Universität Stuttgart
IER Institut für Energiewirtschaft und Rationelle Energieanwendung

BECKER BÜTTNER HELD

Fraunhofer ISI

heimrich+hannot

BORDERSTEP INSTITUT
für Innovation und Nachhaltigkeit
Wissen, das bewegt

dena
Deutsche Energie-Agentur

SDIA

EY

Stand und Entwicklung des Rechenzentrumsstandorts Deutschland

Gutachten im Auftrag des Bundesministeriums für Wirtschaft und Klimaschutz

BMWK-Projekt-Nr.: 115/21-45

Berlin, den 17.01.2025

Impressum

Hauptbearbeitung
Dr. Zhanat Murzakulova, dena, Teamleiterin Energieeffizienz

Hauptbearbeiter der Kapitel
Kapitel I – Borderstep Institut
Kapitel II – Borderstep Institut
Kapitel III – SDIA
Kapitel IV – EY Law und FH ISI
Kapitel V – SDIA und IER
Kapitel VI – dena

Bearbeitende
Deniz Öztürk, dena, Seniorexperte Energieeffizienz
Rafael Noster, dena, Seniorexperte Energieeffizienz
Carla Groß, dena, Seniorexpertin Energieeffizienz
Dr. Lars Sorge, dena, Experte Energieeffizienz
Svitlana Chebotarova, dena, Expertin Energieeffizienz
Simon Kieweg, dena, Experte Digitale Technologien
Max Schulze, SDIA, Geschäftsführender Vorsitzender
Jessica Stüer, SDIA, Projekt-Leitung/Koordinierung
Dr. Ralph Hintemann, Borderstep Institut, Senior Researcher
Simon Hinterholzer, Borderstep Institut, Researcher
Dr. Severin Beucker, Borderstep Institut, Senior Researcher
Prof. Dr. Peter Radgen, IER, Universität Stuttgart, Professor für Effiziente Energienutzung
Benjamin Ott, IER, Universität Stuttgart, Wissenschaftlicher Mitarbeiter Effiziente Energienutzung
Dr. Heike Brugger, Fraunhofer ISI, Leiterin des Geschäftsfelds Energiepolitik
Dr. Markus Fritz, Fraunhofer ISI, Wissenschaftlicher Mitarbeiter
Dr. Tim Mandel, Fraunhofer ISI, Wissenschaftlicher Mitarbeiter
Christine Hohenstein-Bartholl, EY-Law, Rechtsanwältin
Julia Fritz, EY-Law, Rechtsanwältin
Philip Debray, EY-Law, Rechtsanwalt

Review
Steffen Joest, dena, Bereichsleiter Industrie, Mobilität, Energieeffizienz

Formaler Review
Dr. Markus Kachel, bbh, Rechtsanwalt (Kapitel 4.2)
Dr. Carolin Schmidt, bbh, Rechtsanwältin (Kapitel 4.2)

