

# Abschlussbericht

## Hochschule Bielefeld

---

ZE (Zuwendungsempfänger):  
Hochschule Bielefeld (HSBI)  
(vormals FH Bielefeld)

Förderkennzeichen: 13FH074K11

---

Vorhabenbezeichnung: Young Researchers Cloud and Edge Computing Platform for AI

---

Laufzeit des Vorhabens: 01.01.2022 bis 31.03.2024

---

Berichtszeitraum: 01.01.2022 bis 31.03.2024

---

### Teil I: Kurzbericht

Das Projekt „Young Researchers Cloud and Edge Computing Platform for AI“ (kurz: „yourAI“) wurde mit dem Ziel ins Leben gerufen, eine leistungsstarke Cloud- und Edge-Computing-Plattform für KI-Anwendungen zu entwickeln und bereitzustellen. Diese Plattform sollte einerseits aus einem leistungsstarken Rechencluster bestehen (yourAI-Cluster), zum anderen aus zusätzlichen Komponenten zum Edge-Computing. Der Schwerpunkt lag dabei auf dem yourAI-Cluster. Das Vorhaben sollte die Forschung und Lehre an der Hochschule Bielefeld (HSBI) fachbereichsübergreifend unterstützen und insbesondere den wissenschaftlichen Nachwuchs durch den Zugang zu modernen Rechenressourcen fördern. In der ursprünglichen Vorhabenbeschreibung wurden mehrere zentrale Arbeitspakete definiert, die alle wesentlichen Schritte von der Hardwarebeschaffung über die Implementierung der Softwareumgebung bis hin zur Inbetriebnahme und Integration in die bestehende IT-Infrastruktur der Hochschule umfassten. Die Projektziele blieben über die Laufzeit des Projekts unverändert und konnten weitgehend erreicht werden.

An dem technischen Stand, der durch den Data-Analytics-Cluster des Center for Applied Data Science (CfADS) an der HSBI und den HPC-Cluster (HPC: High Performance Computing) des Computational Science Labs an der HSBI definiert war, konnte nur in begrenztem Maße angeknüpft werden. Während diese bestehenden Systeme wertvolle Erfahrungen im Aufbau und Betrieb von Rechenclustern lieferten, unterschied sich der geplante Technologie-Stack des yourAI-Clusters erheblich. Insbesondere die Integration der neuen Infrastruktur in das Hochschul-Netzwerk und das Hochschul-Identitätsmanagement sollte anders erfolgen. Die spezifischen Anforderungen des yourAI-Projekts erforderten den Einsatz und das Erlernen zahlreicher neuer Technologien, die über das bisherige Wissen und die bestehenden Lösungen hinausgingen. Der neue Cluster erforderte daher eine maßgeschneiderte Lösung, die nicht nur leistungsstarke GPUs und FPGAs, sondern auch spezielle Softwareumgebungen wie Jupyter-Notebooks, ein Big-Data-System und einen Ressourcenmanager integrierte. Dieser neue Ansatz ermöglichte es, unterschiedlichste Anforderungen im Bereich von KI-Anwendungen und -Verfahren in einem ganzheitlichen System zu vereinen und so die Grundlage für zukünftige KI-Forschung und -Lehre zu legen.

Das Projekt begann mit der Koordination und Planung (AP1) sowie der Etablierung eines umfassenden Datenschutz- und Sicherheitskonzepts (AP2). In einem Workshop wurden relevante Themen zur DSGVO und IT-Sicherheit adressiert, und technische sowie organisatorische Maßnahmen (TOMs) wurden in der Folge definiert. Diese legten den Grundstein für den sicheren Umgang mit Daten im yourAI-Cluster.

Ein wesentlicher Meilenstein des Projekts war die Beschaffung der benötigten Hardware (AP3). Dieser Prozess war jedoch von Verzögerungen geprägt, unter anderem durch Komplikationen bei der europäischen Ausschreibung von Servern und Storage für den yourAI-Cluster. Trotz dieser Herausforderungen konnte die benötigte Serverhardware schließlich beschafft und auf Funktion getestet werden.

Parallel zur Beschaffung wurden im AP5 die Integration des Clusters in das Hochschul-Netzwerk und die Implementierung eines maßgeschneiderten Benutzermanagementsystems durchgeführt. Diese Integration erforderte eine enge Zusammenarbeit mit der Datenverarbeitungszentrale (DVZ) der Hochschule Bielefeld, um eine sichere Authentifizierung und Autorisierung der Nutzer zu gewährleisten. Hierzu wurden Technologien wie Shibboleth und FreeIPA erfolgreich eingesetzt.

Im nächsten Schritt folgte die Inbetriebnahme der Hardware (AP4), wobei die physische Installation der Server sowie deren Netzwerkanbindung im Mittelpunkt standen. Parallel dazu wurde die Softwareumgebung des Clusters eingerichtet (AP6). Hierzu gehörte die Implementierung von Virtualisierungstechnologien für administrative virtuelle Maschinen, einem verteilten Speichersystem (Ceph) sowie dem Ressourcenmanagementsystem Slurm. Diese Technologien bildeten die Grundlage für den späteren Betrieb des Clusters und ermöglichten eine flexible und effiziente Nutzung durch verschiedene Benutzergruppen.

Im Jahr 2023 lag der Fokus auf der Integration der Edge-Hardware (AP7) und der finalen Inbetriebnahme der Softwareumgebung (AP6). Als Benutzerschnittstelle wurde ein Cluster-Webportal auf der Basis von OpenOnDemand implementiert, das den Nutzern über eine benutzerfreundliche Web-Oberfläche den Zugriff auf die Rechenressourcen des Clusters ermöglicht. OpenOnDemand ist eine Open-Source-Plattform, die es erlaubt, über das Internet auf HPC-Ressourcen zuzugreifen, Batch-Jobs einzureichen, Dateiverwaltung durchzuführen und interaktive Sessions wie Jupyter-Notebooks zu starten.

Im letzten Projektabschnitt wurde eine Testphase (AP8) für den yourAI-Cluster gestartet, um die Funktionalität und Nutzerfreundlichkeit der Plattform zu evaluieren. Diese Testphase, die als „closed beta“ durchgeführt wurde, beschränkte sich jedoch auf einen ausgewählten Nutzerkreis innerhalb der Hochschule. Aufgrund der zeitlichen Verzögerungen und der laufenden Integrationsarbeiten konnte der Cluster innerhalb der Projektlaufzeit noch nicht für alle potenziellen Nutzer zugänglich gemacht werden. Zwischen dem Projektende und dem Berichtsdatum ist fachbereichsübergreifend die Anzahl der Nutzer kontinuierlich erhöht worden, und es wird aktuell schon umfangreich produktiv mit dem yourAI-Cluster gearbeitet, sowohl in Promotionsprojekten als auch in Lehrveranstaltungen.

Zu den wichtigsten wissenschaftlich-technischen Ergebnissen des Projekts zählt die erfolgreiche Entwicklung und Inbetriebnahme der yourAI-Plattform. Diese Plattform unterstützt nun die Entwicklung, Validierung und Anwendung hochperformanter KI-Algorithmen und ist vollständig in die IT-Infrastruktur der Hochschule Bielefeld integriert. Ein zentrales Element der Plattform ist die sichere und effiziente Verwaltung von Daten, die durch ein umfassendes Datenschutz- und Sicherheitskonzept gewährleistet wird.

