

Nachfolgend sind die Bewertungen von drei der verwendeten Quellen aufgeführt. Die Auswahl und Beurteilung erfolgten mithilfe des CRAAP-Tests. Die beiden Internetquellen wurden im Rahmen der Recherchearbeit über Google Scholar, durch Scanning von Abstracts sowie über Connected Graphs gefunden. Zur Organisation und Kommentierung der Literatur kam das Tool Zotero zum Einsatz.

Quelle 1: Booth et al. (2016) – The Craft of Research

Diese Quelle wurde über die Online-Bibliothek der Hochschule gefunden, nachdem ich gezielt nach Büchern zum Thema «Research» gesucht hatte. Ich habe mehrere Werke überflogen und mich schliesslich für dieses entschieden, da es ab Kapitel 3.1 hilfreiche Hinweise zur Entwicklung und Eingrenzung geeigneter Forschungsthemen liefert.

Currency: 4. Auflage (2016), aktuellste Version. Für grundlegende Methodik weiterhin gültig, «zeitlos»

Relevance: Richtet sich an Studierende/Forschende. Niveau und Inhalt geeignet zur Planung & Frageformulierung.

Authority: Autor*innen sind anerkannte Forschende (u.a. University of Chicago). Seriöser akademischer Verlag.

Accuracy: Klar gegliedert, faktenbasiert, gut belegt. Keine sprachlichen oder formalen Mängel.

Purpose: Anleitung zum wissenschaftlichen Arbeiten. Objektiv, neutral, keine wirtschaftlichen Interessen.

Fazit: Für die Erstellung der Forschungsskizze sehr hilfreich und inhaltlich gut geeignet.

Quelle 2: Noy & Zhang (2023) – Productivity Effects of Generative AI

Gefunden über: Google Scholar. Früh in der Recherche entdeckt. Diente zur Beurteilung der Effizienz von KI-Tools und stützt Forschungsfrage 1.

Currency: Publiziert 2023 in Science. Für das schnelllebige Feld der KI nicht mehr «top aktuell», aber weiterhin relevant und zitierfähig. Keine Rückzüge/Korrekturen gefunden.

Relevance: Direkt relevant für das Thema «Zeitersparnis mit KI».

Authority: Autor*innen vom MIT; Fachzeitschrift Science ist renommiert und peer-reviewed.

Accuracy: Empirische Studie, transparente Methodik, statistisch fundiert.

Purpose: Informationsvermittlung. Keine kommerziellen oder ideologischen Absichten erkennbar.

Fazit: Solide, empirisch fundierte Quelle, trotz «Alter» weiterhin gut nutzbar.

Quelle 3: Yetiştiren et al. (2023) – Code Quality of AI Tools

Gefunden über: Connected Papers (Zitationsnetzwerk von Noy & Zhang). Wichtig für Forschungsfragen 2 & 3 zu Codequalität & technischen Schulden.

Currency: arXiv-Paper von 2023. Im aktuellen Forschungskontext nicht mehr ganz neu, aber methodisch weiterhin relevant.

Relevance: Untersucht direkt Codequalität von Copilot, ChatGPT & CodeWhisperer. Inhaltlich daher sehr passend.

Authority: Autor*innen aus dem universitären Umfeld. Veröffentlichung auf arXiv (nicht peer-reviewed), aber in Fachkreisen mehrfach zitiert.

Accuracy: Klare Metriken, strukturierte Darstellung, transparentes Studiendesign.

Purpose: Objektive Vergleichsstudie ohne wirtschaftliche Interessen.

Fazit: Trotz fehlendem Peer Review eine fachlich relevante und methodisch brauchbare Quelle.

5.2 Zugrundeliegende Aufgabenstellung (WIPRO)

Das vorliegende Projekt stammt aus dem Wirtschaftsprojekt-Modul (WIPRO) der Hochschule Luzern – Informatik. Es wird im Herbstsemester 2025 von Samuel Kurmann und mir bearbeitet. Zum Zeitpunkt der Abgabe dieser Forschungsskizze liegt die vollständige Aufgabenstellung noch nicht vor, da diese zum Projektbeginn gemeinsam mit dem Auftraggeber erarbeitet wird.

Typ der Arbeit

Wirtschaftsprojekt (WIPRO)

Titel

HSLU Mobile Apps - Android Jetpack Compose - AI First

Ausgangslage und Problemstellung

Die HSLU Informatik betreibt eigene Mobile-Apps für die Departemente Informatik und Technik & Architektur auf den Plattformen iOS und Android. Im Rahmen einer Forschungsarbeit wurde die Codebasis für iOS stark erweitert, diese Änderungen müssen nun in der Android Jetpack Compose Codebase nachgerüstet werden. Dies soll aber im Rahmen von AI-First Techniken erfolgen und im (sic!) diesem Zusammenhang sollen diverse Möglichkeiten, Ansätze und Tools evaluiert werden. Mit der passendsten Technologie erfolgt dann die Umsetzung.

Ziel der Arbeit und erwartete Resultate

Evaluation AI-Tools für Mobile Programming & Testing Implementation der Features, Testing & Dokumentation Release des neuesten Standes via Gitlab CI/CD & Fastlane in den Google Play Store

Gewünschte Methoden, Vorgehen

Inkrementelles, iteratives, agiles Vorgehen, mit regelmässigen Treffen mit dem Auftraggeber

Kreativität, Varianten, Innovation

Cutting-edge Technologien einsetzen und evaluieren.

Sonstige Bemerkungen (Anforderungen, Vorkenntnisse, ...)

Erfolgreich bestandenes MobPro-Modul.

Organisation

Hochschule Informatik