Auftragsvergabe durch:
Dr. Ron Lipka, BMWK, Referat IIA6

Inhalt

Impressum **2**

Inhalt **3**

Executive Summary **5**

Handlungsoptionen **14**

Einleitung **17**

1 Entwicklungen und Trends im Rechenzentrumsmarkt **22**

1.1 Zusammenfassung 22

1.2 Weltweit 23

1.3 Deutschland 28

1.4 Europäische und internationale Einordnung der deutschen Marktentwicklungen und -trends 33

1.5 Handlungsempfehlungen 36

2 Energiesystem und Rechenzentren **38**

2.1 Zusammenfassung 38

2.2 Energieverbrauch, Energieeffizienz und Klimawirkungen von Rechenzentren 39

2.3 Rechenzentren und das Stromnetz 47

2.4 Rechenzentren und die Abwärmenutzung 59

2.5 Handlungsempfehlungen 66

3 Ökonomie der Rechenzentren **69**

3.1 Zusammenfassung 69

3.2 Wirtschaftliche Bedeutung von digitaler Infrastruktur in Deutschland 69

3.3 Digitale Infrastruktur veredelt Energie und Rohstoff 72

3.4 Zentrale Geschäftsmodelle 74

3.5 Volkswirtschaftliche Effekte 78

3.6 Mechanismen der Spillover-Effekte ... 83
3.7 Beschränkter Marktzugang als Hemmnis ... 85
3.8 KI als Treiber der Nachfrage ... 88
3.9 Fachkräftemangel als Hemmnis ... 88
3.10 Handlungsempfehlungen ... 89
4 Rechenzentren und Regulatorik ... 94
4.1 Zusammenfassung ... 94
4.2 Dimensionen politisch-rechtlicher Rahmenbedingungen ... 95
4.3 Erörterung der Kennzahlen zum Energie- und Ressourceneinsatz ... 105
4.4 Internationale gute Beispiele politisch-rechtlicher Ansätze ... 111
4.5 Handlungsempfehlungen ... 116
5 Blick in die Zukunft der Rechenzentren in Deutschland ... 118
5.1 Zusammenfassung ... 118
5.2 Digitale Dimensionen bei Wirtschaftswachstum und Energiebedarf ... 118
5.3 Digitale Souveränität ... 125
5.4 Technologische Entwicklung und Trends ... 133
5.5 Standortkriterien für Deutschland ... 137
5.6 Handlungsempfehlungen ... 138
6 Übergreifende Handlungsempfehlungen ... 142
6.1 Deutschland braucht eine Rechenzentrumsstrategie ... 142
Abbildungsverzeichnis ... 151
Literatur ... 155

Executive Summary

Mit dem Wachstum der Digitalwirtschaft und der Digitalisierung einher geht ein starker Anstieg des Bedarfs an digitaler Infrastruktur und Rechenzentren (RZ). Rechenzentren haben sich damit zu einer der am schnellsten wachsenden, energieintensiven Branchen in Deutschland entwickelt. Mit mehr als 2.000 Rechenzentren und einer IT-Anschlussleistung von über 2.700 MW ist Deutschland bereits der größte Standort für digitale Infrastruktur in Europa. Die zentrale Lage, politische Stabilität, eine überdurchschnittlich zuverlässige Stromversorgung und einer der größten Internetknoten der Welt machen Deutschland zu einem strategisch vorteilhaften Standort für nationale und internationale Anbieter. Digitalisierungsvorhaben, eine wachsende Digitalwirtschaft und neue Technologien wie zum Beispiel „Künstliche Intelligenz“, sorgen für eine steigende Nachfrage an Rechenleistung, Speicherkapazität und Datenübertragung. Diese Nachfrage lässt sich allein durch Effizienzsteigerungen bei IKT-Equipment und Rechenzentren nicht decken. Um die Nachfrage zu decken, können neue Rechenzentren gebaut, der Nutzungsgrad bestehender Rechenzentren erhöht, oder Rechenkapazität aus dem Ausland bezogen werden. Der Begriff „Rechenzentren“ vereint in dem hier zugrundeliegenden Verständnis sowohl die Gebäude und Gebäudetechnik als auch IT-Infrastruktur und Internetzugang. Neben großen Rechenzentren, die oft von Dienstleistern betrieben werden, gibt es in Deutschland auch ca. 50.000 kleinere IT-Installationen, die insbesondere von mittelständischen Unternehmen betrieben werden.

Seit 2010 hat sich die RZ-Kapazität in Deutschland auf über 2.730 MW im Jahr 2024 mehr als verdoppelt, bis 2030 wird ein weiter beschleunigtes Wachstum auf über 4.800 MW erwartet.

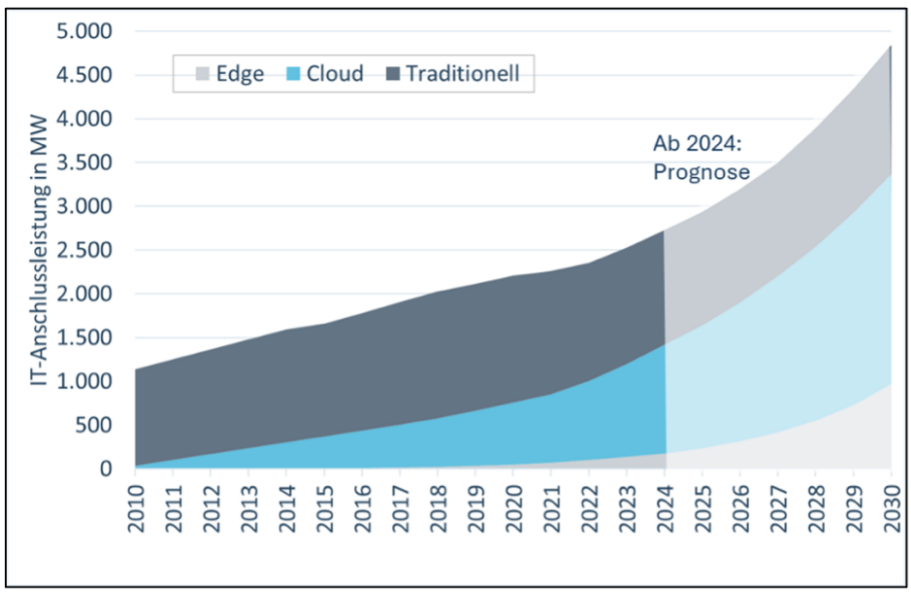


Abbildung 1: Entwicklung der Kapazitäten von Rechenzentren und kleineren IT-Installationen in Deutschland in den Jahren 2010 bis 2024 und Prognose bis 2030. Quelle: Hintemann et al. (2024a)

