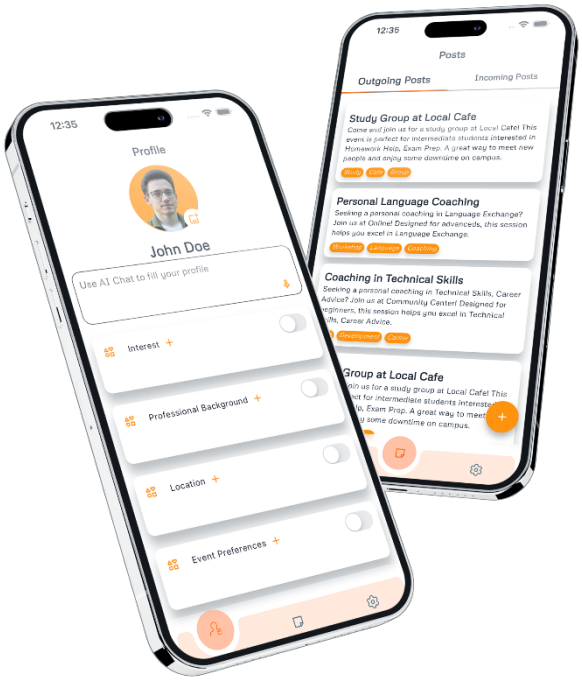
Lucii

link-up . create . interact . inspire



Inhalt

1. Einleitung und Problemstellung

Ziel des InduKo Projekts war es, Innovation durch Kollaboration zu fördern. Damit Kollaboration und Kommunikation überhaupt entstehen kann, müssen aber natürlich zuerst die passenden Personen zueinander gefunden haben.

Die Hochschullandschaft ist geprägt durch eine hohe Diversität: Studierende, Lehrende, Forschende, Verwaltungspersonal, Alumni, Unternehmenspartner sowie externe Expertinnen und Experten bringen jeweils unterschiedliche Kompetenzen, Interessen und Perspektiven mit, die ein enormes Potenzial für (interdisziplinäre) Zusammenarbeit und innovative Lösungsansätze bieten. Dieses Potenzial bleibt allerdings oft ungenutzt, weil geeignete Mechanismen fehlen, um die passenden Personen zusammenzubringen. Vielmehr existieren oft fragmentierte Ansätze, die nur isolierte Teile dieses Netzwerks ansprechen, wodurch viele fruchtbare Kooperationen unentdeckt bleiben. Ein großes Hindernis besteht darin, dass Personen mit ähnlichen Interessen oder komplementären Kompetenzen sich häufig nicht begegnen, selbst wenn sie an derselben Institution tätig sind. Interdisziplinäre Forschungs- und Innovationsprojekte oder studentische Gründungsinitiativen könnten beispielsweise erheblich von gezielten Verknüpfungen profitieren oder Studierende, die sich auf anspruchsvolle Prüfungen vorbereiten, könnten schneller Mitstreiter für spezialisierte Lerngruppen oder Nachhilfe finden. Alumni wiederum, die wertvolle berufliche Kontakte, Mentoring oder Praktikumsplätze vermitteln, stoßen häufig auf organisatorische Hürden, die den Kontakt zu potenziellen Interessierten erschweren. Auch Unternehmenspartner könnten durch eine effizientere Vernetzung leichter Nachwuchstalente identifizieren, deren Profile auf spezifische Positionen zugeschnitten sind. Der Nutzen einer leistungsfähigen Vernetzung reicht weit über die Grenzen einzelner Hochschulen hinaus.

Universitäten, Hochschulen und Forschungseinrichtungen verlassen sich zwar zunehmend auf digitale Plattformen, um den Austausch zwischen Lehrenden, Studierenden und Forschenden aus unterschiedlichen Disziplinen zu erleichtern, herkömmliche Kommunikationstools – etwa E-Mails, zentrale Messaging-Anwendungen oder von Institutionen verwaltete Foren – sind aber oftmals weder ausreichend geeignet noch datenschutzfreundlich und sie setzen den Focus eben viel weniger auf das anfängliche Zusammenführen von Personen, sondern mehr auf Kommunikation, wenn bereits eine Vernetzung stattgefunden hat. Gleichzeitig liegen die Herausforderungen aber auch in der Tatsache begründet, dass zentrale Plattformen in der Regel von den jeweiligen Institutionen betrieben werden, wodurch Teilnehmende die Kontrolle über ihre Daten abgeben. Aus verständlichen Gründen verzichten viele von öffentlichen Einrichtungen betriebenen Plattformen auf sehr weitreichende Profiling-Funktionen, weil die Verwaltung hochsensibler Daten eine Verantwortung und ein Haftungsrisiko birgt, das öffentliche Bildungseinrichtungen weder übernehmen möchten noch können – insbesondere angesichts potenzieller Schwachstellen in der IT-Sicherheit oder möglicher Cyberangriffe. Eine erfolgreiche Vernetzung erfordert aber gerade diese detaillierten Daten über Interessen, Kompetenzen, Forschungsschwerpunkte, Sprachkenntnisse, Erfahrungen und persönliche Eigenschaften der Beteiligten.

Eine weitere Herausforderung ist die Fragmentierung von Datenquellen. Aktuell tauschen Studierende sich z.B. informell in Chatgruppen aus, während Fachschaften und Initiativen ihre Ankündigungen über verschiedene Kanäle verteilen. Alumni-Netzwerke beschränken sich nicht selten auf Kontaktlisten und Unternehmen sind auf individuelle Kooperationen angewiesen, die häufig nicht die Breite des akademischen Netzwerks nutzen. Forschungsprojekte und -ergebnisse werden oft auf verschiedenen Plattformen veröffentlicht, darunter Projekt- und Institutswebseiten, wissenschaftliche Repositorien oder Konferenzseiten. Die Sichtbarkeit von Forschenden ist zusätzlich dadurch eingeschränkt, dass ihre Profile auf unterschiedlichen Plattformen wie Universitäts- oder Forschungswebseiten oder LinkedIn verteilt sind. Diese zersplitterte Informationslandschaft erschwert es, Personen oder Ressourcen effizient zusammenzubringen. Zum Teil werden aktuelle Projekte aber auch aus gutem Grund noch ganz bewusst an keiner Stelle veröffentlicht, weil die Arbeiten sich in einem frühen Stadium befinden. Dennoch könnte es gerade in solchen Fällen häufig sehr hilfreich sein, Mitstreiterinnen oder Mitstreiter für zum Beispiel eine Start-up-Idee oder ein Forschungsprojekt zu gewinnen.

Gleichzeitig herrscht nicht ohne Grund eine natürliche Skepsis gegenüber zentralen Plattformen, die alle Informationen bündeln könnten. Zentrale Lösungen sind anfällig für Datenschutzprobleme, weil eine einzige Verletzung der Sicherheit katastrophale Folgen haben kann. Ein zentrales System ist ohne Frage effizient, birgt jedoch das erhebliche Risiko, sensible Daten an Dritte zu verlieren. Insbesondere die Einbindung moderner KI-Technologien wie Large Language Models verschärft diese Problematik, weil solche Modelle Informationen im Parameterraum speichern und ein nachträgliches Löschen personenbezogener Daten nahezu unmöglich wird.

