



# AnoMed

AnoMed - Sachbericht zum  
Teilvorhaben

# Inhaltsverzeichnis

<b>1</b>	<b>Kurzbericht</b>	<b>2</b>
1.1	Ursprüngliche Aufgabenstellung sowie den wissenschaftlichen und technischen Stand, an den angeknüpft wurde . . . . .	2
1.2	Ablauf des Vorhabens . . . . .	3
1.3	Wesentliche Ergebnisse sowie ggf. die Zusammenarbeit mit anderen Forschungseinrichtungen . . . . .	3
<b>2</b>	<b>Eingehende Darstellung</b>	<b>4</b>
2.1	Ausführliche Darstellung der Ergebnisse . . . . .	4
2.2	Wichtigste Positionen des zahlenmäßigen Nachweises . . . . .	12
2.3	Notwendigkeit und Angemessenheit der geleisteten Projektarbeiten . . . . .	12
2.4	Voraussichtlicher Nutzen, insbesondere die Verwertbarkeit des Ergebnisses . . . .	13
2.5	Während der Durchführung des Vorhabens dem Zuwendungsempfänger bekannt gewordenen Fortschritt auf dem Gebiet des Vorhabens bei anderen Stellen . . . .	14
2.6	Erfolgte oder geplante Veröffentlichungen des Ergebnisses nach Nr. 5 der NKBF/NABF	14

# 1 Kurzbericht

## 1.1 Ursprüngliche Aufgabenstellung sowie den wissenschaftlichen und technischen Stand, an den angeknüpft wurde

Das Kompetenzzentrum AnoMed hat zum Ziel, tragfähige wissenschaftliche Grundlagen und methodische Konzepte für die sichere und datenschutzkonforme Nutzung medizinischer Daten zu entwickeln. Dazu gehören sowohl die methodische Erforschung neuartiger Anonymisierungsverfahren gemeinsam mit der Entwicklung kompatibler Merkmalsextraktionsverfahren, die Erforschung sicherer Datenverarbeitungsverfahren, wie sichere Mehrparteienverfahren, kryptographische Ansätze und hardwarebasierte Sicherheitsmechanismen, die wissenschaftliche Kommunikation in Richtung regulatorischer Behörden und der breiten Öffentlichkeit, die Klärung regulatorischer Anforderungen und die Entwicklung einer Wettbewerbsplattform für einen beschleunigten Technologietransfer. Mit dem ganzheitlichen Ansatz aus Anonymisierung, maschinellem Lernen und medizinischer Domänenexpertise sollte im Rahmen des Projekts eine neue Grundlage für den Transfer dieser, für die sichere Nutzung von Patient\*innendaten notwendigen Werkzeuge gelegt werden.

**Methodische Erforschung neuartiger Anonymisierungsverfahren.** Im Spannungsfeld zwischen dem Schutz von Patient\*innendaten und möglichst effektiver Versorgung hat das Kompetenzzentrum AnoMed das Ziel, Methoden, Bewertungsmaßstäbe und Infrastrukturen zu entwickeln, mit denen Anonymisierungslösungen für medizinische Anwendungen systematisch untersucht, verglichen und weiterentwickelt werden können. Klassische Anonymisierungsverfahren wie K-Anonymität oder einfache Aggregationsansätze haben sich mancherorts etabliert, sind jedoch nachweislich anfällig gegenüber modernen Re-Identifikations- und Inferenzangriffen. Insbesondere für hochdimensionale oder verknüpfte medizinische Datensätze sind sie nur eingeschränkt geeignet. Daher wurde sich bei AnoMed auf die Erforschung von Verfahren mit starken mathematischen Schutzgarantien, insbesondere aus dem Bereich der Differential Privacy, konzentriert. Um die Effektivität von Anonymisierungslösungen zu steigern, sollten kompatible Merkmalsextraktoren entwickelt werden.

**Sichere Datenverarbeitung.** Es sollten grundlegende technische Methoden im Bereich der Datenverarbeitung erforscht werden, wie sichere Mehrparteienverfahren, kryptographische Ansätze und hardwarebasierte Sicherheitsmechanismen.

**Klärung regulatorischer Anforderungen und wissenschaftliche Kommunikation in Richtung regulatorische Behörden.** Auch auf regulatorischer Ebene bestand ein erheblicher Klärungsbedarf. Datenschutzrechtliche Anforderungen waren zwar formuliert, ließen sich jedoch nur eingeschränkt in präzise technische Kriterien oder mathematische Garantien übersetzen. Dies führte zu Unsicherheiten bei der Entwicklung und Bewertung technischer Lösungen und erschwerte eine sachgerechte Einordnung des tatsächlichen Schutzniveaus. AnoMed hat zum Ziel, diese Anforderungen zu übersetzen und die Konzepte aus der Wissenschaft für regulatorische Behörden zu kommunizieren. Bestehende Ansätze sollen dabei in einen kohärenten und interdisziplinären Rahmen überführt werden.

**Wettbewerbsplattform.** Eine Challenge- und Bewertungsplattform soll eine breite Basis für zukünftige Forschung, regulatorische Diskussionen und innovationsorientierte Anwendungen schaffen.

**Wissenschaftskommunikation an die breite Öffentlichkeit.** Im Rahmen des Kompetenzzentrums AnoMed sollten Konzepte entwickelt werden, um Herausforderungen und Konzepte aus der Anonymisierungsforschung effektiv zu kommunizieren.

## 1.2 Ablauf des Vorhabens

Zu Projektbeginn wurden Kommunikations- und Steuerungsstrukturen etabliert und anschließend kontinuierlich betrieben. Dazu gehörten die regelmäßige Mitwirkung im Lenkungsausschuss und die Koordination im Konsortium. Für die Wettbewerbsplattform wurde in den ersten Monaten eine initiale Anforderungsanalyse erstellt und in regelmäßigen Treffen im Laufe des Projektes abgeglichen. Bis zum Projektende wurde damit ein wesentlicher Meilenstein für zukünftigen Technologietransfer erreicht: eine vertikale Demo einer Plattform, auf der Privacy-Challenges und Privacy-Angriffe bereitgestellt und Einreichungen technisch verarbeitet werden können und auf der erste Challenges und erste Angriffe platziert sind. Ergänzend wurde eine Serverstruktur eingerichtet, um Entwicklung und Evaluation in Trusted-Execution-Umgebungen zu ermöglichen. In den wissenschaftlich-technischen Arbeitspaketen wurden im Projektzeitraum eine Reihe an Forschungsrichtungen erfolgreich verfolgt: neue Merkmalsextraktion für medizinische Anwendungen, neue privatsphäre-erhaltende (sog. Differentially Private) Lernverfahren, neue Verfahren zur ausgelagerten sicheren Berechnung von KI-Aufgaben, neuartige Datensyntheseverfahren sowie neue Angriffsverfahren. Flankierend wurden die Arbeiten durch konsortiumsweite Austauschformate getragen. Das zweiwöchentlich stattfindende AnoMed-Seminar, die jährlichen Konsortialtreffen sowie öffentliche Formate dienten dazu, Ergebnisse abzustimmen, Anforderungen nachzuschärfen und den Transfererfolg der entwickelten Komponenten sicherzustellen. Abweichungen ergaben sich insbesondere durch verzögerte Einstellungen sowie durch die notwendige Übernahme von Aufgaben nach dem Ausscheiden der Unternehmenspartner *Ingrano* und *SciEngines*. Dadurch verschob sich der Zeitplan, insbesondere bei der Plattformentwicklung. Diese Verzögerungen wurden organisatorisch, u. a. durch zusätzliche Einstellungen und Priorisierung, adressiert und führten zur Beantragung einer kostenneutralen Verlängerung bis Ende 2025. Die übergeordneten Projektziele wurden in vollem Umfang erreicht.

## 1.3 Wesentliche Ergebnisse sowie ggf. die Zusammenarbeit mit anderen Forschungseinrichtungen

Während der Projektlaufzeit wurden 39 Arbeiten veröffentlicht, von denen der weit überwiegende Teil bereits peer-reviewed ist. Acht davon sind A\*-Publikationen (3x ACM CCS, 1x IEEE S&P, 1x USENIX Security, 1x IJCAI, 1x AAI, 1x CHES), was die hohe wissenschaftliche Qualität der Arbeiten in AnoMed belegt. Alle Arbeiten sind auf der AnoMed-Website aufrufbar. Im Rahmen des regelmäßig stattfindenden *AnoMed-Seminars* wurde sich über aktuelle Fortschritte und zukünftige Herausforderungen ausgetauscht. Außerdem wurden hier die Anforderungen an Challenges und Angriffe besprochen und angepasst. Hinzu kommen zahlreiche Workshops und Treffen, auch mit externen Partnern wie dem *KI-Lab Lübeck*, der *TMF* oder *LifeScienceNord*. Durch das eigenentwickelte Brettspiel *Spurensuche in der KI* wurden die Risiken fehlender Anonymisierung niedrigschwellig kommuniziert. Durch die im Rahmen des Projekts aufgesetzte *Privacy-Challenge-Plattform* können Anonymisierungsstrategien für medizinische Daten gegen aktuelle Angriffe getestet und somit die Nutzbarkeit verschiedener Verfahren evaluiert werden. Nicht zuletzt aufgrund der guten Zusammenarbeit mit unseren Partnern konnten wir mehrere wissenschaftliche Arbeiten veröffentlichen, darunter das A\*-publizierte Paper *DPM*, *DP-Scans* für eine sicherere Datensynthese und *DP-Hype* für eine sichere Hyperparametersuche. Um weitere Aspekte sicheren Lernens zu beleuchten, wurden auch TEE- und MPC-basierte Verfahren erforscht und dabei mehrere neue Verfahren hervorgebracht. Um solche Verfahren, wie ein neues Protokoll für homomorphe Verschlüsselung, auch hardwareseitig zu unterstützen, wurde eine neue quelloffene TEE für FPGA-taugliche RISC-V-Softcores entwickelt. Darüber hinaus wurde das Konzept von *statistical Privacy* eingeführt, Rekonstruktions- und Extraktionsangriffe für verschiedene Modellklassen sowie Side-Channel- und Fault-Angriffe auf Hardware untersucht. Insgesamt wurden acht verschiedene Angriffe erfolgreich implementiert. Parallel wurde die datenschutzrechtliche Arbeit des *ULD* mit technischem Wissen unterstützt.