Der steigende Bedarf an Rechenleistung, Speicherkapazität und Datenübertragung führt zu einem kontinuierlichen Wachstum der Rechenzentrumskapazitäten in Deutschland, das sich in den letzten Jahren sogar weiter beschleunigt hat. Seit 2010 haben sich die Kapazitäten von Rechenzentren und kleineren IT-Installationen in Deutschland auf über 2.730 MW im Jahr 2024 mehr als verdoppelt. Bis zum Jahr 2030 wird mit einem deutlichen weiteren Anstieg der IT-Kapazitäten auf 4.850 MW gerechnet (Bitkom 2024, ähnlich GDA 2024). Setzt sich dieses Wachstum auch danach weiter fort, so werden die Kapazitäten eine Größenordnung von mehr als 12.000 MW im Jahr 2045 erreichen.

Das Wachstum der Kapazitäten findet bei den Rechenzentren mit mehr als 100 kW Anschlussleistung statt. Die Kapazitäten der kleinen IT-Installationen mit unter 100 kW Leistung bleiben dagegen seit Jahren weitgehend konstant. Vor allem die Kapazitäten der Großrechenzentren mit mehr als 5 MW IT-Anschlussleistung wachsen sehr deutlich. Zwischen 2010 und 2024 haben sich die Kapazitäten in diesem Segment fast um den Faktor fünf von 270 MW auf 1.300 MW erhöht. Aktuell gibt es etwa 100 solcher Rechenzentren in Deutschland. Zudem ist bei angekündigten Investitionen ein Trend zu immer größeren Rechenzentren zu beobachten, mit über 100 MW Leistung und teilweise mehreren 100 MW Leistung an einem Standort. Aber auch bei den kleineren Rechenzentren mit IT-Anschlussleistungen zwischen 100 kW und 5 MW ist ein hohes Wachstum zu verzeichnen – zwischen 2010 und 2020 haben sich die Kapazitäten in dieser Größenklasse um den Faktor drei von 210 MW auf 670 MW erhöht.

Weiterhin besteht eine hohe Konzentration von Rechenzentren in Frankfurt/Rhein-Main. Die Hauptstadtregion und das Rheinische Revier entwickeln sich als neue Ausbaugebiete.

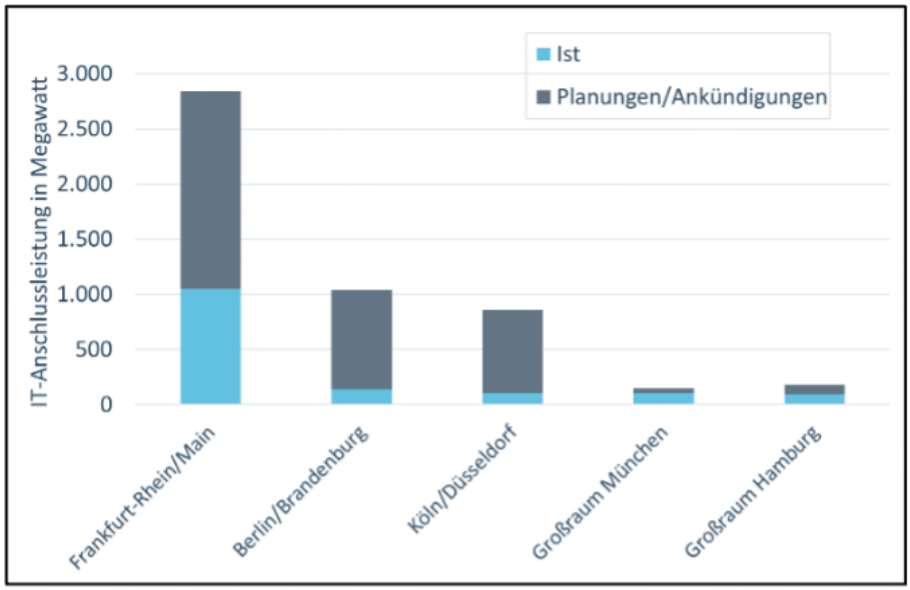


Abbildung 2: Aktuelle und in konkreter Planung befindliche IT-Anschlussleistung von Rechenzentren in verschiedenen deutschen Regionen. Quelle: Hintemann et al. (2024a)