Das Spannungsfeld zwischen dem Bedürfnis nach Vernetzung und den Anforderungen an Datenschutz und Sicherheit ist komplex. Anzustreben wäre eine Lösung, die das enorme Vernetzungspotenzial im Hochschulkontext ausschöpft und gleichzeitig den Datenschutz und die Sicherheit der Beteiligten bestmöglich gewährleistet, um Kommunikation und Kollaboration zu fördern. Bevor wir diese Lösung vorstellen, die im Rahmen eines Teilprojektes des von der Stiftung für Innovation in der Hochschullehre geförderten InduKo-Projekts (Innovation durch Kollaboration) auf Basis des Matrix-Messaging-Protokolls entstand, werden im Folgenden zunächst Aspekte zentraler und dezentraler Lösungsansätze kurz dargestellt. Darauf aufbauend wird das Matrix-Protokoll als geeignete Kommunikationsarchitektur vorgestellt, das sich durch seine dezentrale, offene und föderierte Struktur auszeichnet. Neben dem Matrix-Messaging-Protokoll ist ein weiterer zentraler Baustein unserer Problemlösung das Konzept von AI on Edge, bei dem die KI direkt auf den Endgeräten ausgeführt wird, um den zentralen Datenfluss zu minimieren und die Privatsphäre der Nutzenden zu schützen. Mit der von uns entwickelten mobilen Anwendung „Lucii“ (link-up – create – interact– inspire) wird schließlich veranschaulicht, wie diese Konzepte in die Praxis umgesetzt werden können, um die angesprochenen Herausforderungen mit modernen Tools bestmöglich zu adressieren.

2. Grundlagen

2.1 Zentrale vs. dezentrale Lösungen: Sicherheit, Vertrauen, Effizienz.

Ein zentraler Lösungsansatz des skizzierten Problems würde ein System voraussetzen, das alle relevanten Daten über die Akteurinnen und Akteure an einem Ort sammelt, verarbeitet und verwaltet. In einem solchen Szenario könnten Nutzende detaillierte Profile erstellen, die Angaben zu Interessen, Kompetenzen und anderen persönlichen Eigenschaften enthalten. Diese Profile würden durch eine zentrale Instanz gesammelt, die sie mit Hilfe von KI-Algorithmen analysiert, um passende Personen, Gruppen oder Projekte miteinander zu verbinden. Der Vorteil eines solchen Ansatzes liegt in seiner Effizienz. Zentrale Plattformen können große Datenmengen analysieren, präzise Vorschläge erstellen und durch die Nutzung umfassender Modelle wie Large Language Models ein hochpersonalisiertes Matching ermöglichen. Dank der heute verfügbaren Technologien würde die Umsetzung eines solchen zentralisierten Systems durch zahlreiche bewährte technische Lösungen bzw. Werkzeuge gut unterstützt. Cloud-Computing-Dienste, vortrainierte KI-Modelle und etablierte, sehr leistungsfähige Plattformstrukturen von verschiedenen großen Anbietern bieten eine robuste Grundlage, um solche Systeme mit überschaubarem Aufwand effizient und skalierbar zu implementieren.

Doch die Risiken sind erheblich, eine zentrale Datenhaltung macht das System z.B. interessanter für Angriffe von außen. Cyberkriminelle könnten auf persönliche Informationen zugreifen, diese manipulieren oder für unlautere Zwecke verwenden. Zudem birgt ein solches System auch Risiken durch internen Missbrauch, etwa durch unbefugte Mitarbeiterinnen und Mitarbeiter, die Zugriff auf die sensiblen Daten haben. Ebenso wirft die Speicherung von Daten in zentralisierten KI-Modellen das bereits angesprochene Problem auf: Einmal in ein Modell eingeflossene Daten können nicht ohne Weiteres entfernt werden. Die diffuse Speicherung von Informationen im Parameterraum neuronaler Netzwerke verhindert eine einfache und zielgerichtete Löschung, so wie das bspw. beim Löschen eines Datensatzes aus einer SQL-Datenbank der Fall wäre. Zudem müssen die Anwender und Anwenderinnen dem Plattformbetreiber vertrauen, auch wenn die entsprechenden Datenschutzvereinbarungen sehr gut im Interesse der Nutzenden formuliert sind, dass diese Vereinbarung auch tatsächlich eingehalten wird und keine juristischen Schlupflöcher bietet. Diese fehlende Transparenz und Autonomie untergräbt das Vertrauen in die Plattform und kann zu einer Zurückhaltung bei der Preisgabe von Informationen führen. Das Risiko einer dadurch tendenziell schlechteren Datenqualität ist nicht zu unterschätzen: Wenn Nutzende aufgrund von Datenschutzbedenken bzw. mangelndem Vertrauen zögern vollständige oder korrekte Informationen anzugeben, mindert das die Effektivität des Systems signifikant und kann so zu schlechteren oder falschen Empfehlungen führen.

Zentrale Systeme konzentrieren den gesamten Datenverkehr und die Steuerung an einem Ort, was sie nicht nur anfälliger für Angriffe, sondern auch für Ausfälle macht. Sobald ein Kernsystem gestört ist oder ein Sicherheitsvorfall eintritt, hat dies unmittelbare Auswirkungen auf sämtliche Nutzende und Stakeholder. In einer solchen zentralen Architektur ergibt sich ein Single Point of Failure. Dezentrale Ansätze verteilen dagegen die Verantwortung und lasten das System auf viele Komponenten

verteilt aus, was die Gesamtresilienz erhöht, da kein einzelner Ausfall automatisch zu einem Zusammenbruch des gesamten Ökosystems führt.

Auch aus rechtlicher Perspektive wirft ein zentraler Ansatz komplexe Fragen auf. Es stellt sich die grundsätzliche Frage, wem die in das System eingespeisten Daten „gehören“ bzw. wer die Eigentumsrechte daran hält. Darüber hinaus kann die Verantwortung für Fehler, Datenmissbrauch oder Verstöße gegen rechtliche Vorgaben zu einer erheblichen Belastung werden, die insbesondere öffentliche Einrichtungen davon abhalten könnte, ein zentrales System zu betreiben.

Ein weiterer zentraler Aspekt betrifft die Zugriffskontrolle. In einem komplexen Hochschulumfeld mit zahlreichen Organisationseinheiten, Fachbereichen, Projekten und externen Partnern ist die Festlegung von Rechten, Rollen und Zuständigkeiten eine anspruchsvolle Aufgabe. Die Kontrolle darüber, wer welche Informationen einsehen, verarbeiten und weitergeben darf, gestaltet sich in einem zentralen, monolithischen System i.d.R. ausgesprochen aufwendig. Diese Herausforderung wird noch verschärft, wenn sich die Rahmenbedingungen, personellen Zuständigkeiten oder rechtlichen Vorgaben im Laufe der Zeit ändern.

Insgesamt führt die zentrale Konzentration von Daten zu einem Gefüge, das zwar technisch effizient, gut umsetzbar und leicht skalierbar ist, jedoch in vielerlei Hinsicht erhebliche Risiken, Abhängigkeiten und Unsicherheiten birgt. Gerade im sensiblen Hochschulkontext, der durch hohe Diversität, Autonomie und Innovation gekennzeichnet ist, wiegen diese Nachteile schwer. Letztlich führt die Architektur eines zentralen Systems zu einer inhärenten Schwäche, die sich nicht allein durch technische Sicherheitsmaßnahmen beheben lässt. Ein dezentraler Ansatz löst viele dieser Probleme, indem er die Verarbeitung, Speicherung und Kontrolle personenbezogener Daten vollständig auf die Endgeräte der Nutzerinnen und Nutzer verlagert. Dadurch entfällt z.B. ein zentraler Datenpool als Single Point of Failure, denn es existiert keine konzentrierte Instanz, auf die sich „lohnenswerte“ Angriffe konzentrieren oder deren Ausfall das gesamte System lahmlegen könnte. Zudem behält jede Person die volle Kontrolle über die eigenen Daten, kann Änderungen vornehmen oder Löschungen durchführen, ohne auf eine zentrale Stelle angewiesen zu sein. Damit entfallen nicht nur zentrale Schwachstellen wie die Anfälligkeit für Ausfälle und Angriffe, die begrenzte Datenhoheit, rechtliche Unsicherheiten, komplexe Zugriffsstrukturen sowie das Problem, einmal in ein zentrales LLM eingeflossene personenbezogene Informationen nur schwer entfernen zu können. Vielmehr entsteht ein Umfeld, in dem Datenschutz, Privatsphäre und Autonomie systeminhärent gestärkt werden und damit die Grundlage für ein nachhaltig vertrauenswürdiges und flexibel anpassbares Vernetzungsökosystem geschaffen wird.

