Konsultationsprozess zur nationalen Rechenzentrumsstrategie

Einreichung der Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e. V.

Vorbemerkung

Die resiliente digitale Transformation des Industrie- und Innovationsstandorts Deutschland ist essenziell für seine künftige Wettbewerbsfähigkeit. Diese Transformation kann nur gelingen, wenn Deutschland seine technologischen und industriellen Stärken weiterentwickelt. Entscheidend sind die gezielte Förderung digitaler Schlüsseltechnologien wie Künstliche Intelligenz (KI) und Simulationstechnologie und der beschleunigte Transfer von der Forschung in die Anwendung.

Erst mithilfe von Infrastrukturen wie Rechenzentren können Schlüsseltechnologien gezielt für die Bedürfnisse von Forschung, Wirtschaft und Verwaltung nutzbar gemacht werden. Mit rund 130 000 Vollzeit-Arbeitskräften und zusätzlich rund 80 000 indirekt beschäftigen Arbeitskräften[1] ist die Rechenzentrums-Branche eine wichtige Zukunftsbranche.

Das übergeordnete Ziel einer kohärenten Rechenzentrumspolitik muss sein, dass Europa seine Investitionen in diesem Bereich erhöht, um bei Compute-Ressourcen und -Fähigkeiten auf Augenhöhe mit den USA und China agieren zu können. Auf dieses Ziel sollte die nationale Rechenzentrumsstrategie einzahlen. Dabei ist die Entwicklung Deutschlands zu einem attraktiven Rechenzentrumsstandort die Voraussetzung für technologische Souveränität. Nur eine verringerte Abhängigkeit von US-Hyperscalern durch europäische bzw. deutsche Hard- und Software-Technologien wird es erlauben, Innovationen zu fördern, dabei die Kontrolle über Daten nach EU-Standards zu wahren sowie Kriterien wie Nachhaltigkeit und Energieeffizienz anzuwenden. Zentrale und dezentrale Recheninfrastrukturen spielen dabei gemeinsam eine entscheidende Rolle. Zentrale Recheninfrastrukturen bilden das Rückgrat, da sie künftig schätzungsweise 90 Prozent der Compute-Ressourcen bereitstellen müssen. Dezentrale Infrastrukturen dienen als strategische Ergänzung. Rechencluster sollten beide Säulen berücksichtigen und sich so den Bedarfen der Nutzenden anpassen. Dies muss eine nationale Rechenzentrumsstrategie berücksichtigen.

Gerade jetzt benötigen Wirtschaft, Forschung und Verwaltung einen unbürokratischen und kostengünstigen Zugang zu Recheninfrastrukturen für Schlüsseltechnologien auf verschiedenen Ebenen. Darüber hinaus bedarf es eines Bündels an Anreizen, um den Rechenzentrumsstandort ganzheitlich zu stärken. Hierzu zählen u.a. die Schaffung von wettbewerbsfähigen Strompreisstrukturen, konsequente Förderung von Ökosystemlogiken, Stärkung von europäischen Anbietern von Hard- und Software sowie die Digitalisierung und Vereinfachung von Bau- und Genehmigungsverfahren.

Die Fraunhofer-Gesellschaft begrüßt den offenen Konsultationsprozess zur nationalen Rechenzentrumsstrategie und beteiligt sich gerne an diesem. Als eine der führenden Organisationen für anwendungsorientierte Forschung betreiben die Fraunhofer-Institute über 200 dezentrale Rechenzentren.

[1] Stand: 2022, Quelle: Bitkom »Perspektiven für eine nachhaltige Rechenzentrenwirtschaft bis 2030«

Unsere Bewertung

1. Welche Merkmale und Rahmenbedingungen kennzeichnen aus Ihrer Sicht einen »zukunftsfähigen und leistungsstarken« Rechenzentrumsstandort Deutschland im Jahr 2030?

Ein international wettbewerbsfähiger, zukunftsfähiger, souveräner, leistungsstarker Rechenzentrumsstandort Deutschland im Jahr 2030 zeichnet sich durch folgende Merkmale aus:

Souveränität

Ein souveräner Rechenzentrumsstandort Deutschland zeichnet sich durch den Aufbau und den Betrieb von eigenen Rechenzentren aus. Zentrale und dezentrale Rechenzentren sind die Grundlage, um Schlüsseltechnologien zu beherrschen und die Wertschöpfungskette in Deutschland anzusiedeln. Konkret bedeutet dies, dass in Deutschland wettbewerbsfähige Hardware-/Software-/KI-/Daten-Technologien entwickelt werden müssen, um Infrastrukturen unabhängig von außereuropäischen Unternehmen zu betreiben.