Ein weiterer bedeutender Erfolg war die Implementierung eines innovativen Benutzer- und Ressourcenmanagementsystems. Dies ermöglicht es verschiedenen Nutzergruppen, auf die Plattform zuzugreifen und je nach Bedarf spezifische Ressourcen zu nutzen, was die Flexibilität und Benutzerfreundlichkeit erheblich erhöht.

Das Projekt „yourAI“ konnte die gesetzten Ziele weitgehend erreichen und hat an der Hochschule Bielefeld eine leistungsfähige Infrastruktur etabliert, die sowohl der Forschung als auch der Lehre im Bereich der Künstlichen Intelligenz langfristig zugutekommt. Trotz einiger Verzögerungen im Ablauf konnten alle zentralen Arbeitspakete erfolgreich abgeschlossen werden, und die Plattform ist nun bereit für den breiten Einsatz in der Hochschule.

# Abschlussbericht

## Hochschule Bielefeld

---

ZE (Zuwendungsempfänger):  
Hochschule Bielefeld (HSBI)  
(vormals FH Bielefeld)

Förderkennzeichen: 13FH074K11

---

Vorhabenbezeichnung: Young Researchers Cloud and Edge Computing Platform for AI

---

Laufzeit des Vorhabens: 01.01.2022 bis 31.03.2024

---

Berichtszeitraum: 01.01.2022 bis 31.03.2024

---

## Teil II: Eingehende Darstellung

### 1. Übersicht und Vergleich mit der ursprünglichen Vorhabenbeschreibung

Das Projekt „Young Researchers Cloud and Edge Computing Platform for AI“ (kurz: „yourAI“) wurde mit dem Ziel ins Leben gerufen, eine leistungsstarke Cloud- und Edge-Computing-Plattform für KI-Anwendungen zu entwickeln und bereitzustellen. Diese Plattform sollte einerseits aus einem leistungsstarken Rechencluster bestehen (yourAI-Cluster), zum anderen aus zusätzlichen Komponenten zum Edge-Computing. Der Schwerpunkt lag dabei auf dem yourAI-Cluster. Das Vorhaben sollte die Forschung und Lehre an der Hochschule Bielefeld (HSBI) fachbereichsübergreifend unterstützen und insbesondere den wissenschaftlichen Nachwuchs durch den Zugang zu modernen Rechenressourcen fördern. In der ursprünglichen Vorhabenbeschreibung wurden mehrere zentrale Arbeitspakete definiert, die alle wesentlichen Schritte von der Hardwarebeschaffung über die Implementierung der Softwareumgebung bis hin zur Inbetriebnahme und Integration in die bestehende IT-Infrastruktur der Hochschule umfassten. Die Projektziele blieben über die Laufzeit des Projekts unverändert und konnten weitgehend erreicht werden.

### 2. Durchführung der Arbeiten im Detail

#### 2.1 Arbeitspaket 1: Projektkoordination

Die erfolgreiche Umsetzung des Projekts hing maßgeblich von einer sorgfältigen Planung und Koordination ab. In diesem Arbeitspaket wurden die Projektziele definiert und die notwendigen Maßnahmen für die Erreichung dieser Ziele festgelegt. Um sicherzustellen, dass alle Beteiligten auf dem gleichen Stand waren und die Projektziele effizient erreicht werden konnten, wurden zahlreiche Meetings geplant und durchgeführt. Diese regelmäßigen Besprechungen waren entscheidend, um den Fortschritt des Projekts zu überwachen, potenzielle Herausforderungen frühzeitig zu identifizieren und entsprechende Maßnahmen zu ergreifen. Grundlegende Entscheidungen betrafen z.B. die Software-Architektur des Clusters und die Auswahl des Ressourcenmanagers und anderer kritischer Komponenten. Durch diesen kollaborativen Ansatz konnten wichtige Entscheidungen fundiert getroffen werden, was maßgeblich zum Erfolg des Projekts beitrug.

#### 2.2 Arbeitspaket 2: Datenschutz und Datensicherheit

Ein zentrales Anliegen des Projekts war die Sicherstellung des Datenschutzes und der Datensicherheit, insbesondere im Hinblick auf die strengen Anforderungen der Datenschutz-Grundverordnung

(DSGVO). Dieses Arbeitspaket war von zentraler Bedeutung, da der yourAI-Cluster eine Vielzahl von personenbezogenen Daten verarbeiten soll, die sowohl in der Forschung als auch in der Lehre anfallen. Entsprechend wurden technisch-organisatorische Maßnahmen (TOMs) erarbeitet und deren Umsetzung dokumentiert.

Zu Beginn des Projekts stand die detaillierte Identifizierung aller potenziell betroffenen personenbezogenen Daten im Vordergrund. Diese umfasste nicht nur offensichtliche Informationen wie Namen und E-Mail-Adressen der Nutzer, sondern auch weniger offensichtliche Daten, die im Rahmen von Forschungsprojekten erhoben und verarbeitet werden. Für diese verschiedenen Datentypen wurden spezifische Schutzmaßnahmen entwickelt, die sicherstellen, dass die Vertraulichkeit, Integrität und Verfügbarkeit der Daten zu jeder Zeit gewährleistet sind.

Um die erfolgreiche Implementierung der Maßnahmen sicherzustellen, wurden intern gezielte Maßnahmen und Arbeitsgruppen eingerichtet, um die Einhaltung der Datenschutzanforderungen während der gesamten Projektlaufzeit im Blick zu behalten. Es wurden spezifische Arbeitsgruppen gebildet, die sich auf die Umsetzung der technischen und organisatorischen Maßnahmen konzentrierten und regelmäßig den Fortschritt überprüften.

Zwar war es ursprünglich geplant, ein umfassendes Monitoring- und Reporting-System sowie Schulungen für die Endnutzer des Clusters durchzuführen, diese Maßnahmen konnten jedoch aufgrund von Zeit- und Ressourcenbeschränkungen nicht im Rahmen des Projekts realisiert werden. Stattdessen wurde der Fokus darauf gelegt, die grundlegenden Datenschutz- und Sicherheitsmaßnahmen so robust wie möglich zu gestalten und diese eng in die technische Infrastruktur des Clusters zu integrieren. Dies gewährleistet, dass der yourAI-Cluster dennoch den grundlegenden Anforderungen der DSGVO entspricht und eine sichere Plattform für sensible Forschungs- und Lehrprojekte bietet.

### **2.3 Arbeitspaket 3: Beschaffung**

Die Beschaffung der Hardware stellte eine der größten Herausforderungen des yourAI-Projekts dar und war von entscheidender Bedeutung für den gesamten Projektverlauf. Ursprünglich war geplant, die benötigte Server-Hardware zeitnah zu beschaffen und zu installieren, um damit eine solide Grundlage für die weiteren Arbeiten zu schaffen. Dieser Schritt sollte frühzeitig abgeschlossen werden, um den nachfolgenden Projektphasen genügend Zeit zur Umsetzung zu geben. Allerdings wurde dieser Plan durch eine Reihe unvorhersehbarer Ereignisse erheblich beeinträchtigt.

