

Pressemitteilung

Bayerisches Forschungsinstitut für Digitale Transformation (bidt)

Leonie Liebich

20.01.2025

<http://idw-online.de/de/news845956>

Forschungs- / Wissenstransfer, Forschungsprojekte
Gesellschaft, Informationstechnik, Medien- und Kommunikationswissenschaften, Politik
überregional

bidt Bayerisches Forschungsinstitut
für Digitale Transformation
ein Institut der Bayerischen Akademie der Wissenschaften

Vertrauen und KI – bidt startet neuen Forschungsschwerpunkt und fördert zwölf innovative Projekte

Rekordförderung am bidt: Ein neuer Forschungsschwerpunkt und zwölf innovative Forschungsprojekte gehen im Jahr 2025 an den Start – so viele wie noch nie. Das Bayerische Forschungsinstitut für Digitale Transformation der Bayerischen Akademie der Wissenschaften schärft sein Profil mit dem neuen Forschungsschwerpunkt „Mensch und generative KI: Trust in Co-Creation“. Drei interne und sieben externe Forschungsvorhaben befassen sich mit spannenden Fragen rund um Vertrauen und KI. Zusätzlich nehmen zwei neue Konsortialprojekte ihre Arbeit auf.

Generative KI revolutioniert Lebensbereiche wie Bildung, Arbeit oder Wissenschaft, bringt jedoch auch Herausforderungen mit sich, etwa die Gewährleistung von Qualität und Wahrheitsgehalt. Mit dem neuen Forschungsschwerpunkt widmet sich das Institut der Frage, wann und wie viel Vertrauen in Bezug auf generative KI angebracht ist, sei es bei der Erstellung, Interaktion oder der Bewertung von KI-generierten Produkten. Das bidt fördert ab 2025 gezielt Projekte, die diese Thematik aus verschiedenen Blickwinkeln beleuchten.

Neuer bidt-Forschungsschwerpunkt unter der Leitung von Professorin Hannah Schmid-Petri

Die Leitung des neuen Forschungsschwerpunkts „Mensch und generative KI: Trust in Co-Creation“ übernimmt bidt-Direktorin Hannah Schmid-Petri, Inhaberin des Lehrstuhls für Wissenschaftskommunikation an der Universität Passau. „Vertrauen ist der Dreh- und Angelpunkt für den erfolgreichen Einsatz von generativer KI. Mit unserem neuen bidt-Schwerpunkt wollen wir herausfinden, wie wir die Interaktion von Mensch und KI so gestalten können, dass sie langfristig dem Menschen und der Gesellschaft dienlich ist“, fasst sie zusammen. Schmid-Petri forscht zudem an einem eigenen Projekt rund um Vertrauen und KI. Die drei internen Projekte haben im Jahr 2024 begonnen. Die Forschungsvorhaben im Überblick:

„KI im Journalismus: Der Einfluss generativer KI auf Objektivität und Dialogbereitschaft in der Debatte um Klimaschutz“ | Leitung Prof. Dr. Hannah Schmid-Petri (bidt / Universität Passau)

Das Projekt erforscht, inwiefern KI in journalistischen Berichten dazu beitragen kann, die Annahmefähigkeit von Botschaften zu erhöhen und die sachliche Auseinandersetzung mit Gegenargumenten zu befördern. Zum Projekt: <https://www.bidt.digital/forschungsprojekt/ki-im-journalismus-der-einfluss-generativer-ki-auf-objektivitaet-und-dialogbereitschaft-in-der-debatte-um-klimaschutz-ki-journalismus/>

„Mensch-KI-Co-Creation von Programmcode bei unterschiedlichen Vorkenntnissen – Effekte auf Performanz und Vertrauen“ | Leitung: Prof. Dr. Ute Schmid (bidt / Universität Bamberg)

Dieses Projekt untersucht den Co-Creation-Prozess von Mensch und KI im Kontext der Erstellung von Programmcode, indem es KI-Methoden und Ansätze der experimentellen Kognitionsforschung kombiniert. Zum Projekt: <https://www.bidt.digital/forschungsprojekt/mensch-ki-co-creation-von-programmcode-bei-unterschiedlichen-vorkenntnissen-effekte-auf-performanz-und-vertrauen-pairprog/>

„Rechtsunsicherheit durch generative KI? Reformüberlegungen zur Förderung von Systemvertrauen an Hochschulen“ | Leitung: Prof. Dr. Dirk Heckmann (bidt / TU München)

In dem Projekt soll neben der rechtlichen Analyse im Hochschul- und Prüfungsrecht auch untersucht werden, wie Hochschulen in der Prüfungspraxis auf die Nutzung generativer KI reagieren könnten und wo sie sich anpassen sollten. Zum Projekt: <https://www.bidt.digital/forschungsprojekt/rechtsunsicherheit-durch-generative-ki-reformueberlegung-n-zur-foerderung-von-systemvertrauen-an-hochschulen-akademischer-rahmen-fairer-ki-regulierung-fuer-pruefungen-affaire/>