Sicherheit und Vertraulichkeit müssen konsequent auch den Transport der Daten umfassen, um ein vertrauenswürdiges System zu gewährleisten. Ende-zu-Ende-Verschlüsselung ist hierbei von zentraler Bedeutung, da sie sicherstellt, dass nur die beteiligten Endgeräte Zugriff auf die entschlüsselten Inhalte haben. Selbst Server, die die Übertragung durchführen, dürfen die Daten selbstverständlich nicht einsehen können. Das Matrix-Messaging-Protokoll bietet mit seiner Ende-zu-Ende-

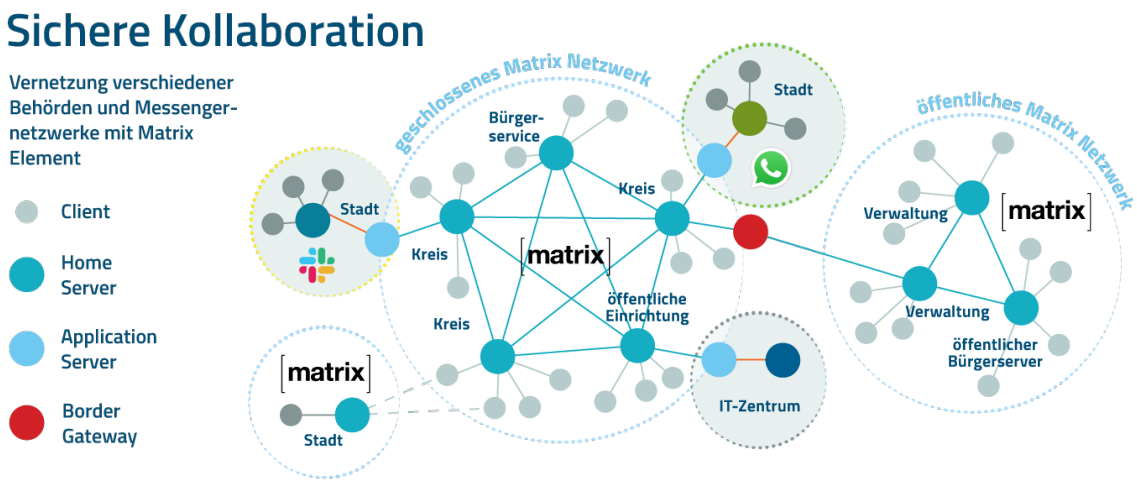
verschlüsselten Übertragung und seiner bewährten Kommunikationsarchitektur eine hervorragende Grundlage für sichere, vertrauenswürdige und skalierbare Vernetzungslösungen.

2.2 Das Matrix-Protokoll

Das Matrix-Protokoll hat sich zu einer etablierten Grundlage für dezentrale, sichere und flexible Echtzeitkommunikation entwickelt. Es bietet eine offene Infrastruktur, die auf föderierten Netzwerken und interoperablen Schnittstellen basiert. Dieser Ansatz fördert Datenschutz, Datenhoheit und eine langfristige Skalierbarkeit. Genau deshalb ist Matrix für Einzelpersonen, Unternehmen, Regierungen sowie Non-Profit-Organisationen gleichermaßen interessant.

Technisch basiert das Protokoll auf einer REST-ähnlichen JSON-API, die den standardisierten Austausch zwischen Clients und Servern sicherstellt. Matrix-Räume fungieren als grundsätzliche Kommunikationseinheiten und können Textnachrichten, Dateien, Sprachanrufe, Videokonferenzen oder Bots aufnehmen. Dank ihrer Flexibilität lassen sich damit einzelne Unterhaltungen, ganze Communitys oder sogar komplette Organisationsstrukturen abbilden. Das Matrix-Messaging-Protokoll bietet Ende-zu-Ende-Verschlüsselung direkt auf den Endgeräten, sodass nur die beteiligten Personen die Inhalte lesen können. Allerdings ist E2EE nicht zwingend für alle Räume voreingestellt. Offene, öffentliche Räume werden teilweise bewusst unverschlüsselt betrieben, damit sie leichter auffindbar sind und mehr Menschen sich unkompliziert beteiligen können.

„Föderation“ ist ein zentrales Element von Matrix. Sie sorgt für Ausfallsicherheit, Skalierbarkeit und verhindert eine Abhängigkeit von zentralen Stellen. Jeder kann selbst entscheiden, ob er einen eigenen Server betreibt oder einen bestehenden nutzt, ohne dadurch die Möglichkeit einzubüßen, mit Teilnehmenden auf anderen Instanzen zu kommunizieren. Für Organisationen ergibt sich daraus eine Plattform, die sowohl intern als auch extern einsetzbar ist. Ein selbstgehosteter Matrix-Server bietet Kontrolle über die eigenen Daten. Sensible Informationen lassen sich in der eigenen IT-Infrastruktur speichern, was vor allem für Unternehmen, Behörden und Non-Profit-Organisationen wichtig ist, die strenge Datenschutzauflagen oder gesetzliche Vorgaben einhalten müssen. Mittlerweile wird Matrix vermutlich auch genau deshalb verstärkt in staatlichen Kontexten eingesetzt. Ein Beispiel ist die französische Regierung, die mit „Tchap“ eine eigene Kommunikationslösung auf Basis von Matrix betreibt. In Deutschland wurde unter der Verantwortung der BWI – einem bundeseigenen IT-Dienstleister, der vor allem für die Bundeswehr und andere Behörden tätig ist – der „BundesMessenger“ ins Leben gerufen. Dieses Werkzeug basiert ebenfalls auf dem Matrix-Protokoll, ist speziell auf die Anforderungen der öffentlichen Verwaltung zugeschnitten und der darauf basierende Bundesmessanger für Behörden ist ohne zusätzliche Gebühren nutzbar. Darüber hinaus nutzt die gematik, die für die Digitalisierung des Gesundheitswesens zuständig ist, das Protokoll für den sogenannten TI-Messenger. Dieses Projekt zielt darauf ab, eine sichere und interoperable Kommunikation im Gesundheitsbereich zu ermöglichen und gleichzeitig den hohen Ansprüchen an Datenschutz und Datensicherheit gerecht zu werden.



Quelle: https://nordeck.net/matrix-loesungen/

2.3 AI on Edge

Leistungsfähige KI-Anwendungen wurden, vor allem aufgrund der benötigten Hardware, lange fast ausschließlich in zentralen Rechenzentren bzw. eben nicht auf mobilen Endgeräten ausgeführt. In den vergangenen Jahren ist jedoch ein Trend hin zu „AI on Edge“ bzw. „Edge AI“ zu beobachten. Darunter versteht man KI-Verarbeitung direkt auf dem Endgerät selbst, also bspw. auf einem Smartphone oder Tablet. Neben Vorteilen, wie bspw. der Fähigkeit zur Offline-Nutzung, ist eine der ganz entscheidenden Triebfedern für diesen Wandel der gestiegene Wunsch nach mehr Privatsphäre: Wenn sämtliche Berechnungen lokal ablaufen, müssen keine sensiblen Daten an entfernte Server geschickt werden, was das Risiko von Datenlecks reduziert und das Vertrauen der Anwenderinnen und Anwender in die jeweilige KI-Anwendung stärkt. Da die Modelle somit direkt auf dem Endgerät laufen, können sie auch besser auf individuelle Vorlieben und Nutzungsgewohnheiten abgestimmt werden, ohne diese Informationen dauerhaft an externe Dienste zu übermitteln. Hinzu kommt, dass lokale Datenverarbeitung spürbar an Sicherheit gewinnt: Selbst bei einer Kompromittierung äußerer Systeme verbleiben die sensiblen Informationen in der Hand der Nutzerin bzw. des Nutzers.