Zu Beginn der Beschaffungsphase im Jahr 2022 sah sich das Projektteam mit der globalen Chipkrise konfrontiert, die zu massiven Lieferengpässen bei der benötigten Hardware führte. Diese Krise führte dazu, dass Lieferzeiten, die normalerweise Wochen in Anspruch nehmen, auf Monate ausgedehnt wurden. Die Verfügbarkeit bestimmter kritischer Komponenten, insbesondere leistungsstarker Prozessoren und spezialisierter Speicherlösungen, war stark eingeschränkt. Diese Lieferengpässe zwangen das Team dazu, alternative Beschaffungswege zu prüfen und ständig den Markt zu überwachen, um die notwendigen Komponenten überhaupt beschaffen zu können. Diese zusätzlichen Bemühungen beanspruchten erheblich mehr Zeit und Ressourcen als ursprünglich eingeplant. Das galt genauso für die Beschaffung der Hardware für das Edge-Computing, die im Jahr 2023 durchgeführt wurde, und für die Beschaffung der FPGA-Karten für den FPGA-Server im yourAI-Cluster, welche 2023 begonnen, aber erst 2024 abgeschlossen werden konnte.

Die europaweite Ausschreibung, die im Jahr 2022 zur Beschaffung der Hardware für den Cluster notwendig war, erwies sich als komplexer und zeitaufwändiger als erwartet. Rechtliche und bürokratische Hürden verzögerten den Prozess. Zudem gab es mehrfach Rückfragen und Einwände von Bietern, die zu einer weiteren Verlängerung der Ausschreibungsfristen führten. Diese Prozesse verlangsamten den gesamten Beschaffungszyklus erheblich, da jede Änderung und jeder Einspruch sorgfältig geprüft und bearbeitet werden musste. In der finalen Phase stellte einer der unterlegenen Bieter einen Nachprü-

fungsantrag bei der Vergabekammer Westfalen. Allein dadurch verlängerte sich das gesamte Verfahren um zwei Monate, so dass der Zuschlag für die Beschaffung der Cluster-Hardware erst im November 2022 erfolgen konnte.

Durch die genannten Herausforderungen verschoben sich nicht nur die Zeitpläne für die Hardwarebeschaffung, sondern auch die nachfolgenden Arbeitspakete. Die Installation und Inbetriebnahme der Hardware, die essenziell für den Beginn der Softwareintegration und die anschließende Testphase war, verzögerte sich erheblich. Diese Verzögerungen hatten direkte Auswirkungen auf die gesamte Projektplanung, da viele der weiteren Aktivitäten auf der Verfügbarkeit der Hardware basierten.

In Anbetracht der begrenzten personellen Ressourcen im Projekt mussten zusätzliche Anpassungen vorgenommen werden. Das Projektteam stand vor der Herausforderung, die verfügbaren Ressourcen so effizient wie möglich zu nutzen, um die Kernziele des Projekts dennoch zu erreichen. Viele Aufgaben, die ursprünglich parallel hätten durchgeführt werden sollen, mussten nun nacheinander erledigt werden, was den Druck auf das Team weiter erhöhte. Einige geplante Erweiterungen und Optimierungen konnten aufgrund der erheblichen Verzögerungen und der damit verbundenen Zeit- und Ressourcenknappheit nicht mehr umgesetzt werden. Der Fokus musste stattdessen auf die zentralen Aufgaben gelegt werden, um in jedem Fall die grundlegenden Funktionen des yourAI-Clusters sicherzustellen.

Die zwei kostenneutralen Projektverlängerungen um insg. neun Monate (ursprünglich geplantes Projektende: 30.06.2023; tatsächliches Projektende: 31.03.2024) sind vor allem durch die Verzögerungen in der Beschaffung der Cluster-Hardware begründet.

Software-Beschaffungen waren im Projekt nur im kleinen Umfang vorgesehen. Diese Beschaffungen wurden kontinuierlich bis in das Jahr 2024 durchgeführt. Leider konnte die geplante EuroPractice-Mitgliedschaft inkl. Beschaffung der Startlizenzen für die über EuroPractice vertriebene Cadence-Software nicht mehr vor Projektende realisiert werden. Auf den Projekterfolg hat das keine negativen Auswirkungen, da diese Beschaffung inzwischen aus anderen Mitteln durchgeführt werden konnte, die bei Projektbeantragung noch nicht zur Verfügung standen.

#### **2.4 Arbeitspaket 4: Inbetriebnahme der Hardware**

Nach der erfolgreichen und zeitaufwendigen Beschaffung der Hardware für den yourAI-Cluster begann die Installation und Inbetriebnahme im Rechenzentrum der Hochschule Bielefeld. Dieser Prozess war ein entscheidender Schritt für das Projekt, da er die physische Grundlage für die gesamte yourAI-Infrastruktur legte. Die Inbetriebnahme der Hardware umfasste mehrere komplexe Phasen, die sorgfältig koordiniert und durchgeführt wurden, um sicherzustellen, dass die neuen Systeme optimal in die bestehende IT-Infrastruktur integriert werden konnten.

Ein zentraler Fokus lag dabei auf der physischen Montage der Server. Diese wurde in speziellen Racks im Rechenzentrum untergebracht, die sowohl die notwendige Kühlung als auch den Strombedarf für die neuen Hochleistungsgeräte bereitstellten. Die korrekte Platzierung der Hardware in den Racks war entscheidend, um eine optimale Luftzirkulation und damit die Kühlung der Komponenten zu gewährleisten. Es wurden spezifische Maßnahmen ergriffen, um eine Überhitzung der Server zu vermeiden, die durch die intensiven Rechenprozesse des Clusters auftreten könnten.

Ein weiterer wichtiger Aspekt war die Stromversorgung in den Racks mit der jeweils notwendigen Leistung. Es war die genaue Berechnung und Planung der Lastverteilung erforderlich. Es musste sorgfältig festgelegt werden, welche Server auf welche Phase in den Racks verteilt werden, um eine Überlastung einzelner Phasen zu vermeiden und die Energieeffizienz zu maximieren. Diese Planung erforderte eine enge Zusammenarbeit mit den technischen Spezialisten der Datenverarbeitungszentrale (DVZ) der Hochschule Bielefeld, um sicherzustellen, dass die Stromversorgung den hohen Anforderungen des Clusters gerecht wird und gleichzeitig die bestehende Infrastruktur des Rechenzentrums optimal genutzt wird.

Neben der physischen Installation war die Netzwerkanbindung der Server von zentraler Bedeutung. Hierfür wurden spezielle Netzwerkinfrastrukturen geschaffen, die es ermöglichen, den Datenverkehr effizient zu verteilen und die Server sicher mit dem restlichen Hochschulnetzwerk zu verbinden. Dies beinhaltete die Einrichtung von redundanten Netzwerkschnittstellen, um Ausfallsicherheit zu gewährleisten, sowie die Implementierung von Hochgeschwindigkeitsverbindungen, um den hohen Datenanforderungen gerecht zu werden. Insbesondere wurde auf die Integration eines Ceph-basierten verteilten Speichersystems geachtet, das eine hohe Verfügbarkeit und Datenintegrität sicherstellt. Dieses System ermöglicht es, Daten automatisch über mehrere Speicherorte zu replizieren, wodurch das Risiko von Datenverlusten minimiert wird.

Ein weiterer wichtiger Aspekt der Inbetriebnahme war die Konfiguration der RAID-Systeme (Redundant Array of Independent Disks). Diese Systeme wurden eingerichtet, um die Datensicherheit zu erhöhen, indem sie redundante Daten auf mehreren Festplatten speichern. Dadurch kann der Ausfall einer einzelnen Festplatte kompensiert werden, ohne dass Daten verloren gehen oder der Betrieb unterbrochen wird. Die Konfiguration der RAID-Systeme erfolgte in enger Abstimmung mit den Anforderungen der spezifischen Anwendungen, die auf dem yourAI-Cluster laufen sollten, um eine optimale Leistung und Zuverlässigkeit sicherzustellen.

