

Künstliche Intelligenz: Datenbasis macht 3,8 Mio. Skills für KI-Agenten erforschbar

Datum 14.10.2026
Subjekt Universität Hohenheim
Region Nemecko
Kategorie IT a komunikace, Technologie, Věda a výzkum

Forschende der Uni Hohenheim entwickeln mit internationalen Partnern frei zugängliche Datengrundlage für mehr Qualität und Sicherheit bei KI-Agenten.

KI-Agenten schreiben und verändern Software, bearbeiten Dokumente und verwalten Dateien. Wie zuverlässig und sicher sie dabei vorgehen, hängt wesentlich von den Anweisungen ab, denen sie folgen - den sogenannten Skills. Ein internationales Forschungsteam mit Beteiligung der Universität Hohenheim hat nun erstmals einen umfassenden offenen Datensatz dieser Anweisungen veröffentlicht: GitSkills umfasst 3,8 Millionen Anweisungsdateien aus mehr als 280.000 öffentlichen GitHub-Projekten. Der Datensatz ermöglicht es erstmals, systematisch zu untersuchen, wie Anweisungen für KI-Agenten entstehen, verbreitet, verändert und gepflegt werden. Die Daten sollen helfen, veraltete, unklare und potenziell unsichere Anweisungen zu erkennen und Prüfverfahren für KI-Agenten weiterzuentwickeln. Der Datensatz findet sich hier: doi.org/10.5281/zenodo.21875637

Immer mehr Software wird mit Hilfe von KI-Agenten geschrieben. Dabei handelt es sich um Programme, die zunehmend eigenständig im Auftrag von Menschen handeln. So beantwortet beispielsweise ein KI-Agent im Kundendienst Anfragen und Reklamationen von Nutzer:innen oder bearbeitet Bestellungen völlig eigenständig.

Um zu erfahren, wie sie bestimmte Aufgaben ausführen sollen, nutzen KI-Agenten sogenannte Skills. Diese Dateien enthalten beispielsweise Informationen zu Werkzeugen, Befehlen, Konventionen und sogar zum Tonfall, den der Agent verwenden soll.

Skills bestimmen, wie KI-Agenten handeln

„Ähnlich wie neue Mitarbeitende eine schriftliche Arbeitsanweisung befolgen, setzen KI-Agenten die Vorgaben eines Skills um.“, erklärt Prof. Dr. Daniel Graziotin, Leiter des Fachgebiets Wirtschaftsinformatik und Software Engineering an der Universität Hohenheim.

„Dabei entscheiden sie selbst, wann ein Skill zu einer Aufgabe passt“, fährt er fort. „Eine automatische Prüfung, ob die ausgewählte Anweisung noch aktuell, eindeutig oder sicher ist, gibt es dabei nicht grundsätzlich.“ Ein Agent erkennt einen veralteten Skill daher nicht zwangsläufig. Er kann ihn ausführen und dabei möglicherweise fehlerhafte oder unsichere Ergebnisse erzeugen, ohne dass es eine Fehlermeldung gibt. Da Entwickler:innen Skills häufig von einem Projekt in ein anderes kopieren, verbreiten sich solche Fehler weiter.

Das ist besonders problematisch, da heutzutage KI-Agenten meistens zur Entwicklung von Software eingesetzt werden. Doch Software verändert sich schnell. Ein Skill kann überholte Befehle, nicht

mehr gültige Einstellungen oder veraltete Vorgehensweisen enthalten. Für Skills gibt es kein zentrales Register und in der Regel auch keine Versionsnummern.

Anweisungen für KI-Agenten: ein unterschätztes Sicherheitsrisiko

Auch sicherheitskritische Befehle können in Skills enthalten sein. Sie können einen Agenten beispielsweise anweisen, Programme aus dem Internet herunterzuladen, Skripte auszuführen oder bestimmte Befehle auf einem Rechner zu starten.

„Wie gut und wie sicher die von KI-Agenten erzeugte Software ist, hängt von den Skills ab, die diese Agenten erhalten. Unternehmen prüfen unbekannte Software, bevor sie sie installieren. Doch ihre KI-Agenten befolgen Anweisungen, die von Unbekannten erstellt und von der Entwicklungsplattform GitHub übernommen wurden – ohne, dass dabei eine automatische Überprüfung erfolgt“, sagt Prof. Dr. Graziotin.

3,8 Millionen Dateien aus mehr als 280.000 Projekten

Ein internationales Forschungsteam hat deshalb den Datensatz GitSkills **zusammengestellt und** veröffentlicht. Er umfasst 3,8 Millionen Skill-Dateien mit einem Gesamtumfang von 44 Gigabyte. Die Dateien wurden im Juli 2026 aus mehr als 280.000 öffentlichen GitHub-Projekten von knapp 200.000 Accounts zusammengetragen.

Neben den vollständigen Texten der Dateien enthält der Datensatz Informationen zu den zugehörigen Projekten, zu vorhandenen Kopien und bei einem großen Teil der Dateien auch zur Änderungshistorie.