Möglich wurde diese Entwicklung durch immer leistungsfähiger werdende moderne Endgeräte, die zunehmend in der Lage sind, selbst komplexe KI-Modelle auszuführen. Sowohl Apple als auch Google investieren bspw. in entsprechende Hardware- und Softwarelösungen, die das maschinelle Lernen auf mobilen Prozessoren ermöglichen. Apple verfolgt diesen Ansatz mit „Apple Intelligence“; diese KI-gestützten Funktionen ermöglichen es, Aufgaben wie das Verfassen und Zusammenfassen von Texten direkt auf dem Gerät durchzuführen. Auch Google hat ebenfalls bedeutende Fortschritte im Bereich Edge AI gemacht. Mit der Einführung von Gemini Nano und dem Google AI Edge SDK können Entwickler generative KI-Funktionen direkt in mobile Anwendungen integrieren. Diese Tools

ermöglichen es ebenfalls, bspw. Aufgaben wie Textumformulierung, intelligente Antworten und Textzusammenfassungen, lokal auf unterstützten Geräten auszuführen. Diese Entwicklungen und das verstärkte Engagement führender Unternehmen verdeutlichen die wachsende Bedeutung von AI on Edge. Da Geräte immer leistungsfähiger werden, ist zu erwarten, dass der Trend zur lokalen Ausführung von KI-Modellen weiter zunimmt und neue Möglichkeiten für personalisierte, sichere Anwendungen eröffnet.

3. Lucii: Link-up – create – interact - inspire

3.1 Überblick – Funktionsweise

Die zuvor beschriebenen Konzepte finden in der mobilen Anwendung Lucii ihre praktische Umsetzung. Lucii wurde speziell entwickelt, um Kollaboration zu fördern, ohne sensible persönliche oder institutionelle Daten an zentrale Server weiterzugeben. Die Architektur sorgt dafür, dass Profile und Berechnungsergebnisse lokal gespeichert werden; kritische Berechnungen finden in geschützten Speicherbereichen statt, in denen weder Verschlüsselungs-Keys noch Zwischenergebnisse von außen eingesehen werden können. Durch den Verzicht auf zentrale Datenspeicherung verringert Lucii das Risiko von Datenlecks, unbefugtem Zugriff und Verstößen gegen Datenschutzrichtlinien wie der DSGVO. Im Kern setzt Lucii auf „lightweight neural network models“, die speziell für ressourcenarme Umgebungen wie Mobilgeräte oder IoT-Knoten optimiert sind. Die Entwicklung verlief schrittweise von „Convolutional Neural Networks (CNNs)“ über anspruchsvollere Transformer-Architekturen „DistilBERT“ bis hin zu optimierten Frameworks wie „Gemini Nano“. Für die Kommunikation nutzt Lucii das Matrix-Messaging-Protokoll, auf diese Weise lassen sich Nachrichten sicher und skalierbar austauschen.

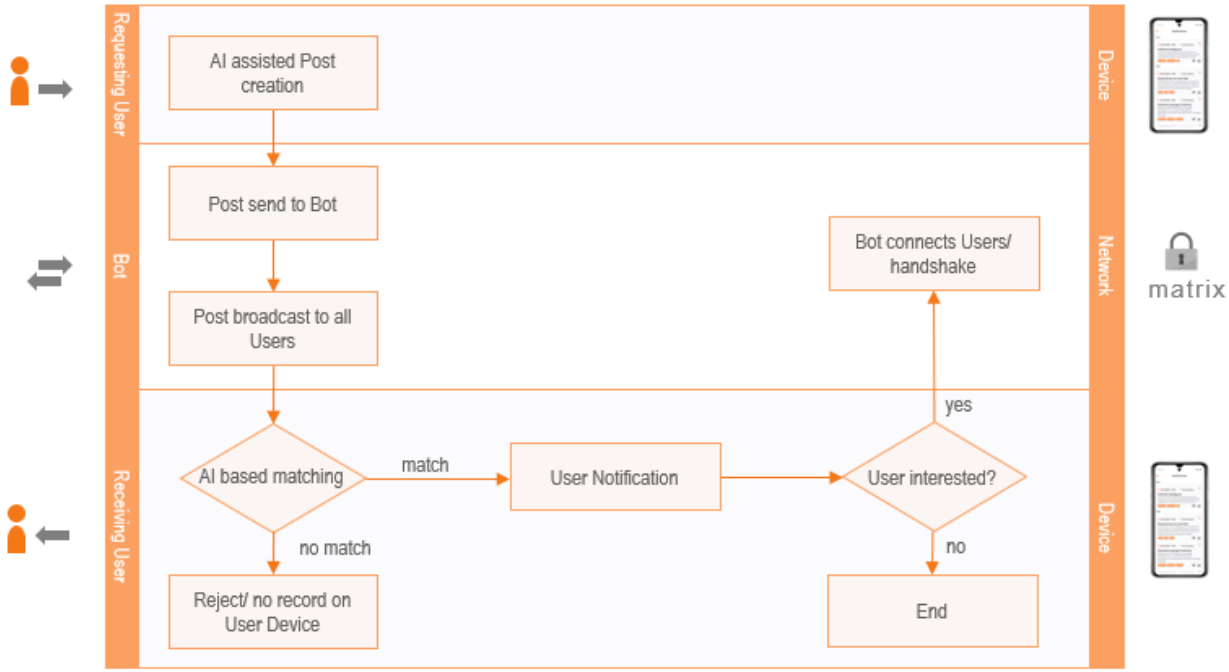
Die praktische Umsetzung des Matchings erfolgt in mehreren Stufen. Nutzende legen zunächst ihre Profildaten lokal auf ihrem Gerät ab, darunter bspw. Kompetenzen, Interessen und Projektschwerpunkte. Auch die Profilerstellung selbst wird dabei durch KI unterstützt – die Nutzenden geben einen Beschreibungstext ein und die KI auf dem Gerät extrahiert die passenden Schlüsselbegriffe; diese Daten können auch sehr sensible Informationen umfassen, weil diese niemals das Gerät verlassen, lediglich sogenannte „Posts“ werden über das Matrix-Protokoll verschlüsselt und anonymisiert ins Netzwerk eingespeist.

Jede Person kann einen Post erstellen, um ein bestimmtes Anliegen zu formulieren – sei es die Suche nach einem Forschungspartner, die Suche einer Lerngruppe, ein Hinweis auf einen Workshop oder das Angebot einer Zusammenarbeit. Diese Posts sind also digitale Einträge, die von Nutzenden erstellt werden, um Anfragen an andere Nutzerinnen und Nutzer zu senden. Sie übernehmen damit sozusagen die Funktion eines Aushangs am schwarzen Brett, nur mit dem entscheidenden

Unterschied, dass sie weder öffentlich einsehbar sind noch persönliche Kontaktinformationen enthalten und mit AI on Edge ein effizientes Matching sichergestellt wird.

Das Matrix-Protokoll sorgt dafür, dass Posts sicher und dezentral verbreitet werden. Ob ein Post zu einer Person passt, entscheidet die lokale KI, die den Post mit den auf dem Gerät gespeicherten Profildaten abgleicht. Dabei kommen die beschriebenen AI-Modelle zum Einsatz, die semantische Zusammenhänge zwischen den Posts und den Interessen der Nutzenden erkennen. Dadurch werden auch thematisch verwandte Einträge erkannt, selbst wenn diese unterschiedlich formuliert sind. Wenn hierbei von der KI eine Übereinstimmung festgestellt wird, erhält die betroffene Person eine Benachrichtigung über den relevanten Post, ohne, dass die Erstellerin oder der Ersteller davon erfährt. Anders als in klassischen Netzwerken oder Plattformen, wo öffentliche Listen durchsucht oder direkte Kontaktanfragen gestellt werden, funktioniert Lucii nach dem Prinzip, dass ausschließlich thematisch relevante Posts per Benachrichtigung angezeigt werden. Erst wenn beide Seiten ihr Interesse bestätigen, können sie schrittweise weitere Informationen in einem anonymen Chat austauschen.

