

### Relevanz-Einordnung

Mit der Veröffentlichung von ChatGPT von OpenAI im Jahr 2022 (Cunningham et al., 2025) wurde eine technologische Wende eingeleitet. Bereits heute vereinfachen und verändern LLM-basierte Applikationen wie ChatGPT von OpenAI, Claude von Anthropic und Gemini von Google viele Tätigkeiten des (Arbeits-)Lebens.

Mit der erwähnten technologischen Wende wurden technologische Assistenten in verschiedenen Formen, zum Beispiel-Chatbots und autonome Agenten zum täglichen Begleiter des Menschen, ob durch eine bewusst durchgeführte Interaktion oder ein im Hintergrund stattfindender, unbewusster Berührungspunkt. Chatbots stellen im Alltag inzwischen eine wichtige Rolle in mehreren Aspekten: Sie unterstützen bei der Informationsbeschaffung, geben praktische Anleitungen und bieten zum Beispiel Hilfe in der Programmierung sowie in Kreativprozessen (Cunningham et al., 2025).

Gemäss «How People use ChatGPT» (Cunningham et al., 2025) hatte OpenAI bereits eine Woche nach der Ankündigung von ChatGPT im November 2022 5 Millionen registrierte Nutzer. Ein Jahr später verzeichnete das Unternehmen 100 Millionen wöchentlich aktive Nutzer (WAU). Bis im Januar 2025 wuchs die Nutzerbasis von ChatGPT auf rund 350 Millionen wöchentlich aktive Nutzer an. Stand Juli 2025 stieg dieser Wert nochmals um mehr als 350 Millionen Nutzer auf über 700 Millionen wöchentlich aktive Nutzer. Das bedeutet zum einen, dass OpenAI seine Nutzerbasis alleine innerhalb von 7 Monaten verdoppeln konnte und die Adoption von solchen digitalen Assistenten im Umfeld des Menschen rasant voranschreitet.

Bei genauerer Betrachtung der Nutzungsdaten von ChatGPT ergeben sich verschiedene Anwendungsfelder der LLM-Lösung. Besonders zwei Anwendungsfelder machen gemeinsam knapp 50% aller Anfragen an die KI-Lösung aus (Cunningham et al., 2025).

«Practical Guidance» (praktische Anleitungen) mit 28,3% und «Seeking Information» (Informationssuche) mit 21,3% aller Anfragen. «Seeking Information» lässt sich weiter unterteilen.

Diese Betrachtung zeigt, dass von den 21,3% über 18% nach «Specific Information» gesucht wird. Bei «Practical Guidance» sind 30% (8.5% von 21.3%) aller Anfragen im

Bereich «How To Advice» verortet (Cunningham et al., 2025).

Diese Tatsache birgt das Potenzial die Informationsbeschaffung für Nutzer zu vereinfachen sowie praktische Tipps und Anleitungen in Eigenregie zu beschaffen. Private und öffentliche Organisationen und Unternehmen stützen ihre Service- und Dienstleistungsangebote mehr und mehr auf KI-gestützte Unterstützungsinstrumente, um genau diesen beiden Bedürfnissen Rechnung zu tragen. Mit massgeschneiderten Lösungen in Form von digitalen Assistenten welche auf bereits vortrainierten Modellen (sogenannte «Foundational Models») basieren und mit unternehmens- und kontextspezifischen Informationen ausgestattet werden, liefern diese Assistenten detaillierte Informationen je nach Umfeld oder Plattform. Dazu zählen zum Beispiel das Sammeln und Vergleichen von Produkt- und Dienstleistungsinformationen oder das Beschaffen von Informationen wie Gesetze, Regularien, Verfahrens- und Prozessdienstleistungen im öffentlichen Bereich. Besonders bei grossen Informationsmengen und einer Vielzahl von Sub-Themen liefern diese Assistenten einen einfachen Zugang zu komplexen Themen oder liefert zentralisiert Informationen, auch wenn diese über verschiedene Inhaltsbereiche- und typen verteilt sind. Mit ihrer relativ jungen (und öffentlichkeitswirksamen) Geschichte ist die generative künstliche Intelligenz, wie viele übergreifende technologischen Veränderungen, einem technologischen- und gesellschaftlichen Adoptionsprozess ausgesetzt. Einen theoretischer Erklärungsansatz dieses Adoptionsprozesses liefert Fred Davis 1987 mit seinem Werk «User acceptance of information systems: the technology acceptance model (TAM)». Während Davis in seiner Arbeit die wahrgenommene Nützlichkeit («Perceived Usefulness») und Einfachheit in der Nutzung («Ease of Use») im Fokus steht und davon die Verhaltensintention («Behaviourial Intention») abgeleitet wird (Davis, 1987), begleitet das Thema generative KI, oder Künstliche Intelligenz im Allgemeinen, der Aspekt des Vertrauens in die Technologie besonders. Neben Nützlichkeit und Einfachheit stellt die Vertrauensfrage den Aspekt dar, ob künstlicher Intelligenz vertrauenswürdig sind. Sämtliche grossen Anbieter wie ChatGPT, Claude von Anthropic, etc. weisen vor- sowie während der Nutzung ausdrücklich darauf hin, dass

die ihre «Foundational Models» und konsequenterweise KI-Assistenten die auf diesen Modellen basieren, fehlerhaft sein können.

Diese Fehleranfälligkeit und zusätzlichen Faktoren wie Angst in verschiedenen Ausprägungen wie «Angst vor Jobverlust durch KI», «Verletzung der Privatsphäre» sowie ethische Bedenken (Li & Huang, 2020), erfordern die Integration und Erfassung von «Vertrauen» als eigenständiges Konstrukt in möglichen theoretischen Modellen. Das bestehende TAM-Modell von Davis (1987) kann durch verschiedene Faktoren wie Vertrauen (XAIT, «Explainable AI Trust»), Intention zur Kollaboration mit KI («Collaborative Intention») und LLM-basierten Metriken («Accuracy Score») erweitert werden. (Baroni et al. (2022), schufen 2022 mit ihrer Erweiterung des TAM-Modells «AI-TAM» in Ihrer Studie «AI-TAM: a model to investigate user acceptance and collaborative intention in human-in-the-loop AI applications» einen neuen Erklärungsbeitrag zum Thema «Vertrauen in KI» (Baroni et al., 2022).

Die theoretische Einbettung sowie das Forschungsdesign orientieren sich an an diesem erweiterten AI-TAM-Modell. In den folgenden Schritten wird die genaue Integration des AI-TAM Modells in das geplante Experiment beschrieben.