Fraunhofer FIT

Institutsteil Wirtschaftsinformatik,
Fraunhofer-Institut für Angewandte
Informationstechnik FIT

SplitNFed

Datenschutzkonforme und effiziente künstliche Intelligenz mit Split & Federated Learning

SplitNFed

Datenschutzkonforme und effiziente künstliche Intelligenz mit Split & Federated Learning

Autoren

Lara Gerlach, Valentin Götz, Dr. Tobias Guggenberger, Prof. Dr. Björn Häckel, Berit Helmus, Dr. Thomas Kreuzer, Marc Principato, Prof. Dr. Nils Urbach, Nina Weber, Florian Weiß, Domenique Zipperling

Über das Fraunhofer FIT

Unsere Ambition ist es, Themen der Wirtschaftsinformatik inhaltlich wie methodisch umfassend auf höchstem Niveau abzudecken. Gemeinsam mit unseren Partner*innen aus Wirtschaft und Gesellschaft erarbeiten wir auf Basis unserer fachlichen und technischen Expertise innovative Lösungen für individuelle Probleme. Unsere Lösungen betrachten dabei sowohl alle Ebenen der Unternehmensarchitektur integriert als auch die Einbettung von Unternehmen in digitale Wertschöpfungsnetze. Zudem bieten wir Impulse für Digitalisierungsstrategien und transformative Veränderungsprozesse in Unternehmen.

Danksagung

Dieser Artikel wurde durch das Bayerische Staatsministerium für Wirtschaft, Landesentwicklung und Energie im Rahmen des Projekts “Fraunhofer Blockchain Center (20-3066-2-6-14)” gefördert. Wir danken an dieser Stelle für die Unterstützung.

Gefördert durch

Bayerisches Staatsministerium für
Wirtschaft, Landesentwicklung und Energie

Kontakt:

Tobias Guggenberger | tobias.guggenberger@fit.fraunhofer.de

Fraunhofer-Institut für Angewandte Informationstechnik FIT

Institutsteil Wirtschaftsinformatik

Wittelsbacherring 10

95444 Bayreuth

Disclaimer

Dieses Whitepaper wurde vom Fraunhofer-Institut für Angewandte Informationstechnik FIT nach bestem Wissen und unter Einhaltung der nötigen Sorgfalt erstellt. Fraunhofer FIT, seine gesetzlichen Vertreter*innen und/oder Erfüllungsgehilfen übernehmen keinerlei Garantie dafür, dass die Inhalte dieses Whitepapers gesichert, vollständig für bestimmte Zwecke brauchbar oder in sonstiger Weise frei von Fehlern sind. Die Nutzung dieses Whitepapers geschieht ausschließlich auf eigene Verantwortung. In keinem Fall haften das Fraunhofer FIT, seine gesetzlichen Vertreter*innen und/oder Erfüllungsgehilfen für jegliche Schäden, seien sie mittelbar oder unmittelbar, die aus der Nutzung des Whitepapers resultieren.

Empfohlen Zitierweise

Gerlach, L., Götz, V., Guggenberger, T., Häckel, B., Helmus, B., Kreuzer, T., Principato, M., Urbach, N., Weber, N., Weiß, F., und Zipperling, D. (2025). *SplitNFed: Datenschutzkonforme und effiziente künstliche Intelligenz mit Split & Federated Learning.* Fraunhofer FIT Institutsteil Wirtschaftsinformatik, Augsburg / Bayreuth.

Aktueller Stand: Oktober 2025

Highlights

- **Einführung in verteiltes Lernen**
  In diesem Whitepaper führen wir in das Konzept des verteilten Lernens ein, mit besonderem Fokus auf Split Learning und Federated Learning. Unser Ziel ist es, Entscheidungsträgern die praktische Anwendung und Umsetzung dieser Technologien zu erleichtern.
- **Vorstellung des Ansatzes SplitNFed**
  In diesem Whitepaper stellen wir den Ansatz SplitNFed vor, der die Stärken von Split Learning und Federated Learning vereint. Unser Ansatz ermöglicht ein datensicheres, effizientes und skalierbares Training von Künstlicher Intelligenz.
  - **Mehrwert: Volle Datenhoheit**
    Mit SplitNFed behalten Unternehmen die Kontrolle über sensible Daten, da diese während des gesamten Trainings der Künstlichen Intelligenz lokal bleiben.
  - **Mehrwert: Effiziente Ressourcennutzung**
    Mit SplitNFed nutzen Unternehmen ihre vorhandenen lokalen Rechenressourcen effizient. Sie senken Betriebskosten und optimieren die Auslastung ihrer bestehenden IT-Infrastruktur.
  - **Mehrwert: Skalierbarkeit auch bei begrenzter Netzwerkinfrastruktur**
    Mit SplitNFed binden Unternehmen Standorte mit geringer Bandbreite nahtlos in das KI-Training ein. Dadurch schaffen sie eine wirtschaftlich tragfähige Grundlage für skalierbare Anwendungen von Künstlicher Intelligenz.
- **Hohe Relevanz von SplitNFed für datenintensive, regulierte Branchen**
  Mit SplitNFed können Unternehmen Künstliche Intelligenz auch in stark regulierten und datenintensiven Branchen wie dem Finanz- und Gesundheitswesen sicher und effizient einsetzen. Es hilft ihnen dabei, Anforderungen an Datenschutz, Effizienz und Verfügbarkeit zu erfüllen.

Management Summary

Der Einsatz von Künstlicher Intelligenz ist heutzutage ein entscheidender Treiber für Wettbewerbs- und Innovationsfähigkeit von Unternehmen. Europäische Unternehmen schöpfen das Potenzial bisher jedoch nicht vollständig aus, da sie zwei wesentlichen Herausforderungen gegenüberstehen: Erstens gefährdet die hohe Abhängigkeit von Cloud-Anbieter*innen, die häufig in den USA ansässig sind, die souveräne Durchführung. Zweitens hemmt die ineffiziente Nutzung lokaler Ressourcen die wirtschaftliche Nutzung.

Damit Europa seine Potenziale nutzen kann, ist der Aufbau skalierbarer, datensouveräner Infrastrukturen für das Training von Künstlicher Intelligenz notwendig. Unser Ansatz „SplitNFed" adressiert diese Herausforderungen. SplitNFed erweitert das Verfahren des Split Learning gezielt um Mechanismen des Federated Learning. Beim Training Künstlicher Intelligenz mit SplitNFed verbleiben Daten vollständig im Unternehmen, während lediglich modellbezogene Informationen übertragen werden. Dazu ermöglicht SplitNFed die Einbindung vorhandener lokaler Rechenressourcen.

Durch diese Vorteile werden Unternehmensdaten effektiv geschützt, lokale Hardware-Ressourcen optimal eingebunden und technologische Abhängigkeiten reduziert. SplitNFed ermöglicht stabile, skalierbare Trainingsprozesse auch bei eingeschränkter Netzwerkinfrastruktur und schafft so die Grundlage für eine wirtschaftliche und zukunftsfähige Nutzung von Künstlicher Intelligenz.

Dieses Whitepaper richtet sich an CEOs, CTOs, Informationstechnologie-Architekt*innen und -Entwickler*innen sowie an alle interessierten Leser*innen. Es hilft, die strategische Bedeutung souveräner Infrastrukturen von Künstlicher Intelligenz einzuordnen und konkrete Handlungsoptionen für den Aufbau eigener verteilter Systeme zu entwickeln. Mit SplitNFed stellen wir einen Ansatz vor, mit dem Unternehmen ihre Daten beim Training von Künstlicher Intelligenz wirksam schützen, bestehende Ressourcen effizienter nutzen und technologische Unabhängigkeit nachhaltig ausbauen können. Das Verfahren kann in bestehende Informationstechnologie-Landschaften integriert werden und zeigt konkrete Wege zu wirtschaftlicher und regulatorisch sicherer Nutzung von Künstlicher Intelligenz auf.

Abkürzungsverzeichnis

FIT	Fraunhofer-Institut für Angewandte Informationstechnik
IT	Informationstechnologie
KI	Künstliche Intelligenz

Abbildungsverzeichnis

Inhaltsverzeichnis

1

Einleitung

1 Einleitung

1.1 Chance & Herausforderung der Künstlichen Intelligenz für europäische Unternehmen

Künstliche Intelligenz (KI) entwickelt sich zunehmend zur Schlüsseltechnologie für Unternehmen, indem sie die Grundlage für eine dynamische Innovationsfähigkeit schafft (Martínez et al. 2024; Singla et al. 2024), die Entlastung von Fachkräften unterstützt (McKinsey&Company 2023) und zur Erhöhung der unternehmerischen Resilienz beiträgt (Martínez et al. 2024). Doch viele europäische Unternehmen nutzen diese Potenziale nicht, da sie mit zwei zentralen Herausforderungen konfrontiert sind: **Datenschutz** und **Rechenkapazität.**

Damit Unternehmensdaten beim KI-Training geschützt bleiben, benötigen europäische Unternehmen eine vertrauenswürdige Umgebung. Viele Unternehmen verfügen dabei nicht über die technischen Voraussetzungen, KI-Anwendungen unabhängig zu betreiben, sodass Cloud-Infrastrukturen zur bevorzugten Lösung für speicher- und rechenintensive Trainingsprozesse geworden sind (Allam 2023). Über 65 % der weltweiten Cloud-Infrastrukturen werden jedoch von den drei führenden US-amerikanischen Anbietern Microsoft, Amazon und Google kontrolliert (Canalys 2025). Diese Konzentration auf US-amerikanische Cloud-Anbieter*innen schränkt nicht nur die Auswahlmöglichkeiten für Unternehmen ein und birgt Datenschutz- sowie Compliance-Risiken (Alex Mathew 2024), sondern gefährdet auch Europas digitale Souveränität.

Gleichzeitig bleiben unternehmenseigene Infrastrukturen der Informationstechnologie (IT) außerhalb zentralisierter Cloud-Plattformen bislang weitgehend ungenutzt. So verfügen viele Organisationen über lokal betriebene Server- und Rechensysteme, deren durchschnittliche Auslastung unter 40 % liegt (Hintermann et al. 2024). Dieses ungenutzte Potenzial bleibt jedoch weitgehend isoliert, da es an technischen und organisatorischen Voraussetzungen fehlt, um lokal verfügbare Ressourcen standortübergreifend nutzbar zu machen. Dadurch entsteht ein scheinbarer Mangel an Rechenkapazität, obwohl physisch ausreichende Ressourcen vorhanden wären.

Vor diesem Hintergrund stellen wir in diesem Whitepaper unser Verfahren SplitNFed vor, ein bisher im Forschungskontext entwickeltes, noch nicht praktisch erprobtes Konzept, das gezielt die begrenzte Nutzbarkeit verteilter Rechenressourcen sowie datenschutzbedingte Einschränkungen beim KI-Training adressiert.