Um die europäische Resilienz auszubauen, sollten bei der Hardware jährlich steigende Anteile europäischer Hardware-Komponenten vorgesehen werden. Dies bedarf insbesondere in Deutschland einer verstärkten Förderung dieser Entwicklung in innovative Start-ups – neben der Entwicklungsförderung u.a. durch frühzeitige, einfache und gegebenenfalls geförderte Beschaffungen. Drittstaaten haben gezeigt, dass es durch die konsequente Nutzung eigener Technologien über mehrere Generationen möglich ist, zu den marktführenden Unternehmen aufzuschließen.

Bau und Genehmigung

Bau- und Genehmigungen für Rechenzentren dürfen kein bürokratisches Hindernis mehr darstellen. Deshalb muss der schnelle Aufbau neuer Rechenzentren mit kurzen Genehmigungs- und Bauzeiten erlaubt sowie aktiv gefördert werden - sei es für kommerzielle Rechenzentren der Industrie oder innerhalb öffentlicher Einrichtungen wie Forschungsorganisationen und Universitäten. Insbesondere müssen der Aufbau und der Betrieb von Rechenzentren in Deutschland für kommerzielle Anbieter durch den Abbau marktwirtschaftlicher, regulatorischer und infrastruktureller Hürden attraktiv und wettbewerbsfähig sein.

Energie-, Kosten-, Flächeneffizienz

Der Rechenzentrumsstandort Deutschland zeichnet sich durch konsequente Ausrichtung auf Energieeffizienz und Klimaneutralität aus. Investitionen in Rechenzentren werden unter den Gesichtspunkten 1. Bedarf der Nutzenden, 2. Innovationskraft, 3. Gesamtlogik des Ökosystems sowie 4. Kosten-, Energie- und Flächeneffizienz geplant und umgesetzt. Das lokale Potenzial für erneuerbare Energieversorgung, die Auswirkungen auf das Stromnetz und die Möglichkeiten der Abwärmenutzung sind Planungskriterien für einen Rechenzentrumsstandort.

Ökosystem

Rechenzentren müssen zusammen mit Akteuren aus Forschung und Wirtschaft konsequent in einem integrierten Ökosystem gedacht werden, in dem Open Source als Innovationstreiber dient, während gleichzeitig Closed-Source-Modelle für kritische Anwendungen und unternehmerische Differenzierung Platz haben. Dabei erfordert Innovation eine enge Zusammenarbeit von Industrie, Forschung und Rechenzentrumsbetreibern, damit Investitionen in Infrastruktur unmittelbar in Wertschöpfung für Schlüsseltechnologien übersetzt werden.

Zugänglichkeit

Der Rechenzentrumsstandort Deutschland zeichnet sich durch die Wettbewerbsfähigkeit von öffentlichen sowie privaten Rechenzentren aus. Diese ist künftig nicht allein durch modernste Hardware bestimmt, sondern ebenso durch offene und leicht zugängliche Entwicklungsumgebungen. Niedrige Eintrittsbarrieren sind von zentraler Bedeutung, um Industrie, Forschung, Mittelstand und Start-ups gleichermaßen den Zugang zu innovativen Recheninfrastrukturen zu ermöglichen. Auf diese Weise entsteht ein innovationsfreundliches Ökosystem, das Wertschöpfung in Deutschland sichert und digitale Abhängigkeiten nachhaltig verringert.

Deutschland in Europa

Die nationale Rechenzentrumsstrategie und andere nationale Initiativen sowie Gesetze dürfen nicht für sich stehen, sondern müssen eng verzahnt mit europäischen Initiativen zur Datensouveränität gedacht werden. Nur durch diese Anbindung lässt sich sicherstellen, dass Industrie, Wissenschaft und Gesellschaft ihre Daten vertrauensvoll nutzen können, ohne dabei Kontrolle oder Unabhängigkeit an außereuropäische Anbieter abzugeben. Der Ausbau und Betrieb von Rechenzentren in Deutschland leistet damit nicht nur einen Beitrag zur Stärkung der digitalen Infrastruktur, sondern zugleich zur technologischen Souveränität Europas.