Zusätzlich zur physischen Installation und Netzwerkanbindung wurde ein umfassendes Virtualisierungs-Framework implementiert. Auf den Servern wurden Virtualisierungstechnologien wie oVirt für die Verwaltung der virtuellen Maschinen (VMs) eingesetzt. Diese VMs sind entscheidend für die flexible Nutzung der Hardware-Ressourcen, da sie es ermöglichen, verschiedene Betriebssysteme und Anwendungen parallel auf denselben physischen Maschinen zu betreiben.

Abschließend wurden umfangreiche Tests durchgeführt, um die Funktionsfähigkeit der installierten Hardware und der konfigurierten Systeme zu überprüfen. Diese Tests umfassten sowohl die Überprüfung der Netzwerkkonnektivität als auch die Validierung der Virtualisierungsumgebung und der Speicherlösungen. Es wurden spezifische Testfälle definiert, um sicherzustellen, dass alle Komponenten nahtlos zusammenarbeiten und die geplanten Leistungsziele erreichen. Diese Testphase war entscheidend, um mögliche Schwachstellen frühzeitig zu identifizieren und zu beheben, bevor der Cluster in den produktiven Betrieb überging.

Durch die erfolgreiche Durchführung dieser umfassenden Installations- und Inbetriebnahmearbeiten konnte die Hardware des yourAI-Clusters voll funktionsfähig in Betrieb genommen werden. Dies legte den Grundstein für die weitere Implementierung der Softwareumgebung.

## **2.5 Arbeitspaket 5: Integration ins Hochschulnetz und Benutzermanagement**

Die Integration des yourAI-Clusters in die bestehende IT-Infrastruktur der Hochschule Bielefeld stellte eine komplexe und kritische Phase des Projekts dar. Um sicherzustellen, dass der Cluster nahtlos in das Hochschulnetzwerk eingebunden und für alle vorgesehenen Nutzergruppen zugänglich war, war eine enge Zusammenarbeit mit der Datenverarbeitungszentrale (DVZ) der Hochschule unerlässlich. Dieses Arbeitspaket umfasste die Entwicklung und Implementierung einer maßgeschneiderten Netzwerkarchitektur sowie eines umfassenden Benutzermanagementsystems, das sowohl den hohen Sicherheitsanforderungen als auch den Bedürfnissen der Nutzer gerecht wurde.

Ein zentraler Bestandteil der Integration war die Entwicklung einer flexiblen und sicheren Netzwerkarchitektur, die es ermöglicht, verschiedene Nutzergruppen sicher und effizient auf den Cluster zugreifen zu lassen. Hierfür wurden mehrere Netzwerke konfiguriert und mit spezifischen VLANs (Virtual Local Area Networks) versehen, um den Datenverkehr zwischen den unterschiedlichen Nutzergruppen und Anwendungen zu segmentieren und abzusichern. Diese VLANs wurden so gestaltet, dass sie sowohl dem internen Hochschulnetz als auch externen Partnern den Zugang zum Cluster ermöglichen, ohne die Sicherheit oder die Performance zu beeinträchtigen.

Ein besonderer Fokus lag auf der Implementierung einer Authentifizierungs- und Autorisierungslösung, die auf den Technologien Shibboleth und FreeIPA basiert. Das zentrale Identity Management (IdM)-System der Hochschule, das von der Datenverarbeitungszentrale (DVZ) betrieben wird, stellt für den yourAI-Cluster die Authentifizierungsdaten bereit. In diesem Kontext wurde unser Shibboleth-Dienst als Service Provider (SP) konfiguriert, der das DVZ-IdM-System abfragt, um die Authentifizierung der Nutzer zu ermöglichen. Diese Konfiguration ermöglicht es den Nutzern, sich mit ihren bestehenden Hochschulzugangsdaten am Cluster anzumelden, ohne dass zusätzliche Anmeldedaten verwaltet werden müssen. Dies erleichterte den Zugang erheblich und gewährleistete gleichzeitig eine hohe Sicherheit, da die Authentifizierung zentral über das etablierte IdM-System der Hochschule erfolgt.

Die Rolle von FreeIPA war es, die Authentifizierung und Autorisierung der Nutzer innerhalb des Clusters zu verwalten. FreeIPA bietet eine zentrale Benutzerverwaltung, die es ermöglicht, spezifische Rechte und Ressourcen für unterschiedliche Nutzergruppen zu konfigurieren. Diese Konfigurationen wurden so ausgearbeitet, dass sie den jeweiligen Anforderungen der verschiedenen Nutzergruppen entsprechen – von Studierenden, die Zugang zu Basisressourcen benötigen, bis hin zu Forschenden, die Zugriff auf spezialisierte Rechenressourcen und Anwendungen benötigen. Die granularen Zugriffsrechte wurden dabei über Gruppenrichtlinien und Rollenmanagement gesteuert, was eine flexible Anpassung an wechselnde Anforderungen gewährleistete.

Darüber hinaus wurden umfangreiche Maßnahmen zur Sicherstellung der Netzwerksicherheit ergriffen. Dazu gehörte die Implementierung einer Firewall-Konfiguration, die den Zugriff auf den Cluster streng kontrollierte und unerlaubte Zugriffsversuche verhinderte. Zudem wurden mehrere Sicherheitszonen innerhalb des Netzwerks eingerichtet, die den Zugang zu kritischen Systemen und Daten beschränkten. Diese Zonen basieren auf dem Prinzip der minimalen Rechtevergabe, wodurch sichergestellt wird, dass Nutzer nur auf die Ressourcen zugreifen können, die sie tatsächlich benötigen.

Die Integration des Clusters umfasste auch die Einbindung der Hardware in die bestehende Infrastruktur des Rechenzentrums. Hierbei war es notwendig, die Netzwerkschnittstellen der Server präzise zu konfigurieren, um eine reibungslose Kommunikation mit den anderen Systemen im Rechenzentrum zu gewährleisten. Dabei wurde besonderes Augenmerk auf die Konfiguration der Netzwerkadapter und deren Zuordnung zu den entsprechenden VLANs gelegt, um den Datenverkehr effizient zu leiten und Engpässe zu vermeiden.

Ein weiterer wichtiger Aspekt war die Einrichtung von NTP-Servern (Network Time Protocol) innerhalb des Clusters, um die Synchronisation der Systemzeiten sicherzustellen. Die korrekte Zeitsynchronisation ist essenziell für die Ausführung verteilter Anwendungen und die Konsistenz von Logdateien, was wiederum entscheidend für die Fehlersuche und -behebung im Clusterbetrieb ist. Diese NTP-Server wurden so konfiguriert, dass sie mit den zentralen Zeitsystemen der Hochschule synchronisiert werden und damit eine einheitliche Zeitbasis im gesamten Netzwerk gewährleisten.

Die Umsetzung des Benutzermanagements beinhaltete zudem die Einrichtung eines webbasierten Portals, das den Nutzern einen einfachen Zugang zum Cluster ermöglichte. Über dieses Portal, das auf der OpenOnDemand-Plattform basiert, konnten die Nutzer nicht nur auf die Rechenressourcen des Clusters zugreifen, sondern auch ihre Jobs verwalten, interaktive Sessions wie Jupyter-Notebooks starten und Dateien hochladen oder herunterladen. OpenOnDemand bietet eine intuitive Benutzeroberfläche, die die Nutzung des Clusters erheblich erleichtert und somit die Akzeptanz und Effizienz der Nutzer steigert.