## 2 Eingehende Darstellung

### 2.1 Ausführliche Darstellung der Ergebnisse

Als Teil und Leitung des Kompetenzcluster AnoMed hat die Universität zu Lübeck 39 Arbeiten veröffentlicht, von denen der weit überwiegende Teil bereits peer-reviewed ist. Acht davon sind A\*-Publikationen (3x ACM CCS, 1x IEEE S&P, 1x USENIX Security, 1x IJCAI, 1x AAAI, 1x CHES), was die hohe wissenschaftliche Qualität der Arbeiten in AnoMed belegt. Alle Arbeiten werden in Kapitel 2.6 aufgelistet und sind darüber hinaus auf der AnoMed-Website aufrufbar.

**Projekt- und Clustermanagement sowie Lenkungsausschuss (AP 0.1):** Die sachgerechte und vollständige Kommunikation sowie die fristgerechte Abarbeitung der Arbeitspakete wurden durch Projektleitung und die UTK in enger Absprache mit der Projektleitung, welche Teil der UzL ist, sichergestellt. Weiterhin hat die Projektleitung im Lenkungsausschuss des Forschungsnetzwerks Anonymisierung zeitweise moderiert und mitgewirkt. Die Projektleitung hat einen maßgeblichen inhaltlichen Beitrag bei der Planung der beiden AnoSiDat-Konferenzen (2024 und 2025) getätigt. Außerdem hat die UzL einen Arbeitskreis des Forschungsnetzwerks Anonymisierung koordiniert, indem ein Positionspapier für Privacy-Forschung für das BMBF geschrieben und abgeschickt wurde, das von dem BMBF im Anschluss an die AnoSiDat 2024 angefragt wurde. Gemeinsam mit der UTK hat die Projektleitung das regelmäßig stattfindende AnoMed-Seminar über die gesamte Projektlaufzeit hinweg inhaltlich und organisatorisch begleitet. Das AnoMed Seminar hat sich als Instrument für eine kontinuierliche Anforderungsanalyse und als Austauschformat zwischen Forschung und Anwendung etabliert.

**Anforderungsanalyse (AP 0.2):** Im Rahmen des von der UTK und UzL organisierten AnoMed Seminars und auf den Konsortialtreffen hat die UzL regelmäßig eine Anforderungsanalyse angestoßen, um die Anforderungen mit dem im Laufe des Projekts wachsenden Verständnis synchron zu halten. In wiederholten Vorträgen über z.B. die AnoMed Wettbewerbsplattform oder aktuellen Forschungsergebnissen wurden Kommentare eingeholt. Damit wurde Raum geschaffen, um Korrekturen vorzunehmen. Die UzL hat diesen Raum genutzt, um eine Vielzahl an eigenen Forschungsarbeiten vorzustellen. Darüber hinaus wurden in regelmäßigen bilateralen Besprechungen die Anforderungsanalysen erhoben und abgeglichen.

**Verknüpfung mit KI-Lab (AP 1.5):** Von 2020 bis 2021 wurde das BMBF-Projekt KI-Lab gefördert, welches einen Grundstein für eine KI-Recheninfrastruktur an der UzL gelegt hat. Basierend aus den Erfahrungen aus dem KI-Lab Projekt haben wir das Clustermanagement neu und professioneller aufgesetzt. Zusätzlich haben wir mit den Fördermitteln aus dem AnoMed Kompetenzzentrum die Rechenressourcen um eine Größenordnung verstärkt, sodass wir im Laufe des Projekts nun eine KI-Recheninfrastruktur mit 39 GPUs und mehr als 2,5 TB VRAM für Forschungszwecke und Unternehmenskollaborationen aufgebaut haben. Ziel war es insbesondere, perspektivisch auch Unternehmen eine Zusammenarbeit an KI-Anwendungsprojekten unter datenschutzkonformen Rahmenbedingungen zu ermöglichen. Die geschaffene Anbindung bildet damit eine wichtige Grundlage für anwendungsnahe KI-Entwicklung, Technologietransfer sowie zukünftige Kooperationen mit Forschung und Wirtschaft.

**Challenge-Plattform: Infrastruktur (AP 1.1, AP 1.2, AP 1.4, AP 4.1, AP 4.2):** Eine Multi-Container-Webanwendung wurde als vertikaler Prototyp realisiert, der eine produktionsähnliche Umgebung ermöglicht. Erste Beiträge der AnoMed-Partner wurden integriert. Eine Vorabversion ist unter <https://gpus.its.uni-luebeck.de/anomed/7a2b10f0/challenges/> erreichbar. Zusätzlich wurde ein Server mit SGX-Support eingerichtet. Der Prototyp für die AnoMed-Plattform deckt bereits den wesentlichen Funktionsumfang ab. Neben beispielhaften

Privacy-Challenges wurden auch authentische Beiträge verschiedener AnoMed-Partner integriert. Für die Challenge ECG Age wurde von dem Konsortium auch schon erste Beiträge eingereicht, u.a. das unten vorgestellte Verfahren S-BDT aus dem Konsortium. Im Rahmen des AP 1.2 hat die UzL einen Demonstrator zur Evaluation nicht öffentlicher Daten entwickelt. Nach einem initialen Setup, basierend auf Intel-SGX-fähiger Hardware und einem Setup basierend auf AMD SEV-SNP[62], wurde durch Nvidia mit nvtrust eine Möglichkeit geschaffen, KI-Beschleuniger in den vertraulichen Bereich prozessorbasierter vertrauenswürdiger Ausführungsumgebungen (TEEs) zu integrieren. Ein Großteil der AnoMed-Arbeiten fokussiert sich auf rechenintensive KI-basierte Verfahren, die von einer derartigen Integration stark profitieren. Daher wurde die Integration von KI-Beschleunigern in die TEE als neues Ziel in die Arbeiten dieses Themenkomplexes aufgenommen. Da nvtrust auf Prozessorseite Intel TDX oder AMD SEV voraussetzt, wurde ein entsprechendes Hardware-Setup beschafft und eingerichtet (Intel(R) Xeon(R) 6972P / 4x NVIDIA H200 NVL). Basierend auf diesem Hardware-Setup hat die UzL einen Demonstrator, die AnoMed Secure API, entwickelt, der die Nutzung einer NVIDIA H200 NVL innerhalb der vertraulichen Ausführungsumgebung einer Intel-TDX-VM ermöglicht. Neben dem sicheren Zugang über SSH ermöglicht die VM zur anwenderfreundlichen Nutzung zudem den Zugang mittels einer grafischen Remote-Desktop-Lösung (X2Go Graphical Access). Zur Evaluation des Demonstrators wurden Inferenz- und Trainingsworkloads auf dem NIH-ChestX-ray8-Datensatz mit und ohne GPU-Support durchgeführt. Im Rahmen der Evaluation wurde zudem eine Sicherheitsanalyse der Single-Stepping-Gegenmaßnahmen von Intel TDX durchgeführt und Schwachstellen [63] gefunden und publiziert. Zum Projektabschluss ist die Entwicklung des Demonstrators vollständig abgeschlossen, einzig die Attestierung der CPU/GPU-Umgebung konnte nicht vollständig implementiert werden, da auf die Rückmeldung der Hardwarehersteller gewartet werden muss. Die UzL steht in Kontakt zu den Hardwareherstellern und plant, diese Arbeiten im Rahmen von AnoMed 2 zu beenden. Das produktive Hosting der Challenge-Plattform erfolgt auf einem hochverfügbaren virtualisierten System, das vom AI-LAB des ITSC administriert wird. GPU-Ressourcen werden innerhalb virtueller Maschinen und mit der NVIDIA-vGPU-Technologie bereitgestellt. Neben der höheren Verfügbarkeit ermöglicht diese Technik die Separierung von Workloads und die gleichzeitige effiziente Nutzung der GPU-Ressourcen. Für AP 4.1 haben wir eine Testumgebung zur Evaluation von Deanonymisierungsverfahren entwickelt. Daraus gewonnene Daten und Erkenntnisse wurden für eine Demo der Plattform genutzt, die unter <https://gpus.its.uni-luebeck.de/anomed/7a2b10f0/> für interne Zwecke und zur Abstimmung mit Projektpartnern veröffentlicht wurde. Wir planen, diese Plattform inkrementell für Parteien außerhalb des Projekts zu öffnen. In dem Folgeprojekt AnoMed-RZ soll diese Plattform weiterentwickelt und öffentlich verfügbar und beworben werden.

**Challenge-Plattform: Challenges und Analysen (2.1, 2.2, 2.3, 2.4, 3.2):** Das ursprüngliche Ziel, eine realistische und reproduzierbare Privacy-Challenge auf Basis medizinisch relevanter Videodaten zu konzipieren, wurde erreicht. Es wurden mehrere Datensätze (Toronto NeuroFace [5], UNBC Pain Database [33], CelebA [32]) untersucht, medizinische Anforderungen analysiert und potenzielle Angriffsszenarien definiert. Für die final ausgewählte Challenge auf Basis des CelebA-Datensatzes wurden die Daten vollständig aufbereitet und dokumentiert sowie Referenzimplementierungen bereitgestellt, die eine hohe Klassifikationsperformance erzielen. Bei den Referenzimplementierungen handelt es sich um EfficientNet-B0- und ResNet18-Modelle für die Klassifikation der Attribute „Bags\_Under\_Eyes“, „Eyeglasses“, „Mouth\_Slightly\_Open“, „Narrow\_Eyes“, „Bags\_Under\_Eyes“ und „Smiling“, die sich durch unterschiedliche Klassenimbilanzen und semantische Feinheiten auszeichnen. Die Trainings- und Evaluationspipeline wurde vollständig dokumentiert und erlaubt eine standardisierte Bewertung mittels Accuracy und F1-Score. Ein zentrales Ziel bestand in der Aufbereitung, Standardisierung und nachhaltigen Bereitstellung eines selbst erhobenen multimodalen Datensatzes zum gesunden Altern, Hörerleben und

Kognition. Dieses Ziel wurde vollständig erreicht. Der Datensatz wurde umfassend vorbereitet, qualitätsgesichert und gemäß dem BIDS-Standard (Brain Imaging Data Structure) strukturiert. Für jede Versuchsperson wurden konsistente Datenordner angelegt, die Verhaltensdaten, Fragebogenantworten, demographische Informationen sowie rohe und vorverarbeitete EEG- und fMRT-Daten enthalten. Ergänzend wurden qualitätsrelevante Metadaten integriert, um eine systematische Auswahl und Nachnutzung zu ermöglichen. Die EEG-Vorverarbeitung umfasste unter anderem Filterung, Normalisierung und Verfahren zur Dimensionalitätsreduktion. Das vollständige Datenset wurde abschließend auf der Open Science Framework (OSF)-Plattform veröffentlicht und umfassend dokumentiert (Themenbereich *Datensatzaufbereitung und Standardisierung*, UAP 2.1.10).