Zukunftsfähige Ausrichtung

Zukunftsfähige Konzepte, wie das Edge-Computing Continuum[2], müssen ausreichend Platz in der grundsätzlichen Ausrichtung des Rechenzentrumsstandorts finden. Denn bei diesem Konzept werden durch die nahtlose Integration von Edge- und Cloud-Computing Speicherkapazitäten und Rechenleistungen automatisch an dem Ort genutzt, der am effizientesten und ökonomisch sinnvollsten ist. Kapazitäten für Recheninfrastrukturen müssen einerseits bis zu einem bestimmten Grad skalierungsfähig sein, aber anderseits auch den Anforderungen an geringe Latenz gerecht werden. Das Edge-Computing Continuum vereint diese realen Bedarfe und stärkt damit sowohl resiliente Strukturen durch verteilte Kapazitäten als auch die Umsetzung souveräner Anwendungen vor Ort bzw. im örtlichen Umfeld (on-prem bzw. near-prem).

Deutschland darf sich außerdem nicht in theoretischen Debatten verlieren, während USA und China längst in Richtung AGI (Artificial General Intelligence) investieren. Nur wenn ein Zusammenschluss aus Industrie, Wissenschaft und Politik zügig und entschlossen handelt, kann Deutschland im Jahr 2030 international auf Augenhöhe agieren.

2. Welche zentralen Herausforderungen und Chancen sehen Sie für den Rechenzentrumsstandort Deutschland in den kommenden Jahren?

Aus Sicht der Fraunhofer-Gesellschaft gilt es vor allem folgende zentrale Herausforderungen anzugehen:

Ausrichtung an Bedarfen und Schlüsseltechnologien

Eine zentrale Herausforderung für den Rechenzentrumsstandort Deutschland stellt die Ausrichtung an den unterschiedlichen Bedarfen der Nutzenden und der unterschiedlichen Schlüsseltechnologien dar.

Ein Schwerpunkt für Rechenzentren ist ihre hohe Bedeutung für KI-getriebene Wertschöpfung. Um mit der Dynamik der KI-Entwicklung Schritt zu halten, sind umsetzungsnahe Lösungen erforderlich. Diese Lösungen bestehen neben den anderen High-Performance Computing (HPC)-Anforderungen und bilden eine Erweiterung dieser. Die Innovationsfähigkeit im KI-Bereich erfordert einen agilen, niedrigschwelligen Zugang zu leistungsfähigen KI-Ressourcen für Forschung und Industrie. Die Rechencluster müssen sich an den unterschiedlichen Bedarfen der Nutzenden orientieren. Für das Trainieren sehr großer Modelle sind Rechencluster derzeit in der Größenordnung von 100 000 GPUs notwendig (AI Gigafactories), die für Industrie und Forschung genutzt werden sollen. Daneben benötigt es AI Factories und regionale AI Satelliten bis zu 5 000 GPU[3]. Darunter braucht es KI-spezialisierte Rechencluster für Forschung, Entwicklung und Lehre (2 000 bis 3 000 KI-Beschleuniger) und schließlich kommerzielle Rechenzentren mittlerer und kleinerer Kapazität. Insbesondere das Finetuning trainierter Modelle wird großflächig Anwendung in der Industrie finden, sodass auch der Zugang zu kleineren, dezentralen KI-Clustern essenziell ist. Damit die Investitionen ihr Ziel erreichen, müssen sie in eine Infrastruktur fließen, die speziell auf KI ausgerichtet ist.

Die größere Herausforderung sind die technische Konzeption der neuen Rechenzentren und die damit verbundenen Kostenstrukturen. Heute werden typische Infrastrukturen fünf bis sieben Jahre oder länger betrieben. Das kollidiert mit der sehr schnellen Entwicklung von Hardware und Kommunikationsnetzwerken sowie von neuen Technologien wie speziellen Compute-Acceleratoren und mittelfristig Neuromorphic- und Quantencomputing. Ebenso schnell und algorithmisch verläuft die Entwicklung im Bereich KI: All das wird gehemmt durch eine frühzeitige Festlegung auf spezifische Hardware in großem Umfang.

[2] https://www.cit.fraunhofer.de/de/ueber-uns/Edge-Cloud-Continuum.html

[3] Dies ist auch im EU AI Continent Action Plan vorgesehen.