Von KI in Wahlen bis AI-Aided Design – Förderung von sieben Forschungsvorhaben aus ganz Bayern

Die sieben externen Forschungsvorhaben wurden im Rahmen der bidt-Jahresausschreibung ausgewählt und starten im ersten Quartal 2025. Sie sind alle an bayerischen Hochschulen und Forschungseinrichtungen angesiedelt und haben sich in einem kompetitiven Verfahren aufgrund ihrer exzellenten Qualität, Aktualität und Relevanz durchgesetzt.

„Die große Resonanz auf unsere Jahresausschreibung hat uns überwältigt. Sie zeigt, dass unser Forschungsthema Vertrauen und KI den Zahn der Zeit trifft. Wir gratulieren den Projekten, die aus einem Pool von 147 Bewerbungen hervorstechen konnten und erhoffen uns von ihnen wertvolle Impulse und Handlungsempfehlungen für eine gemeinwohlorientierte digitale Zukunft“, sagt Christoph Egle, Geschäftsführer des bidt. Dies sind die externen Forschungsvorhaben:

„Generative Artificial Intelligence in Elections: Uses, Preferences, and Trust“ | Prof. Dr. Andreas Jungherr (Universität Bamberg)

Die Forschenden untersuchen, wie deutsche Parteien generative KI nutzen, und welchen Einfluss dies auf Wahlkampagnen und das Vertrauen der Öffentlichkeit hat. Damit siedelt das Projekt an der Schnittstelle von KI, Demokratie und politischem Wettbewerb an. Zum Projekt: <https://www.bidt.digital/forschungsprojekt/generative-kuenstliche-intelligenz-im-wahlkampf-anwendungen-praeferenzen-und-vertrauen-ai-wahlkampf/>

„Selbstreguliert und kompetent mit generativer KI interagieren: Diagnostik und Förderung“ | Prof. Dr. Marion Händel (Hochschule Ansbach)

Das Projekt befasst sich mit der Interaktion von Studierenden mit KI und beleuchtet das selbstregulierte Lernen. Es zielt darauf ab, Diagnoseverfahren und Fördermaßnahmen zu entwickeln, die zu einem reflektierten Umgang mit generativer KI beitragen. Zum Projekt: <https://www.bidt.digital/forschungsprojekt/selbstreguliert-und-kompetent-mit-generativer-ki-interagieren-diagnostik-und-foerderung-sekoki/>

„Algorithmische Darstellungsverzerrungen aus Userperspektive: Bewertung, Auswirkungen, Interventionen“ | Prof. Dr. Markus Appel (Universität Würzburg)

Die mit generativer KI erstellten Texte und Bilder stellen Frauen und Minderheiten oft unausgewogen dar. Ziel des Projekts ist es, dieses Phänomen, seine Auswirkungen und mögliche Gegenmaßnahmen besser zu verstehen. Zum Projekt: <https://www.bidt.digital/forschungsprojekt/algorithmische-darstellungsverzerrungen-aus-userperspektive-bewertung-auswirkungen-interventionen-adubai/>

„Wie „menschlich“ muss Gerechtigkeit sein? Psychologische Determinanten von Vertrauen in Ko-Kreation im Kontext Rechtssystem“ | Prof. Dr. Friederike Funk (LMU München)

Das Projekt analysiert psychologische Prozesse, die das Vertrauen in die Zusammenarbeit mit generativer KI fördern oder hemmen. Im Fokus stehen die richterliche Urteilsfindung, die Zuverlässigkeit von Zeugenaussagen und die Identifizierung von Tatverdächtigen. Zum Projekt: <https://www.bidt.digital/forschungsprojekt/wie-menschlich-muss-gerechtigkeit-sein-psychologische-determinanten-von-vertrauen-in-ko-kreation-mit-generativer-ki-im-kontext-rechtssystem-gerechte-ki/>

„Trustworthy Generative AI Copilots for Data Analytics in Business Decision-Making“ | Prof. Dr. Ulrich Gnewuch (Universität Passau)

Wie treffen Nutzende bei der Interaktion mit einem KI-Copiloten Entscheidungen? Das Projekt untersucht die Gestaltung vertrauenswürdiger generativer KI-Copiloten und deren Einfluss auf die Geschäftsentscheidungen von Nutzenden ohne technischen Hintergrund. Zum Projekt: <https://www.bidt.digital/forschungsprojekt/vertrauenswuerdi-ge-generative-ki-copiloten-fuer-datenanalysen-in-geschaefentlichen-entscheidungsprozessen-genaicopilot/>