Ein weiterer Schwerpunkt lag im Themenbereich *EEG-Privacy-Challenges*. Auf Basis des aufbereiteten Datensatzes wurde eine Privacy-Challenge im Bereich der EEG-Signalanalyse konzipiert, die privacy-relevante Signalmerkmale wie Phase, Frequenz und Sensorort berücksichtigt. Nach einer umfassenden Literaturrecherche wurden geeignete Zielgrößen definiert und als Benchmarks dokumentiert. Es wurden mehrere wissenschaftlich relevante und zugleich privacy-sensitive Regressionsaufgaben umgesetzt, darunter die Vorhersage der Leistung in einer Höraufgabe sowie die Abschätzung von Hörsensitivität und Griffkraft in Abhängigkeit vom Alter. Als Evaluationsmetrik wurde das Bestimmtheitsmaß  $R^2$  verwendet. Zusätzlich wurde eine Baseline-Lösung ohne Privacy-Mechanismen implementiert, um Privacy-Risiken systematisch bewerten und privacy-erhaltende Ansätze vergleichbar machen zu können (UAP 2.1.11).

Im Themenbereich *Analyse und Merkmalsextraktion* wurden selbsterfasste EEG-Daten systematisch im Zeit- und Frequenzbereich analysiert, um relevante Charakteristika für die Entwicklung von Merkmalsextraktionsverfahren zu identifizieren. Ergänzend erfolgte eine Fusion der EEG-Daten mit fMRT-Daten, um konsistente multimodale Trainingsdatensätze zu erstellen. Aufbauend auf diesen Analysen wurden datenschutzkonforme Merkmalsextraktionsverfahren entwickelt und implementiert, darunter temporale Response-Funktionen zur Analyse der neuronalen Sprachverarbeitung sowie Alpha-Power-Analysen (8–12 Hz) als Indikator für Aufmerksamkeitszustände. Für fMRT-Daten wurden funktionelle Konnektivitätsanalysen durchgeführt. Die Implementierungen erfolgten in MATLAB und Python und wurden in die AnMed-Plattform integriert.

Es wurde im Rahmen von AP 2.3 ein spezialisiertes Lernverfahren entwickelt, um gleichzeitig Zieleigenschaften zu extrahieren und die Rückschlüsse auf ungewollte zusätzlichen Informationen zu verhindern. Die Strategie wurde sowohl für das Lernen von Repräsentationen als auch für die direkte Klassifikation angewandt. Während die erste Anwendung robuste Nutzbarkeit selbst bei verrauschten Eingabedaten bewahren kann, eignet sich die zweite besonders gut, um aus sehr kleinen Datensätzen zu generalisieren. Erste Ergebnisse deuten darauf hin, dass sich die beiden Spezialisierungen zu einer noch robusteren Methode kombinieren lassen könnten. Bei einer der beliebtesten Methoden zu Merkmalsextraktion, neuronalen Netzwerken, wird in der praktischen Anwendung häufig unterschätzt, dass die Ausgabe auch ungewollte zusätzliche Informationen enthalten kann. Um die Relevanz spezialisierter Methoden zu unterstreichen, wurde eine komplexe Pipeline entworfen und realisiert, welche aus der Ausgabe eines regular trainierten neuronalen Netzes eine gute Annäherung an die originale Eingabe erzeugt, und damit Rückschlüsse auf sensible Daten (z.B. die Identität einer Person) demonstriert.

**Entwurf einer API für hardwarebeschleunigte Algorithmen und Entwicklung einer Trusted-Execution-Erweiterung (TEE) für FPGA-Beschleuniger (AP 1.3, AP 3.11):** Nach dem Wegfall von SciEngine als Unterauftragnehmer wurden die Arbeiten für AP 1.3 und AP 3.11 durch die UzL übernommen. Dabei wurden die Anforderungen an das System neu spezifiziert:

Ein zentraler Virtualisierungsserver dient als Ausgangsplattform, um diese mit mehreren FPGAs zu verbinden. Da der Server direkt mit den FPGAs verbunden sein muss, konnte nicht

auf bestehende Infrastruktur zurückgegriffen werden. Die FPGA-Remote-Plattform basiert auf einem Multi-Container-Design, bei dem jedem FPGA und jedem Concurrent-User ein eigener Container zugeordnet ist. Durch eine webbasierte Buchungsplattform werden die Container miteinander mithilfe eines virtuellen Overlays verbunden. Dieses Design ermöglicht den exklusiven Zugang eines FPGAs durch einen User und trennt gleichzeitig den Userspace. Innerhalb des Projektes konnten verschiedene Use-Cases ausgemacht werden, welche unterschiedliche Anforderungen an die FPGA-Plattform stellen: Wearable-Anwendungen zielen auf kleine und energieeffiziente FPGAs, wohingegen kryptografische Anwendungen eher auf leistungsfähige FPGAs zielen. Daher wurden mehrere verschiedene FPGAs beschafft, um diese Use-Cases abzudecken. Die Virtualisierungslösung besteht aus einer kombinierten Proxmox- (Hypervisor) und Docker-Lösung. Dabei hat die UzL sowohl die IT-Infrastruktur als auch das User Front-End entsprechend den Anforderungen umgesetzt.

In AP 3.11 konzentrierte sich die UzL auf die Implementation eines Trusted Execution Environment (TEE) auf einem FPGA-tauglichen RISC-V-Softcore. Ziel ist der Schutz sensibler, durch Diagnostik erzeugter medizinischer Daten. Durch die Integration des TEE mit homomorpher Verschlüsselung (HE) wird eine sichere Isolation von Code und Daten erreicht, sodass diese vor allen auf demselben Chip ausgeführten und potenziell nicht vertrauenswürdigen Software-Schichten geschützt sind. Die quelloffene RISC-V-Architektur ist flexibel anpassbar und für den Einsatz in FPGA-basierten Umgebungen besonders geeignet. Sie eignet sich daher gleichermaßen für Prototyping und Betrieb. Somit sind robuste Vertraulichkeit, Integrität und Herkunftsprüfbarkeit von Patientenakten gewährleistet bei gleichzeitiger Konformität mit Gesundheitsvorschriften.

**Privatsphäre-erhaltendes Lernen und Datensynthese (AP 3.2-3.10,3.12)** Anonymitätsbewahrende Methoden für Gaussian Process Learning und Baum-basiertes Lernen wurden erfolgreich weiterentwickelt. Fortschritte bei der Synthese privatsphäreschützender Daten wurden erzielt, mit optimierten PRM-Methoden und Privacy-Utility-Verbesserungen. Ein sicheres Framework für ausgelagerte CNN-Inferenz wurde entwickelt, und Privacy-Mechanismen für interaktive Datenbankabfragen wurden mathematisch analysiert.

**Gradient-boosting Decision Trees (AP 3.4):** Wir haben ein innovatives Framework für privatsphäre-erhaltendes baum-basiertes Lernen entwickelt, S-BDT [49], und auf der A\* Konferenz *ACM SIGSAC Conference on Computer and Communications Security* erfolgreich veröffentlicht und vorgestellt.

**Clustering (AP 3.6):** Wir haben ein Clustering-Verfahren für sensible Daten was auf gezielter Separation basiert entwickelt, DPM [31] und auf der A\* Konferenz *ACM SIGSAC Conference on Computer and Communications Security* erfolgreich veröffentlicht und vorgestellt. Das Clusteringverfahren war geplant als Basis für Datensynthese mit PRMs. In einer Folgearbeit haben wir die theoretischen Garantien dieses Ansatzes näher untersucht und konnten allgemeine Schranken für die Nutzbarkeit des Algorithmus zeigen sowie vereinfachte Schranken für konkrete Verteilungen. Diese Arbeit wurde bei arXiv veröffentlicht [56].

**Datensynthese via Gaussian-Mixtures (DP-Scans, AP 3.7):** Wir haben erfolgreich ein neuartiges privatsphäre-erhaltendes Verfahren zur Datensynthese entwickelt. In einem ersten Schritt wurde ein auf Gaussian-Mixtures basierender Ansatz untersucht, wobei sich zeigte, dass die Verwendung quadratischer Messgrößen, insbesondere der Kovarianz, zu einem unzureichenden Privacy-Utility-Trade-off führt. Zur Adressierung dieses Problems wurde in einem zweiten Schritt ein alternativer Ansatz entwickelt, bei dem die Datensynthese auf Basis von Häufigkeitsstatistiken erfolgt. Konkret wurden Häufigkeiten über eine Vielzahl zufälliger Korrelationen berechnet und anschließend mittels Gradient Descent-Verfahren Gaussian-Mixtures interpoliert, die an die zugrunde liegenden Daten angepasst sind. Mit diesem Verfahren konnte State-of-the-Art-Performance auf gängigen Benchmark-Datensätzen und -Problemstellungen erzielt werden. Die Ergebnisse finden sich in der aktuellen Version der Publikation ([28]).

**Verteiltes nichtinteraktives Lernen (AP 3.8):** In dem Bereich zur sicheren verteilten Berechnung haben vorangegangene Arbeiten [25] große Teile der geplanten Aufgaben erledigt. In einem nächsten Schritt haben wir uns der nächsten Herausforderung in der Praxis gestellt: Eine sichere Hyperparametersuche. Eine unvorsichtige Hyperparametersuche kann zur Preisgabe von schützenswerten Daten führen. Wir haben uns daher auf die Lösung des Problems der privatsphäre-erhaltenden und verteilten Hyperparametersuche fokussiert. Dieses Problem ist fundamental wichtig, da Hyperparameter gerade bei KI-Verfahren wichtig für eine gute Performance sind und bisherige Arbeiten meist nur einzelne Spezialfälle behandelt haben. Ein allgemeines Verfahren haben wir daraufhin in [30] entwickelt. Privacy-Garantien konnten dabei bewiesen werden. Eine empirische Evaluation zeigt, dass der Privacy-Utility Trade-off im Vergleich zu vorherigen Arbeiten deutlich verbessert werden konnte. Zum Einsatz kamen dabei unter anderem Protokolle zur sicheren Addition.