Energie- und Stromnetz

Eine zentrale Herausforderung für den Rechenzentrumsstandort Deutschland stellt die Beschränkung der Sektorenkopplung auf Strom/Wärme/Mobilität/Industrie dar. Diese sollte anders als bisher Recheninfrastrukturen mitdenken. Häufig sind verschiedene Player für Strombereitstellung, Planung/Aufbau Rechenzentrum, Betrieb Rechenzentren und Kühlung/Abwärmenutzung involviert. Hier kommt es auf eine sinnvolle und umfängliche Vernetzung an. Darüber hinaus sollten neue Geschäftsmodelle gefördert und Anreize geschaffen werden, damit Abwärmenutzung auch wirtschaftlich attraktiv umgesetzt werden kann.

Eine zentrale Herausforderung für den Rechenzentrumsstandort Deutschland stellt der Konflikt von Rechenzentren als unverzichtbare digitale Infrastruktur einerseits und ihrem hohen Energiebedarf andererseits dar. Derzeit beschränken begrenzte Netzanschlusskapazitäten und lange Genehmigungsdauern den Bau von Rechenzentren. Dies führt zur Herausforderung für die effiziente Erneuerung und Konsolidierung zahlreicher dezentraler Rechenzentren. Gerade hierin liegt jedoch auch eine Chance: Moderne Rechenzentren können Flexibilität (insbesondere im Strom- und Wärmesektor) bereitstellen und so aktiv zum Gelingen der Energiewende beitragen – sowohl technisch als auch wirtschaftlich. Entscheidend ist dabei ihre optimale Einbindung in die zentralen Infrastrukturen und Marktmechanismen der Energieversorgung. Parallel dazu gilt es, das Energiemanagement konsequent zu optimieren und Rechenzentren stärker auf Flexibilisierung auszurichten.

Zukunftsfähige Rechenzentrumsstandorte erfordern wettbewerbsfähige Strompreisstrukturen: Industriestrompreise für Großverbraucher sollten durch langfristige Lieferverträge und mit erneuerbaren Energieerzeugern unterstützt werden. Insbesondere lokale Energieerzeugung spart zusätzlich Kosten. Die bestehende Privilegierung für stromintensive Betriebe sollte gezielt für nachhaltige Rechenzentren erweitert werden.

Rechenzentren sollten sich im Stromnetz als Flexibilität einbringen (Batteriespeicher, Energiemanagement, Dynamische Strompreise). Die effektive Abwärmenutzung und Kühlung sind Kriterien für die Standortauswahl, die durch hochleistungsfähige Glasfaseranbindung gestützt werden muss. Vorhandene Standorte der HPC Infrastruktur (GCS, NHR Zentren) sollten ausgebaut werden (Personal und Administration mit Erfahrung bereits vorhanden).

Ökosystem

Heute bestehen die relevanten Sektoren Infrastruktur, Forschung und Industrie noch weitgehend isoliert, obwohl nur durch ihre Integration die Entwicklung von wettbewerbsfähigen Large Industry Models – das bedeutet branchen- und unternehmensspezifische KI-Modelle im Sinne des IPCEI AI, die Aufgaben von hoher Komplexität lösen können – möglich wird.

Fachkräfte

Fachkräfte sind Mangelware, so auch für den Rechenzentrumsstandort Deutschland. Dies gilt es politisch anzupacken und die richtigen Weichen zu stellen. Dazu gehören die Weiterqualifizierungen für den Bereich Rechenzentren. Zudem benötigt diese Branche neben Fachkräften mit Programmierkenntnissen auch Arbeitskräfte, die physische Arbeit vor Ort fachgerecht leisten können. MINT-Fächer[4] in Schule und Studium sowie duale Studiengänge sollten weiterhin gefördert werden. Wichtig sind außerdem attraktive Rahmenbedingungen für Fachkräfte im Bereich Rechenzentrum, HPC, Cloud- und KI-Infrastruktur. Infrastruktur wird derzeit nicht als Expertise gesehen – IT weder als Wissenschaft noch als hochtechnisch. Problematisch ist außerdem, dass die Gehälter deutlich unter denen in den USA liegen und deshalb oft nicht wettbewerbsfähig sind (Stichwort: Besserstellungsverbot).