Durch dieses System ersetzt Lucii herkömmliche Methoden der Kontaktaufnahme durch ein intelligentes, anonymes und dezentrales Matching über Posts. Es verbindet das in die Jahre gekommene Prinzip eines Schwarzen Bretts mit modernen KI-Technologien und verschlüsselter Kommunikation, sodass Nutzende zielgerichtete Verbindungen knüpfen können, ohne persönliche Informationen preisgeben zu müssen.



Matching – Schematische Darstellung

Lucii veranschaulicht damit, wie sich fortschrittliche Technologien wie das Matrix-Protokoll und On-Device-KI wie bspw. DistilBERT oder Gemini Nano erfolgreich kombinieren lassen, um ein datenschutzfreundliches, sicheres und effizientes Vernetzungstool im Hochschulumfeld zu schaffen.

3.2 Entwicklungsphasen – KI-Modelle

Der Entwicklungsprozess der App Lucii verlief wie bereits erwähnt über mehrere Stufen. Jede Phase baute auf den in der vorherigen Phase gewonnenen Erkenntnissen auf und behebt die dort erkannten Schwachstellen. Auf diese Weise konnten Genauigkeit, Geschwindigkeit (Latenz) und Speicherbedarf schrittweise optimiert werden. Dieses Kapitel beschreibt die jeweiligen Modellaufbauten, Trainingsmethoden und Optimierungsmaßnahmen. Zudem werden die Kennzahlen vorgestellt, die zur Bewertung und zum Vergleich der einzelnen Entwicklungsstufen herangezogen wurden. Für das Training und die Evaluierung des Modells wurden synthetische Datensätze verwendet, die mit GPT-3 generiert wurden. Diese simulieren vielfältige Nutzerprofile und Ereignisse, sodass keine echten persönlichen Daten erfasst, gespeichert oder verarbeitet werden mussten. Die Bewertung fokussiert auf Metriken, die die Leistungsfähigkeit des Systems in der Echtzeit-Profilerstellung und der Zuordnung von Ereignissen zu Usern („matching“) messen. Der F1-Score wird genutzt, um das Gleichgewicht zwischen Präzision (Precision) und Vollständigkeit (Recall) zu erfassen. Recall misst dabei speziell die Fähigkeit des Systems, alle relevanten Ereignisse korrekt zu identifizieren. Die Latenz wird analysiert, um sicherzustellen, dass die Systemleistung den Anforderungen für den Echtzeitbetrieb entspricht. Zusätzlich wird eine Ähnlichkeitsmatrix verwendet, um die Genauigkeit der Erstellung von Nutzerprofilen und des Matching zu bewerten.

Phase 1: Grundlegendes CNN-Modell

Im ersten Schritt wurde ein Basis-CNN entwickelt, das grundlegende Aufgaben der Nutzerprofilierung und Ereigniszuordnung erfüllt. Obwohl Convolutional Neural Networks typischerweise in der Bildverarbeitung eingesetzt werden, können sie mithilfe von Faltungskernen, die auf Wortembeddings angewendet werden, auch relevante Textmuster erkennen.

In dieser Phase kam ein synthetischer Datensatz zum Einsatz, der unter anderem Nutzerinteressen, Standortvorlieben, Veranstaltungstypen und berufliche Spezialisierungen abbildete. Um aussagekräftige Sprachrepräsentationen zu erhalten, wurden vortrainierte Embeddings (z. B. Word2Vec) genutzt. Zu den wichtigsten Schritten der Datenvorverarbeitung zählten die Tokenisierung, das Entfernen von Stoppwörtern sowie Stemming bzw. Lemmatisierung.

Die CNN-Architektur umfasste eine Embedding-Schicht, die den Text in Feature-Vektoren überführte, gefolgt von zwei Convolutional Layers mit ReLU-Aktivierung, die lokale Textmuster in den tokenisierten Sequenzen extrahierten. Anschließend kamen Pooling-Schichten zum Einsatz, um die

Dimensionen zu reduzieren und relevante Merkmalsausprägungen hervorzuheben. Eine oder mehrere vollständig vernetzte Schichten (Fully Connected Layers) dienten der eigentlichen Klassifikation und gaben über eine Softmax-Aktivierung Wahrscheinlichkeiten für unterschiedliche Klassen aus.

Architektur des in Phase 1 verwendeten CNN Modells

Trainiert wurde das Modell mit dem Adam-Optimizer und der kategorischen Cross-Entropy als Verlustfunktion. Zudem wurde ein Early-Stopping-Mechanismus implementiert, um Overfitting zu verhindern. Die Hyperparameter (z. B. Anzahl und Größe der Filter, Lernrate, Batch-Größe) wurden gezielt feinabgestimmt, was letztlich zu einer moderaten Genauigkeit führte.

Category	**Accuracy**
Interests	84%
Location Preferences	78%
Event Preferences	80%
Professional Background	79%
Social Preferences	82%

Phase 1 CNN Performance Metrics

Der Rechenaufwand und die Modellgröße blieben auf einem akzeptablen Niveau. Die durchschnittliche Inferenzzeit betrug 1,2 Sekunden; die Modellgröße lag bei 5 MB, wodurch es mit mittelklassigen Geräten kompatibel war. Allerdings hatte das Modell Schwierigkeiten mit nuancierten natürlichen Spracheingaben und zeigte eine begrenzte Anpassungsfähigkeit.

Diese erste Version diente als Referenzpunkt. Die hierbei erkannten Defizite im Umgang mit komplizierten sprachlichen Strukturen gaben den Ausschlag, in der nächsten Phase auf ein leistungsfähigeres Verfahren zu wechseln, das Kontextinformationen besser abbilden kann.

Phase 2: DistilBERT-Ansatz mit verbesserter Sprachverarbeitung

Um die in Phase 1 identifizierten Einschränkungen des CNN-Modells zu überwinden, wurde in Phase 2 DistilBERT eingeführt – eine komprimierte, transformerbasierte Architektur, die auf BERT (Bidirectional Encoder Representations from Transformers) aufbaut. DistilBERT bietet einen guten Kompromiss zwischen Rechenaufwand und semantischer Genauigkeit und ermöglicht eine deutlich verbesserte Erkennung von komplexen Spracheingaben direkt auf dem Endgerät.

Im Gegensatz zum vorherigen CNN-Ansatz erfordert DistilBERT eine aufwendigere Vorverarbeitung. Die Eingangsdaten werden zunächst per WordPiece-Tokenisierung zerlegt, wobei spezielle Symbole wie [CLS] und [SEP] Satz- bzw. Abschnittsgrenzen markieren. Anschließend berechnet DistilBERT kontextuelle Einbettungen, die bidirektionale Abhängigkeiten erfassen. Eine nachgeschaltete Klassifikationsschicht (eine oder mehrere Fully Connected Layers) verarbeitet diese Einbettungen und gibt Klassifikationswerte (Logits) für die Vorhersage aus.

Architektur des in Phase 1 verwendeten DistilBERT-Modells

Zu Beginn wurde DistilBERT auf synthetischen Datensätzen feinjustiert, die mithilfe von GPT-3 generiert wurden und vielfältige Profile von Forschenden sowie Eventbeschreibungen simulieren. Um das Modell für den Einsatz auf mobilen und eingebetteten Geräten zu verkleinern, kamen Pruning- und Quantisierungstechniken zum Einsatz, wodurch die Modellgröße von rund 330 MB auf etwa 30 MB reduziert werden konnte. Anschließend erfolgte die Konvertierung ins TFLite-Format für ressourcenschonendere Inferenz.

Trotz dieser Optimierungen sind die Rechenanforderungen von DistilBERT höher als beim ursprünglichen CNN, was besonders auf Geräten mit geringer Leistungsfähigkeit spürbar ist.