1.2 SplitNFed als Ausgangspunkt für eine effiziente und souveräne KI-Infrastruktur in Europa

Stellen Sie sich vor, mehrere europäische Krankenhäuser möchten gemeinsam ein KI-Modell zur Früherkennung seltener Krankheitsverläufe auf Basis medizinischer Bilddaten entwickeln. Dabei sehen sie sich mit einer doppelten Herausforderung konfrontiert. Zum einen sind Bilddaten im Gesundheitswesen äußerst sensibel und ihre Vertraulichkeit und Integrität müssen jederzeit gewahrt bleiben. Eine Verarbeitung in einer US-amerikanischen Cloud würde dies gefährden, da

Patientendaten dort nicht mehr dem alleinigen Schutz des europäischen Datenschutzrechts unterliegen und potenziell ohne Zustimmung offengelegt werden könnten (Alex Mathew 2024). Zum anderen verfügen viele Kliniken nicht über eine IT-Infrastruktur, die die hohen Anforderungen für das Training komplexer KI-Modelle erfüllt. Es wird daher eine Lösung benötigt, die sowohl den Schutz sensibler Daten sicherstellt als auch den begrenzten technischen Ressourcen der Krankenhäuser Rechnung trägt.

An dieser Stelle setzt das Verfahren SplitNFed an, welches eine Lösungsmöglichkeit für die beiden Herausforderungen darstellt. **SplitNFed ist ein verteiltes und kollaboratives Verfahren des maschinellen Lernens**, das die Stärken von Split Learning und Federated Learning vereint. Es ist ein skalierbarer Ansatz, der ein **effizientes und datenschutzkonformes Training** von KI-Modellen ermöglicht. Mit SplitNFed werden die erste und letzte Schicht des Modells lokal trainiert, wodurch die sensiblen Trainingsdaten das lokale System nicht verlassen. Gleichzeitig kann die Verarbeitung der verbleibenden Modellteile auf verschiedenen Servern stattfinden. Diese Aufteilung ermöglicht es, Berechnungen auf mehrere Geräte zu verteilen und somit effizienter zu trainieren, ohne die sensiblen Daten einem erhöhten Risiko auszusetzen. Die Kommunikation zwischen den Geräten im Trainingsprozess beschränkt sich hierbei auf ein Minimum, wodurch die Anforderungen an die Netzwerkinfrastruktur gering bleiben. So können auch Unternehmen und Einrichtungen mit begrenzter Bandbreite unter vollständiger Wahrung der Datenhoheit in die verteilten Trainingsprozesse eingebunden werden.

Das Verfahren ist **für jedes europäische Unternehmen von Interesse**, da es eine sichere und effiziente Nutzung von KI ermöglicht. Besonders gut eignet sich SplitNFed für **datenintensive, regulierte Branchen**, in denen Datenschutz, Effizienz und Verfügbarkeit höchste Priorität haben, wie beispielsweise im Finanz- und Gesundheitswesen. Wir schaffen mit SplitNFed die Grundlage für eine zukunftsweisende KI-Entwicklung in Europa und unterstützen **CEOs und CTOs, IT-Architekt*innen und -Entwickler*innen** dabei, die strategische Bedeutung souveräner Infrastrukturen von KI einzuordnen und konkrete Handlungsoptionen für den Aufbau eigener verteilter Systeme zu entwickeln.

Im weiteren Verlauf dieses Whitepapers stellen wir zunächst den Status Quo des KI-Trainings in Europa dar und beleuchten Herausforderungen, mit denen europäische Unternehmen beim KI-Training konfrontiert sind (Kapitel 2). Anschließend stellen wir bestehende Ansätze des verteilten Lernens, namentlich Split Learning und Federated Learning, vor. Darauf aufbauend präsentieren wir das Verfahren SplitNFed (Kapitel 3) und zeigen mögliche Anwendungsszenarien auf (Kapitel 4). Abschließend zeigen wir strategische sowie operative Wettbewerbsvorteile für Unternehmen auf (Kapitel 5) und reflektieren das Konzept (Kapitel 6).

2

Status Quo

2 Status Quo

2.1 Zwischen Cloud-Abhängigkeit und ungenutzten Potenzialen

Die KI-Wertschöpfung klassischer zentraler Modelle, lässt sich grundlegend in zwei eng miteinander verknüpfte Teilbereiche gliedern: **Datenspeicherung** und **Datenverarbeitung**. Derzeit greifen europäische Akteur*innen in beiden Bereichen überwiegend auf außereuropäische Cloud-Infrastrukturen zurück, zum Beispiel von US-Anbieter*innen. Dies gefährdet die digitale Souveränität Europas. Darunter versteht man die Fähigkeit, digitale Technologien, Infrastrukturen und Daten selbstbestimmt kontrollieren zu können. Diese Elemente prägen moderne Gesellschaften grundlegend (Pohle und Thiel 2020). Dazu erhöht die zentrale Speicherung und Verarbeitung sensibler Daten auf global zugänglichen Infrastrukturen das Risiko von Cyberangriffen und erschwert die Transparenz über Sicherheitsmaßnahmen (Brender und Markov 2013). Erschwerend kommt hinzu, dass der CLOUD Act US-Behörden den Zugriff auf Daten von US-Cloud-Anbieter*inne erlaubt, selbst wenn diese außerhalb der USA gespeichert sind, womit die die Vertraulichkeit der Daten potenziell gefährdet wird (Rojszczak 2020). In Folge behindert die Abhängigkeit von außereuropäischen Cloud-Infrastrukturen eine breite Nutzung von KI durch europäische Akteure sowie Akteurinnen und unterstreicht den akuten Handlungsbedarf. Abbildung 1 zeigt hierbei aktuelle Lösungsansätze zur Stärkung der digitalen Souveränität in Europa auf.

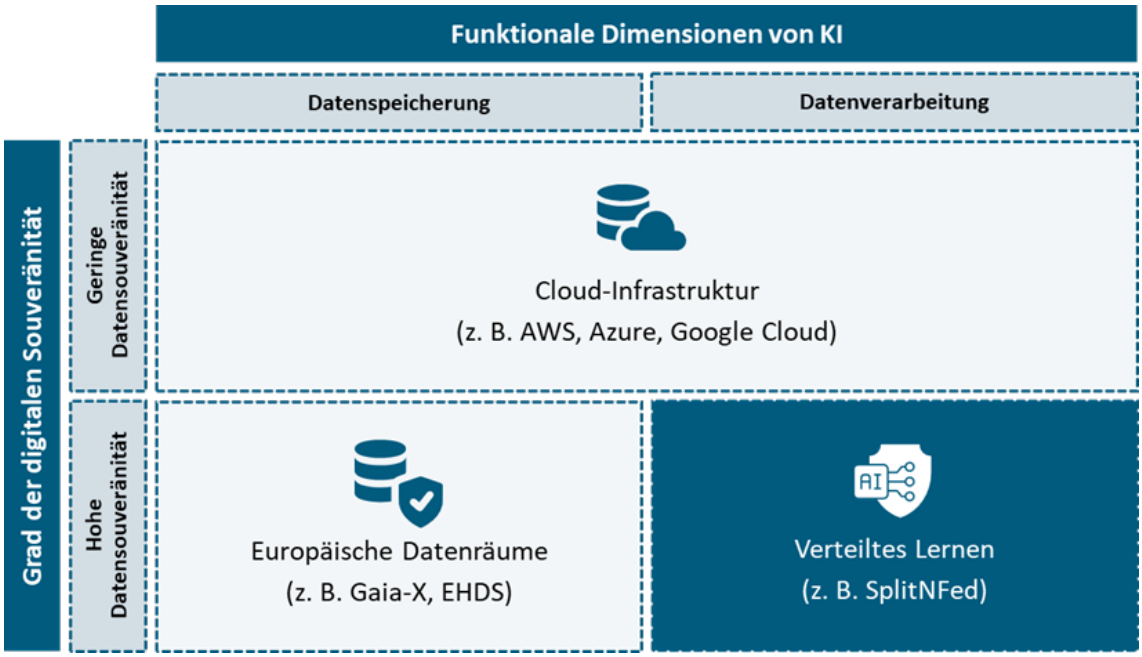

Abbildung 1: Ansätze zur Erreichung digitaler Souveränität in der KI-Wertschöpfung nach Datenspeicherung und -verarbeitung

Der erste Bereich, die *Datenspeicherung*, wird zunehmend durch den Aufbau **europäischer Datenräume** adressiert. Initiativen wie Gaia-X, aus Wirtschaft, Forschung und Politik, zielen auf

den Aufbau föderierter, transparenter und datenschutzkonformer Dateninfrastrukturen. Dabei stehen europäische Werte wie Interoperabilität, Portabilität und Rechtssicherheit im Vordergrund (Gaia-X European Association for Data and Cloud AISBL 2023; Schmitz und Mitrovic 2024). In der Praxis stoßen solche Initiativen jedoch an Grenzen, insbesondere in Anwendungsszenarien, in denen mehrere Unternehmen gemeinsam KI-Modelle trainieren wollen. Selbst föderierte Infrastrukturen setzen häufig voraus, dass Daten zumindest logisch geteilt oder zentral verfügbar gemacht werden. Doch genau dies scheitert oft an datenschutzrechtlichen Vorgaben, unternehmensinternen Compliance-Anforderungen oder auch am fehlenden Vertrauen zwischen den Beteiligten. Eine zentrale oder geteilte Datenspeicherung ist in solchen Fällen für Unternehmen keine realistische Option.

Vor diesem Hintergrund gewinnen **verteilte Lernverfahren** an Bedeutung. Diese ermöglichen das gemeinsame Training von Modellen, ohne dass Rohdaten zwischen den Unternehmen ausgetauscht oder zentral gespeichert werden müssen. Zwei etablierte Ansätze dieser Form der *Datenverarbeitung* sind **Federated Learning** und **Split Learning**. Beide Ansätze stärken die Datensouveränität und eröffnen neue Möglichkeiten für kollaboratives KI-Training. Jedoch bringen auch verteilte Lernverfahren distinkte Nachteile mit sich, die einen Einsatz erschweren (Kairouz et al. 2021):

- **Räumliche Verteilung der Rechenstandorte**: Hohe Latenzen im Netzwerk und begrenzte Bandbreiten erschweren die Synchronisation von Modellparametern und machen die Kommunikation zum zentralen Engpass.
- **Technologische Heterogenität der beteiligten Systeme**: Unterschiedliche Software-Stacks, Betriebssysteme und Hardwarekonfigurationen führen zu Inkompatibilitäten, die den Einsatz einheitlicher Trainingsarchitekturen erschweren und den Integrationsaufwand erhöhen.
- **Vielfalt der beteiligten Akteure und Akteurinnen**: Unterschiedliche Interessenlagen, Sicherheitsstandards und regulatorische Anforderungen der Akteure erschweren eine koordinierte Umsetzung. Dies betrifft insbesondere unternehmensübergreifende Kontexte, in denen der Aufbau vertrauenswürdiger, verteilter Lernumgebungen zusätzliche Abstimmungsprozesse erfordert.