Das Wachstum der Rechenzentrumsbranche fokussiert sich auf wenige Regionen in Deutschland. Vor allem der Großraum Frankfurt/Rhein-Main mit dem Netzwerkknoten DE-CIX ist für die Ansiedlung von Rechenzentren attraktiv. Aktuell werden auch in der Hauptstadtregion mehrere neue Groß-Rechenzentren geplant. Das Milliarden-Engagement von Microsoft im Rheinischen Revier und der damit verbundene Bau mehrerer Groß-Rechenzentren lassen vermuten, dass auch in dieser Region

in Zukunft die Rechenzentrumskapazitäten deutlich wachsen werden, da bestehende Rechenzentren die Ansiedlung weiterer Rechenzentren anziehen. Die räumliche Konzentration von Rechenzentren und Datennetzen kann mit dem Phänomen der „data gravity“ erklärt werden. Damit ist gemeint, dass große Datenmengen eine „Anziehungskraft“ entwickeln, die weitere Daten anzieht. Begründet wird das Phänomen beispielsweise damit, dass sich große Datenmengen nur noch mit hohem Aufwand verschieben lassen und viele Anwendungen, die auf die Daten zugreifen, geringe Latenzen erfordern. Verstärkt wird die Konzentration von Großrechenzentren in bestimmten Regionen zusätzlich dadurch, dass große Cloud-Anbieter aus Gründen der Verfügbarkeit meist drei oder mehr Rechenzentren in einer Region aufbauen. Sollte ein Rechenzentrum ausfallen, kann der Cloud-Dienst in der Region weiterhin aus den anderen Rechenzentren angeboten werden.

Deutschland ist weiterhin größter Rechenzentrumsstandort in Europa. China und USA beschleunigen aber den Ausbau von digitaler Infrastruktur im eigenen Land.

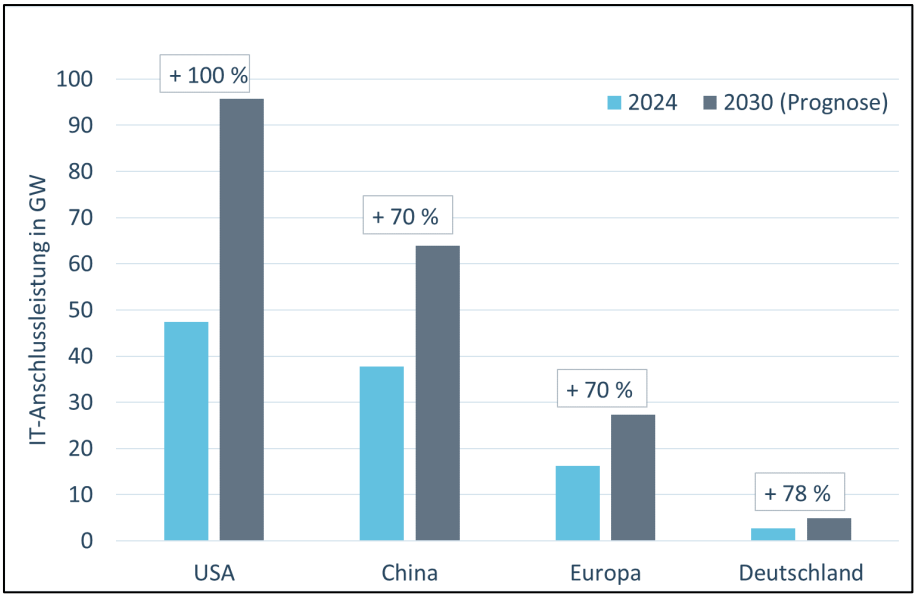


Abbildung 3: Abschätzung der Entwicklung der RZ-Kapazitäten in den USA, in China, Europa und Deutschland 2024 und 2030. Quelle: Hintemann et al. (2024a)

Gemessen an den Kapazitäten aller Rechenzentren und kleineren IT-Installationen ist Deutschland der größte Rechenzentrumsstandort in Europa. Dies ist vor allem auf die starke Wirtschaftskraft Deutschlands zurückzuführen. Insbesondere im Vergleich zu Großbritannien konnte diese Position in den letzten Jahren etwas verbessert werden. Vergleicht man allerdings die Entwicklung der Rechenzentrumskapazitäten in Europa mit anderen Weltregionen, so fallen Deutschland und Europa voraussichtlich in Zukunft weiter zurück (Abbildung 3). Bereits in der Vergangenheit hat der Anteil Deutschlands an den weltweiten IT-Kapazitäten abgenommen. Zwischen 2015 und 2024 sank der Anteil der in den Rechenzentren in Deutschland installierten Server am Weltmarkt von 3,5 % auf knapp 2,5 % (Bitkom 2024).