Category	Accuracy
Interests	92%
Location Preferences	88%
Event Preferences	90%
Professional Background	90%
Social Preferences	87%

Phase 2 (DistilBERT) Performance Metrics

Die Übereinstimmung stieg auf 99,5 % und die durchschnittliche Inferenzzeit betrug 2,5 Sekunden. Obwohl das Modell eine höhere Genauigkeit erzielte, erschwerten seine höheren Rechenanforderungen den Einsatz auf leistungsschwächeren Geräten und die statischen Aktualisierungen schränkten die Echtzeit-Anpassungsfähigkeit ein. Trotz der klaren Fortschritte im Hinblick auf Sprachverständnis und Klassifikationsgüte zeigen sich damit neue Grenzen bei den Anforderungen an Rechenleistung und Speicherbedarf, die in den folgenden Entwicklungsschritten weiter optimiert werden sollen.

Phase 3: Optimierung mit dem Gemini Nano Model

Phase 3 beschäftigte sich mit der Optimierung der Lucii-Anwendung für den Einsatz unter realen Bedingungen. Dabei wurde Gemini Nano, ein für Edge-Computing entwickeltes Leichtgewichtsmodell, integriert, um eine hohe Klassifikationsgenauigkeit bei reduzierten Rechenanforderungen und kürzerer Latenzzeit zu gewährleisten. Das Ziel bestand darin, das KI-Modell auf mobilen Geräten zuverlässig auszuführen, ohne dabei Abstriche bei der Profilierungsgenauigkeit in Kauf nehmen zu müssen. Gemini Nano vereint konvolutionale und transformerbasierte Komponenten in einer hybriden Architektur, was eine effiziente Textklassifikation ermöglicht. Durch Quantisierung und Pruning ließ sich die Modellgröße auf 10 MB reduzieren, während Knowledge Distillation mit DistilBERT als Teacher-Modell und Gemini Nano als Student zum Einsatz kam und mithilfe von Kreuzentropie und Kullback-Leibler-Divergenz gesteuert wurde.

Architektur des in Phase 1 verwendeten DistilBERT-Modells

Im Rahmen dieser Phase wurde außerdem ein interaktiver Feedback-Mechanismus eingeführt, der Nutzerinnen und Nutzern erlaubt, ihre Profile mithilfe natürlicher Spracheingaben zu verfeinern. Die auf dem Matrix-Protokoll basierende sichere Kommunikation stellt sicher, dass Matching-Ergebnisse in Echtzeit übermittelt werden können.

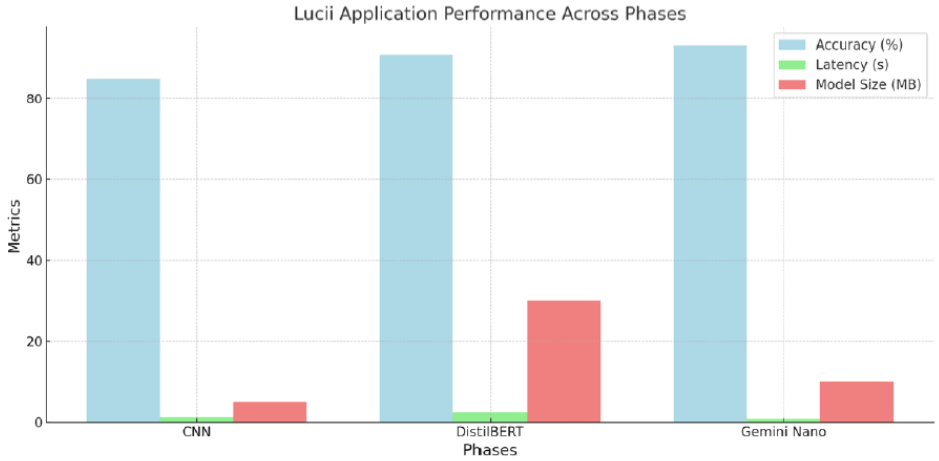
Category	Accuracy
Interests	95%
Location Preferences	93%
Event Preferences	94%
Professional Background	92%
Social Preferences	91%

Phase 3 Gemini Nano Performance Metrics

Die Match-Gewissheit stieg auf 99,56 %, während die durchschnittliche Inferenzzeit auf 0,8 Sekunden sank. Diese Ergebnisse zeigen, dass Gemini Nano die Herausforderungen der vorherigen Phasen erfolgreich bewältigt und auch auf mobilen Geräten eine hohe Leistung bietet.

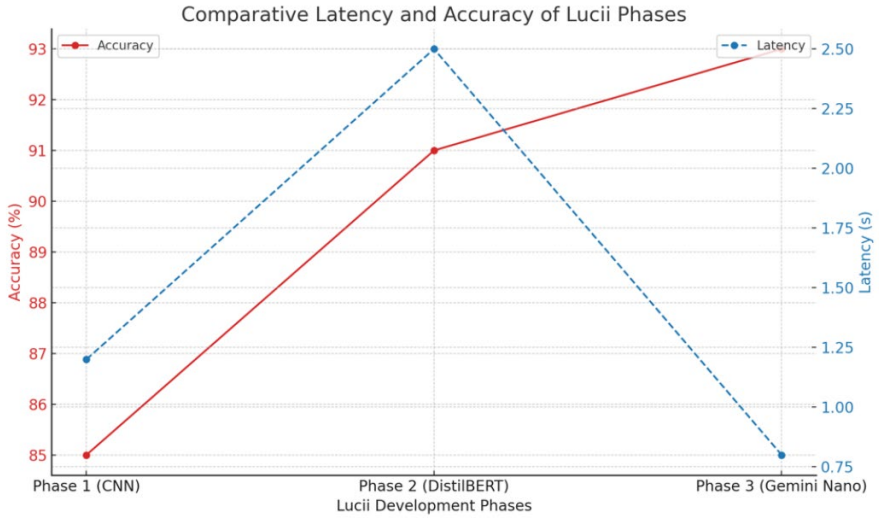
Ergebnisse über die Entwicklungsphasen hinweg

Das Gemini-Nano-Modell zeigte das beste Gleichgewicht zwischen Genauigkeit und Effizienz und eignet sich daher besonders für den Einsatz direkt auf dem Gerät.



Metric	Phase 1 (CNN)	Phase 2 (DistilBERT)	Phase 3 (Gemini Nano)
Accuracy	85%	91%	93%
Latency (s)	1.2	2.5	0.8
Model Size (MB)	5	30	10

Die experimentelle Evaluation des Frameworks belegt dessen Effektivität durch eine hohe Genauigkeit und geringere Latenzzeiten.