Um diese Herausforderungen zu adressieren, wurde das Verfahren **SplitNFed** entwickelt. Es **kombiniert die Stärken von Federated und Split Learning** und überwindet deren technischen und organisatorischen Limitationen. SplitNFed bietet eine skalierbare Lösung für eine souveräne, verteilte Verarbeitung sensibler Daten in Europa. Durch seinen modularen Aufbau, die Reduktion zentraler Abhängigkeiten und die Möglichkeit zur Integration unterschiedlicher Sicherheits- und Datenschutzanforderungen erleichtert SplitNFed die Zusammenarbeit heterogener Akteur*innen, auch über Organisationsgrenzen hinweg. Damit besteht ebenfalls die Möglichkeiten verteilte Rechenkapazitäten auszureizen und Ineffizienzen in IT-Infrastrukturen zu beseitigen (Hintermann et al. 2024).

Zum besseren Verständnis des Verfahrens werden im Folgenden zunächst die zugrunde liegenden Ansätze Federated Learning und Split Learning vorgestellt, bevor anschließend die Funktionsweise und die Mehrwerte von SplitNFed erläutert werden.

2.2 Split und Federated Learning als Wegbereiter*in datensouveräner KI-Systeme

2.2.1 Neuronale Netze: Die Grundlage für effizientes Lernen

Split Learning und Federated Learning zielen darauf ab, das Training großer KI-Modelle sowohl datenschutzkonform als auch effizient zu gestalten (Thapa et al. 2020). Sie basieren in ihrer technischen Umsetzung auf **neuronalen Netzen**. Um die Funktionsweise dieser Verfahren sowie ihre Potenziale und Schwächen besser zu verstehen, ist ein grundlegendes Verständnis des Aufbaus und des Trainings neuronaler Netze erforderlich.

Neuronale Netze stellen eine zentrale Klasse von Algorithmen im übergeordneten Bereich des maschinellen Lernens dar (Goodfellow et al. 2016). **Maschinelles Lernen** umfasst dabei Methoden, bei denen Modelle durch das Erkennen statistischer Muster in Daten selbstständig lernen, anstatt auf fest programmierte Regeln angewiesen zu sein. Innerhalb dieses Spektrums wurden neuronale Netze speziell dafür entwickelt, komplexe, nichtlineare Zusammenhänge in großen Datenmengen zu erfassen und abzubilden. Inspiriert von den neuronalen Strukturen des menschlichen Gehirns bestehen sie aus zahlreichen miteinander verbundenen Einheiten (Neuronen), die Informationen verarbeiten, abstrahieren und Muster identifizieren. Diese vernetzten Systeme ermöglichen es, umfangreiche Datenmengen zu analysieren und komplexe Zusammenhänge aufzudecken, wodurch sie zu den leistungsstärksten und am weitesten verbreiteten Algorithmen im Maschinellen Lernen zählen.

Abbildung 2 veranschaulicht die grundlegende Funktionsweise: Neuronale Netze verarbeiten eingehende Informationen Schicht für Schicht (Forward Pass) und erstellen daraus zunächst eine Vorhersage. Diese wird mit dem tatsächlichen Ergebnis verglichen, woraufhin das Netz seine Gewichtungen anpasst (Backward Pass). Mit jeder zusätzlichen Trainingsinstanz verbessert sich so die Genauigkeit des Modells. Auf diese Weise entstehen skalierbare Entscheidungshilfen, die klassischen regelbasierten Systemen oft überlegen sind. Diese Methode erlaubt es, komplexe Muster zu erlernen und bildet die Grundlage für die Entwicklung fortschrittlicher KI-Lösungen.

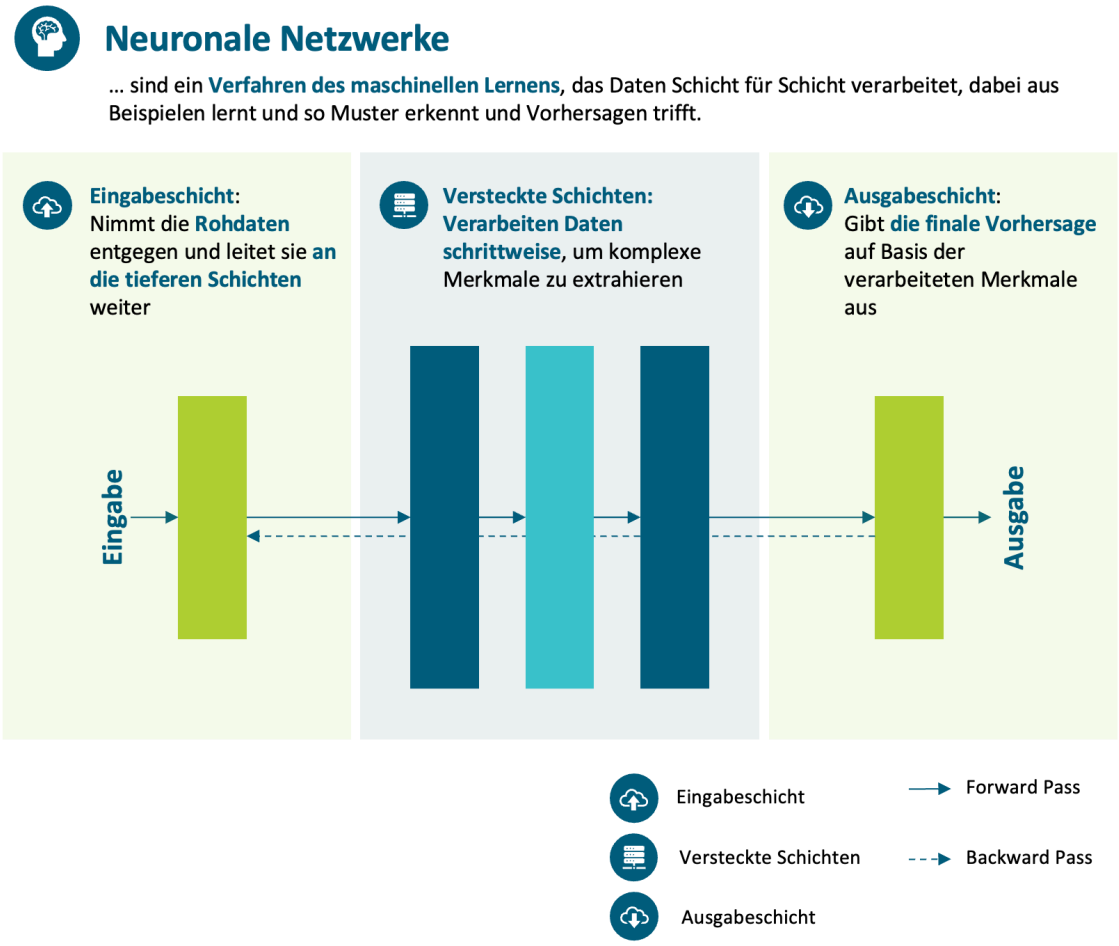

Abbildung 2: Neuronale Netzwerke

2.2.2 Split Learning: Effizientes Lernen bei sensiblen Daten und knappen Ressourcen

Split Learning wurde 2018 am MIT Media Lab vorgestellt (Vepakomma et al. 2018) und richtet sich insbesondere an **Szenarien mit begrenzten lokalen Rechenressourcen und strengen Datenschutzanforderungen**. Das Verfahren wurde ursprünglich für den Gesundheitsbereich entwickelt, in dem Vorschriften wie der Health Insurance Portability and Accountability Act in den USA strenge Anforderungen an die zentrale Verarbeitung sensibler Daten stellen (Vepakomma et al. 2018). Um diese Anforderungen einzuhalten, ermöglicht Split Learning die Aufteilung eines neuronalen Netzwerks in separate Segmente, die jeweils lokal trainiert werden. Dadurch verbleiben die Rohdaten stets auf dem lokalen System, sodass verschiedene Institutionen kooperativ Modelle trainieren können, ohne vertrauliche Informationen preiszugeben.

Das grundlegende Prinzip von Split Learning besteht darin, dass ein neuronales Netzwerk in mehrere Teilbereiche aufgeteilt wird, die jeweils getrennt trainiert werden. Anders als beim herkömmlichen zentralen Training werden somit nicht alle Schichten des Modells lokal an einem Ort verarbeitet (Vepakomma et al. 2018). In der dargestellten Konfiguration (vgl. Abbildung 3) trainiert die Dateninhaber*in zunächst die ersten, kleineren Schichten lokal auf eigener Hardware. Die daraus resultierenden Zwischenergebnisse werden anschließend verschlüsselt an einen

oder mehrere externe Akteure bzw. Akteurinnen übermittelt, die als Eigentümer*in von Rechenressourcen die größeren Modellsegmente ohne Zugriff auf die Rohdaten verarbeiten. Abschließend werden die Ausgaben dieser zentral trainierten Schichten zurück an den ursprünglichen Dateninhaber*in übermittelt, der die finalen Modellschichten lokal trainiert und mögliche Feinanpassungen im Rahmen der Rückpropagation vornimmt. Primär wird die Datensicherheit hierbei dadurch gewährleistet, dass Rohdaten die lokale Umgebung nie verlassen, sondern nur Zwischenberechnungen (Aktivierungen und Gradienten) weitergegeben werden. Die Verschlüsselung stellt darauf aufbauend sicher, dass selbst die Weitergabe von Zwischenberechnungen zwischen den Parteien des Trainingsprozesses nicht durch Dritte abgefangen werden kann.

Verschiedene Varianten, wie etwa die Peer-to-Peer-Architektur, ermöglichen es mehrere Recheneinheiten anstelle eines einzigen zentralen Cloud-Servers in den Trainingsprozess einzubeziehen. Neben den Vorteilen des datenschutzkonformen Trainings, ermöglicht diese Architektur ein **effizientes verteiltes Lernen mit geringeren Speicher- und Rechenanforderungen** auf den lokalen Geräten, da jeweils nur ein Teil des Modells lokal verarbeitet wird. Dies bietet nicht nur Einrichtungen im Gesundheitswesen, sondern auch kleinen und mittelständischen Unternehmen einen entscheidenden Vorteil beim Training von KI-Modellen. Ein konkretes Beispiel ist ein verteiltes Prognosemodell für den Krankheitsverlauf von Krebspatient*innen: Mehrere onkologische Abteilungen trainieren gemeinsam ein neuronales Netz, indem sie nur die ersten und letzten Schichten lokal berechnen, während die rechenintensiven Zwischenschichten auf zentralen Servern ausgeführt werden, ohne dass patientenbezogene Daten die Kliniken verlassen (vgl. Abschnitt 5).