**Ausgelagertes Rechnen (AP 3.9):** Im Rahmen des AP 3.9 wurden vom ITS anonymisierende Lernverfahren basierend auf TEEs und MPC-Protokollen entwickelt und untersucht. Beginnend mit einer Performanz- und Nutzbarkeits-Evaluation verfügbarer SMPC-Protokolle hat das ITS das GNNP-Framework [4] als vielversprechenden Kandidaten für eine Weiterentwicklung in Kombination mit einer TEE identifiziert. GNNP diente anschließend als Grundlage für die Entwicklung von Dash [52]. In einer weiteren Performanz- und Nutzbarkeits-Evaluation verfügbarer TEE-Protokolle wurde das Slalom-Protokoll [60] mit seiner Unterteilung in Offline- und Online-Phase als vielversprechende Grundlage für zukünftige Entwicklungen identifiziert. Aufbauend auf Slalom wurde die Optimierung „Carnival“ [15] entwickelt, die unter Ausnutzung der Pseudozufälligkeit des Teilsummenproblems die Vorverarbeitungsphase erheblich beschleunigt. Im Rahmen einer Machbarkeitsstudie zu anonymisierenden Lernverfahren wurde das Framework „Dash“ [52] zur sicheren ausgelagerten CNN-Inferenz mittels arithmetischer Garbled Circuits und TEEs entwickelt. Um größere Modelle effektiv nutzen zu können, konnte ein Skalierungs-Gadget für die Chinesische-Restsatz-Repräsentation in arithmetischen Garbled Circuits entwickelt werden. Zur Unterstützung der Skalierbarkeit wurde Dash mit dem Konzept der LabelTensoren entwickelt, das die Ausnutzung paralleler Hardware erleichtert. In einem ersten Schritt wurde Dash zur Ausführung auf Multicore-CPU implementiert. Aufbauend auf Dash hat das ITS anschließend das Skalierungs-Gadget in Dash weiter beschleunigt und als Erweiterung *ReDash* [44] implementiert und veröffentlicht.

Neben den auf arithmetischen Garbled Circuits basierenden Entwicklungen hat das ITS Anwendungen von vollständig homomorpher Verschlüsselung (engl. Fully Homomorphic Encryption, FHE) untersucht und ein neues Protokoll „Silenzio“ [51] zum nicht-interaktiven Auslagern von MLP-Training entwickelt und implementiert. Bei der Ausführung von FHE-Workloads auf herkömmlichen Hardware-Architekturen kommt es zu erheblichen Latenzen bei Speicherezugriffen. Ausgehend von dieser Einschränkung hat das ITS eine beschleunigte Implementierung *DRAMatic* für In-Memory-Verarbeitung entwickelt. Im Rahmen des AP 3.10 wurde *Dash* um eine parallele GPU-Implementierung der Evaluations-Phase ergänzt. Die anschließende Evaluation dieser Erweiterung zeigte eine erhebliche Beschleunigung der Evaluation aller CNN-Schichten, teilweise mit einem Speedup-Faktor  $> 200$  im Vergleich zur parallelen CPU-Implementierung.

**Anonymisierende, verteilte, hardwarebeschleunigte Lernverfahren (AP 3.10):** Die UzL untersuchte die Verwendung von systolischen Arrays (SAA) als NTT-Beschleuniger und fand eine optimale Hardwarearchitektur für alle Problemgrößen. Die Analyse des Entwurfsraums führte zu einer neuen Konfiguration für einen effizienten zweidimensionalen NTT-Beschleuniger, der die Ausführung anderer Workloads nicht beeinträchtigt. Unsere Ergebnisse zeigen, dass ein optimaler systolischer Array-Beschleuniger für die 22-nm-Technologie eine Fläche von  $53,04 \text{ mm}^2$  benötigt. Die Beschleunigerimplementierung kann NTT effizient auf ein Polynom mit 4096 32-Bit-Ganzzahlkoeffizienten anwenden und benötigt dafür 3296 Zyklen und 1794,92 nJ.

UzL hat eine effiziente Lösung für rekonfigurierbare Edge-Hardware (FPGA) vorgestellt, die SAA nutzt und einen neuartigen Modulo-Beschleuniger-IP (MIP) als Pipeline-Vektorprozessor vorschlägt. Zur Validierung wurde die FPGA-basierte NTT-Implementierung systematisch mit drei verschiedenen Architekturkonfigurationen: einem RISC-V-Softcore, einem RISC-V-Softcore mit SAA und einem RISC-V-Softcore mit SAA und MIP. Die Ergebnisse zeigen, dass SAA die Latenz im Vergleich zu einem RISC-V-Softcore um den Faktor 38,22 verbessert. MIP zusätzlich zu SAA verbessert die Latenz nochmals um einen Faktor 2,61, was zu einer Gesamtverbesserung von 99,63 für einen RISC-V-Softcore mit beiden Beschleunigern führt.

Darüber hinaus hat die UzL Microsoft SEAL (Simple Encrypted Arithmetic Library) getestet, eine Open-Source-Bibliothek für homomorphe Verschlüsselung (HE), die von Microsoft Research entwickelt wurde. SEAL implementiert modernste gitterbasierte homomorphe Verschlüsselungsverfahren, vorwiegend BGV. Insgesamt ermöglicht Microsoft SEAL die Überprüfung von datenschutzfreundlichen Hintergrundverifizierungspipelines, die gut mit modernen Datenschutzbestimmungen und Zero-Trust-Sicherheitsprinzipien übereinstimmen.

**Angreifer mit beschränktem Hintergrundwissen (AP 3.12):** Ziel dieses Arbeitspakets ist die Entwicklung, Erweiterung und Analyse von Modellen, die – im Gegensatz zum klassischen Differential-Privacy-Modell (DP) [22] – nicht die Worst-Case-Annahme eines Angreifers mit vollständigem Hintergrundwissen zugrunde legen, sondern von einer gewissen Unsicherheit bezüglich der zugrunde Datenbank ausgehen. Für viele praktische Anwendungen erweist sich das Differential-Privacy-Modell als zu pessimistisch, da es einen unnötigen Verlust an Datenqualität verursacht. Im technischen Report Statistical Privacy [12] wurde zunächst der Begriff der Statistical Privacy eingeführt und klar vom Konzept der Differential Privacy abgegrenzt. In *Improving Statistical Privacy by Subsampling* [9] wurde ein allgemeines Framework entwickelt, um den Einfluss von Subsampling-Mechanismen auf beliebige Abfragen systematisch zu untersuchen, darunter Poisson-Sampling, Sampling ohne Zurücklegen sowie Sampling mit Zurücklegen (SmZ). Außerdem wurde der genaue Zusammenhang zwischen Statistical Privacy und  $f$ -Differential Privacy [20] mathematisch analysiert.

Um aus diesen Vorarbeiten Kompositionsergebnisse herzuleiten, mussten Methoden entwickelt werden, um mit dem wachsenden Wissen eines Angreifers über mehrere Abfragen umzugehen. Zu diesem Zweck wurde in [11] das in [9] entwickelte Framework erweitert, indem Partitionsmechanismen untersucht wurden. Der vorgeschlagene Mechanismus wurde zudem mittels Simulationen mit aktuellen Kompositionsergebnissen aus der Differential Privacy verglichen [24]. Dabei zeigte sich insbesondere für kleine Datenbanken und eine geringe Anzahl von Abfragen ein deutlicher Vorteil gegenüber herkömmlichen Kompositionsmethoden.

**Feinjustierung von Merkmalsextraktoren (AP 3.2), Gaussian Process Learning (AP 3.3) und Segmentierungsverfahren (AP 3.5 & 4.4):** Wir haben die anonymitätsbewahrende Feinjustierung von Merkmalsextraktoren im Bereich des DP-Finetunings zahlreiche parameter-effiziente Ansätze untersucht, darunter LoRA-ähnliche Methoden, Random Gradient Projection, Second-Order-Optimierung und Sparsity-Promotion. Die Ergebnisse zeigten, dass bestehende DP-SGD-basierte Verfahren [64, 65] einen hohen Reifegrad besitzen und dass zusätzliche PEFT-Methoden unter DP-Bedingungen häufig keinen konsistenten Mehrwert liefern. Für Gauß-Prozess-Klassifikation wurde der Bolt-on DP-SGD Ansatz eingesetzt. Dabei wird der Posterior approximiert und mit der Newton-Methode optimiert. Wir haben notwendige Einschränkungen bezüglich der Sensitivität und Konvexität des Problems abgeleitet, aber diese Forschungsrichtung war nicht zielführend, deswegen haben wir dazu keine Publikation verfasst. Bezüglich der medizinischen Bildsegmentierung hatten wir das Ziel, Verfahren zur medizinischen Bildsegmentierung mit starker Differential Privacy zu entwickeln. Die medizinische Bildsegmentierung ist allerdings inzwischen ein sehr ausgereiftes Forschungsfeld: Es existieren zahlreiche leistungsstarke Medical Foundation Models sowie umfangreiche öffentlich verfügbare Datensätze, die bereits exzellente Ergebnisse liefern. Daher haben wir uns stattdessen auf die

datenschutzkonforme Synthese von Daten mittels Random Projections konzentriert und dafür vielversprechende Resultate erreicht (siehe DP-Scans).

**Angriffe (AP 2.4, AP 4.2- 4.7):** Im Zuge von AnMed wurden mehrere Angriffsstrategien implementiert und evaluiert. Um die Sicherheit zukünftiger Veröffentlichungen aus der elektronischen Patientenakte (EPA) zu testen, wurden Rekonstruktionsangriffe gegen simulierte EPA-Daten gefahren. Diese Angriffe wurden zuvor genutzt, um den US-Census von 2010 anzugreifen ([19]). Die Ergebnisse unserer Experimente zeigen hier, dass eine Veröffentlichung von EPA-Daten ohne Schutzmechanismen die Rekonstruktion von 6% der Bevölkerung zulässt. Zusätzlich wurde durch die erstellte Simulation von EPA-Daten eine Grundlage geschaffen, an der zukünftige Angriffe und Verteidigungsstrategien evaluiert werden können. Darüber hinaus wurden mehrere Forschungsprojekte zur Angreifbarkeit von Decision Trees und Random Forests im Rahmen von AP 4.3: Angriffe auf Gaussian Processes und GBDTs durchgeführt. In einem ersten Projekt [47] wurde die Übertragung eines existierenden State-of-the-Art Data-Reconstruction-Angriffs (LiRA, [18]), der ursprünglich für neuronale Modelle entwickelt wurde, auf Decision Trees und Random Forests umgesetzt und systematisch untersucht. Die Anwendung des Angriffs in diesem neuen Modellkontext lieferte dabei grundlegende Erkenntnisse über den Zusammenhang zwischen der Angreifbarkeit einzelner Datenpunkte und der Verteilung ihrer Modelloutputs, wenn diese Modelle, sogenannte Shadow-Models, unterschiedlich trainiert wurden. Diese Ergebnisse sind nicht nur für den konkreten Anwendungsfall von Bedeutung, sondern liefern auch Impulse für zukünftige Forschung zur Privatsphäre maschinell gelernter Modelle im Allgemeinen. Ein weiteres zentrales Ergebnis des Projekts ist die erneute Identifikation von Overfitting als primäres Privacy-Risiko auch bei Decision-Tree- und Random-Forest-Klassifikatoren. Daran anschließend wurde in einem zweiten Projekt ein neuartiger gradientenbasierter Data-Reconstruction-Angriff für Gradient-Boosted Decision Trees (GBDTs) entwickelt und analysiert [21]. Hierbei gelang ein erster, vielversprechender Schritt, um gradientenbasierte Angriffstechniken auf ein ursprünglich nicht-differenzierbares Trainingsverfahren zu übertragen. Dieser bildet die Grundlage für weitere Forschungsarbeiten zur Verstärkung un stetigen Modellverhaltens durch den Einsatz geeigneter Soft-Approximationen sowie durch speziell angepasste Lossfunktionen, die das Training dieses Angriffes ermöglichen. Die Arbeit zeigt, dass gradientenbasierte Angriffsmethoden grundsätzlich auch auf Modelle und Trainingsverfahren anwendbar sind, die keine explizite Differenzierbarkeit aufweisen, und erweitern damit den Anwendungsbereich dieser Angriffsklasse erheblich.