Ein wesentlicher Grund für diesen Rückgang ist, dass Deutschland hinsichtlich der Nutzung von Cloud-Diensten ein Importland ist. Mit dem Trend zu immer mehr Anwendungen im Bereich der Künstlichen Intelligenz (KI) scheint sich das Wachstum der Rechenzentrumskapazitäten insbesondere in den USA und in China noch einmal deutlich zu beschleunigen.

Die Energieeffizienz der Rechenzentrumsinfrastruktur und Gebäudeausrüstung stagniert derzeit auf globaler Ebene.

In Deutschland konnte über die letzten 10 Jahre ein tendenziell positiver Trend mit einer Reduktion des Kennwerts „PUE" (Power Usage Effectiveness) auf aktuell durchschnittlich 1,46 festgestellt werden. Weltweit stagniert seit 2018 die Effizienz bei der technischen Gebäudeausrüstung von Rechenzentren. Der PUE-Wert liegt seit 2018 im Bereich von 1,55 bis 1,59. Das bedeutet, dass für 1 kW Rechenleistung ca. 0,55 bis 0,59 kW für die Kühl- und Stromversorgung sowie weitere Gebäudetechnik benötigt werden.

In der Vergangenheit konnten gerade durch relativ einfach umsetzbare Maßnahmen, wie eine Einhausung des Kalt- oder Warmganges, eine Anhebung der Raumtemperatur, oder durch den Einsatz von Verdunstungskühlung, der Energiebedarf gerade für die Kühlung deutlich reduziert werden. Um hier den PUE aber noch weiter reduzieren zu können, sind weitere Maßnahmen, wie z. B. eine Flüssigkeitskühlung, erforderlich. Weitere Einflussfaktoren auf den PUE sind die klimatische Lage (je niedriger die Außentemperatur, desto höher das Potenzial für freie Kühlung) oder auch die installierte IT-Kapazität (je größer, desto effizienter) sowie der vergangene Zeitraum seit der Inbetriebnahme (je neuer, desto moderner und energieeffizienter).

Führende Unternehmen im Markt bauen Rechenzentren mit PUE-Werten im Bereich von 1,2 bis 1,3 – auch in Deutschland. Dies zeigt, dass die vorhandenen Potenziale der Gebäudetechnik bislang nicht vollständig genutzt werden. Hier muss aber zwischen den Möglichkeiten im Bestand und Neubau differenziert werden. Mit dem Energieeffizienzgesetz (EnEfG) wurde in Deutschland für neue RZ, die ab Juli 2026 in Betrieb gehen, ein PUE von 1,2 als Grenzwert festgelegt, der zu einer deutlichen Absenkung der durchschnittlichen PUE-Werte führen wird. Dies gilt aber nur unter der Annahme, dass weiterhin neue Rechenzentren in Betrieb gehen, die gleichzeitig über eine hohe Leistungskapazität verfügen. Auch in China wurde für neue Rechenzentren ein PUE von 1,3 in gemäßigten und 1.25 in kühleren Regionen als gesetzliche Vorgabe ab dem Jahr 2025 eingeführt[1] (Xie et al., 2024).

Die Entwicklung der IT-Infrastruktur selbst (Server, Storage und Netzwerktechnik) ist vor allem geprägt durch die steigende Dichte von integrierten Schaltkreisen. Durch stetige Miniaturisierung, welche unter anderem durch moderne Lithographieverfahren ermöglicht wird, stellen Chips pro Watt Leistungsaufnahme weiterhin steigende Rechenperformance zur Verfügung. In einigen Segmenten der Chipherstellung ist bereits eine Verlangsamung der Miniaturisierung zu erkennen, welche sich auch mit den schon lange diskutierten physikalischen Grenzen des Gesetzes von Moore begründen lässt. Spezielle Rechenkerne (z. B. Tensor Cores) für spezielle Anwendungen, wie beispielsweise maschinelles Lernen, haben jedoch auch in den letzten Jahren noch große Performancesprünge ermöglicht. Dass die Performancegewinne bei vielen Chips nicht nur durch Miniaturisierung erreicht werden, ist daran erkennbar, dass die elektrische Leistungsaufnahme von Chips in den letzten Jahren massiv angestiegen ist und auch in Zukunft ein weiterer Anstieg zu erwarten ist – dies hat natürlich Implikationen für den Gesamtenergieverbrauch der Rechenzentren. Aber auch die Gestaltung der Kühlinfrastruktur ist betroffen, wo enorme Leistungsdichten ein Absenken der Kühltemperaturen erfordern oder ggf. eine Flüssigkeitskühlung unverzichtbar wird.