„AI-Aided Design: Generative AI as a Co-Creation Tool“ | Prof. Dr.-Ing. Klaus Diepold (TU München)

Dieses Projekt erforscht, wie generative KI eingesetzt werden kann, um Designerinnen und Designer beim Entwurf von 3D Objekten zu unterstützen. Ziel ist ein co-kreativer Prozess zwischen Software und Mensch. Zum Projekt: <https://www.bidt.digital/forschungsprojekt/ki-gestuetzter-entwurf-generative-ki-als-werkzeug-zur-co-creation-ai-ad/>

„Human-Centered Specification-Driven Software Engineering with Generative AI“ | Prof. Dr. Albrecht Schmidt (LMU München)

Das Projekt soll die digitale Souveränität von Individuen und Gesellschaft stärken. Es basiert auf einem spezifikationsgetriebenen Ansatz für die Softwareentwicklung ohne Programmierkenntnisse und soll einen breiteren Zugang zur Gestaltung digitaler Technologien ermöglichen. Zum Projekt: <https://www.bidt.digital/forschungsprojekt/menschzentrierte-spezifikationsgetriebene-softwareentwicklung-mit-generativer-ki-ai-spec2sw/>

Zwei weitere Konsortialprojekte rund um Deepfakes in der Strafverfolgung und autoritäre KI

Ergänzend zu den Projekten des Forschungsschwerpunkts wurden im Rahmen der Jahresausschreibung zwei weitere Konsortialprojekte aus Bayern ausgewählt, die 2025 anlaufen werden. Die Vorhaben verstärken vorhandene Forschungsfelder und erweitern das Themenportfolio des bidt.

„For the Greater Good? Deepfakes in der Strafverfolgung“ | Prof. Dr. Lena Kästner (Universität Bayreuth), Prof. Dr. Niklas Kühl (Universität Bayreuth), Prof. Dr. Christian Rückert (Universität Bayreuth), OStA Miriam Margerie (ZAC NRW)
Der Einsatz moderner KI birgt auch für die Strafverfolgung enormes Potenzial. Doch unter welchen Umständen und in welchem Maße ist der Einsatz von Deepfakes gesellschaftlich akzeptabel? Das Projekt bezieht Strafrecht und Strafverfolgung, Wirtschaftsinformatik und Philosophie ein. Zum Projekt: <https://www.bidt.digital/forschungsprojekt/for-the-greater-good-deepfakes-in-der-strafverfolgung-fogg/>

„Autoritäre KI: Wie große Sprachmodelle (LLMs) an Russlands Propaganda angepasst werden“ | Prof. Dr. Florian Toepfl (Universität Passau), Prof. Dr. Andreas Jungherr (Universität Bamberg), Prof. Dr. Florian Lemmerich (Universität Passau)
Das Projekt wirft zwei Fragen auf: Wie und mit welchen Konsequenzen werden LLMs unter strenger Aufsicht und Zensur im heutigen Russland entwickelt? Und welche Auswirkungen haben autoritäre Daten, wenn sie in demokratische LLM-gestützte Systeme eingespeist werden? Zum Projekt: <https://www.bidt.digital/forschungsprojekt/autoritaere-ki-wie-grosse-sprachmodelle-llms-an-russlands-propaganda-angepasst-werden-ai-prop/>

Pressekontakt:

Leonie Liebich

Referentin Dialog

Tel.: +49 89 540 235 631

E-Mail: presse@bidt.digital

Über das bidt:

Das Bayerische Forschungsinstitut für Digitale Transformation (bidt) ist ein Institut der Bayerischen Akademie der Wissenschaften. Es trägt dazu bei, die Entwicklungen und Herausforderungen des digitalen Wandels besser zu verstehen. Damit liefert es die Grundlagen, um die digitale Zukunft im Dialog mit der Gesellschaft verantwortungsvoll und gemeinwohlorientiert zu gestalten. Das bidt fördert herausragende interdisziplinäre Forschung und liefert als Think

Tank Entscheidungsträgern in Politik und Wirtschaft evidenzbasierte Empfehlungen. Forschung findet am Institut im offenen Dialog zwischen Wissenschaft, Politik, Wirtschaft und Gesellschaft statt.

wissenschaftliche Ansprechpartner:

Dr. Christoph Egle

Geschäftsführer

Tel. +49 89 540 235 610

E-Mail: christoph.egle@bidt.digital

URL zur Pressemitteilung: <https://www.bidt.digital/forschungsprojekte/> Übersicht der bidt-Forschungsprojekte

URL zur Pressemitteilung: <https://www.bidt.digital/generative-ki/> Mehr zum Forschungsschwerpunkt "Mensch und generative Künstliche Intelligenz: Trust in Co-Creation"