[4] MINT: Mathematik, Informatik, Naturwissenschaften und Technik

3. Welche Rahmenbedingungen sollten aus Ihrer Sicht wie verändert werden, um Rechenzentrumsinvestitionen zu fördern und Innovation zu ermöglichen?

In Deutschland gibt es viele hemmende Faktoren, die sich negativ auf den Bau von und die Investition in Rechenzentren auswirken. Besonders gilt es aus Sicht der Fraunhofer-Gesellschaft folgende Punkte politisch anzugehen:

Konsequente Ausrichtung an Innovationen

Um Rechenzentrumsinvestitionen zu fördern und Innovation zu ermöglichen, empfehlen wir die Etablierung eines Expert*innen-Gremiums. Dieses sollte beim Bau von staatlich geförderten Rechenzentren beratend sicherstellen, dass die mit der Investition verknüpften Investitionsversprechen von Industrie, F&E-Einrichtungen und Infrastruktur eingehalten werden. Im Mittelpunkt der Entscheidungsgrundlage sollten marktwirtschaftliche Zahlen (KPIs) stehen. Auf diese Weise soll vermieden werden, dass bspw. eine Gigafactory ausschließlich für Inferenz oder andere Ziele wie 3-D-Simulationen verwendet wird. Es ist wichtig, dass das Gremium unabhängig besetzt wird, u. a. mit Expert*innen aus Wissenschaft und Industrie, die sich persönlich durch Innovation ausgezeichnet haben.

Finanzielle und regulatorische Anreize

Förderlich für den Bau von Rechenzentren sind finanzielle und regulatorische Anreize. Es bedarf staatlicher Mischfinanzierungsmodelle (Public-Private-Partnerships), Steuervergünstigungen und staatlicher Garantien, um Investitionsrisiken abzufedern. Auch können Vergünstigungen für Open Source-Modelle und KI-Datensätze Anreize setzen.

Darüber hinaus sollten Anreize zur Nutzung europäischer Lösungen geschaffen werden, insbesondere im öffentlichen Sektor. Dies schafft Marktanreize für unternehmerisches Risiko in diesem Bereich, ohne die es nie gleichwertige Konkurrenz geben wird. Denkbar wäre eine Zertifizierung von Unternehmen, ein Gütesiegel analog zu »lokalen Inhaltsstoffen«.

Bürokratieabbau und schnelle Netzanschlüsse

Eine zentrale Aufgabe wird es sein, dass staatliche Genehmigungsprozesse schneller werden müssen. Deutschland gehört aktuell zu den Ländern mit den längsten Netzanschlusszeiten für Rechenzentren – bis zu sieben Jahre. Das ist mit einem zukunftsfähigen Standort nicht vereinbar. Proaktive Bauleitplanung durch Kommunen sollten erlaubt und gefördert werden (analog zu Industrie: Kommunen haben Interesse und sind abhängig von Steuereinnahmen). Vorab-Eignungsprüfungen und die Ausweisung spezieller Baugebiete für Rechenzentren inklusive Planung/Eignungsprüfung, Stromversorgung sowie redundantem Anschluss Glasfaser an Backbones beschleunigen Bauprozesse drastisch. Darüber hinaus sollte die rechtliche Lage bzgl. Immissions- und anderen Umweltschutzauflagen vorab zugesichert werden.

Grüne Rechenzentrumsinfrastruktur

Rechenzentren müssen dort entstehen, wo ausreichend erneuerbare Energien verfügbar sind. Hier sind Standortentscheidungen an die Energie- und Klimapolitik zu koppeln und neue Geschäftsmodelle zu unterstützen, welche die Energieeffizienz von Rechenzentren verbessern (Stichwort: Green Computing), eine Flexibilisierung der Rechenressourcen mit den markt- und netzdienlichen Mechanismen des Strommarkts verbinden und eine Nutzung der Abwärme auch wirtschaftlich attraktiv machen. Ebenso ist für die Vorrangplanung der Wärmenetze zu sorgen, ohne die eine nachhaltige Betriebsweise nicht möglich sein wird.

4. Welche Rolle sollte der Staat bei der Entwicklung einer souveränen und resilienten Recheninfrastruktur einnehmen?

Um eine souveräne und resiliente Recheninfrastruktur in Deutschland zu fördern, muss der Staat aus Sicht der Fraunhofer-Gesellschaft folgende Rolle einnehmen:

Innovationsermöglicher

Wichtig ist, dass zunächst verlässliche und innovationsfördernde Rahmenbedingungen hergestellt werden, die den Aufbau und Betrieb von Rechenzentrum für kommerzielle Anbieter attraktiv machen. Den Aufbau der Infrastruktur (Strom- und Glasfasernetze) muss der Staat ständig vorantreiben und zeitnah umsetzen.