Trotz der Komplexität der Aufgaben und der Herausforderungen, die sich im Zuge der Integration ergaben, konnte die Einbindung des yourAI-Clusters in die IT-Infrastruktur der Hochschule erfolgreich abgeschlossen werden. Die erstellte Netzwerkarchitektur und das implementierte Benutzermanagementsystem bieten nun eine solide Grundlage für den Betrieb des Clusters und ermöglichen den Nutzern einen sicheren, effizienten und komfortablen Zugang zu den Rechenressourcen. Nur ein selbstgestecktes Ziel konnte leider während der Projektlaufzeit aufgrund zu hoher technischer Hürden (gerade auch

bzgl. IT-Sicherheit) nicht umgesetzt werden: Die automatische Erstellung eines Nutzeraccounts auf dem yourAI-Cluster, sobald sich ein Hochschulangehöriger erstmalig dort einzuloggen versucht. Aktuell müssen Nutzeraccounts weiterhin auf Anfrage durch den Administrator erstellt werden.

## **2.6 Arbeitspaket 6: Inbetriebnahme der Software**

Die Inbetriebnahme der Softwareumgebung stellte einen weiteren zentralen Meilenstein des yourAI-Projekts dar und war entscheidend dafür, den Cluster in eine voll funktionsfähige Forschungs- und Lehrplattform zu verwandeln. Nachdem die Hardware erfolgreich installiert und in das Hochschulnetzwerk integriert worden war, begann die Implementierung mehrerer spezialisierter Softwarelösungen, die eine flexible und leistungsstarke Umgebung für die Entwicklung und Ausführung von KI-Anwendungen bieten sollten. Der weit überwiegende Teil der verwendeten Software ist Open Source, weshalb im Projekt nur vergleichsweise geringe Ausgaben für Softwarebeschaffungen angefallen sind.

Ein wichtiger Bestandteil dieser Softwareumgebung war die Implementierung eines Virtualisierungsclusters, der auf dem Hypervisor oVirt basiert. oVirt ist eine Open-Source-Virtualisierungsplattform, die es ermöglicht, mehrere virtuelle Maschinen (VMs) auf den physischen Servern des Clusters zu betreiben. Diese VMs spielen eine zentrale Rolle für die Flexibilität und Sicherheit des Clusters, da sie sowohl für administrative Aufgaben als auch für spezielle Nutzerbedarfe eingesetzt werden.

Für administrative Zwecke wurde oVirt so konfiguriert, dass es eine zentrale Admin-VM bereitstellt. Diese Admin-VM dient als sichere Ausgangsbasis für Administratoren, die sich ausschließlich von dieser VM aus auf die anderen Server im Cluster als Admin verbinden können. Diese Architektur sorgt für eine strikte Kontrolle und Isolation der administrativen Zugänge, was die Sicherheit und Integrität des gesamten Clusters erheblich erhöht. Die Admin-VM ermöglicht es, administrative Aufgaben zentral zu steuern und zu überwachen, ohne dass direkte Zugriffe auf die physischen Server notwendig sind.

Ein weiterer wichtiger Bestandteil der Virtualisierungsumgebung ist die Bereitstellung einer SSH-Login-Node-VM. Diese VM dient als dedizierter Einstiegspunkt für Nutzer, die über SSH auf die Rechenressourcen des Clusters zugreifen möchten. Der SSH-Login-Node bietet eine sichere und kontrollierte Umgebung, in der Nutzer ihre Jobs einreichen und verwalten können, bevor diese auf den eigentlichen Rechenknoten ausgeführt werden. Durch die Trennung des SSH-Zugangs von den Rechenknoten wird die Sicherheit und Stabilität des Clusters erhöht, da der direkte Zugriff auf die sensiblen Ressourcen des Clusters auf diese Weise beschränkt bleibt.

Zusätzlich wurde eine VM zur Bereitstellung von OpenOnDemand eingerichtet. OpenOnDemand ist eine benutzerfreundliche Web-Oberfläche, die es Nutzern ermöglicht, auf die HPC-Ressourcen des Clusters zuzugreifen und ihre Arbeitsabläufe effizient zu verwalten. Diese Plattform bietet Funktionen wie das Einreichen von Batch-Jobs, die Verwaltung von Dateien und den Start interaktiver Sessions wie Jupyter-Notebooks direkt im Browser. Die OpenOnDemand-VM wurde so konfiguriert, dass sie nahtlos in die Clusterinfrastruktur integriert ist, um den Nutzern eine einfache und intuitive Möglichkeit zu bieten, auf die Rechenressourcen zuzugreifen, ohne dass tiefgehende technische Kenntnisse erforderlich sind.

Darüber hinaus wurde oVirt auch eingesetzt, um speziellen Nutzergruppen eine eigene VM anzubieten, insbesondere in Fällen, in denen dies aufgrund von Datenschutzerfordernissen wie der DSGVO im Rahmen von Forschungsprojekten erforderlich ist. Diese VMs können von den Nutzern selbst über das benutzerfreundliche oVirt-VM-Portal verwaltet werden. Das Portal ermöglicht es den Nutzern, ihre VMs eigenständig zu starten, zu stoppen und zu konfigurieren, ohne auf direkte Unterstützung durch die IT-Abteilung angewiesen zu sein. Dies ist besonders nützlich für Forschungsprojekte, die eine isolierte Umgebung benötigen, in der die Nutzer ihre eigenen Software-Stacks und Betriebssysteme betreiben können, ohne die Ressourcen anderer Nutzergruppen zu beeinträchtigen.

Die Konfiguration von oVirt umfasste auch die Einrichtung von Netzwerkverbindungen und Speicherpools, um sicherzustellen, dass die VMs effizient und sicher mit den übrigen Komponenten des Clusters

interagieren können. Insgesamt trägt die Implementierung von oVirt im yourAI-Cluster dazu bei, eine flexible, sichere und benutzerfreundliche Umgebung zu schaffen, die sowohl den administrativen Anforderungen als auch den spezifischen Bedürfnissen der Nutzer gerecht wird. Diese Virtualisierungsplattform ermöglicht es, verschiedene Workloads effizient zu verwalten und bietet den Nutzern die notwendige Autonomie, um ihre Projekte erfolgreich umzusetzen.

Ein ursprünglich geplanter Bestandteil der Softwareumgebung war die Installation und Konfiguration des Hadoop-Ökosystems, das für die Verarbeitung großer Datenmengen in verteilten Systemen genutzt wird. Hadoop sollte auf mehreren Knoten des Clusters eingerichtet werden, um parallele Datenverarbeitungsaufgaben zu unterstützen, die für KI-Anwendungen, insbesondere im Bereich Big Data, von zentraler Bedeutung sind. Allerdings ergaben sich während der Projektlaufzeit unerwartete Entwicklungen, die eine Änderung dieses Plans erforderlich machten.

Die Einstellung der Weiterentwicklung von Ambari, einer wichtigen Managementplattform für Hadoop-Umgebungen, sowie das Verschwinden von Hortonworks, einem der Hauptanbieter von Hadoop-Distributionen, stellten erhebliche Herausforderungen dar. Diese Veränderungen führten dazu, dass die Installation und der Betrieb von Hadoop/HDFS deutlich komplizierter und weniger zuverlässig wurden. In Anbetracht dieser Entwicklungen entschied sich das Projektteam, von der ursprünglich geplanten Implementierung von Hadoop abzusehen.