[1] siehe auch https://www.nature.com/articles/s41599-024-02963-0

In Nachfrage-Szenarien ist mit einem starken Anstieg des Energieverbrauchs von Rechenzentren zu rechnen – in Deutschland und weltweit.

Der Ausbau der Rechenzentrums-Infrastrukturen ist weltweit mit einem deutlich steigenden Bedarf an Strom verbunden. Trotz neuer Technologien und Innovationen in der Branche (bspw. neue Kühl-Systeme sowohl auf Gebäude- als auch auf Serverebene), die zur Senkung des Strombedarfs führen können, ist auch in Zukunft mit einem weiterhin deutlich steigenden Strombedarf für digitale Infrastruktur zu rechnen. Die steigende Größe der Rechenzentren – bei neuen Projekten jetzt teilweise mehrere hundert MW – führt gleichzeitig zu einem zunehmenden Bedarf an größeren Stromanschlüssen, teilweise sogar direkt auf Ebene der Übertragungsnetze.

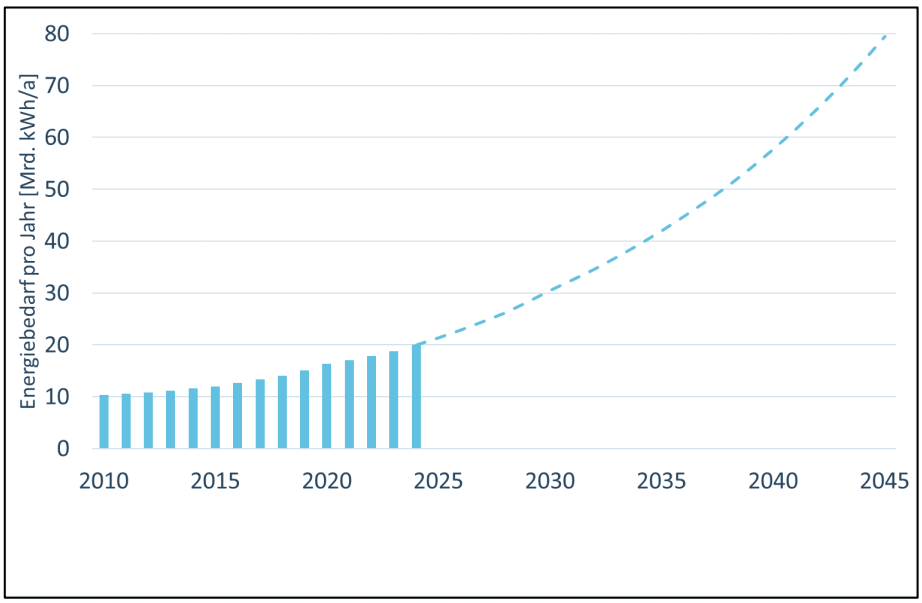


Abbildung 4: Entwicklung des Strombedarfs der Rechenzentren und kleineren IT-Installationen in Deutschland in den Jahren 2010 bis 2024 und Langfristprognose bis 2045. Quelle: Berechnungen Borderstep

In Deutschland sind deutlich steigende Rechenzentrumskapazitäten zu erwarten und damit, dass auch der Strombedarf weiter ansteigen wird. Im Jahr 2024 haben Rechenzentren und kleinere IT-Installationen in Deutschland einen Strombedarf von 20 TWh und damit ca. 4 % des Bruttostromverbrauchs von 517 TWh (2023). Im Jahr 2030 werden voraussichtlich 31 TWh Strom für die in Deutschland betriebenen Rechenzentren benötigt. Ohne die Einführung des EnEfG wäre der Stromverbrauch im Jahr 2030 noch um 2,5 TWh höher. Bis zum Jahr 2045 kann der Stromverbrauch – geht man von einem gleichbleibenden Wachstum aus – auf knapp 80 TWh ansteigen und damit ca. 6 % des voraussichtlichen Bruttostromverbrauchs (gemäß Langfristszenario O45-Strom) ausmachen. Gegenüber dem Jahr 2024 ist das ein Anstieg um fast 400 %. Die stromverbrauchsbedingten CO_2-Emissionen betrugen im Jahr 2024 etwa 6,5 Mio. Tonnen und werden auf etwa 4,5 Mio. Tonnen im Jahr 2030 sinken, bedingt durch die Änderung des Strommixes in Deutschland.