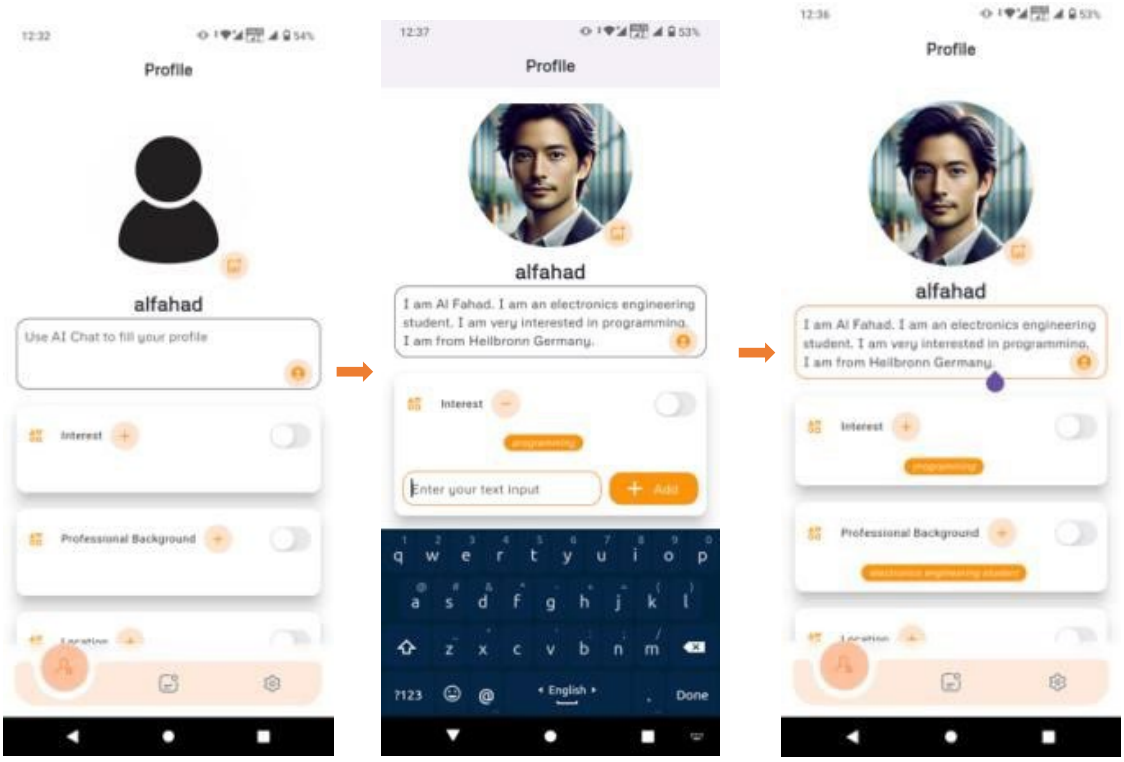
4. Lucii: Umsetzung und User Journey

Die iterative Entwicklung von Luciis KI-Modellen, am Ende mit der Implementierung von Gemini Nano, ermöglicht es dem System, direkt auf den Endgeräten der Nutzerinnen und Nutzer ein datenschutzfreundliches und gleichzeitig insgesamt treffsicheres, reaktionsschnelles Event-Matching bereitzustellen. Parallel dazu wurden die Kommunikationsinfrastruktur und die Sicherheitsmaßnahmen weiterentwickelt; zudem wurde ein interaktiver Feedback-Prozess eingeführt. In kontrollierten Tests wurde zudem die föderierte Kommunikation über das Matrix-Protokoll erprobt. Externe Datenübertragungen erfolgen dabei über SSL und HTTPS, während anonymisierte Tokens und temporäre IDs sicherstellen, dass keine persönlich identifizierbaren Informationen (PII) preisgegeben werden.

In der finalen Version von Lucii umfasst die User Journey alle im Kern wesentlichen Funktionen: Vom Anlegen eines Profils über die Erstellung von Posts und eine interaktive Nachrichtenverwaltung werden sämtliche Schritte eines realitätsnahen Workflows abgedeckt.

Profilerstellung

Die Benutzerinnen und Benutzer beginnen damit, ihre Interessen, ihren beruflichen Hintergrund, ihren Standort und ihre Vorlieben in natürlicher Sprache einzugeben. Gemini Nano verarbeitet diese Angaben lokal, extrahiert relevante Schlüsselwörter und speichert sie sicher ab.



Leeres Profil *Texteingabe/* Schlüsselworterkennung Profil

Die Genauigkeit bei der Schlüsselworterkennung und Kategorisierung lag im Durchschnitt bei 93 % und entspricht damit den Leistungsdaten des finalen Gemini-Nano-Modells. Auf mittelklassigen Smartphones blieb die Latenz dabei unter einer Sekunde, was die für eine nahtlose Benutzererfahrung erforderliche Reaktionsfähigkeit gewährleistet.

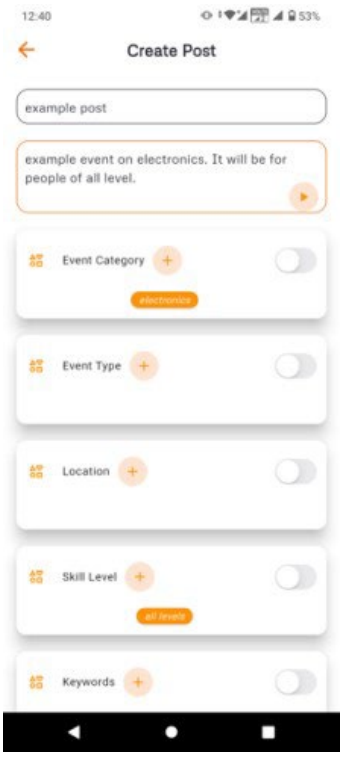
Der linke Screenshot zeigt das leere Profil mit den vordefinierten Kategorien wie bspw. Interessen und beruflicher Hintergrund. In der Mitte wird deutlich, wie das System durch die Eingabe eines Textes, zum Beispiel einer bereits vorhandenen Beschreibung aus einem Jobportal die man per Copy und Paste einfügen könnte, automatisch relevante Schlüsselwörter direkt auf dem Gerät extrahiert. Dieses Vorgehen ist nicht nur eine Arbeitserleichterung, sondern es führt insbesondere zu konsistenteren Eingaben und einer höheren Datenqualität. Tippfehler oder unklare Formulierungen durch manuelle Eingaben können so durch diese Art der KI-basierten, systematischen Kategorisierung weitgehend vermieden werden; das ist ein bedeutender Vorteil gegenüber rein manuellen Eingabeverfahren und trägt zur Zuverlässigkeit des gesamten Systems bei. Profilinformationen können temporär über entsprechende Schalter auch deaktiviert werden, falls derzeit ein Matching im Blick auf diese Kategorien nicht gewünscht wird.

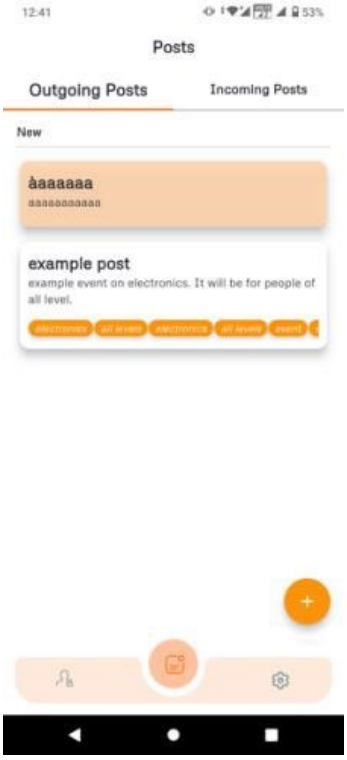
Profilaktualisierungen, die durch eine JWT-Authentifizierung gesichert sind, verhindern unautorisierte Änderungen. Experimente haben gezeigt, dass eine erneute Modellinferenz nach der Hinzufügung von Schlüsselwörtern aktualisierte Empfehlungen in weniger als einer Sekunde liefert, was eine Echtzeitanpassung ermöglicht.

Erstellung von ausgehenden Posts

Mit sogenannten Posts veröffentlichen Nutzerinnen und Nutzer Beiträge, die wie ein digitaler Aushang auf einer virtuellen Pinnwand fungieren, mit dem Unterschied, dass diese Informationen nicht von allen eingesehen werden können. Mit Posts können Mitstreiterinnen und Mitstreiter für z.B. Start-ups, Forschungsprojekte, Studiengruppen oder andere Initiativen gesucht werden – etwa, wenn spezielle Interessen, Expertisen oder auch gesundheitliche Einschränkungen vorliegen und Personen auf der Suche nach Kollaboration oder Kontakt sind. Diese Posts werden bei der Erstellung, wie das bei den Profilinformationen auch der Fall ist, automatisch verschlagwortet. Anstelle von Benutzer-IDs kommen anonymisierte Tokens zum Einsatz, um persönliche Informationen und die Identität zu schützen.