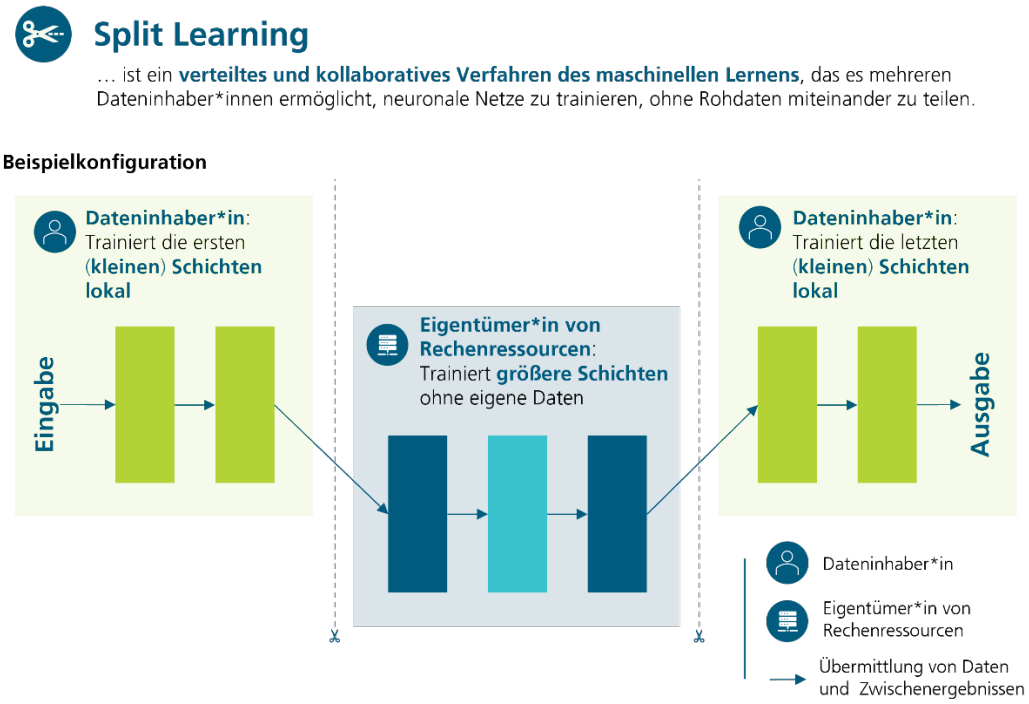

Abbildung 3: Split Learning

Die Methode reduziert damit deutlich die Anforderungen an lokale Hardware, bringt jedoch einen **erhöhten Kommunikationsaufwand** mit sich, da in jeder Iteration die Zwischenaktivierungen zwischen den beteiligten Einheiten übertragen werden müssen. Außerdem bringt das klassische Split Learning wenig Flexibilität mit. So muss sich vor Trainingsbeginn auf eine konkrete, starre Zuordnung der jeweiligen Modellteile auf die einzelnen, verteilten Hardware-Komponenten auf Basis unterschiedlicher Rechen- und Speicherressourcen geeinigt werden. Diese Back-End-Struktur der Split-Learning-Architektur kann nicht flexibel und laufend auf dynamisch wechselnde Infrastrukturen angepasst werden.

2.2.3 Federated Learning: Lokales Modelltraining mit globalem Nutzen

Federated Learning wurde 2016 von Google Research entwickelt, um den Herausforderungen herkömmlicher zentralisierter Trainingsansätze zu begegnen und richtet sich insbesondere an **Szenarien mit geringer Fehlertoleranz**. Beim Training moderner neuronaler Netze ist es in solchen Kontexten häufig nicht praktikabel, sämtliche Daten zentral zu bündeln, etwa aufgrund von Datenschutzvorgaben oder fehlender Infrastruktur. Wie in Abbildung 4 veranschaulicht, zielt das Verfahren darauf ab, das gesamte Modelltraining lokal auf den jeweiligen Geräten oder Servern der beteiligten Dateninhaber*in durchzuführen. Jeder Akteur bzw. Akteurin trainiert eine vollständige Kopie des Modells mit seinen eigenen Daten.

Anschließend werden lediglich die aktualisierten Modellgewichte an einen zentralen Server gesendet, der diese aggregiert, mittelt und die konsolidierte Version zurück an die Akteur*innen verteilt (McMahan et al. 2017). Auf diese Weise bleibt die **Datenhoheit** gewahrt, da die sensiblen Rohdaten das Unternehmen oder den Ursprungsort nicht verlassen. Gleichzeitig wird der Kommunikationsaufwand im Vergleich zu zentralisierten Trainingsansätzen reduziert, da keine vollständigen Datensätze übertragen werden. Der Ansatz eignet sich insbesondere für **Anwendungen mit datenreichen Akteur*innen und ausreichender lokaler Rechenleistung**, wie etwa im Finanz- oder Gesundheitswesen. Ein konkretes Beispiel ist die Verarbeitung medizinischer Bilddaten zur Unterstützung der Krebsdiagnose: Kliniken können jeweils lokal KI-Modelle auf MRT- oder CT-Scans trainieren, ohne Patientendaten mit anderen Einrichtungen oder einem zentralen Server teilen zu müssen (vgl. Abschnitt 5).

Trotz dieser Vorteile stößt Federated Learning in der Praxis an Grenzen. Die lokale Verarbeitung erfordert **hohe Rechen- und Speicherkapazitäten**, da jeder Teilnehmende das vollständige neuronale Netz eigenständig trainieren muss (Caldas et al. 2018). Für Organisationen mit begrenzten technischen Ressourcen kann dies eine erhebliche Hürde für die erfolgreiche Umsetzung darstellen. Darüber hinaus wirken sich Unterschiede in der vorhandenen Infrastruktur, etwa in Bezug auf Netzwerkkonnektivität oder Hardwareleistung, negativ auf die Stabilität und Effizienz des Trainingsprozesses aus.

Federated Learning

… ist ein **verteiltes und kollaboratives Verfahren des maschinellen Lernens.** Es ermöglicht mehreren Dateninhaber*innen, eigene Modelle lokal zu trainieren und dabei nur Modellparameter mit einem zentralen Server auszutauschen, ohne Rohdaten miteinander zu teilen.

Beispielkonfiguration

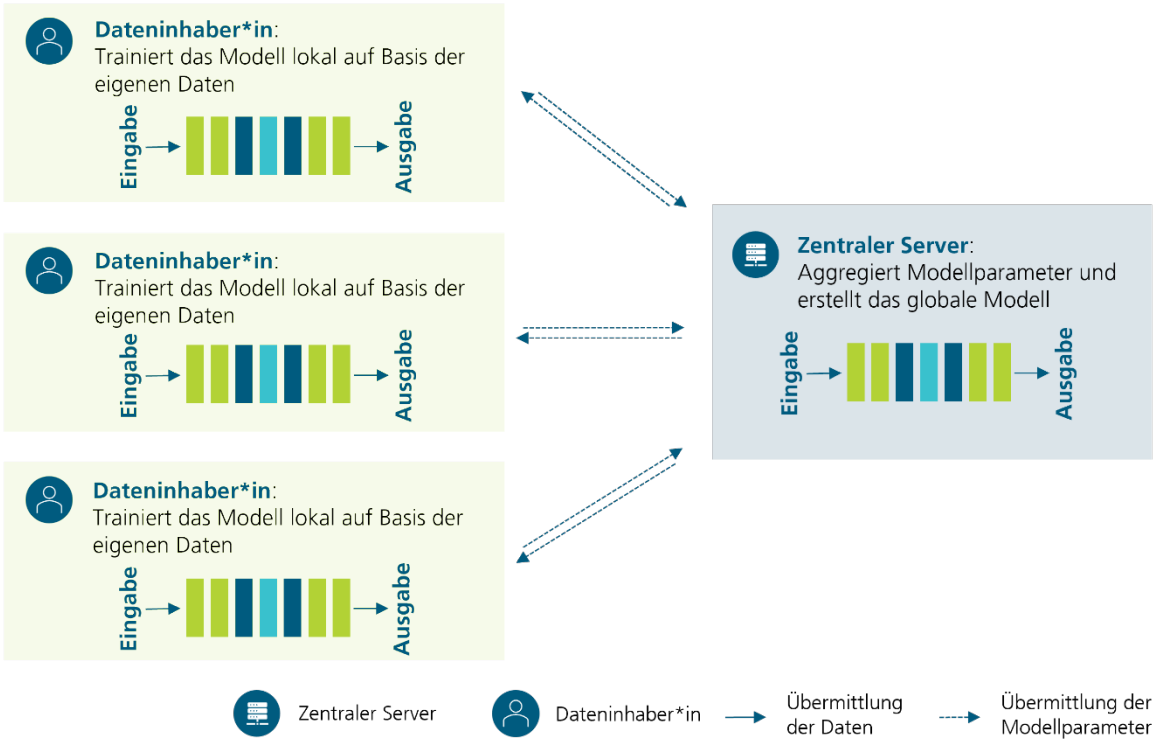

Abbildung 4: Federated Learning

3

SplitNFed

3 SplitNFed

3.1 Die Architektur von SplitNFed als ein integratives Verfahren für effizientes und souveränes KI-Training

SplitNFed ist ein verteiltes und kollaboratives Verfahren des maschinellen Lernens, das auf den Prinzipien des Split Learning basiert und gezielt um Mechanismen des Federated Learning erweitert wurde. Dadurch vereint der Ansatz die jeweiligen Stärken beider Verfahren in einem skalierbaren System, das ein effizientes und datenschutzkonformes Training von KI-Modellen ermöglicht. Die Architektur von SplitNFed ist entlang zweier Ebenen organisiert: einer Organisationsebene und einer technischen Trainingsebene. Abbildung 5 veranschaulicht das Verfahren.

Auf der **Organisationsebene** steuert eine zentrale Orchestrationseinheit den Trainingsprozess. Er übernimmt die Zuweisung und Koordination der Serverressourcen und optimiert die Layer-Zuordnung gemäß den jeweiligen Hardwareeigenschaften. Diese Koordinationsfunktion stammt konzeptionell aus dem Federated Learning und wird in SplitNFed weiterentwickelt, um auch in heterogenen Systemlandschaften einen effizienten und sicheren Ablauf sicherzustellen.