Stattdessen wurde Ceph als Speicherlösung eingesetzt, das sich als flexiblere und zukunftssichere Alternative erwies. Ceph wurde gewählt, da es eine hochverfügbare, skalierbare und redundante Speicherlösung darstellt, die den Anforderungen an Datensicherheit und -integrität gerecht wird. Ceph ermöglicht es, Daten automatisch über mehrere Speicherorte im Cluster zu replizieren, wodurch ein Höchstmaß an Ausfallsicherheit gewährleistet wird. Dies ist besonders wichtig für KI-Anwendungen, die oft große Datenmengen verarbeiten und speichern müssen. Die Konfiguration von Ceph umfasste die Einrichtung von OSDs (Object Storage Daemons), die Verwaltung der MONs (Monitor Daemons) und die Optimierung der Netzwerkeinstellungen, um die maximale Leistung und Zuverlässigkeit des Speichersystems zu erreichen. Zusätzlich wurde Ceph in die Virtualisierungsumgebung integriert, um den VMs einen schnellen und zuverlässigen Zugriff auf den Speicher zu ermöglichen.

Neben der Virtualisierung und dem Speichersystem spielte das Ressourcenmanagementsystem Slurm eine zentrale Rolle in der Softwareumgebung. Slurm (Simple Linux Utility for Resource Management) ist ein weit verbreitetes Open-Source-Job-Management-System, das speziell für den Einsatz in großen Rechenclustern entwickelt wurde. Die Implementierung von Slurm im yourAI-Cluster ermöglichte eine effiziente Verwaltung der Rechenressourcen, indem es die Aufgaben der Nutzer automatisch auf die verfügbaren Serverknoten verteilte. Slurm wurde so konfiguriert, dass es die unterschiedlichen Anforderungen der Nutzergruppen berücksichtigt, einschließlich der Priorisierung von Aufgaben, der Verwaltung von Warteschlangen und der Zuweisung von Ressourcen wie CPU-Kernen, Arbeitsspeicher und GPUs. Diese Konfiguration stellte sicher, dass die leistungsstarken Rechenressourcen des Clusters optimal genutzt werden und die Ausführung von KI-Anwendungen so effizient wie möglich erfolgt.

Ursprünglich war geplant, Kubernetes als Orchestrierungstool für Container-Anwendungen auf allen Compute-Nodes des Clusters zu implementieren. Die Idee bestand darin, statische Slurm-Node-Container innerhalb von Kubernetes zu betreiben, um so eine flexible und skalierbare Umgebung für die Verwaltung der Workloads zu schaffen. Diese Integration sollte die Entwicklung und Bereitstellung von Microservices und anderen containerisierten KI-Anwendungen ermöglichen und gleichzeitig die Effizienz der Ressourcenverwaltung erhöhen.

Allerdings stellte sich während der Projektlaufzeit heraus, dass die Zeit aufgrund der bereits beschriebenen Verzögerungen und Herausforderungen nicht ausreichte, um dieses Konzept vollständig umzusetzen. Deshalb wurde entschieden, Kubernetes nicht auf den Compute-Nodes auszurollen. Stattdessen wurde Slurm direkt auf den Compute-Nodes als bare-metal-System installiert und betrieben. Diese Entscheidung ermöglichte es, die Rechenressourcen des Clusters optimal zu nutzen und gleichzeitig

die geplante Funktionalität von Slurm zu gewährleisten, auch wenn die angedachte Flexibilität durch Kubernetes nicht vollständig realisiert werden konnte.

Abschließend wurde die gesamte Softwareumgebung in umfangreichen Tests geprüft, um sicherzustellen, dass alle Komponenten nahtlos zusammenarbeiten und die erforderliche Leistung erbringen. Diese Tests umfassten die Validierung der Netzwerk- und Speicherperformance, die Überprüfung der Ressourcenzuweisung durch Slurm und die Funktionalität der Virtualisierungslösungen. Besonderes Augenmerk wurde auf die Interoperabilität zwischen den verschiedenen Softwarekomponenten gelegt, um sicherzustellen, dass die Nutzer nahtlos zwischen verschiedenen Tools und Plattformen wechseln können, ohne auf Hindernisse zu stoßen.

Durch die erfolgreiche Implementierung und Integration dieser spezialisierten Softwarelösungen bietet der yourAI-Cluster nun eine leistungsstarke und flexible Umgebung, die den Anforderungen moderner KI-Anwendungen gerecht wird. Diese Umgebung ermöglicht es Forschenden und Studierenden, innovative Projekte durchzuführen, die auf modernster Technologie basieren, und legt den Grundstein für zukünftige Entwicklungen im Bereich der Künstlichen Intelligenz an der Hochschule Bielefeld.

## **2.7 Arbeitspaket 7: Integration der Edge-Hardware**

Die Integration der Edge-Hardware war ursprünglich als ein weiteres Ziel des Projekts geplant. Allerdings führten die bereits erwähnten Lieferengpässe und die anhaltende Chipkrise dazu, dass dieses Arbeitspaket nicht vollständig abgeschlossen werden konnte. Zwar wurden alle wesentlichen Komponenten der Edge-Infrastruktur beschafft, jedoch konnte die geplante Entwurfsraumexploration nur eingeschränkt realisiert werden. Dies führte dazu, dass die ursprünglich angestrebte einfache Skalierbarkeit der Infrastruktur für KI-Anwendungen nicht in vollem Umfang erreicht wurde. Aktuell ist der Toolstack, der den Nutzern der yourAI-Infrastruktur für die Edge-Komponenten zur Verfügung steht, noch separiert vom Toolstack auf dem yourAI-Cluster.

## **2.8 Arbeitspaket 8: Test und Dokumentation**

Im letzten Abschnitt des Projekts wurde eine umfassende Testphase eingeleitet, um die Funktionalität, Stabilität und Benutzerfreundlichkeit der yourAI-Plattform zu evaluieren. Diese Phase war entscheidend, um sicherzustellen, dass die entwickelten Systeme und Technologien den hohen Anforderungen gerecht werden und zuverlässig im produktiven Betrieb eingesetzt werden können. Die Testphase wurde als "closed beta" konzipiert und zielte darauf ab, eine ausgewählte Gruppe von Nutzern innerhalb der Hochschule Bielefeld in die Nutzung der Plattform einzuführen und deren Feedback zu sammeln.

Die Entscheidung, die Testphase als "closed beta" durchzuführen, ermöglichte es, den Umfang der Tests auf eine überschaubare Nutzergruppe zu beschränken. Dies hatte den Vorteil, dass potenzielle Probleme und Schwachstellen in einer kontrollierten Umgebung identifiziert und behoben werden konnten, bevor die Plattform einem breiteren Publikum zugänglich gemacht wurde. Die Auswahl der Nutzer erfolgte auf Basis ihrer spezifischen Anforderungen und ihrer Erfahrung mit Hochleistungsrechnen (HPC). Dadurch konnte sichergestellt werden, dass das Feedback sowohl von technisch versierten Nutzern als auch von solchen, die weniger Erfahrung mit HPC-Umgebungen haben, berücksichtigt wurde.

