



Vertraulichkeitsklassen für Informationen im Kontext des Einsatzes von KI-Systemen

| Klassifizierung | ÖFFENTLICH | INTERN | VERTRAULICH | STRENG VERTRAULICH |
|--|---|---|--|---|
| Kürzel | C1 | C2 | C3 | C4 |
| Farbcode | WHITE | GREEN | AMBER | RED |
| Beschränkung | Informationen, die ohne Einschränkung geteilt werden dürfen, insbesondere auch mit der allgemeinen Öffentlichkeit. | Informationen, die nur innerhalb der Universität und gegebenenfalls mit Partnern geteilt werden dürfen. | Informationen, die nur innerhalb der Universität und gegebenenfalls mit Partnern nach dem Prinzip der Erforderlichkeit (need-to-know) geteilt werden dürfen. | Informationen, die nur innerhalb einer explizit definierten Gruppe geteilt werden dürfen. |
| Schadenspotential | Eine unautorisierte Offenlegung dieser Informationen hätte keine Auswirkungen auf betroffene Personen, Prozesse, Systeme oder Einrichtungen und keinen Folgen für die Universität Stuttgart. | Eine unautorisierte Offenlegung dieser Informationen hätte sehr geringe bis geringe Auswirkungen auf betroffene Personen und/oder Prozesse, Systeme oder Einrichtungen mit minimalen bis kaum spürbaren Folgen für die Universität Stuttgart. | Eine unautorisierte Offenlegung dieser Informationen hätte mäßige bis erhebliche Auswirkungen auf betroffene Personen und/oder Prozesse, Systeme oder Einrichtungen mit wahrnehmbaren bis ernsthaften Folgen für die Universität Stuttgart. | Eine unautorisierte Offenlegung dieser Informationen hätte große bis sehr große Auswirkungen auf betroffene Personen und/oder Prozesse bzw. Systeme mit schwerwiegenden bis katastrophalen Folgen für die Universität Stuttgart. |
| Personenbezogene Daten | Personenbezogene Daten, die von den Betroffenen frei zugänglich gemacht wurden. | Personenbezogene Daten, deren unsachgemäße Handhabung zwar keine besondere Beeinträchtigung erwarten lässt, die aber von den Betroffenen nicht frei zugänglich gemacht wurden. | Personenbezogene Daten, deren unsachgemäße Handhabung den Betroffenen in seiner gesellschaftlichen Stellung oder in seinen wirtschaftlichen Verhältnissen beeinträchtigen könnte („Ansehen“). | Personenbezogene Daten, deren unsachgemäße Handhabung Gesundheit, Leben oder Freiheit des Betroffenen beeinträchtigen könnte und/oder den Betroffenen in seiner gesellschaftlichen Stellung oder in seinen wirtschaftlichen Verhältnissen erheblich beeinträchtigen könnten („Existenz“). |
| Beispiele | <ul style="list-style-type: none"> • Öffentlich zugängliche Forschungs- und Publikationsinformationen (z.B. Abstracts, Vorträge, Konferenzbeiträge). • Marketing- und PR-Material (Flyer, Broschüren, Pressemitteilungen) • Öffentliche Informationen auf der Universitätswebsite. | <ul style="list-style-type: none"> • Interne Rundschreiben, organisatorische Dokumente der Fachbereiche ohne vertrauliche Personenbezüge. • Nicht-öffentliche Lehrmaterialien, die jedoch keine sensiblen Personendaten beinhalten. • Erste konzeptionelle Forschungsansätze (ohne sensible Details). • Urheberrechtlich geschütztes Material | <ul style="list-style-type: none"> • Personaldaten (z.B. Gehaltsinformationen, Bewerbungen), Studierenden daten (Leistungsnachweise, Prüfungsdaten). • Noch unveröffentlichte Forschungsdaten • Interne Protokolle von Gremien und Ausschüssen, die sensible Themen enthalten. | <ul style="list-style-type: none"> • Medizinische oder psychologische Daten von Proband*innen, Patient*innen oder Studierenden. • Details zu laufenden Drittmittelprojekten mit strengem Patentschutz, Vertraulichkeitsvereinbarungen oder hohen Reputationsrisiken. • Geschäftsgeheimnisse in kooperativen Forschungsprojekten mit Unternehmen (z.B. geheime technische Daten, Prototypen, neue Verfahren). |