Die **Trainingsebene** folgt der Struktur des Split Learning, nach welcher die Modellverarbeitung in drei Schritte aufgeteilt wird. Zunächst erfolgt beim Akteur bzw. der Akteurin die lokale Vorverarbeitung der Eingangsdaten. Hier werden die ersten Schichten des neuronalen Netzes verarbeitet, wodurch sichergestellt wird, dass Rohdaten das Unternehmen nicht verlassen. Anschließend werden verschlüsselte Zwischenaktivierungen an Server übermittelt, die die rechenintensiven mittleren Schichten des Modells trainieren. Diese Verarbeitung erlaubt es, auch Geräte mit geringerer Rechenleistung in den Trainingsprozess einzubinden. Nach Abschluss dieser Phase werden die Outputs der Server zurück an den ursprünglichen Akteur*in übergeben, wo die finalen Modellschichten trainiert und zusammengeführt werden. Dieser finale Aggregationsschritt ermöglicht die Integration lokaler Besonderheiten in das globale Modell und stärkt die Modellanpassung an kontextspezifische Anforderungen.

Ergänzt wird diese Struktur durch ein zentrales Prinzip aus dem Federated Learning. Die ausgelagerten Modellsegmente werden nicht nur einmalig verteilt, sondern zusätzlich auf mehreren zentralen Servern redundant gespiegelt (Trainingsebene). Da die übergreifende Steuerung dieses verteilten Trainings orchestriert erfolgt (Organisationsebene), wird die Konsistenz der Modellsegmente gewährleistet und die Abstimmung zwischen den beteiligten Servern effizient gestaltet.

Durch die Kombination dieser Komponenten entsteht ein Verfahren, das nicht nur **souverän** und **effizient** ist, sondern sich zugleich durch hohe technische **Resilienz**, **Skalierbarkeit** und **Integrationsfähigkeit** auszeichnet. SplitNFed adressiert damit zentrale Anforderungen europäischer Unternehmen an Datenschutz, Skalierbarkeit und Ressourceneffizienz beim KI-Training. Es überwindet zentrale Schwächen bestehender Ansätze und bildet eine robuste Grundlage für moderne, verteilte KI-Anwendungen.

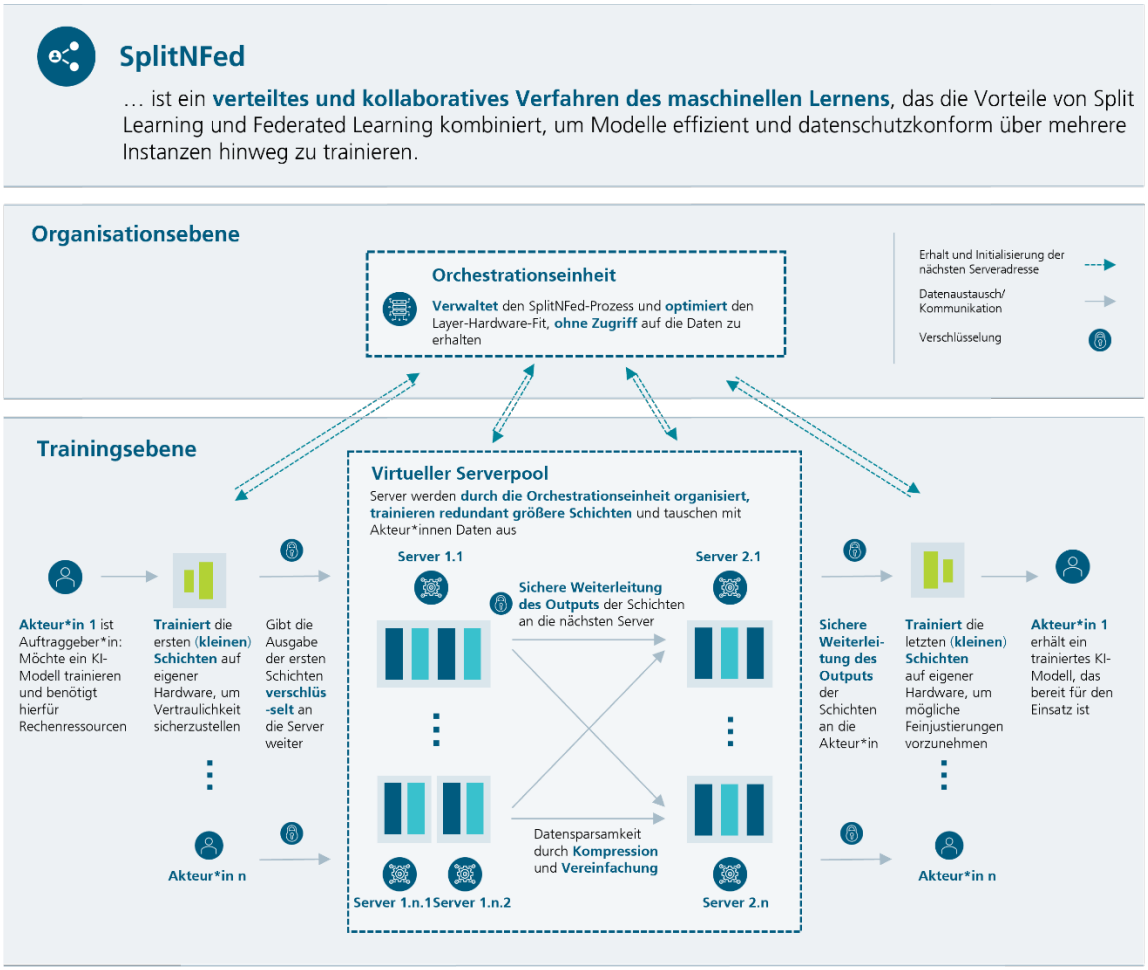

Abbildung 5: SplitNFed

3.2 Effizientes und systemrobustes KI-Training mit SplitNFed

Die Architektur von SplitNFed ermöglicht eine ressourceneffiziente Verteilung der Rechenlast, optimiert die Kommunikation zwischen den Beteiligten und gewährleistet eine hohe Systemstabilität. Diese Aspekte werden im Folgenden näher erläutert:

Ein zentrales Merkmal von SplitNFed ist die **gezielte und ressourceneffiziente Verteilung der Modellarchitektur** auf unterschiedliche Hardwareeinheiten. Durch die dreiteilige Struktur (vgl. Abbildung 5) lassen sich rechenintensive Zwischenschichten auf Hochleistungsserver auslagern, während initiale und finale Modellkomponenten weiterhin bei der Datenhalter*in verbleiben. Während beim Federated Learning jede Akteur*in das vollständige neuronale Netz lokal verarbeiten muss, beschränkt sich die Rechenlast bei SplitNFed auf einen kleinen Modellabschnitt. Dadurch können auch Geräte mit begrenzten Ressourcen in den Trainingsprozess eingebunden werden. Die gezielte Layer-Verteilung bringt jedoch nicht nur Flexibilität, sondern erfordert auch eine präzise Abstimmung. Die unterschiedlichen Schichten eines neuronalen Netzwerks haben variierende Rechen- und Speicheranforderungen, die stark von der zugrunde liegenden Hardware abhängen. Diese Herausforderung wird bei SplitNFed durch einen Optimierungsalgorithmus adressiert. Der Algorithmus löst ein mehrdimensionales Entscheidungsprob-

lem, indem er unter Berücksichtigung von Modellarchitektur, Netzwerk-Latenz, Hardwareverfügbarkeit und Datenschutzanforderungen die ideale Aufteilung der Modellsegmente bestimmt. Die Koordination durch die zentrale Orchestrationseinheit stellt sicher, dass jede Recheneinheit entsprechend ihrer Leistungsfähigkeit und Auslastung optimal eingesetzt wird. Diese dynamische Layer-Verteilung erlaubt eine flexible Anpassung an heterogene IT-Landschaften und minimiert den Ressourceneinsatz, ohne die Qualität des Trainingsprozesses zu beeinträchtigen. Zudem eröffnet sie durch die gezielte Vermeidung ineffizienter Auslastung das Potenzial, Trainingsprozesse zu beschleunigen und Einsparungen in der Ressourcennutzung zu realisieren.

Darüber hinaus geht SplitNFed mit einer hohen Kommunikationslast einher, die sich aus dem regelmäßigen Austausch zwischen den beteiligten Recheneinheiten ergibt. Um diese Herausforderung zu adressieren, integriert SplitNFed zusätzliche Methoden, um den Kommunikationsaufwand gezielt zu optimieren, mit möglichst geringem Einfluss auf die Qualität des Modells. Zum Einsatz kommen Verfahren wie Kompression, Sparsifikation und eine dynamische Anpassung der Kommunikationsfrequenz, etwa durch vermehrt lokale Trainingsrunden. So wird bei der Kompression die zu übertragende Datenmenge bspw. durch Quantisierung oder Huffman-Codierung reduziert, während bei der Sparsifikation die Datenmenge reduziert wird, indem etwa nur die wichtigsten Informationen (z. B. nur die größten Gradienten) übertragen werden. Auf diese Weise lässt sich die Menge der zu übertragenden Daten verringern, was die Netzwerklast senkt und auch den Trainingsprozess insgesamt effizienter und skalierbarer macht. Gerade in verteilten Umgebungen mit begrenzter Bandbreite wird somit ein stabiler Trainingsprozess sowie eine skalierbare Zusammenarbeit zwischen unterschiedlich ausgestatteten Partner*innen ermöglicht. Besonders bei komplexen Architekturen mit Millionen von Parametern spielt dies eine zentrale Rolle, da die dabei erzeugten Zwischenaktivierungen große Datenmengen umfassen können.

Ein weiterer zentraler Vorteil von SplitNFed ist die hohe **Systemrobustheit**, die den zuverlässigen Betrieb auch in dynamischen und heterogenen IT-Umgebungen sicherstellt. Durch die Kombination aus verteilter Modellarchitektur und föderierter Koordination entsteht ein Trainingssystem, das auch bei schwankenden Ressourcen oder temporären Ausfällen einzelner Komponenten stabil bleibt. Die Teilmodelle können parallel in mehrfacher Ausführung bei verschiedenen Servern trainiert werden, sodass einzelne Knoten bei Bedarf nahtlos kompensiert werden können, wie in Abbildung 5 durch die parallele Verteilung der Zwischenschichten veranschaulicht. Die zentrale Orchestrationseinheit überwacht dabei fortlaufend die Verfügbarkeit und Auslastung aller Beteiligten und passt die Ressourcenzuweisung dynamisch an. Auf diese Weise bleibt der Trainingsprozess auch unter anspruchsvollen technischen Bedingungen stabil und effizient. SplitNFed bildet somit die Grundlage für eine robuste und flexible Lernumgebung, die selbst in komplexen Anwendungsszenarien zuverlässig einsetzbar ist.