Ein zentrales Element dieser Testphase war die Nutzung des Cluster-Webportals OpenOnDemand. OpenOnDemand, eine Open-Source-Plattform, bietet eine benutzerfreundliche Web-Oberfläche, die es den Nutzern ermöglicht, bequem über den Browser im VPN der Hochschule Bielefeld auf die HPC-Ressourcen des yourAI-Clusters zuzugreifen. Über das Portal konnten die Nutzer Batch-Jobs einreichen, Dateien verwalten und interaktive Sessions wie Jupyter Notebooks starten. Die Implementierung von OpenOnDemand spielte eine Schlüsselrolle in der Testphase, da sie den Zugang zu den komplexen Rechenressourcen erheblich vereinfachte und die Akzeptanz der Plattform bei den Nutzern förderte.

Während der Testphase wurden verschiedene Aspekte der Plattform ausführlich getestet, darunter die Leistung und Stabilität der Rechenressourcen, die Effizienz der Jobverwaltung über Slurm, die Zuverlässigkeit der Speicherlösungen sowie die Benutzerfreundlichkeit des OpenOnDemand-Portals. Die Tests umfassten auch die Interoperabilität zwischen den verschiedenen Softwarekomponenten, wie die nahtlose Integration von Ceph als verteiltem Speichersystem und die Nutzung von Slurm für die Ressourcenverwaltung. Besonderes Augenmerk wurde auf die Sicherheit der Plattform gelegt, insbesondere auf die Implementierung der Authentifizierungs- und Autorisierungsmechanismen über Shibboleth und FreeIPA.

Um ein möglichst breites Spektrum an Anwendungsszenarien abzudecken, wurden die Nutzer der closed beta dazu ermutigt, unterschiedliche Arten von Workloads auf der Plattform zu testen. Dies reichte von einfachen Batch-Jobs bis hin zu komplexen, parallelisierten Rechenaufgaben, die die Leistungsfähigkeit und Skalierbarkeit des Clusters auf die Probe stellten. Die Nutzer erhielten die Möglichkeit, ihre Erfahrungen direkt an das Projektteam weiterzugeben, was wertvolle Einblicke in die tatsächliche Nutzung und potenzielle Optimierungsmöglichkeiten lieferte.

Die während der Testphase gewonnenen Erkenntnisse waren von großem Wert für die Weiterentwicklung der yourAI-Plattform. Es wurden mehrere Optimierungsmaßnahmen identifiziert, die in zukünftige Versionen der Plattform einfließen werden. Dazu gehören unter anderem Verbesserungen in der Benutzerführung des OpenOnDemand-Portals, Optimierungen und Erweiterung der Nutzungs-Dokumentation, und Anpassungen in der Konfiguration von Slurm, um die Ressourcenzuweisung noch effizienter zu gestalten.

Parallel zur Testphase wurde eine umfassende Dokumentation erstellt, die alle Aspekte der Plattform, von der Hardwarearchitektur über die Softwarekonfiguration bis hin zu den Benutzerhandbüchern, abdeckt. Diese Dokumentation dient nicht nur als Referenz für die zukünftige Wartung und Weiterentwicklung der Plattform, sondern auch als Schulungsressource für neue Nutzer und Administratoren. Die Dokumentation wurde in enger Abstimmung mit den Nutzern der closed beta erstellt, um sicherzustellen, dass sie praxisnah ist und die tatsächlichen Bedürfnisse der Nutzer reflektiert.

Insgesamt hat die Test- und Dokumentationsphase gezeigt, dass die yourAI-Plattform eine leistungsfähige und flexible Lösung für die Anforderungen der Hochschule Bielefeld darstellt. Die gesammelten Rückmeldungen und die identifizierten Optimierungsmöglichkeiten bilden eine solide Grundlage für den weiteren Ausbau und die breite Einführung der Plattform. Der erfolgreiche Abschluss dieser Phase markiert einen wichtigen Meilenstein im Projekt.

## **2.9 Arbeitspaket 9: Roll Out**

Zum Projektende am 31.03.2024 konnte der Roll Out noch nicht durchgeführt werden, da noch die in Abschnitt 2.8 beschriebene Testphase lief. Seitdem ist fachbereichsübergreifend die Anzahl der Nutzer kontinuierlich erhöht worden, und es wird schon umfänglich produktiv mit dem yourAI-Cluster gearbeitet, sowohl in Promotionsprojekten als auch in Lehrveranstaltungen.

## **3. Zahlenmäßiger Nachweis und Verwendungszweck**

Die Zuwendungsmittel wurden im Rahmen des Projekts in erster Linie für die Beschaffung der erforderlichen Hardware sowie für die Implementierung und Integration der Softwareumgebung verwendet. Aufgrund der unvorhergesehenen Verzögerungen und der damit verbundenen Kostensteigerungen war es notwendig, einige der ursprünglich geplanten Investitionen anzupassen. Die größten Kostenpositionen entfielen auf die Anschaffung der Server-Hardware sowie Personalmittel. Diese Ausgaben waren notwendig, um die Zielsetzung des Projekts, nämlich die Bereitstellung einer leistungsstarken KI-Infrastruktur, zu erreichen.

Die wichtigsten Positionen im Einzelnen:

Kategorien 0812 und 0817 (Personalmittel): Es wurden insg. 24 Personenmonate für einen wissenschaftlichen Mitarbeiter finanziert (Vollzeit; anfangs in TVL-Gruppe E11, später in E13).

Kategorie 0831 (Gegenstände bis 800€): Hierunter fiel ein großer Teil der Edge-Hardware, wobei die Beschaffungsliste an die Marktlage angepasst werden musste, so dass Nachfolge- bzw. Ersatzprodukte beschafft wurden. Die Softwarebeschaffungen fielen alle in Kategorie 0831.

Kategorie 0850 (Gegenstände ab 800€): Die Ausgaben in Kategorie 0850 umfassten ca. 80% des Projektvolumens. Hierüber wurden in erster Linie die Server- und Storage-Komponenten für den yourAI-Cluster beschafft (2022/23: 4 GPU-Server mit je 4 Nvidia A100, 1 FPGA-Server, 2 Virtualisierungs-Server, 5 Data-Nodes, 1 File-Server, 1 Backup-Server, Switches; 2023/2024: 7 FPGA-Karten für den FPGA-Server). Außerdem entfielen Edge-Komponenten im Wert von gut EUR 8.000 auf die Kategorie 0850.

#### **4. Notwendigkeit und Angemessenheit der geleisteten Projektarbeiten**

Die durchgeführten Projektarbeiten waren in jeder Hinsicht notwendig und angemessen, um die anspruchsvollen Ziele des yourAI-Projekts zu erreichen. Die Herausforderungen, die sich insbesondere durch die globale Chipkrise und die damit verbundenen Lieferengpässe und durch den komplexen Beschaffungsprozess für die Cluster-Hardware ergaben, machten es erforderlich, die schon erwähnten zwei kostenneutralen Projektverlängerungen um insg. neuen Monate zu beantragen. Damit verbunden wurden Einsparungen bei der Hardwarebeschaffung genutzt, um die Personalmittel für die wissenschaftliche Mitarbeiterstelle von zunächst geplanten 18 Personenmonaten auf 24 Personenmonate aufzustocken. Dieser Personaleinsatz war definitiv notwendig, um alle geleisteten Arbeiten durchführen zu können. Aus der Beschreibung der Arbeitspakete wird ersichtlich, wie komplex das Projekt letztendlich war.

Trotz dieser Herausforderungen konnte das Projekt insgesamt erfolgreich umgesetzt werden, und die erzielten Ergebnisse sind für die Hochschule von großer Bedeutung.