Für die Entwicklung einer souveränen und resilienten Recheninfrastruktur ist die Unterstützung des Staates unerlässlich. Die Bereitstellung von Mitteln für strategische Rechenzentrumsinfrastruktur, bspw. aus Sondervermögen, kann entscheidend zur Wettbewerbsfähigkeit des Standorts beitragen. Eine innovative Möglichkeit wäre, dass Länder oder der Bund Rechenzentrums-Gebäude inklusive Internetverbindung und Stromversorgung bereitstellen, welche von Forschungseinrichtungen und anderen öffentlichen Organisationen genutzt werden können. Dies würde die Einstiegshürden für diese Organisationen senken und gleichzeitig eine effiziente Nutzung der Ressourcen ermöglichen.

Der Staat sollte bei der Entwicklung einer souveränen und resilienten Recheninfrastruktur weiterhin folgende Funktionen ausfüllen:

1. Etablierer eines Ökosystems zwischen Industrie, Forschung und Rechenzentrumsbetreibern, um eine verzahnte Entwicklung von Innovation aus Forschung und Industrie (z.B. wettbewerbsfähige Basismodelle, KI-Datensätze) und Infrastruktur eng abzustimmen.
2. Ankerkunde, indem er Nachfrage aus Verwaltung, Forschung und öffentlicher Hand bündelt und so Planungssicherheit schafft.
3. Förderer von Open Source, da nur offene Modelle eine breite Innovationsbasis für Wissenschaft und Mittelstand ermöglichen.

5. Gibt es konkrete Maßnahmen oder Best Practices aus Ihrer Praxis/Erfahrung, die in die Strategie aufgenommen werden sollten?

Gigafactory-Initiative als Best Practices etablieren

AI Gigafactories bieten Deutschland die Chance, die Grundlage für digitale Souveränität zu schaffen. Sie sind ein entscheidender Schritt, um die Abhängigkeit von außereuropäischen Anbietern zu reduzieren und den Industriestandort durch die schnelle Umsetzung von KI-Anwendungen zu stärken. Darüber hinaus fördern sie das Ökosystem, indem sie Start-ups, Forschung und Unternehmen Zugang zu modernster Infrastruktur und Rechenleistung ermöglichen, Innovationen beschleunigen und Talente anziehen. Perspektivisch kann mit Gigafactories die Entwicklung eigener Hardware-Kompetenzen vorangetrieben werden, wodurch die Grundlage für eine unabhängige und wettbewerbsfähige europäische KI-Industrie geschaffen wird. Mit einer AI Gigafactory setzt Deutschland zudem ein Signal für seinen technologischen Führungsanspruch und gestaltet damit aktiv den Weg hin zu europäischer Technologie und offenen Standards.

GPU-Cluster für KI-Forschung

Um die KI-Innovationskraft zu stärken, ist der Aufbau dedizierter GPU-Cluster für KI-Forschung erforderlich. Diese sollten den Bedarf mehrerer Forschungseinrichtungen decken, wodurch sie Synergien schaffen. Forschungseinrichtungen mit zahlreichen dezentralen Rechenzentren können durch Konsolidierung zu wenigen überregionalen Standorten erhebliche Effizienzgewinne erzielen. Der Aufbau mehrerer überregionaler Rechenzentren spart Kosten und steigert auch die Leistungsfähigkeit und Energieeffizienz.

Swiss AI Initiative

Als Best Practice stellt die Swiss AI Initiative dar, wie staatlich unterstützte, offene Infrastruktur zu einem Innovationshub werden kann. Dies ist ein Gemeinschaftsprojekt von ETH Zürich und EPFL im Rahmen des Swiss National AI Institute (SNAI) und verbindet Spitzenforschung mit nationaler Infrastruktur: Auf dem CSCS-Supercomputer Alps werden offene Foundation-Modelle entwickelt und veröffentlicht. Über eine strategische Partnerschaft mit Swisscom (SNAI-Mitglied) wird die Brücke in die Industrie geschlagen – inklusive Telco-Plattform-Know-how und sicheren Betriebsumgebungen.