| KI-Systeme Kriterien: <ul style="list-style-type: none"> • Vertragliche Absicherung • Hosting • Training durch Anbieter • Accountbindung | Nutzung von KI-Systemen ohne Einschränkung erlaubt. Hosting in Cloud (EU / Nicht-EU) sowie On-Premise zulässig. Trainingsnutzung der Eingaben durch Anbieter ist akzeptabel. Externe Accountbindung in eigenem Ermessen. | Nutzung erlaubt, sofern keine Nutzung der Eingabedaten für Trainingszwecke. Hosting auf On-Premise oder in EU-Cloud zulässig. Nutzung von Nicht-EU-Clouds nur mit vertraglicher Absicherung (z. B. DPA + SCC). Externe Accountbindung nur wenn keine langfristige Protokollierung und keine Auswertung. | Nutzung nur mit vertraglicher Zusicherung, dass Eingabedaten nicht für Trainingszwecke verwendet werden. Hosting nur On-Premises oder in EU-Cloud mit starker vertraglicher Absicherung (kein Zugriff Dritter, Ende-zu-Ende-Verschlüsselung). Nutzung von Nicht-EU-Clouds nicht zulässig. Interne Zugangsbeschränkung auf Universitätsangehörige. Externe Accountbindung nur mit pseudonymisiertem Zugang. | Nutzung von KI-Systemen nur auf isolierten On-Premises-Systemen mit vollständiger Kontrolle. Cloud-Nutzung (auch EU) i.d.R. nicht erlaubt. Eingaben dürfen unter keinen Umständen für Trainingszwecke verwendet werden – vertragliche Absicherung zwingend erforderlich. Nachvollziehbarkeit intern erwünscht oder sogar notwendig, z. B. um Missbrauch zu verhindern. |
| Beispiele für KI-Tools | <ul style="list-style-type: none"> • ChatGPT Free / Plus • MS Copilot mit privatem Account | <ul style="list-style-type: none"> • Uni-Tool „RAI“ mit HAWKI 2 Oberfläche und OpenAI Modellen über Azure • MS Copilot mit Uni-MS-Account | <ul style="list-style-type: none"> • Uni-Tool „RAI“ mit HAWKI 2 Oberfläche und GWDG Modellen | <ul style="list-style-type: none"> • Lokale Open-Source-Modelle wie LLaMA, Mistral (momentan noch nicht verfügbar) |

Legende / Kurzvorgaben

- **On-Premises** = Systeme in Uni-Kontrolle (Rechenzentrum, dedizierte GPU-Server).
- **EU-Cloud** = Cloud-Region innerhalb EWR + EU-DSGVO-Konformität.
- **Nicht-EU-Cloud** = jede Cloud-Region außerhalb EWR. Hier sind zusätzlich Standardvertragsklauseln (SCC oder gleichwertige Garantien erforderlich; bei C3/C4 nicht zulässig).
- **DPA** = Data-Processing-Agreement; **SCC** = Standard Contractual Clauses.
- **Accountbindung: Accountbindung** bezeichnet die Zuordnung von Nutzereingaben zu einem **personenbezogenen Konto** beim Anbieter eines KI-Systems (z. B. Log in mit E-Mail-Adresse oder Uni-Account). Dabei kann der Anbieter nachvollziehen, **welche Person welche Inhalte** eingegeben hat.
- **Interne Nachvollziehbarkeit:** Eine **lokale interne Nachvollziehbarkeit** (z. B. Log in bei On-Prem-Systemen) ist **nicht gleichzusetzen** mit externer Accountbindung und kann bei höheren Vertraulichkeitsstufen sogar erforderlich sein.