Für die Nutzung von Abwärme aus Rechenzentren, z. B. zur Dekarbonisierung von Wärmenetzen, ergeben sich damit erhebliche Potenziale. Geht man davon aus, dass die neuen gesetzlichen Vorgaben des EnEfG von den neu errichteten Rechenzentren eingehalten werden, so könnte im Jahr 2030 etwa 1 TWh Wärme aus Rechenzentren beim Endverbraucher genutzt werden. Im Jahr 2035

ständen bereits mehr als 3 TWh Wärme aus Rechenzentren zur Verfügung. Bis zum Jahr 2045 würde sich die nutzbare Abwärmemenge auf etwa 10 TWh pro Jahr erhöhen.

Aus der Perspektive des Stromnetzes in Deutschland sollten Rechenzentren gezielt dort angesiedelt werden, wo eine hohe Erzeugungskapazität von erneuerbaren Energien besteht.

Rechenzentren wurden in der Vergangenheit meist an Standorten geplant, die sich an Internetknoten und Kundenbedürfnissen orientierten. Zunehmend spielt die Verfügbarkeit von grünem Strom eine wichtige Rolle bei der Standortwahl großer Rechenzentren. Neben dem Hotspot Frankfurt am Main werden aktuell Standorte wie Berlin/Brandenburg oder das Rheinische Revier immer wichtiger, aber auch nördlichere Bundesländer sind zunehmend interessant für Investoren und Betreiber.

Diese Entwicklung ist aus der Perspektive des Stromnetzes zu begrüßen und sollte möglichst unterstützt werden. Sobald Stromerzeugungskapazitäten und Stromsenken näher beieinander sind, kann das umliegende Übertragungsnetz entlastet werden. Aktuell stellt vor allem der Überschuss an Windstrom in Norddeutschland und der Energieverbrauchsschwerpunkt im Großraum Frankfurt am Main eine deutliche Belastung des Stromnetzes dar. Über Flexibilisierung von Rechenzentren, zum Beispiel durch die Verwendung von Notstromaggregaten, Batterie-Systemen und anderen Pufferspeichertechnologien, kann der Mehrwert von Rechenzentren für den Strommarkt und das Stromnetz in Deutschland und über Abwärmeabgabe in lokale Fernwärmenetze der Mehrwert für eine Region erhöht werden. Hierzu müssen aber gleichzeitig die rechtlichen Voraussetzungen erfüllt sein.

Rechenzentren leisten insbesondere im Zusammenspiel mit Digitalwirtschaft und IT-Dienstleistungen einen relevanten volkswirtschaftlichen Beitrag.

Ein großer Teil der Wertschöpfung, die aus der RZ-Leistung geschaffen wird, entsteht in den Dienstleistungen, z. B. durch IT-Beratungen, Managed Service Anbieter (MSP) und Cloud-Dienste. Einem Arbeitsplatz für Wartung und Betrieb eines 1 MW Rechenzentrums stehen 20 Arbeitsplätze bei einem MSP gegenüber, der aus der entstehenden Rechenkapazität des Rechenzentrums Dienstleistungen anbietet.

Aus der Perspektive der Gewerbesteuer kann ein Rechenzentrum für eine kleine Region durchaus einen signifikanten Beitrag leisten (Beispiel Hattersheim am Main). Jedoch kann der Hauptsitz eines verteilten Co-Location- oder Cloud-Anbieters nicht über alle Rechenzentrumsstandorte verteilt werden, was der anteilsmäßigen Verteilung der Steuereinnahmen nach betroffenen Regionen und beanspruchter Landflächen und Strominfrastruktur im Wege steht. In einer neuen Studie zu „Spillover-Effekten“ von Rechenzentren (IW CONSULT, 2024) wird auf die dezentrale Unternehmensverteilung und Struktur in Deutschland verwiesen, die als Stärke betrachtet wird. Um die in der Studie beschrieben Effekte zu realisieren, reicht jedoch der Ausbau von Rechenzentren allein nicht aus, sondern es müssen Ökosysteme in der Region geschaffen werden, die aus IT-Bildungseinrichtungen, IT-Dienstleistern und Digitalunternehmen bestehen. Diese Organisationen können gemeinsam neue IT-Kompetenzen schaffen und mit lokalen Arbeitsplätzen die Rechenkapazität „veredeln“ und damit nachhaltig digitale Wertschöpfung in der Region entwickeln.

Ohne die Produktion von digitalen Produkten und IT-Dienstleistungen auch in die Region zu verlagern, wird lediglich Fläche und Strom in Rechenkapazität veredelt, was nur beschränkt Arbeitsplätze (ca. 6) pro MW RZ-Leistung schafft und meist geringe Steuereinnahmen bringt. Andere Spillover-Effekte, die zum Beispiel während des Baus entstehen, sind hauptsächlich temporär.