Durch die gezielte Kombination der Stärken von Split Learning und Federated Learning entsteht ein kollaboratives Ökosystem, das sowohl die Leistungsfähigkeit als auch die Flexibilität im Umgang mit heterogenen Hardwareplattformen optimiert. SplitNFed ermöglicht damit eine datenschutzkonforme, ressourceneffiziente und skalierbare Lösung für das Training großer Modelle maschinellen Lernens, die sich für europäische Unternehmen als zukunftsweisend herausstellen kann.

4

SplitNFed in der Praxis

4 SplitNFed in der Praxis

Split Learning, Federated Learning und SplitNFed unterscheiden sich nicht nur hinsichtlich ihrer technischen Architektur, sondern vor allem in Bezug auf ihre Eignung für unterschiedliche Anwendungsszenarien. Die drei Verfahren stellen keine trennscharfen Alternativen dar, sondern müssen als **Teil eines Lösungsspektrums** verstanden werden. Je nach Kontext entfalten ihre jeweiligen Stärken und Schwächen eine unterschiedliche Relevanz, sodass die Wahl des geeigneten Ansatzes maßgeblich von den konkreten Anforderungen an Datenschutz, Modellarchitektur und Infrastruktur abhängt.

Zur Veranschaulichung dieser Unterschiede werden im Folgenden exemplarische Anwendungsszenarien skizziert, die aufzeigen, in welchen Konstellationen welches Verfahren besondere Vorteile bietet. Die in Abbildung 6 dargestellte Übersicht dient als Orientierungsrahmen für die Auswahl geeigneter Ansätze in Abhängigkeit vom jeweiligen Anwendungsumfeld. Berücksichtigt werden dabei zentrale Entscheidungskriterien wie Fehlertoleranz, Skalierbarkeit, Kommunikationseffizienz, Individualisierung und Ressourceneffizienz.

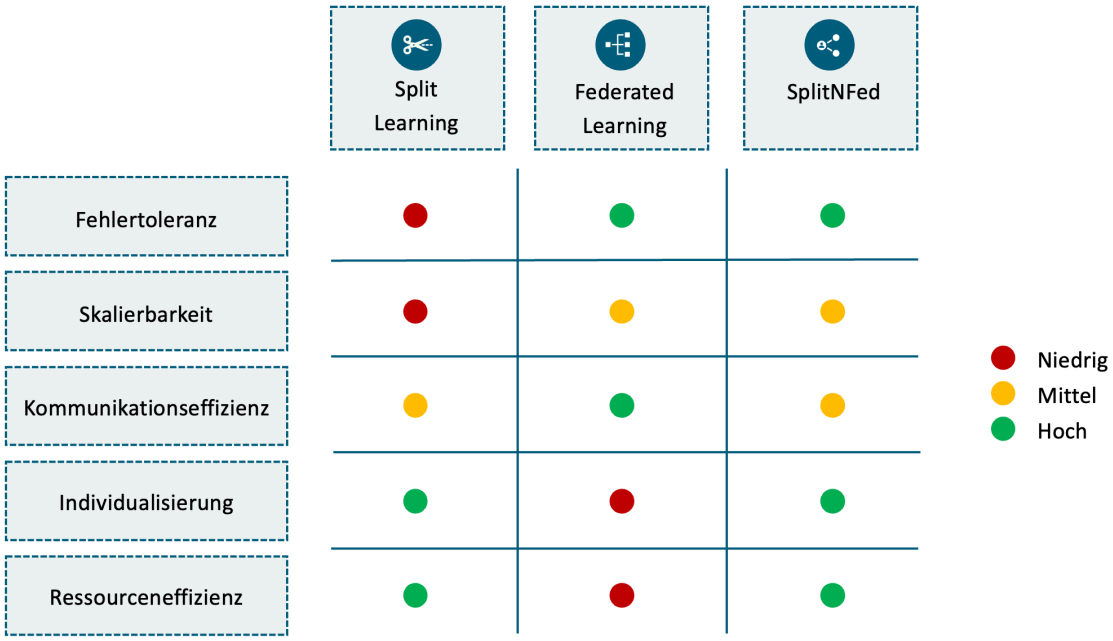

Abbildung 6: Vergleich zentraler Eigenschaften verteilter Lernverfahren

Split Learning eignet sich vor allem für Szenarien, in denen **individuelle Modelle für verschiedene Akteur*innen erforderlich sind, deren Daten jedoch strukturelle Ähnlichkeiten aufweisen**. Dies ist der Fall, wenn sich in Daten trotz gemeinsamer Muster lokale Unterschiede erkennen lassen, etwa, weil ähnliche Prozesse, Abläufe oder Inhalte vorliegen. Die Individualisierung ist kein nachgelagerter Schritt, sondern integraler Bestandteil der Modellarchitektur. Dies ist zum Beispiel dann sinnvoll, wenn sich die Rahmenbedingungen wie Risikoprofile oder

infrastrukturelle Gegebenheiten deutlich unterscheiden. Ein Anwendungsbeispiel aus dem Gesundheitsbereich wäre ein Prognosemodell zur Verlaufsschätzung bei Krebspatient*innen in mehreren onkologischen Abteilungen. Die Trennung des Modells in lokale und zentrale Komponenten erlaubt es, auf standortspezifische Besonderheiten wie verfügbare Medizintechnik oder organisatorische Abläufe einzugehen und zugleich gemeinsame Muster zu identifizieren. Die standortspezifischen Besonderheiten werden berücksichtigt, während die geteilten Schichten übergreifende Muster, etwa häufige Komplikationen, identifiziert und nutzbar macht.

Federated Learning ist besonders dann vorteilhaft, wenn mehrere Akteure bzw. Akteurinnen ein gemeinsames Modell entwickeln möchten, das auf einer **homogenen Datenbasis beruht und universell für alle Beteiligten nutzbar ist, unter gleichzeitig hohen Datenschutzanforderungen**. Bei Federated Learning verbleiben die Daten vollständig bei den beteiligten Institutionen, während lediglich die Modellgewichte ausgetauscht und zentral aggregiert werden. Ein Beispiel aus dem Gesundheitsbereich wäre der Fall zweier Universitätskliniken in München und Berlin, die gemeinsam ein KI-Modell zur Erkennung von Knochenbrüchen auf Basis von Röntgenbildern trainieren möchten. Da sowohl die Datenstruktur als auch das Klassifikationsziel nahezu identisch sind, profitieren sie von einem gemeinsamen Modell. Federated Learning ermöglicht es den Kliniken, ihre komplementären Datenbestände zu nutzen, um ein robustes und leistungsfähiges KI-Modell zu entwickeln, das auf einer größeren und diversifizierten Datenbasis basiert. Dies verbessert die Generalisierbarkeit des Modells, da es aus den Erfahrungen beider Kliniken lernt. Ein wesentlicher Vorteil dieses Ansatzes ist, dass die Kliniken sensible Patientendaten nicht austauschen müssen, wodurch die hohen Datenschutzanforderungen im Gesundheitssektor gewahrt bleiben. Die beteiligten Universitätskliniken können somit ein einheitliches Modell entwickeln, ohne die Integrität und Vertraulichkeit der Patientendaten zu gefährden. Eine nachträgliche Individualisierung des globalen Modells durch Feinabstimmung auf lokale Gegebenheiten ist dabei grundsätzlich möglich, stellt jedoch keinen integralen Bestandteil des Verfahrens dar.

SplitNFed verbindet die Prinzipien von Split Learning und Federated Learning und entfaltet sein volles Potenzial insbesondere in **Umgebungen mit fragmentierter Infrastruktur und hohen Datenschutzanforderungen**. Ein Anwendungsbeispiel aus dem Gesundheitsbereich ist ein Netzwerk bayerischer Universitätskliniken, das gemeinsam patientenspezifische KI-Modelle zur Therapieoptimierung entwickeln möchte. Aufgrund regulatorischer Anforderungen dürfen die Rohdaten die Klinikserver nicht verlassen und insbesondere nicht auf Cloud-Dienste großer Anbieter übertragen werden. Gleichzeitig verfügen einige Kliniken nur über begrenzte Rechenressourcen. SplitNFed ist in dieser Situation besonders geeignet, da es nicht nur die Verteilung des Modells über mehrere Institutionen ermöglicht, sondern auch die rechenintensiven Zwischenschichten flexibel auf eine Vielzahl unterschiedlich leistungsfähiger Server verteilt. Das Verfahren ermöglicht in diesem Fall eine sichere und skalierbare Kooperation. Die Zwischenschichten des Modells werden dynamisch auf verfügbare interne Ressourcen verteilt, während die zentrale Orchestrationseinheit die Trainingsprozesse stabil koordiniert. Durch die föderierte Koordination bleibt die Trainingsrobustheit erhalten.

Ein weiterer denkbarer Anwendungsfall ergibt sich auf institutioneller Ebene: Ein Ministerium plant den Aufbau einer Intermediationsebene, die es verschiedenen Akteur*innen, wie Kliniken,

Forschungseinrichtungen oder Unternehmen, erlaubt, ihre Ressourcen bedarfsgerecht miteinander zu teilen. SplitNFed eignet sich in besonderem Maße für ein solches Szenario, da die Architektur eine dynamische Verknüpfung heterogener IT-Infrastrukturen erlaubt, ohne dass sensible Daten zentral gespeichert oder weitergegeben werden müssen. Der Vorteil dieses Ansatzes steigt mit der Größe und Diversität des Netzwerks, doch auch kleinere Kooperationsverbünde profitieren bereits von der strukturellen Flexibilität. Bestehende Institutionen können ihre bislang ungenutzten Rechenkapazitäten effizient einbringen und einen Beitrag zur kollaborativen Entwicklung verteilter KI-Modelle leisten. Gleichzeitig bleibt die Entscheidungshoheit stets bei den Datenhaltenden, unabhängig davon, ob die Plattform öffentlich oder privat organisiert ist. Das Anwendungsbeispiel zeigt, dass SplitNFed die technologische Grundlage für eine souveräne, wirtschaftlich tragfähige und skalierbare Infrastruktur verteilter KI in institutionellen Ökosystemen schaffen kann.

Anhand der dargestellten Anwendungsfälle wird deutlich, dass sich die drei Ansätze jeweils durch spezifische Stärken auszeichnen, die in Abhängigkeit vom Anwendungskontext unterschiedlich relevant sind. Während Federated Learning maximale Modellkonsistenz bei minimalem Datenaustausch ermöglicht, erlaubt Split Learning eine stärkere Individualisierung unter Wahrung der Datenhoheit. SplitNFed erweitert diesen Handlungsspielraum, indem es auch unter anspruchsvollen regulatorischen und infrastrukturellen Bedingungen eine kollaborative Nutzung von KI-Modellen ermöglicht. Die Wahl des passenden Verfahrens richtet sich wesentlich nach dem konkreten Anwendungskontext, insbesondere im Hinblick auf die Anforderungen an Datenschutz, technische Infrastruktur und Modellarchitektur.