Zusätzlich wurden mögliche Angriffsvektoren und Datenschutzrisiken im medizinischen bzw. neuropsychologischen Kontext untersucht. Hierzu wurden Kovarianz- und Kovariationsstrukturen in hochdimensionalen Datensätzen analysiert. Aus multiplen bivariaten Regressionsmodellen konnten zusätzliche Modellparameter extrahiert werden, die gezielte Rekonstruktionsangriffe mittels Gradient-Descent-Verfahren ermöglichten. Die Angriffe zur Datenrekonstruktion zeigten eine hohe Erfolgsquote von insgesamt rund 90 % teilweise oder nahezu vollständig rekonstruierter Datensätze und verdeutlichen die erheblichen Risiken bei der Veröffentlichung von Regressionsmodellen im neurophysiologischen Kontext (AP 2.4). Die entsprechenden Ergebnisse und Erkenntnisse wurden in einer Ausarbeitung festgehalten [29]. Auch wurden Schwachstellen in großen Sprachmodellen untersucht. In der Arbeit „PromptPirate“ [42] konnte eine Schwachstelle in Zufallsgeneratoren in weitverbreiteten Machine-Learning-Bibliotheken gefunden werden. Anschließend konnte das ITS zeigen, dass diese Schwachstelle ausgenutzt werden kann, um die Prompts bei der Bildgenerierung mittels Diffusionsmodellen zu stehlen. Aufbauend auf PromptPirate und in Zusammenarbeit mit Wissenschaftlern der Universitäten in Bochum und Ankara (Türkei) hat das ITS sichere und performante Zufallsgeneratoren für den Einsatz in Machine-Learning-Pipelines entwickelt [1]. Als weiteres großes Ziel wurden auch hardwarenahe Angriffe untersucht. Nach einer einführenden Evaluation existierender Seitenkanal-Angriffe entwickelte die UzL in einem ersten Schritt einen Seitenkanal-Angriff mittels EM-Strahlung zur Extrahie-

rung der Modellarchitektur von einem mobilen Hardware-Beschleuniger (Intel Neural Compute Stick). Anschließend entwickelte die UzL in Zusammenarbeit mit Wissenschaftlern der Universität Birmingham einen Fault-Angriff *BarkBeetle* [61] zur Extraktion von Entscheidungsbäumen von einem eingebetteten System (Raspberry Pi RP2350).

Im Rahmen einer Fallstudie entwickelte die UzL eine Plattform für elektromagnetische Fehlerinjektionsangriffe [7]. Hierbei wurde aufgrund einschlägiger Machbarkeitsstudien zunächst das kryptographische Verfahren BIKE angegriffen. Darauf aufbauend wurde im Rahmen einer Masterarbeit eine Softwareplattform für Rowhammer-Angriffe konzipiert. Diese Arbeit mündete in *SLasH-DSA* [8]. In dieser Veröffentlichung wird neben der Softwareplattform ein Angriff gegen das Post-Quantum Signaturverfahren SLH-DSA präsentiert. Dieser Angriff bricht die Authentizitätsgarantien des Verfahrens durch einen Universal-Forgery-Angriff.

**Testsuite von Anonymisierungsverfahren (AP 4.2):** Im Rahmen der Entwicklung der AnoMed Wettbewerbsplattform wurde im Rahmen einer CICD Pipeline eine Testsuite entwickelt, welche die Anonymisierungsverfahren prüft.

**Datenschutzrisiken (AP 4.9)** Wir haben gemeinsam mit den Projektpartnern ULD und UHH eine Arbeit verfasst, welche sowohl die Terminologie von Abstreitbarkeitsgarantien (AP 4.9.5) mit den Begriffen von der k-Anonymitätsforschung erklärt, als auch Strategien zur Nutzbarmachung personenbezogener Gesundheitsdaten (AP 4.9.7) und rechtliche Strategien zur Nutzbarmachung personenbezogener Gesundheitsdaten im Kontext von Restrisiken (AP 4.9.8) aufzeigt [46]. Zusätzlich haben wir in wöchentlichen Besprechungen Rückmeldungen für die zahlreichen Beiträge des Projektpartners ULD gegeben (siehe <https://anomed.de/publikationen/reports>), welche Terminologien für Anonymisierung, Pseudonymisierung, und Abstreitbarkeit, formale Darstellungen von Zustandsdiagrammen aus rechtlicher Perspektive, eine erschöpfende Taxonomie von juristischen Anonymitätsfällen und rechtliche Strategien für den Fall einer und den Umgang mit einer unerwarteten Re-Identifizierung von Daten umfasst.

**Workshops, Schulungen und Konferenzen (AP 5.1):** Das AnoMed-Seminar fand wie geplant alle 14 Tage statt und lud alle Mitarbeiter des AnoMed-Projekts, und Partner, sowie dem AnoMed-Cluster assoziierte Projekte ein. Das Seminar wurde hybrid durchgeführt, um allen die Teilnahme zu ermöglichen. Darüber hinaus wurden die Seminare aufgezeichnet und auf der AnoMed-Homepage veröffentlicht, um das Wissen und die Erkenntnisse mit allen zu teilen. Dieses wird auch in AnoMed-II fortgeführt. Konsortialtreffen wurden im September 2023 und 2024 in Lübeck organisiert und abgehalten. Im Rahmen des öffentlichen KI MED CONNECT Kongresses, wo auch das Konsortialtreffen stattfand, konnten Problemstellungen und Lösungen ausgetauscht werden. Der Wirtschaft und anderen nicht dem Cluster zugehörigen Projekten konnten neueste Forschungsergebnisse präsentiert und in einer Postersession diskutiert werden. Im Oktober 2025 fand abschließend das Treffen aller Clusterpartner „AnoSiDat – Anonymisierung für eine sichere Datennutzung“ in Berlin statt, um Vertretern der Politik, Presse und Wirtschaft die Ergebnisse zu präsentieren und zu diskutieren. Zusätzlich wurde an Workshops mit der TMF, LifeScienceNord, dem HIC Lunch Pitch vom Zentrum für Transfer und Austausch und der Lübeck Summer Academy on Medical Technology teilgenommen. Das Ziel hiervon ist es, dass Anonymisierung und sichere Datennutzung nicht ein abgekapseltes Spezialgebiet bleiben, sondern von allen Datennutzern und Verarbeitern als notwendiges Werkzeug angesehen werden. In der Praxis wollen wir einen offenen Austausch über bestehende und potenzielle Schwächen etablieren. Zusätzlich wurden selbst Schulungen und Workshops angeboten. So unterstützte AnoMed bei der *Summer School on real-world crypto and privacy 2025* sowie der *Summer School on Artificial Intelligence and Cybersecurity*. Im November 2023, 2024 und 2025 wurde an den KI-Erlebnistagen in Lübeck teilgenommen, welche sich an die Öffentlichkeit

richten und Forschung niederschwellig dargestellt wird. Weitere niederschwellige Angebote waren ein Workshop mit KiTa-Angestellten um auch diese im Umgang mit sensiblen Daten zu schulen.

Auf allen diesen Veranstaltungen wurde das im Rahmen von AnoMed entwickelte Brettspiel *Spurensuche in der KI* präsentiert, welches niederschwellig die Risiken von unsicher trainierten Modellen verdeutlicht. Eine digitale Version kann auf der AnoMed Website gespielt werden: <https://anomed.de/brettspiel>. Durch die sehr positive Rückmeldung wird zurzeit ein weiteres Spiel entwickelt, welches die Vorteile von differential privacy vermitteln soll. Dieses wird voraussichtlich im Rahmen von AnoMed-II fertiggestellt. Auch auf Fachkonferenzen war das Cluster stark vertreten. Durch die zahlreichen Veröffentlichungen, war AnoMed unter anderem auf der TCHES 2025, dem Leuven Hardware Summit for Computing on Encrypted Data (COED) 2024 Secure Computation Industry Day 2025, dem HiPEAC25 sowie dem International Symposium on Applied Reconfigurable Computing ARC 2025 vertreten.

## 2.2 Wichtigste Positionen des zahlenmäßigen Nachweises

Die wichtigsten Positionen des zahlenmäßigen Nachweises fallen mit rund 41% auf den Aufbau eines KI-Rechenzentrums für Sichere Datennutzung, mit besonders investiven Beschaffungen sowie technischer Infrastruktur, und mit rund 52% auf die Personalkosten. Weitere Ausgaben sind u.a. für studentische Hilfskräfte und Konferenzteilnahmen angefallen.