Die Zahl der Dateien verdeutlicht zugleich, wie stark Anweisungen kopiert werden: Die 3,8 Millionen Dateien lassen sich auf rund 1,9 Millionen unterschiedliche Inhalte reduzieren. Damit sind etwa die Hälfte der erfassten Dateien bytgenaue Kopien anderer Dateien. Veraltete oder unsichere Anweisungen können sich dadurch weiterverbreiten, ohne dass eine Stelle für ihre Aktualisierung verantwortlich ist.

GitSkills macht Skills erstmals systematisch untersuchbar

Bislang mussten Forschende, die untersuchen wollten, wie Skills verfasst, kopiert und gepflegt werden, auf wenige Einzelbeispiele zurückgreifen. Der neue Datensatz ermöglicht dagegen Analysen über die gesamte Sammlung hinweg. „GitSkills ermöglicht es Forschenden erstmals, Skills systematisch zu analysieren,“ so Prof. Dr. Graziotin.

So lässt sich beispielsweise untersuchen, wie schnell sich das Format verbreitet, welche Skills besonders häufig kopiert werden, wie stark sich Kopien im Laufe der Zeit verändern und wie oft sie veralten. Ebenso können Forschende prüfen, ob Kopien zusätzliche Befehle oder Netzwerkzugriffe enthalten, die im ursprünglichen Skill nicht vorgesehen waren.

Auch außerhalb der Softwareentwicklung gewinnen solche Skills an Bedeutung. Nach einer aktuellen Auswertung des Forschungsteams stammt etwa ein Viertel der erfassten Skills aus anderen Bereichen; dazu gehört insbesondere die Medizin. Damit eröffnet der Datensatz Perspektiven für Untersuchungen, wie KI-Agenten in unterschiedlichen Fachgebieten Anweisungen

nutzen und welche Anforderungen sich daraus an deren Qualität und Sicherheit ergeben.

Datengrundlage für bessere Prüfverfahren

GitSkills behebt jedoch keine Fehler in einzelnen Skilldateien. Der Datensatz schafft eine Grundlage, um solche Probleme systematisch zu erforschen und geeignete Prüfverfahren zu entwickeln. Sie könnten vor unklaren Beschreibungen, veralteten Befehlen oder auffälligen Veränderungen zwischen Original und Kopie warnen.

„Für die Entwicklung solcher Werkzeuge sind reale Beispiele erforderlich – darunter auch schlecht formulierte, veraltete oder möglicherweise manipulierte Dateien“, sagt Graziotin. Da GitSkills Zusammenhänge zwischen Dateien erfasst, können Forschende problematische Anweisungen bis zu ihrem Ursprung zurückzuverfolgen.

International frei zugänglich für Forschung und Praxis

GitSkills steht als Open Data kostenfrei zur Verfügung. Eine Creative Commons Attribution- Lizenz erlaubt die Nutzung und Weiterverwendung der Daten, einschließlich kommerzieller Anwendungen, sofern die Quelle angegeben wird.

GitSkills konzentriert sich dabei nicht auf ein einzelnes Produkt, sondern auf Skills, die in unterschiedlichen KI-Entwicklungsumgebungen zum Einsatz kommen. Das zugrunde liegende Skill-Format wurde im Oktober 2025 von Anthropic eingeführt und als offene Spezifikation veröffentlicht. Inzwischen verwenden es verschiedene Werkzeuge für KI-Agenten, darunter OpenAI Codex, Googles Gemini CLI, GitHub Copilot, Cursor und weitere Anwendungen.

Die Organisatoren der International Conference on Mining Software Repositories (MSR 2027) haben GitSkills als einen von zwei Datensätzen für ihre „Mining Challenge“ ausgewählt. Die internationale Konferenz zur Erforschung von Daten aus Softwareprojekten findet im April 2027 in Dublin statt. Forschungsteams aus aller Welt werden die Datensätze unabhängig voneinander analysieren und ihre Ergebnisse vergleichen.

Dadurch entsteht eine gemeinsame Grundlage, um unterschiedliche Untersuchungsmethoden und Befunde zur Qualität, Verbreitung und Sicherheit von Skills zu bewerten. Bereits im ersten Monat wurde GitSkills fast 4.000-mal heruntergeladen.

An der Entwicklung von GitSkills beteiligt sind Forschende der Universität Hohenheim, des University College London (UCL) in Großbritannien und der Universität Cagliari in Italien. Zum Hohenheimer Team gehört auch Postdoktorand Dr. Matteo Vaccargiu. Die Projektleitung liegt bei Dr. Giuseppe Destefanis vom UCL.

Weitere Informationen

Publikation:

G. Destefanis, D. Graziotin, M. Vaccargiu, and M. Ortu, “GitSkills: A Dataset of Agent Skills on GitHub” in Proceedings of the 24th IEEE/ACM International Conference on Mining Software Repositories (MSR '27). IEEE, Piscataway, NJ, USA, 2027. To appear. Available for free: <https://arxiv.org/abs/2608.10906>

Link zu GitSkills auf Zenodo:

<https://doi.org/10.5281/zenodo.21875637>

https://www.uni-hohenheim.de/pressemitteilung?cHash=1c8ede64e1f199e00d22a674c962580d&tx_ttnews%5Btt_news%5D=70795