#### **5. Voraussichtlicher Nutzen und Verwertbarkeit der Ergebnisse**

Die Ergebnisse des yourAI-Projekts sind von erheblichem Nutzen für die Forschung und Lehre im Bereich der Künstlichen Intelligenz (KI) an der Hochschule Bielefeld. Es wurde eine leistungsfähige Plattform geschaffen, die es ermöglicht, eine breite Palette von KI-Anwendungen zu unterstützen, von maschinellem Lernen und Datenanalyse bis hin zu komplexen Simulationsrechnungen und tiefen neuronalen Netzwerken. Diese Infrastruktur ermöglicht dem wissenschaftlichen Nachwuchs an der Hochschule Bielefeld, mit aktuellen KI-Verfahren, die teilweise extremen Ressourcenbedarf haben, zu arbeiten und diese z.B. in Master- und Promotionsarbeiten weiterzuentwickeln. Dies fördert nicht nur die Qualität der wissenschaftlichen Arbeit, sondern stärkt auch die Position der Hochschule als Standort für angewandte KI.

Ein besonderer Vorteil der yourAI-Plattform ist ihre Flexibilität und Skalierbarkeit, die es ermöglicht, unterschiedliche Nutzergruppen mit ihren jeweiligen spezifischen Anforderungen zu bedienen. Studierende können die Plattform nutzen, um praktische Erfahrungen im Umgang mit großen Datenmengen und komplexen Algorithmen zu sammeln, während Promovierende von den leistungsstarken Rechenkapazitäten profitieren, um anspruchsvolle wissenschaftliche Fragestellungen zu untersuchen. Durch die Integration verschiedener Softwarelösungen und Tools, wie etwa JupyterServer und OpenOnDemand, wird den Nutzern zudem eine benutzerfreundliche Umgebung geboten, die den Zugang zu den Ressourcen vereinfacht und die Effizienz der Arbeit erheblich steigert.

In der näheren Zukunft ist geplant, die yourAI-Plattform weiter zu optimieren und innerhalb der Hochschule noch stärker zu bewerben, wobei schon jetzt die GPU-Server mehr als ausgelastet sind. Darüber hinaus wird die Integration neuer Softwarelösungen dazu beitragen, die Plattform noch leistungsfähiger und vielseitiger zu machen. Geplant ist unter anderem die Implementierung von Tools für die verteilte Datenverarbeitung und -speicherung, die es den Nutzern ermöglichen, noch größere Datenmengen effizient zu verwalten und zu analysieren. Dabei wird die Nutzung von Ceph als zentralem Speichersystem weiter ausgebaut, um eine optimale Performance und Datenintegrität sicherzustellen. Auch die Weiterentwicklung der vorhandenen Virtualisierungs- und Containerisierungsinfrastruktur, beispielsweise durch die spätere Einführung von Kubernetes, wird ein wichtiger Schritt sein, um die Flexibilität der Plattform weiter zu erhöhen und die Bereitstellung neuer Dienste zu erleichtern.

Ein weiteres Ziel ist die Nutzung der yourAI-Infrastruktur in Kooperationsprojekten mit externen Partnern (z.B. andere Hochschulen und Forschungseinrichtungen, aber auch Industrie), in denen der wissenschaftliche Nachwuchs an der Hochschule Bielefeld beteiligt ist. So kann die Plattform als Knotenpunkt für gemeinsame Forschungsprojekte dienen, die über institutionelle Grenzen hinweg durchgeführt werden. Dies fördert nicht nur den wissenschaftlichen Austausch, sondern eröffnet auch neue Möglichkeiten für interdisziplinäre Forschung, die von den umfangreichen Rechenressourcen und den spezialisierten Tools der yourAI-Plattform profitieren kann.

Neben den technologischen Weiterentwicklungen wird auch die Schulung und Unterstützung der Nutzer weiterhin ein zentrales Thema bleiben. Geplant sind regelmäßig stattfindende Workshops und Schulungen, die den Nutzern helfen sollen, die umfangreichen Möglichkeiten der Plattform optimal zu nutzen. Diese Veranstaltungen werden nicht nur die technischen Aspekte der Plattform abdecken, sondern auch die Anwendung der KI-Tools in spezifischen Forschungs- und Lehrkontexten thematisieren. Dadurch wird sichergestellt, dass die yourAI-Plattform nicht nur technologisch auf dem neuesten Stand bleibt, sondern auch den Nutzern die bestmögliche Unterstützung bietet.

Ein besonders vorteilhafter Aspekt der yourAI-Plattform ist die Möglichkeit, einzelne Server, die im Rahmen spezifischer Folgeprojekte angeschafft wurden, nahtlos in den Cluster zu integrieren. Diese Server können dann während der Laufzeit des jeweiligen Projekts exklusiv den Projektbeteiligten (Mitarbeiter, Professoren, externe Partner, ...) zur Verfügung gestellt werden. Dies bietet der Hochschule Bielefeld einen erheblichen Mehrwert, da es den Forschenden ermöglicht, ihre Rechenressourcen gezielt auf die Bedürfnisse ihres Projekts abzustimmen und gleichzeitig von der robusten Infrastruktur und den fortschrittlichen Verwaltungsmöglichkeiten des Clusters zu profitieren. Ein weiterer bedeutender Vorteil ist, dass die Projektbeteiligten sich voll und ganz auf die inhaltlichen Aspekte ihres Projekts konzentrieren können, ohne sich um die Administration der Hardware kümmern zu müssen. Dadurch können die Forscher ihre Energie und Ressourcen effektiv auf die wissenschaftlichen Ziele ihres Projekts richten, während sie gleichzeitig sicher sein können, dass die technische Seite durch die stabile und gut verwaltete Plattform des Clusters zuverlässig unterstützt wird. Nach Abschluss eines Projekts können die Server wieder in den allgemeinen Pool des Clusters integriert werden, wodurch die Nachhaltigkeit der Investitionen erhöht und die Gesamtkapazität des Clusters langfristig erweitert wird. Dieses Modell der dynamischen Ressourcenzuweisung stärkt die Forschungskapazitäten der Hochschule und unterstützt die erfolgreiche Durchführung hochspezialisierter Forschungsprojekte.

Zusammenfassend lässt sich sagen, dass die Ergebnisse des yourAI-Projekts bereits jetzt von erheblichem Nutzen sind und das Potenzial haben, in Zukunft noch weiter ausgebaut zu werden. Die entwickelte Infrastruktur bildet das Rückgrat für moderne KI-Forschung und -Lehre an der Hochschule Bielefeld und wird durch die geplanten Erweiterungen und Optimierungen noch leistungsfähiger und vielseitiger werden. Diese Entwicklungen werden dazu beitragen, die Hochschule als renommierten Standort für KI-Forschung weiter zu etablieren und den Studierenden und Forschenden Zugang zu den besten verfügbaren Ressourcen und Tools zu geben.

## **6. Bekannter Fortschritt auf dem Gebiet des Vorhabens bei anderen Stellen**

Während der Durchführung des Projekts wurde kein bedeutender technologischer Fortschritt auf dem Gebiet des Vorhabens bei anderen Stellen bekannt, der die Zielsetzung oder Durchführung des yourAI-Projekts beeinflusst hätte. Vielmehr konnten durch die enge Zusammenarbeit mit anderen Forschungsgruppen und externen Partnern Synergien genutzt werden, die zur erfolgreichen Umsetzung des Projekts beitrugen.

## **7. Veröffentlichungen**

Da es sich bei dem yourAI-Projekt um ein reines Infrastrukturprojekt gehandelt hat, sind keine Veröffentlichungen geplant gewesen.