Verzeichnis der Mitwirkenden

Dr. Michael Andres
Fraunhofer-Institut für Angewandte Informationstechnik FIT

Dr. Nicolas Flores-Herr
Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS

Michael Fritz
Fraunhofer Cluster of Excellence Cognitive Internet Technologies CCIT

Dr. Gerrit Füldner
Fraunhofer-Institut für Solare Energiesysteme ISE

Norman Gerhardt
Fraunhofer-Institut für Energiewirtschaft und Energiesystemtechnik IEE

André Gemünd
Fraunhofer-Institut für Algorithmen und Wissenschaftliches Rechnen SCAI

Dr. Hubert Krammer
Zentrale der Fraunhofer-Gesellschaft

Dr. Jens Krüger
Fraunhofer-Institut für Techno- und Wirtschaftsmathematik ITWM

Dr. Juliane Lutz
Zentrale der Fraunhofer-Gesellschaft

Dr. Elisabeth Peinsipp-Byma
Zentrale der Fraunhofer-Gesellschaft

Heinrich Pettenpohl
Fraunhofer-Institut für Software- und Systemtechnik ISST

Pierre Prasuhn
Zentrale der Fraunhofer-Gesellschaft

Prof. Dr. Anita Schöbel
Fraunhofer-Institut für Techno- und Wirtschaftsmathematik ITWM

Prof. Dr. Marc Alexander Schweitzer
Fraunhofer-Institut für Algorithmen und Wissenschaftliches Rechnen SCAI

Prof. Dr. Ingo Weber
Zentrale der Fraunhofer-Gesellschaft

Prof. Dr.-Ing. Christof Wittwer
Fraunhofer-Institut für Solare Energiesysteme ISE

Über die Fraunhofer-Gesellschaft

Die Fraunhofer-Gesellschaft mit Sitz in Deutschland ist eine der führenden Organisationen für anwendungsorientierte Forschung. Im Innovations-prozess spielt sie eine zentrale Rolle – mit Forschungsschwerpunkten in zukunftsrelevanten Schlüsseltechnologien und dem Transfer von Forschungsergebnissen in die Industrie zur Stärkung unseres Wirtschafts-standorts und zum Wohle unserer Gesellschaft. Seit ihrer Gründung als gemeinnütziger Verein im Jahr 1949 nimmt sie eine einzigartige Position im Wissenschafts- und Innovationssystem ein.

Knapp 32 000 Mitarbeitende an 75 Instituten und selbstständigen Forschungseinrichtungen in Deutschland erarbeiten das jährliche Finanz-volumen von 3,6 Mrd. €. Davon entfallen 3,1 Mrd. € auf das zentrale Geschäftsmodell von Fraunhofer, die Vertragsforschung. Im Vergleich zu anderen öffentlichen Forschungseinrichtungen bildet die Grundfinanzierung durch Bund und Länder lediglich das Fundament des jährlichen Forschungs-haushalts. Sie ist die Basis für wegweisende Vorlaufforschung, die in den kommenden Jahren für Wirtschaft und Gesellschaft bedeutend wird. Das entscheidende Alleinstellungsmerkmal ist der hohe Anteil an Wirtschafts-erträgen, der Garant ist für die enge Zusammenarbeit mit Wirtschaft und Industrie und die stetige Marktorientierung der Fraunhofer-Forschung: 2024 beliefen sich die Wirtschaftserträge auf 867 Mio. € des laufenden Haushalts. Ergänzt wird das Forschungsportfolio durch im Wettbewerb eingeworbene öffentliche Projektmittel, wobei eine ausgewogene Balance zwischen öffentlichen und wirtschaftlichen Erträgen angestrebt wird.

Kontakt

Herausgeber
Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e. V.
Hansastraße 27 c, 80686 München
https://www.fraunhofer.de

Ansprechpersonen
Pierre Prasuhn
Referent Wissenschaftspolitik
Abteilung Wissenschaftspolitik
Tel. +49 30 688 3759-1607
pierre.prasuhn@zv.fraunhofer.de

Prof. Dr. Ingo Weber
Direktor Digitalisierung und IuK-Infrastruktur
Tel. +49 89 1205 3050
ingo.weber@zv.fraunhofer.de

© Fraunhofer-Gesellschaft e. V., München 2025