## 2.3 Notwendigkeit und Angemessenheit der geleisteten Projektarbeiten

Die sichere Nutzung schützenswerter medizinischer Daten ist notwendig, um sowohl die Patientenversorgung als auch die Forschung zu fördern und somit Innovationspotenziale in der Gesundheitsbranche zu heben. Frühere Verfahren für die sichere Nutzung von schützenswerten Daten führten zu großen Datenschutzrisiken, die von engagierten, ressourcenreichen Angreifern ausgenutzt werden können. Somit ist die Entwicklung neuartiger, sicherer Datennutzungsverfahren absolut unabdingbar. Die in AnoMed erforschten Methoden zur sicheren Datennutzung verringern dieses Risiko signifikant in mehreren Anwendungsbereichen. Ein zentraler Aspekt der Projektarbeit war die Auseinandersetzung mit dem Zielkonflikt zwischen Datenschutz und Nützlichkeit. Auch rückblickend erweist sich dieser Fokus als von besonderem gesellschaftlichem Interesse, da maximaler Privatsphäre-Schutz und maximale Datenqualität nicht gemeinsam erreichbar sind. Die Arbeiten in AnoMed haben diesen Trade-off systematisch analysiert und damit eine realistische Grundlage für die Bewertung von Anonymisierungslösungen geschaffen. Dies war eine wesentliche Voraussetzung, um belastbare Aussagen über Eignung, Grenzen und Einsatzbedingungen unterschiedlicher Verfahren treffen zu können. Darüber hinaus hat das Projekt die Risiken adressiert, die sich aus der Nutzung ausgelagerter und verteilter Datenverarbeitung ergeben. Die zunehmende Komplexität solcher Systeme führt zu zusätzlichen Angriffsflächen und neuen Bedrohungsszenarien, die mit klassischen Sicherheitsannahmen nicht ausreichend erfasst werden. Die in AnoMed durchgeführten Analysen haben diese Risiken aufgezeigt und die praktischen Grenzen bestehender Sicherheitsansätze transparent gemacht, anstatt sie durch vereinfachende Annahmen zu relativieren. Ein weiterer wichtiger Aspekt lag in der Adressierung regulatorischer Unsicherheiten. Unterschiedliche Begriffsverwendungen und Auslegungen im Datenschutzrecht lassen sich nicht unmittelbar in technische Garantien übersetzen. Die interdisziplinäre Herangehensweise in AnoMed hat dazu beigetragen, diese Spannungen sichtbar zu machen und eine sachgerechte Einordnung technischer Ergebnisse zu ermöglichen. Damit wurde das Risiko reduziert, technische Lösungen rechtlich zu überinterpretieren oder umgekehrt regulatorische Anforderungen technisch verkürzt umzusetzen. Ein zentrales Ergebnis des Projekts ist die Entwicklung der Challenge-Plattform, die als Infrastruktur für den Transfer der neuen Erkenntnisse und Methoden in die medizintechnische Community dient. Diese Plattform ermöglicht nicht nur die reproduzierbare Evaluation von Anonymisierungslösungen,

sondern schafft auch einen standardisierten Rahmen für den Vergleich und die Weiterentwicklung von Verfahren. Durch die Integration medizinisch motivierter Privacy-Challenges wurde eine objektive, praxisnahe Bewertung unterschiedlicher Ansätze unter realistischen Bedingungen möglich. Die Plattform unterstützt damit nachhaltig die Forschung, indem sie Fortschritte messbar macht und als Referenz für zukünftige Projekte dient. Zur Wissenschaftskommunikation und Sensibilisierung für Datenschutzrisiken wurden im Projekt innovative Formate wie das AnoMed-Brettspiel entwickelt. Dieses spielerische Format vermittelt auch nicht-technischen Zielgruppen die Herausforderungen und Risiken von Anonymisierungsverfahren und fördert das Verständnis für die Komplexität des Themenfelds. Durch den Einsatz in Workshops und öffentlichen Veranstaltungen konnte das Bewusstsein für sichere Datennutzung gestärkt und ein breiterer Diskurs angeregt werden. Schließlich zeigt sich die Angemessenheit der Projektarbeit auch im Umgang mit Erwartungen an Verwertbarkeit und Anwendbarkeit. AnoMed war bewusst als wissenschaftlich fundiertes Vorhaben mit evaluativem Charakter angelegt. Durch die Entwicklung von Bewertungsmetriken, Challenges und transparenten Vergleichsformaten wurden Fortschritte nachvollziehbar gemacht, ohne eine vorschnelle Operationalisierung der Ergebnisse zu forcieren. Zusammenfassend lässt sich festhalten, dass die Projektarbeit in AnoMed in der Rückschau notwendig war, um zentrale Risiken der sicheren Gesundheitsdatennutzung systematisch zu erfassen und zu analysieren, und angemessen, weil Umfang, Methodik und Zielsetzung der Arbeiten in einem ausgewogenen Verhältnis zur Komplexität und Sensitivität des Themenfeldes standen. Die erzielten Ergebnisse – insbesondere die Challenge-Plattform und die Wissenschaftskommunikationsformate – bilden eine belastbare Grundlage für weiterführende Forschung, regulatorische Einordnung und nachgelagerte Innovationsprozesse.

## **2.4 Voraussichtlicher Nutzen, insbesondere die Verwertbarkeit des Ergebnisses**

Die Förderung des AnoMed Kompetenzzentrum zeigt, dass eine gezielte Förderung zu großem Fortschritt führen kann. In AnoMed haben wir mit acht A\* Publikationen und insgesamt 39 Peerreviewed Publikationen den Stand der Wissenschaft signifikant vorangebracht und viele wissenschaftliche Untersuchungen angestoßen, die in naher Zukunft zu einem verbesserten Verständnis für Lösungen für sichere Datennutzung führen werden. Die wissenschaftlich-technischen Projektergebnisse sind über AnoMed hinaus nutzbar: Die erarbeiteten Grundlagen zu Differential Privacy, Utility-Privacy-Tradeoffs, Merkmalsextraktion sowie zu sicheren ausgelagerten Berechnungen bleiben auf neue Datentypen und Anwendungsszenarien übertragbar. Die Challenge-Plattform ermöglicht auch nach Projektende eine standardisierte Evaluation neuer Verfahren und bildet damit eine langfristige Transfer- und Benchmark-Infrastruktur. Darüber hinaus schafft die interdisziplinäre Übersetzung regulatorischer Anforderungen in technische Schutzgarantien eine übertragbare Brücke zwischen Recht und Technik, die in Leitlinien, Schulungen und der Aufsichtspraxis weiterwirken kann. Die wirtschaftliche Verwertbarkeit der erzielten Ergebnisse ist im Kontext einer Universität als Forschungseinrichtung primär indirekt, aber substanziell. Im Projekt wurden grundlegende Methoden, Konzepte, Software-Prototypen und Evaluationsinstrumente für Anonymisierung und sichere Datennutzung entwickelt, die eine belastbare Basis für eine wirtschaftliche Verwertung durch Dritte darstellen. Durch standardisierte Bewertungsverfahren und demonstratorische Implementierungen werden Entwicklungsrisiken reduziert und die Überführung in marktfähige Produkte insbesondere für KMU und Start-ups erleichtert. Zugleich stärken die Arbeiten nachhaltig die Schnittstelle zwischen Datenschutz, IT-Sicherheit und medizinischer Datenverarbeitung, die für eine datengetriebene Stärkung der Patientenversorgung und entsprechende Geschäftsmodelle im Gesundheitswesen zentral ist. Die hohe Anschlussfähigkeit der Ergebnisse wird durch die Folgefinanzierung von AnoMed-II bestätigt. Die entwickelten Methoden, Plattformen und Kooperationsstrukturen können in der zweiten Phase direkt weiterentwickelt, auf komplexere Datentypen übertragen und gezielt für anwendungsnahe Szenarien (z. B. klinische Bild- und Biosignaldaten, verteilte

Lernverfahren) ausgebaut werden. Die AnoMed Wettbewerbsplattform senkt die Eintrittsbarriere für Innovationen, erhöht die Planungssicherheit für Transferaktivitäten und stellt sicher, dass die aufgebauten Kompetenzen, Infrastrukturen und Netzwerke über die Förderperiode hinaus wirksam bleiben.

## **2.5 Während der Durchführung des Vorhabens dem Zuwendungsempfänger bekannt gewordenen Fortschritt auf dem Gebiet des Vorhabens bei anderen Stellen**

Während der Durchführung des Projekts wurden neue Ansätze zur Generierung synthetischer relationaler Daten veröffentlicht [16, 17]. Da diese Ansätze jedoch keine PRMs verwenden, wie es in unserer Arbeit der Fall ist, hatten sie keinen Einfluss auf unsere Fortschritte. Parallel dazu wurde an anderen Stellen weiter erforscht, welche Eigenschaften einer Fehlerfunktion allgemein zur Robustheit gegen ungewollte Zusatzinformationen beitragen können. Bereits etablierte Annahmen über die Vorteile großer *soft* Margins wurden dabei vertieft und konnten auch in unserer Forschung bestätigt werden. Neue Erkenntnisse zu vorteilhaften geometrischen Eigenschaften im Repräsentationsraum ließen sich bisher nicht direkt von der allgemeinen Klassifikation auf unser spezifisches Szenario übertragen.

Zudem zeigte sich während der Projektlaufzeit, dass die einheitliche und standardisierte Veröffentlichung hochdimensionaler Forschungsdatensätze zunehmend an Bedeutung gewinnt. Besonders in den Neuro- und Kognitionswissenschaften setzt sich der Trend zu etablierten, interoperablen Datenstandards durch, um Transparenz, Nachnutzbarkeit und Vergleichbarkeit komplexer Datensätze zu gewährleisten. In diesem Kontext wurde das Brain Imaging Data Structure (BIDS)-Format kontinuierlich weiterentwickelt und über klassische neurophysiologische Daten hinaus erweitert, etwa um Bewegungs- und Verhaltensdaten, was die Abbildung multimodaler Experimente deutlich erleichtert. Diese Entwicklungen bestätigten den in unserem Projekt verfolgten Ansatz, komplexe Datensätze frühzeitig standardisiert aufzubereiten und öffentlich zugänglich zu machen.

Nach der Preprint-Veröffentlichung von Silenzio [51] erschien eine Arbeit [53], die denselben Use-Case behandelt (nicht interaktives Auslagern des Trainings neuronaler Netze). Diese Arbeit nutzt jedoch andere Techniken, weshalb das ITS einen detaillierten Benchmark-Vergleich plant. Ebenfalls nach Abschluss unserer Arbeiten an DRAMatic, aber noch vor der Veröffentlichung der Ergebnisse, wurde eine weitere Arbeit [6] mit demselben Ziel veröffentlicht – der Beschleunigung von HE auf UPMEM PIM. Auch hier verfolgt das ITS einen anderen Ansatz und bereitet einen umfassenden Benchmark-Vergleich vor.