5

Wettbewerbsvorteile durch SplitNFed

5 Wettbewerbsvorteile durch SplitNFed

Der vorgestellte SplitNFed-Ansatz kombiniert die Prinzipien von Federated Learning und Split Learning zu einer skalierbaren Architektur für datensouveränes, verteiltes Modelltraining. Aufbauend auf dieser technologischen Grundlage richtet sich der Fokus im Folgenden auf den unternehmerischen Nutzen.

Abbildung 7 zeigt, wie SplitNFed zentrale technologische Komponenten miteinander verbindet, um strukturelle Herausforderungen in konkrete operative Mehrwerte zu überführen. Die Kombination aus effizienter Kommunikation, adaptiver Ressourcenverteilung und föderierter Aggregation ermöglicht die Integration in heterogene IT-Landschaften, reduziert die Anforderungen an Netzwerkinfrastrukturen und erhöht die Ausfallsicherheit im laufenden Trainingsbetrieb. Somit werden die Voraussetzungen für einen breit skalierbaren, wirtschaftlich tragfähigen und technologisch souveränen Einsatz von KI im Unternehmen geschaffen.

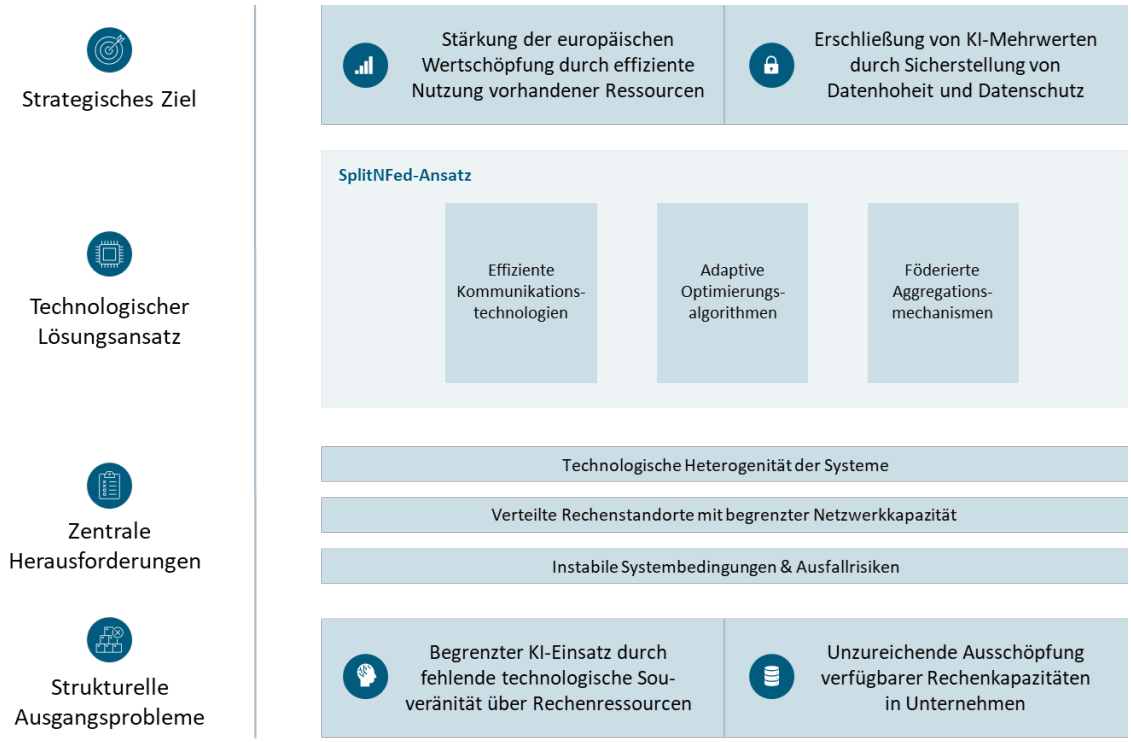

Abbildung 7: Operative Mehrwerte durch SplitNFed im technologischen Wirkzusammenhang

5.1 Stärkung der Souveränität, Datenhoheit und Effizienz

Unternehmen können vom Einsatz verteilter Lernverfahren erheblich profitieren, da die gemeinsame Entwicklung mit Partner*innen den Zugriff auf eine größere und vielfältigere Datenbasis ermöglicht, wodurch sich die Qualität der trainierten Modelle signifikant steigern lässt. Der Aufbau entsprechender KI-gestützter Systeme erfordert jedoch in der Regel eine Infrastruktur, die stark auf meist außereuropäische Cloud-Anbieter*innen angewiesen ist, was Abhängigkeiten schafft und die Kontrolle über die eigenen Daten und Modelle einschränkt (van der Vlist et al. 2024). SplitNFed begegnet dieser Problematik durch die Bereitstellung einer unabhängigen, verteilten Trainingsinfrastruktur, die ohne zentrale Cloud-Dienste auskommt. Ein konkreter Mehrwert für Unternehmen liegt daher in der **Wiedererlangung digitaler Souveränität**: Die Rohdaten verbleiben im Unternehmen und werden nicht an andere Parteien gegeben, wie es beim klassischen zentralen Trainingsprozess der Fall ist. Damit verringert SplitNFed das Risiko einer langfristigen Abhängigkeit an einen bestimmten Anbieter („Vendor Lock-in"). Es schafft die Grundlage für eine eigenständige, strategisch gestaltbare Daten- und KI-Politik.

Gleichzeitig stellt der Schutz von Unternehmensdaten, sowohl im Sinne gesetzlicher Vorgaben als auch im Hinblick auf unternehmenskritische Informationen, eine zentrale Hürde für den Einsatz zentralisierter KI-Modelle dar (Rojszczak 2020). In vielen Branchen besteht nicht nur die Pflicht, personenbezogene Daten regulatorisch zu schützen, sondern auch ein berechtigtes Eigeninteresse, strategisch relevante Unternehmensdaten vor externem Zugriff zu bewahren. SplitNFed begegnet dieser Problematik durch eine verteilte Trainingsarchitektur, bei der Rohdaten ausschließlich lokal verarbeitet werden und das System lediglich verschlüsselte Zwischenwerte und aggregierte Modellparameter überträgt. Dadurch verbleiben **sämtliche sensible Daten dauerhaft in der Datenhoheit** des jeweiligen Unternehmens. Neben der Einhaltung gesetzlicher Datenschutzvorgaben entsteht so ein effektiver Schutz unternehmerischer Werte. Dies ermöglicht nicht nur eine rechtskonforme, sondern auch strategisch kontrollierte Nutzung sensibler Daten für KI-Anwendungen – insbesondere in regulierten Branchen wie der Gesundheitsversorgung, der industriellen Fertigung oder der Finanzbranche, in denen Informationsschutz und Innovationsdruck in besonderer Weise zusammentreffen. Gleichzeitig ist dieser Schutzmechanismus branchenübergreifend relevant, da nahezu alle Unternehmen auf sensible Informationen angewiesen sind, etwa im Hinblick auf Managementstrategien, Kundenbeziehungen oder proprietäre Datenmodelle.

Eine weitere strategische Herausforderung beim KI-Training liegt in der ineffizienten Nutzung vorhandener IT-Ressourcen. Lokale Server oder Edge-Devices sind häufig nur gering ausgelastet, während externe Cloud-Leistungen eingekauft werden müssen (b-com 2025). SplitNFed adressiert diese Problematik durch eine dynamische Allokation der Trainingslast. Modellschichten werden so verteilt, dass auch intern verfügbare Recheneinheiten sinnvoll eingebunden werden. Der Mehrwert für Unternehmen liegt in einer deutlich verbesserten Ressourcennutzung. Unternehmen können vorhandene Infrastruktur aktiv in die Wertschöpfung integrieren, senken laufende Kosten für externe Rechenleistung und erzielen eine höhere Kapitalrendite aus getätigten IT-Investitionen. Damit wird nicht nur die **Effizienz gesteigert**, sondern auch die wirtschaftliche Nachhaltigkeit der eigenen IT-Strategie gestärkt.

5.2 Stärkung von Integrationsfähigkeit, Netzwerkeffizienz und Systemresilienz

In der praktischen Umsetzung von KI-Projekten stellt die technische Heterogenität vieler Unternehmensinfrastrukturen eine große Herausforderung dar. Unterschiedliche Betriebssysteme, Hardwarekonfigurationen, Frameworks oder Sicherheitsvorgaben erschweren die Integration verteilter Systeme. SplitNFed begegnet dieser Komplexität mit einer adaptiven Architektur und standardisierten Schnittstellen, die die zugrundeliegenden Systemunterschiede abstrahieren. Für Unternehmen ergibt sich daraus ein zentraler Vorteil: KI kann ohne aufwändige und teure Umstrukturierung oder Neuanschaffung durch eine hohe **Integrationsfähigkeit** eingeführt werden. Bestehende Geräte, Softwareumgebungen und Netzwerke lassen sich weiterhin nutzen und zugleich können neue Partner*innen oder Standorte unkompliziert integriert werden. Dies ermöglicht eine schrittweise, risikoarme Einführung von KI-Technologie in komplexe oder gewachsene IT-Landschaften – insbesondere dort, wo einheitliche Systeme nicht realisierbar sind.

Auch in verteilten IT-Umgebungen mit begrenzter Netzwerkkapazität lassen sich durch SplitNFed moderne KI-Trainingsprozesse effizient umsetzen, selbst an internationalen Produktionsstandorten oder in mobilen Anwendungsszenarien. Da keine sensiblen Rohdaten, sondern ausschließlich komprimierte Modellparameter und Zwischenberechnungen übertragen werden, bleibt die Netzwerklast selbst bei komplexen Trainingsprozessen moderat. Zwar erreicht SplitNFed im Hinblick auf die **Netzwerkeffizienz** nicht die Kompaktheit klassischer Federated-Learning-Verfahren, reduziert jedoch gegenüber Split Learning den Overhead deutlich, insbesondere durch die gezielte Zentralisierung der rechenintensiven Zwischenschichten. Auf diese Weise lassen sich auch Standorte mit begrenzter Bandbreite gleichberechtigt in verteilte KI-Trainingsprozesse einbinden, ohne dass Infrastrukturengpässe entstehen. Dies senkt nicht nur die Betriebskosten, sondern erhöht auch die Nutzbarkeit vorhandener Ressourcen.