Experimentelle Auswertungen zeigten, dass das automatische Tagging in über 90 % der getesteten Fälle präzise mit den Interessen übereinstimmt, was verdeutlicht, dass das von Gemini Nano bereitgestellte semantische Verständnis eine hohe Genauigkeit liefert.

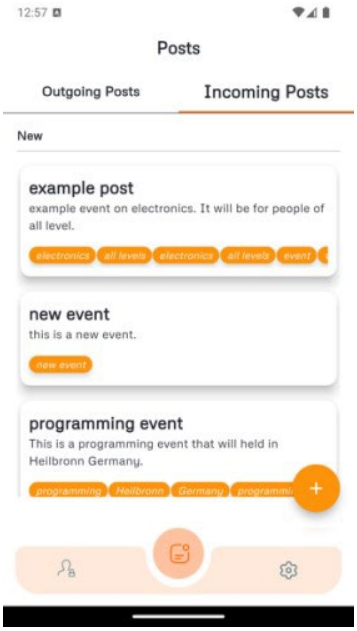




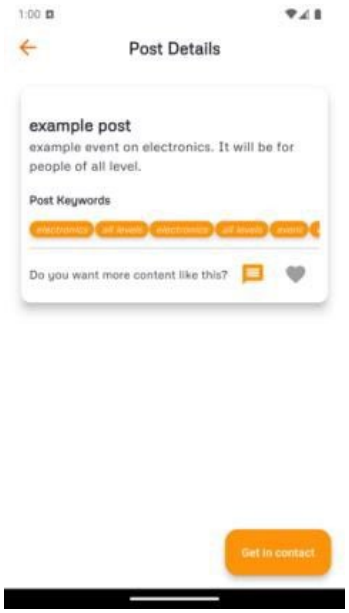
„Create New Post“ *Texteingabe/* Schlüsselworterkennung „Outgoing Posts“

Eingehende Posts

Nur wenn die lokale KI auf dem Gerät eine Übereinstimmung zwischen einem Post und den im Profil hinterlegten Informationen erkennt, wird auf den Endgeräten ein eingehender Post angezeigt. Nutzende können die neu eingehenden Beiträge prüfen und Veranstaltungen oder Kooperationsmöglichkeiten erkennen, die zu ihren Profilen passen. Als entscheidendes Erfolgskriterium dient hier die Relevanz: In über 92 % der durchgeführten Simulationen wurden die vorgeschlagenen Beiträge als zutreffend und kontextuell passend zu den jeweiligen Interessen beurteilt.



„Incoming Posts“



„Post Details“

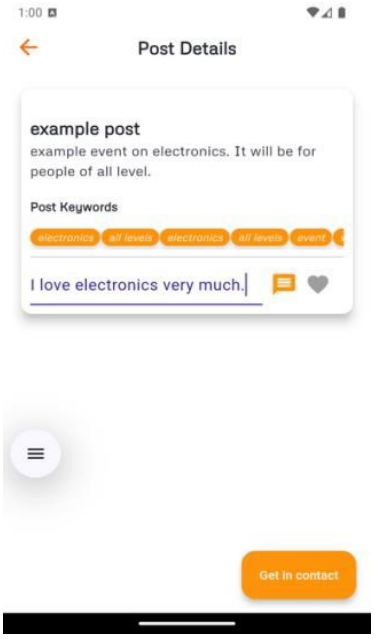
Anonyme Kontaktaufnahme per Chat

Wenn eine Nutzerin oder ein Nutzer einen eingehenden Beitrag für interessant hält, besteht zunächst die Möglichkeit einer anonymen Kontaktaufnahme per Chat, um gegebenenfalls weitere Details zu klären, bevor tatsächliche Kontaktdaten ausgetauscht werden. Für private Unterhaltungen zwischen anonymen Teilnehmenden wird eine stabile Verschlüsselung und ein tokenbasiertes Identitätssystem verwendet. Die Latenz beim Aufbau der Ende-zu-Ende-verschlüsselten Kommunikationskanäle blieb durchgängig niedrig, mit durchschnittlichen Übertragungszeiten von unter 200 Millisekunden.

Möglichkeit, anonyme Nachrichten auszutauschen

Feedback und Empfehlungsverfeinerung

Die Nutzerinnen und Nutzer haben die Möglichkeit, natürlichsprachliches Feedback zu hinterlassen, um zukünftige Empfehlungen zu verfeinern. Diese Option geht über die herkömmliche "Gefällt mir" oder "Gefällt mir nicht" Kennzeichnung hinaus und ermöglicht weitgehendere Modellverbesserungen. Gemini Nano verarbeitet diese Eingaben lokal und aktualisiert dabei seine internen Repräsentationen, ohne dass die Daten extern preisgegeben werden.



Feedback und Empfehlungsverfeinerung

Sicherheitstest

Um sicherzustellen, dass Datenschutz- und Sicherheitsanforderungen eingehalten werden, wurde eine Reihe von Penetrationstests und Schwachstellenscans durchgeführt. Dabei wurden die SSL/TLS-Konfigurationen überprüft, wobei eine 100 % sichere Übertragung bestätigt werden konnte. JWT-Token wurden in isolierten Umgebungen Replay-Angriffen ausgesetzt, ohne dass unautorisierte Profiländerungen protokolliert wurden. Der Einsatz anonymisierter Tokens verhinderte zudem, dass in simulierten Angriffen personenbezogene Informationen (PII) preisgegeben wurden. Darüber hinaus zeigte die Backend-Validierung im Bereich der Ereignisverarbeitung keinerlei Hinweise auf Injektions- oder DoS-Schwachstellen. Durch die konsequente Nutzung von HTTPS für alle externen Anfragen konnte ein unerlaubtes Auslesen personenbezogener Daten effektiv verhindert werden. In der finalen Implementierung führt somit keiner der Angriffsversuche zu einer Beeinträchtigung der Anonymität oder Integrität der Daten.

5. Fazit und Ausblick

Die Projektergebnisse zeigen, dass die innovative Verknüpfung des Matrix-Messaging-Protokolls mit AI on Edge die Vernetzung von Personen im Hochschulbereich gezielt verbessern kann, ohne den Datenschutz zu gefährden. Klassische Plattformen, die auf große Datenmengen und cloudbasierte Verarbeitung setzen, bieten hier keine zufriedenstellende Lösung, da sie Abhängigkeiten schaffen, Sicherheitsrisiken bergen und oft nicht mit datenschutzrechtlichen bzw. ethischen Anforderungen vereinbar sind. Das hier vorgestellte dezentrale Lösungskonzept zeigt eine innovative Alternative auf. Durch die Kombination des Matrix-Protokolls für sichere, föderierte Kommunikation mit AI on Edge entsteht ein System, das Datenschutz und Nutzerfreundlichkeit vereint. Lucii zeigt, dass intelligente, KI-gestützte Matching-Verfahren auch ohne zentrale Datenhaltung Anwenderfreundlich möglich sind. Anonyme Kommunikation, sichere Verschlüsselung und die Möglichkeit, durch natürlichsprachliches Feedback die Empfehlungsqualität zu verbessern bieten zusätzliche innovative Verbesserungen für die Vernetzung im Hochschulumfeld.

Mit Lucii wurde die Machbarkeit und das Potenzial einer solchen Lösung demonstriert, ein größerer Test außerhalb der Entwicklung, also ein Praxistest mit einer breiteren Nutzergruppe, ist erforderlich, um gezielte Verbesserungen vorzunehmen und das System zu einer stabilen, einsatzfähigen Lösung weiterzuentwickeln. In zukünftigen Arbeiten könnte es auch darum gehen, Lucii in verschiedenen Anwendungsszenarien zu testen – etwa in einem Hochschulverbund, mit Kooperationsunternehmen, gemeinnützigen Organisationen oder Forschungsnetzwerken. Fortschritte in der Hardware mobiler Geräte, effizientere Modellkomprimierung und die Integration von erklärbaren KI-Ansätzen bieten künftig weitere Möglichkeiten zur Verbesserung des Systems.