## **2.6 Erfolgte oder geplante Veröffentlichungen des Ergebnisses nach Nr. 5 der NKBF/NABF**

- Towards Privacy and Utility in Tourette TIC Detection Through Pretraining Based on Publicly Available Video Data of Healthy Subjects [14]
- Efficient Detection of Exchangeable Factors in Factor Graphs [35]
- Colour Passing Revisited: Lifted Model Construction with Commutative Factors [40]
- Distributed DP-Helmet: Scalable Differentially Private Non-interactive Averaging of Single Layers [25]
- DPM: Clustering Sensitive Data through Separation [31]
- S-BDT: Distributed Differentially Private Boosted Decision Trees [49]
- Slalom at the Carnival: Privacy-preserving Inference with Masks from Public Knowledge [15]

- SNPGuard: Remote Attestation of SEV-SNP VMs Using Open Source Tools [62]
- TDXdown: Single-Stepping and Instruction Counting Attacks against Intel TDX [63]
- Efficient Detection of Commutative Factors in Factor Graphs [34]
- Differentially Private Inductive Miner [55]
- Lifting factor graphs with some unknown factors for new individuals [37]
- Dash: Accelerating Distributed Private Convolutional Neural Network Inference with Arithmetic Garbled Circuits [52]
- Statistical Privacy [13]
- Optimizing Systolic Array-based NTT Accelerators [54]
- MammothDP: Differentially Private Boosted Decision Trees, hyperparameter-free and ready for Trusted Hardware [26]
- DP-HYPE: Distributed Differentially Private Hyperparameter Search [30]
- Attacks and Remedies for Randomness in AI: Cryptanalysis of PHILOX and THREE-FRY [1]
- BarkBeetle: Stealing Decision Tree Models with Fault Injection [61]
- Non-omniscient backdoor injection with one poison sample: Proving the one-poison hypothesis for linear regression, linear classification, and 2-layer ReLU neural networks [48]
- DRAMatic Speedup: Accelerating HE Operations on a Processing-in-Memory System [27]
- Silenzio: Secure Non-Interactive Outsourced MLP Training [51]
- Prompt Pirates Need a Map: Stealing Seeds helps Stealing Prompts [42]
- TDXploit: Novel Techniques for Single-Stepping and Cache Attacks on Intel TDX [50]
- Improving Statistical Privacy by Subsampling [10]
- Approximate Lifted Model Construction [39]
- Towards Explainability of Approximate Lifted Model Construction: A Geometric Perspective [58]
- StaRAI: From a Probabilistic Propositional Model to a Highly Compressed Probabilistic Relational Model [23]
- Compression Versus Accuracy: A Hierarchy of Lifted Models [57]
- SLasH-DSA: Breaking SLH-DSA Using an Extensible End-To-End Rowhammer Framework [8]
- ReDASH: Fast and efficient Scaling in Arithmetic Garbled Circuits for Secure Outsourced Inference [43]
- Lifted Model Construction without Normalisation: A Vectorised Approach to Exploit Symmetries in Factor Graphs [36]
- Estimating Causal Effects in Partially Directed Parametric Causal Factor Graphs [41]

- Evaluating the Impact of Face Anonymization Methods on Computer Vision Tasks: A Trade-Off Between Privacy and Utility [59]
- Towards Privacy-Preserving Relational Data Synthesis via Probabilistic Relational Models [38]
- Revealing Unintentional Information Leakage in Low-Dimensional Facial Portrait Representations [3]
- Understanding Differential Privacy in Terms of Crowd Size Reduction [46]
- Mixnets on a Tightrope: Quantifying the Leakage of Mix Networks Using a Provably Optimal Heuristic Adversary [45]
- Neutralizing Unwanted Dependencies Using Nearest-Neighbor Probability Estimation [2]

## Referenzen:

- [1] Jens Alich u. a. „Attacks and Remedies for Randomness in AI: Cryptanalysis of PHILOX and THREEFRY“. In: *IACR Cryptol. ePrint Arch.* (2025), S. 2161. URL: <https://eprint.iacr.org/2025/2161>.
- [2] Kathleen Anderson und Thomas Martinetz. *Neutralizing Unwanted Dependencies Using Nearest-Neighbor Probability Estimation*. 2025. URL: [https://anomed.de/wp-content/uploads/2025/03/Anderson\\_TNNLS\\_25-1.pdf](https://anomed.de/wp-content/uploads/2025/03/Anderson_TNNLS_25-1.pdf).
- [3] Kathleen Anderson und Thomas Martinetz. „Revealing Unintentional Information Leakage in Low-Dimensional Facial Portrait Representations“. In: *International Conference on Artificial Neural Networks*. 2024. URL: <https://arxiv.org/abs/2503.09306>.
- [4] Marshall Ball u. a. „Garbled Neural Networks are Practical“. In: *IACR Cryptol. ePrint Arch.* (2019), S. 338. URL: <https://eprint.iacr.org/2019/338>.
- [5] Andrea Bandini u. a. „A New Dataset for Facial Motion Analysis in Individuals With Neurological Disorders“. In: *IEEE Journal of Biomedical and Health Informatics* 25.4 (2021), S. 1111–1119. DOI: 10.1109/JBHI.2020.3019242.
- [6] Tathagata Barik u. a. „Long Integer NTT Execution on UPMEM-PIM for 128-bit Secure Fully Homomorphic Encryption“. In: *Future Generation Computer Systems* (2026), S. 108386.
- [7] Jeremy Boy, Jack Mähl und Robin Sehm. *I Want To Fault My BIKE: On the Feasibility of Electromagnetic Fault Injection Attacks Against The BIKE Cryptosystem*. [https://www.its.uni-luebeck.de/fileadmin/files/documents/CS\\_JeremyBoy\\_JackMaehl\\_RobinSehm\\_I\\_Want\\_To\\_Fault\\_My\\_BIKE.pdf](https://www.its.uni-luebeck.de/fileadmin/files/documents/CS_JeremyBoy_JackMaehl_RobinSehm_I_Want_To_Fault_My_BIKE.pdf). 2024.
- [8] Jeremy Boy u. a. „SLasH-DSA: Breaking SLH-DSA Using an Extensible End-To-End Rowhammer Framework“. In: *2nd Microarchitecture Security Conference ( $\mu$ ASC '26)*. 2026. URL: <https://arxiv.org/abs/2509.13048>.
- [9] Dennis Breutigam und Rüdiger Reischuk. *Improving Statistical Privacy by Subsampling*. arXiv:2504.11429 [cs.CR]. 2025. arXiv: 2504.11429.
- [10] Dennis Breutigam und Rüdiger Reischuk. *Improving Statistical Privacy by Subsampling*. <https://anomed.de/wp-content/uploads/2025/03/2503-SamplingStatisticalPrivacy.pdf>. 2025.
- [11] Dennis Breutigam und Rüdiger Reischuk. *Parallel Composition for Statistical Privacy*. arXiv:2602.09627 [cs.CR]. 2026. arXiv: 2602.09627.
- [12] Dennis Breutigam und Rüdiger Reischuk. *Statistical Privacy*. arXiv:2501.12893 [cs.CR]. 2025. arXiv: 2501.12893.
- [13] Dennis Breutigam und Rüdiger Reischuk. *Statistical Privacy*. 2025. arXiv: 2501.12893 [cs.CR]. URL: <https://arxiv.org/abs/2501.12893>.
- [14] Nele Brügge u. a. „Towards Privacy and Utility in Tourette TIC Detection Through Pre-training Based on Publicly Available Video Data of Healthy Subjects“. In: *ICASSP*. 2023. URL: <https://ieeexplore.ieee.org/document/10095309>.
- [15] Ida Bruhns u. a. „Slalom at the Carnival: Privacy-preserving Inference with Masks from Public Knowledge“. In: *IACR Communications in Cryptology* 1.3 (7. Okt. 2024). ISSN: 3006-5496. DOI: 10.62056/akp-49qgqx.
- [16] Kuntai Cai, Xiaokui Xiao und Graham Cormode. „PrivLava: Synthesizing Relational Data with Foreign Keys under Differential Privacy“. In: *ACM Manag. Data* 1.2 (Juni 2023). DOI: 10.1145/3589287. URL: <https://doi.org/10.1145/3589287>.

- [17] Kuntai Cai, Xiaokui Xiao und Yin Yang. „PrivPetal: Relational Data Synthesis via Permutation Relations“. In: *Proc. ACM Manag. Data* 3.3 (Juni 2025). DOI: 10.1145/3725341. URL: <https://doi.org/10.1145/3725341>.
- [18] Nicholas Carlini u. a. „Membership Inference Attacks From First Principles“. In: *2022 IEEE Symposium on Security and Privacy (SP)*. 2022, S. 1897–1914. DOI: 10.1109/SP46214.2022.9833649.
- [19] Travis Dick u. a. „Confidence-ranked reconstruction of census microdata from published statistics“. In: *Proceedings of the National Academy of Sciences* 120.8 (2023), e2218605120. URL: <https://www.pnas.org/doi/abs/10.1073/pnas.2218605120>.
- [20] J. Dong, A. Roth und W. Su. „Gaussian Differential Privacy“. In: *J. Royal Statistical Society Series B: Statistical Methodology* (2022).
- [21] Roni Dräther, Marven Kummerfeldt und Esfandiar Mohammadi. *Towards Gradient-Based Data Reconstruction Attacks on Non-Differentiable Models: A Study on GBDTs*. [https://www.its.uni-luebeck.de/fileadmin/files/theses/MA\\_RoniDraether\\_GBDTAttacksMLE.pdf](https://www.its.uni-luebeck.de/fileadmin/files/theses/MA_RoniDraether_GBDTAttacksMLE.pdf). 2026.
- [22] C. Dwork. „Differential Privacy“. In: *Automata, Languages and Programming*. 2006, S. 1–12. ISBN: 978-3-540-35908-1.
- [23] Marcel Gehrke und Malte Luttermann. „StaRAI: From a Probabilistic Propositional Model to a Highly Compressed Probabilistic Relational Model (Extended Abstract)“. In: *Joint Proceedings of the ECSQARU 2025 Workshops and Tutorials*. HAL Open Science, 2025, S. 71–74.
- [24] Peter Kairouz, Sewoong Oh und Pramod Viswanath. „The Composition Theorem for Differential Privacy“. In: *IEEE Trans. Inf. Theor.* 63.6 (Juni 2017), S. 4037–4049. ISSN: 0018-9448. DOI: 10.1109/TIT.2017.2685505.
- [25] Moritz Kirschte u. a. „Distributed DP-Helmet: Scalable Differentially Private Non-interactive Averaging of Single Layers“. In: *CoRR* abs/2211.02003 (2022). DOI: 10.48550/ARXIV.2211.02003. URL: <https://arxiv.org/abs/2211.02003>.
- [26] Moritz Kirschte u. a. *MammothDP: Differentially Private Boosted Decision Trees, hyperparameter-free and ready for Trusted Hardware*. 2025. URL: <https://anomed.de/wp-content/uploads/2025/03/MammothDP-3.pdf>.
- [27] Niklas Klinger u. a. *DRAMatic Speedup: Accelerating HE Operations on a Processing-in-Memory System*. 2026. arXiv: 2602.12433 [cs.CR]. URL: <https://arxiv.org/abs/2602.12433>.
- [28] Marven Kummerfeldt u. a. *DPScans: Differentially Private Data Synthesis using Scans*. 2026. URL: <https://anomed.de/wp/wp-content/uploads/2026/02/dpsscans1.pdf>.
- [29] Marven Kummerfeldt u. a. *Reconstructing Missing Data: Privacy Risks in Regression Models for Human Data*. 2025. URL: [https://anomed.de/wp-content/uploads/2025/03/reconstructing\\_missing\\_data.pdf](https://anomed.de/wp-content/uploads/2025/03/reconstructing_missing_data.pdf).
- [30] Johannes Liebenow, Thorsten Peinemann und Esfandiar Mohammadi. „DP-HYPE: Distributed Differentially Private Hyperparameter Search“. In: *arXiv preprint arXiv:2510.04902* (2025).
- [31] Johannes Liebenow u. a. „DPM: Clustering Sensitive Data through Separation“. In: *CCS*. 2024.
- [32] Ziwei Liu u. a. *Deep Learning Face Attributes in the Wild*. 2015. arXiv: 1411.7766 [cs.CV]. URL: <https://arxiv.org/abs/1411.7766>.