SplitNFed ist außerdem darauf ausgelegt, Trainingsprozesse auch unter instabilen Systembedingungen zuverlässig fortzuführen. Durch redundantes Layer-Training auf mehreren verteilten Recheneinheiten könnte unabhängig von der Verfügbarkeit einzelner Komponenten eine **hohe Systemresilienz** erreicht werden. Die zentrale Orchestrationseinheit erkennt Ausfälle automatisch und verteilt die Rechenlast umgehend neu. Daraus ergibt sich für Unternehmen eine höhere Verfügbarkeit und Ausfallsicherheit des KI-Trainings. Der Betrieb bleibt auch bei technischen Störungen stabil, Trainingsfortschritte gehen nicht verloren und es entstehen keine zusätzlichen Koordinationsaufwände. Besonders in produktionskritischen oder sicherheitsrelevanten Bereichen stellt dies einen wesentlichen Mehrwert dar – sowohl im Hinblick auf Betriebsstabilität als auch auf Planbarkeit.

6

Reflexion

6 Reflexion

Durch die gezielte Kombination der Stärken von Split Learning und Federated Learning schließt SplitNFed die Lücke zwischen strengen Datenschutzanforderungen und dem steigenden Bedarf an leistungsfähigen KI-Trainingsumgebungen in Europa. Rohdaten verbleiben in den jeweiligen Organisationen, während ausschließlich modellbezogene Informationen ressourceneffizient ausgetauscht werden. Die dreiteilige Modellstruktur verteilt rechenintensive Schichten auf Server, bindet lokale Hardware ein und minimiert damit sowohl Netzwerklast als auch externe Abhängigkeiten.

Die vorangegangenen Ausführungen zeigen deutlich, dass SplitNFed nicht nur die digitale Souveränität stärkt, sondern auch wirtschaftliche Vorteile erschließt. Unternehmen können bislang ungenutzte Server- und Edge-Kapazitäten einbinden, externe Cloud-Abhängigkeiten reduzieren und Trainingsprozesse selbst bei begrenzter Bandbreite stabil betreiben. Gleichzeitig erlaubt die Architektur eine flexible Integration heterogener IT-Landschaften und schafft so die Grundlage für skalierbare, unternehmensübergreifende Kooperationen.

Mit dem Übergang in die Pilotierungsphase rückt nun die praxisnahe Validierung des Ansatzes in den Mittelpunkt. Geplante Schritte reichen von der Erhebung konkreter Nutzeranforderungen über die prototypische Implementierung bis hin zur systematischen Evaluierung unter realen Bedingungen.

Wir möchten im Rahmen der Pilotierungsphase gemeinsam mit Unternehmen eine auf spezifische Bedürfnisse zugeschnittene Lösung entwickeln und realitätsnahe Tests durchzuführen. Durch die flexible Anpassbarkeit ist SplitNFed neben den bereits dargestellten Beispielen aus dem medizinischen Kontext in unterschiedlichen betrieblichen Bereichen einsetzbar. Die Einbindung von Unternehmen in die Phase der Entwicklung von Prototypen trägt nicht nur zur wissenschaftlichen Forschung bei, sondern ermöglicht es auch, praxisnahe Erkenntnisse zu gewinnen, die direkt in die Technologieoptimierung einfließen. Im Rahmen zukünftiger Projekte soll untersucht werden, wie sich die Skalierbarkeit in heterogenen IT-Umgebungen weiter optimieren lässt und welche neuen Anwendungsszenarien sich durch die Kombination mit zusätzlichen Lernverfahren ergeben. Die gewonnenen Erkenntnisse sollen die Grundlage für praxisorientierte Konzepte und skalierbare Implementierungsstrategien bilden, die eine nachhaltige Verbreitung datensouveräner KI-Lösungen fördern.

Der Einbezug von Unternehmen aus der Praxis ist ein zentraler Erfolgsfaktor für die Weiterentwicklung und Validierung von SplitNFed. Durch die aktive Beteiligung können reale Anforderungen frühzeitig adressiert, praxisnahe Erkenntnisse gesammelt und technische Anpassungen direkt in heterogenen IT-Umgebungen erprobt werden. Unternehmen erhalten so frühzeitig Zugang zu einer Schlüsseltechnologie für datensouveräne KI, stärken ihre digitale Selbstbestimmung und reduzieren ihre Abhängigkeit von externen Plattformanbieter*innen. Die Möglichkeit, unter Wahrung der Vertraulichkeit kollaborativ KI zu nutzen, schafft zudem einen strategischen Wettbewerbsvorteil. Europäische Unternehmen können sich so als Vorreiter*in im Aufbau souveräner, skalierbarer KI-Infrastrukturen positionieren und langfristig ihre Innovationsfähigkeit sichern.

7 Literaturverzeichnis

Alex Mathew (2024): Cloud Data Sovereignty Governance and Risk Implications of Cross-Border Cloud Storage.

Allam, Karthik (2023): Adoption of Artificial Intelligence in Cloud Computing. In: *IJCTT* 71 (6), S. 91–95. DOI: 10.14445/22312803/IJCTT-V71I6P116.

b-com (2025): Optimize your data centers: Use your dormant IT resources | b-com. Online verfügbar unter https://b-com.com/en/optimize-et-decarbonize-your-processes-with-augmented-intelligence/optimize-your-dormant-it-resources, zuletzt aktualisiert am 02.06.2025, zuletzt geprüft am 02.06.2025.

Brender, Nathalie; Markov, Iliya (2013): Risk perception and risk management in cloud computing: Results from a case study of Swiss companies. In: *International Journal of Information Management* 33 (5), S. 726–733. DOI: 10.1016/j.ijinfomgt.2013.05.004.

Caldas, Sebastian; Konečny, Jakub; McMahan, H. Brendan; Talwalkar, Ameet (2018): Expanding the Reach of Federated Learning by Reducing Client Resource Requirements. Online verfügbar unter http://arxiv.org/pdf/1812.07210v2.

Canalys (2025): Marktanteile der führenden Unternehmen am Umsatz im Bereich Cloud Computing weltweit im 4. Quartal 2024 [Graph]. Statista. Online verfügbar unter https://de.statista.com/statistik/daten/studie/150979/umfrage/marktanteile-der-fuehrenden-unternehmen-im-bereich-cloud-computing/, zuletzt geprüft am 25.02.2025.

Gaia-X European Association for Data and Cloud AISBL (2023): Gaia-X. Together towards a federated and secure data infrastructure. Online verfügbar unter https://gaia-x.eu/, zuletzt geprüft am 16.04.2025.

Goodfellow, Ian; Courville, Aaron; Bengio, Yoshua (2016): Deep learning. Cambridge, Massachusetts: The MIT Press (Adaptive computation and machine learning). Online verfügbar unter https://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=2565107.

Hintermann, Ralph; Hinterholzer, Simon; Progni, Kejsi (2024): Bitkom-Studie Rechenzentren in Deutschland: Aktuelle Marktentwicklungen – Stand 2024. Borderstep Institut. Berlin. Online verfügbar unter https://www.bitkom.org/sites/main/files/2024-11/241121-studie-rechenzentrums-markt.pdf.

Kairouz, Peter; McMahan, H. Brendan; Avent, Brendan; Bellet, Aurélien; Bennis, Mehdi; Nitin Bhagoji, Arjun et al. (2021): Advances and Open Problems in Federated Learning. In: *FNT in Machine Learning* 14 (1–2), S. 1–210. DOI: 10.1561/2200000083.

Martínez, Marco Antonio Díaz; Salinas, Reina Verónica Román; Hernández, Santos Ruíz; Ruíz Domínguez, Herson Santos; Zubirías, Gabriela Cervantes; Morales Rodríguez, Mario Alberto (2024): Artificial intelligence an essential factor for the benefit of companies: systematic review

of the literature. In: *Cogent Engineering* 11 (1), Artikel 2380344. DOI: 10.1080/23311916.2024.2380344.

McKinsey&Company (2023): Studie: Generative KI kann zum Produktivitätsbooster werden. Online verfügbar unter https://www.mckinsey.de/news/presse/genai-ist-ein-hilfsmittel-um-die-produktivitaet-zu-steigern-und-das-globale-wirtschaftswachstum-anzukurbeln, zuletzt geprüft am 12.03.2025.

McMahan, Brendan; Moore, Eider; Ramage, Daniel; Hampson, Seth; Arcas, Blaise Aguera y. (2017): Communication-Efficient Learning of Deep Networks from Decentralized Data. In: Aarti Singh und Jerry Zhu (Hg.): Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, Bd. 54: PMLR (Proceedings of Machine Learning Research), S. 1273–1282. Online verfügbar unter https://proceedings.mlr.press/v54/mcmahan17a.html.

Pohle, Julia; Thiel, Thorsten (2020): Digital sovereignty. In: *Internet Policy Review* 9 (4). DOI: 10.14763/2020.4.1532.

Rojszczak, Marcin (2020): CLOUD act agreements from an EU perspective. In: *Computer Law & Security Review* 38, S. 105442. DOI: 10.1016/j.clsr.2020.105442.

Schmitz, Anna-Raphaela; Mitrovic, Marco (2024): Gaia-X: Souveräne Dateninfrastruktur für Europa. Bayerisches Forschungsinstitut für Digitale Transformation. Online verfügbar unter https://www.bidt.digital/phaenomene/gaia%E2%80%91x-souveraene-dateninfrastruktur-fuer-europa/, zuletzt aktualisiert am 18.09.2024, zuletzt geprüft am 16.04.2025.

Singla, Alex; Sukharevsky, Alexander; Yee, Lareina; Chui, Michael; Hall, Bryce (2024): The state of AI in early 2024. Gen AI adoption spikes and starts to generate value. Hg. v. Heather Hanselman. QuantumBlack; AI by McKinsey; McKinsey Digital. Online verfügbar unter https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai#/, zuletzt geprüft am 28.02.2025.

Thapa, Chandra; Chamikara, M. A. P.; Camtepe, Seyit; Sun, Lichao (2020): SplitFed: When Federated Learning Meets Split Learning.

van der Vlist, Fernando; Helmond, Anne; Ferrari, Fabian (2024): Big AI: Cloud infrastructure dependence and the industrialisation of artificial intelligence. In: *Big Data & Society* 11 (1), Artikel 20539517241232630. DOI: 10.1177/20539517241232630.

Vepakomma, Praneeth; Gupta, Otkrist; Swedish, Tristan; Raskar, Ramesh (2018): Split learning for health: Distributed deep learning without sharing raw patient data. In: *32nd Conference on Neural Information Processing Systems (NIPS 2018)*. DOI: 10.48550/arXiv.1812.00564.

Fraunhofer FIT

Institutsteil Wirtschaftsinformatik,
Fraunhofer-Institut für Angewandte
Informationstechnik FIT

Kontakt

Fraunhofer-Institut für Angewandte Informationstechnik FIT
Institutsteil Wirtschaftsinformatik
Wittelsbacherring 10
95444 Bayreuth

Telefon +49 921 55-4710
info@fit.fraunhofer.de
www.wi.fit.fraunhofer.de