- [33] Patrick Lucey u. a. „Painful data: The UNBC-McMaster shoulder pain expression archive database“. In: *2011 IEEE International Conference on Automatic Face & Gesture Recognition (FG)*. 2011, S. 57–64. DOI: 10.1109/FG.2011.5771462.
- [34] Malte Luttermann, Johann Macheimer und Marcel Gehrke. „Efficient Detection of Commutative Factors in Factor Graphs“. In: *Proceedings of the Twelfth International Conference on Probabilistic Graphical Models (PGM-2024)*. PMLR, 2024, S. 38–56.
- [35] Malte Luttermann, Johann Macheimer und Marcel Gehrke. „Efficient Detection of Exchangeable Factors in Factor Graphs“. In: *Proceedings of the Thirty-Seventh International Florida Artificial Intelligence Research Society Conference (FLAIRS-2024)*. **Best Student Paper**. Florida Online Journals, 2024.
- [36] Malte Luttermann, Ralf Möller und Marcel Gehrke. „Lifted Model Construction without Normalisation: A Vectorised Approach to Exploit Symmetries in Factor Graphs“. In: (2025), 46:1–46:17.
- [37] Malte Luttermann, Ralf Möller und Marcel Gehrke. „Lifting Factor Graphs with Some Unknown Factors for New Individuals“. In: *International Journal of Approximate Reasoning* 179 (2025), S. 109371.
- [38] Malte Luttermann, Ralf Möller und Mattis Hartwig. „Towards Privacy-Preserving Relational Data Synthesis via Probabilistic Relational Models“. In: *Proceedings of the Forty-Seventh German Conference on Artificial Intelligence (KI-2024)*. Springer, 2024, S. 175–189.
- [39] Malte Luttermann u. a. „Approximate Lifted Model Construction“. In: (2025), S. 9077–9085.
- [40] Malte Luttermann u. a. „Colour Passing Revisited: Lifted Model Construction with Commutative Factors“. In: *Proceedings of the Thirty-Eighth AAAI Conference on Artificial Intelligence (AAAI-2024)*. AAAI Press, 2024, S. 20500–20507.
- [41] Malte Luttermann u. a. „Estimating Causal Effects in Partially Directed Parametric Causal Factor Graphs“. In: *Proceedings of the Sixteenth International Conference on Scalable Uncertainty Management (SUM-2024)*. Springer, 2024, S. 265–280.
- [42] Felix Mächtle u. a. „Prompt Pirates Need a Map: Stealing Seeds helps Stealing Prompts“. In: *CoRR* abs/2509.09488 (2025). URL: <https://doi.org/10.48550/arXiv.2509.09488>.
- [43] Felix Maurer, Jonas Sander und Thomas Eisenbarth. *Accelerated Garbled Scaling of Residue Representations with Applications to Secure Machine Learning*. [https://www.its.uni-luebeck.de/fileadmin/files/theses/MA\\_Felix\\_Maurer\\_ReDash.pdf](https://www.its.uni-luebeck.de/fileadmin/files/theses/MA_Felix_Maurer_ReDash.pdf).
- [44] Felix Maurer, Jonas Sander und Thomas Eisenbarth. „ReDASH: Fast and Efficient Scaling in Arithmetic Garbled Circuits for Secure Outsourced Inference“. In: *Applied Cryptography and Network Security Workshops - ACNS 2025 Satellite Workshops: AIHWS, AIoTS, QSHC, SCI, PrivCrypt, SPIQE, SiMLA, and CIMSS 2025, Munich, Germany, June 23-26, 2025, Revised Selected Papers, Part III*. Hrsg. von Mark Manulis. Bd. 15655. Lecture Notes in Computer Science. Springer, 2025, S. 43–51. DOI: 10.1007/978-3-032-01823-6\_3. URL: [https://doi.org/10.1007/978-3-032-01823-6\\_3](https://doi.org/10.1007/978-3-032-01823-6_3).
- [45] Sebastian Meiser u. a. „Mixnets on a Tightrope: Quantifying the Leakage of Mix Networks Using a Provably Optimal Heuristic Adversary“. In: *2025 IEEE Symposium on Security and Privacy (SP)*. 2025, S. 4457–4475. DOI: 10.1109/SP61157.2025.00233.
- [46] Esfandiar Mohammadi u. a. „Understanding Differential Privacy in Terms of Crowd Size Reduction“. In: *Technical Report / Working Paper* (2026). 14 pages; online PDF. URL: [https://anomed.de/wp/wp-content/uploads/2026/02/The\\_promise\\_of\\_k\\_anonymity-14.pdf](https://anomed.de/wp/wp-content/uploads/2026/02/The_promise_of_k_anonymity-14.pdf).

- [47] Julius Benedict Niehoff u. a. *Extending the Likelihood Ratio Attack to Random Forest Models*. [https://www.its.uni-luebeck.de/fileadmin/files/theses/BA\\_JuliusNiehoff\\_LiRA\\_RF.pdf](https://www.its.uni-luebeck.de/fileadmin/files/theses/BA_JuliusNiehoff_LiRA_RF.pdf). 2025.
- [48] Thorsten Peinemann u. a. *Non-omniscient backdoor injection with one poison sample: Proving the one-poison hypothesis for linear regression, linear classification, and 2-layer ReLU neural networks*. 2026. arXiv: 2508.05600 [cs.LG]. URL: <https://arxiv.org/abs/2508.05600>.
- [49] Thorsten Peinemann u. a. „S-BDT: Distributed Differentially Private Boosted Decision Trees“. In: *Proceedings of the 2024 on ACM SIGSAC Conference on Computer and Communications Security*. CCS '24. Salt Lake City, UT, USA: Association for Computing Machinery, 2024, S. 288–302. ISBN: 9798400706363. DOI: 10.1145/3658644.3690301. URL: <https://doi.org/10.1145/3658644.3690301>.
- [50] Fabian Rauscher u. a. „TDXploit: Novel Techniques for Single-Stepping and Cache Attacks on Intel TDX“. In: *Proceedings of the 34th USENIX Security Symposium (USENIX Security '25)*. USENIX Security Symposium. Long Presentation. Seattle, WA, USA: USENIX Association, 2025, S. 1207–1222. URL: <https://www.usenix.org/conference/usenixsecurity25/presentation/rauscher>.
- [51] Jonas Sander und Thomas Eisenbarth. „Silenzio: Secure Non-Interactive Outsourced MLP Training“. In: *CoRR* abs/2504.17785 (2025). DOI: 10.48550/ARXIV.2504.17785. arXiv: 2504.17785. URL: <https://doi.org/10.48550/arXiv.2504.17785>.
- [52] Jonas Sander u. a. „Dash: Accelerating Distributed Private Convolutional Neural Network Inference with Arithmetic Garbled Circuits“. In: *IACR TCHES* 2025.1 (2025).
- [53] Thomas Schneider, Huan-Chih Wang und Hossein Yalame. „HE-SecureNet: An Efficient and Usable Framework for Model Training via Homomorphic Encryption“. In: *IACR Cryptol. ePrint Arch.* (2025), S. 1591. URL: <https://eprint.iacr.org/2025/1591>.
- [54] Eike Schultz u. a. „Optimizing Systolic Array-based NTT Accelerators“. In: *Authorea Preprints* (2025).
- [55] Max Schulze u. a. „Differentially Private Inductive Miner“. In: *2024 6th International Conference on Process Mining (ICPM)*. 2024, S. 89–96. DOI: 10.1109/ICPM63005.2024.10680684.
- [56] Yara Schütt und Esfandiar Mohammadi. *Understanding the Theoretical Guarantees of DPM*. 2025. URL: [https://anomed.de/wp-content/uploads/2025/03/DPM\\_utilityAnalysis.pdf](https://anomed.de/wp-content/uploads/2025/03/DPM_utilityAnalysis.pdf).
- [57] Jan Speller u. a. „Compression versus Accuracy: A Hierarchy of Lifted Models“. In: *Proceedings of the Twenty-Eighth European Conference on Artificial Intelligence (ECAI-2025)*. IOS Press, 2025, S. 5051–5058.
- [58] Jan Speller u. a. „Towards Explainability of Approximate Lifted Model Construction: A Geometric Perspective“. In: *Proceedings of the Eleventh Workshop on Formal and Cognitive Reasoning (FCR-2025)*. CEUR, 2025, S. 41–56.
- [59] Roland Stenger u. a. „Evaluating the Impact of Face Anonymization Methods on Computer Vision Tasks: A Trade-Off Between Privacy and Utility“. In: *IEEE Access* 13 (2025). DOI: 10.1109/ACCESS.2024.3519441.
- [60] Florian Tramèr und Dan Boneh. „Slalom: Fast, Verifiable and Private Execution of Neural Networks in Trusted Hardware“. In: *7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019*. OpenReview.net, 2019. URL: <https://openreview.net/forum?id=rJVorjCckQ>.

- [61] Qifan Wang u. a. „BarkBeetle: Stealing Decision Tree Models with Fault Injection“. In: *CoRR* abs/2507.06986 (2025). DOI: 10.48550/ARXIV.2507.06986. arXiv: 2507.06986. URL: <https://doi.org/10.48550/arXiv.2507.06986>.
- [62] Luca Wilke und Gianluca Scopelliti. „SNPGuard: Remote Attestation of SEV-SNP VMs Using Open Source Tools“. In: *EuroS&PW*. 2024. DOI: 10.1109/EuroSPW61312.2024.00026.
- [63] Luca Wilke, Florian Sieck und Thomas Eisenbarth. „TDXdown: Single-Stepping and Instruction Counting Attacks against Intel TDX“. In: *CCS*. 2024. URL: <https://dl.acm.org/doi/epdf/10.1145/3658644.3690230>.
- [64] Da Yu u. a. „Differentially private fine-tuning of language models“. In: *International Conference on Learning Representations (ICLR)*. 2022.
- [65] Da Yu u. a. „Large Scale Private Learning via Low-Rank Reparametrization“. In: *Proceedings of the International Conference on Machine Learning (ICML)*. 2021.