Schätzungen aus der Branche zeigen einen Fokus auf Produktivitätssteigerung, wie zum Beispiel, dass die durch den Einsatz von KI gesteigerte unternehmerische Produktivität sich in ein gesamtwirtschaftliches Produktivitätswachstum übersetzt, welches Wertschöpfungspotenziale in Höhe von 330 Milliarden Euro realisieren kann, sofern über 50 % der Unternehmen in Deutschland KI einsetzen (IW Consult, 2024).

Durch den Einsatz von KI entstehen gleichzeitig Kosten für Rechenleistung und IT-Dienstleistungen, durch die mindestens 30 - 50 % der geschätzten zusätzlichen Wertschöpfung an Dienstleister und IT-Infrastruktur-Anbieter übertragen werden. Hier bietet sich eine Chance kleinere, regionale Rechenzentren zusammen mit digitalen Kompetenz- und Dienstleistungszentren zu schaffen, die mittelständischen Unternehmen in der Region beim Einsatz der neuen Technologien helfen können. Regionale Anbieter sind volkswirtschaftlich relevant, da sie die unternehmerische Produktivität in der Region steigern, Arbeitsplätze und Kompetenzen schaffen. Als Alternative können Unternehmen auf IT- und Cloud-Dienstleistungen von internationalen Unternehmen setzen, was jedoch zu einem Verlust von Arbeitsplätzen und Know-how in der deutschen Wirtschaft führen kann. Der Fokus sollte auf der Entwicklung regionaler Ökosysteme liegen, die sowohl souverän als auch resilient und nachhaltig agieren können.

Die derzeitige Regulatorik für einen energieeffizienten Betrieb von Rechenzentren führt zu Rechtsunsicherheiten. Auch gibt es aktuell keine verbindlichen Vorgaben für eine örtliche bzw. überörtliche integrale und strukturierte Planung für geeignete RZ-Flächen.

Mit Blick auf die gebotene weitere Ansiedlung von Rechenzentren in Deutschland und die Wettbewerbsfähigkeit des Standorts Deutschland hat die derzeit bestehende Regulatorik für Rechenzentren, die u. a. Abwärme nutzen müssen, insbesondere wegen langer und wenig strukturierter Genehmigungsprozesse eine eher hemmende Wirkung. Dies gilt auch im Rahmen der Bauleitplanung, aufgrund einer nicht verbindlich vorgegebenen strukturierten und integralen Planung von geeigneten Flächen, welche zugleich die notwendige Anbindung insbesondere an Wärme- und Stromnetze umfasst. Hier wären im Rahmen von Genehmigungsprozessen die Einführung von verbindlichen Strukturen zwischen Genehmigungsbehörden und weiteren Fachbehörden sowie eine zeitliche Straffung durch Bearbeitungsfristen der im Baugenehmigungsprozess beteiligten Behörden begrüßenswert.

Auch die verbindliche Einführung einer integralen Planung und Ausweisung von geeigneten Flächen für die Ansiedlung von regionalen Rechenzentren, welche auch das Vorhandensein oder den Ausbau von Wärme- und Stromnetzen beinhaltet, sowie die verbindliche Vorgabe von Kommunikationsformaten wären zur Beschleunigung von Genehmigungsprozessen förderlich und würden zu entsprechender Planungssicherheit beitragen. Weiterhin werden Rechenzentrumsbetreibern hohe Aufwände in Bezug auf Meldepflichten von Energieeffizienzkennzahlen unterschiedlichen Umfangs und auf Basis unterschiedlicher Rechtsgrundlagen abverlangt. Auch könnten Verantwortlichkeiten zwischen Rechenzentrumsbetreibern und Wärmenetzbetreibern nach dem EnEfG besser voneinander abgegrenzt werden, um Unsicherheiten in der Rechtsanwendung zu vermeiden. Förderlich wären insoweit eine Straffung, Vereinheitlichung und Bündelung von Meldepflichten mit Nachhaltigkeitsbezug. Ebenfalls hilfreich wäre im Rahmen des EnEfG eine klare Abgrenzung von Verantwortlichkeiten und die Aufnahme von Legaldefinitionen und Regelbeispielen im Zusammenhang mit der Abwärmenutzung durch Rechenzentren.

Mit einer klaren, langfristigen Vision für die digitale Infrastruktur in Deutschland lässt sich die Nachfrage nach Rechenzentren weiter stimulieren.

Deutschland ist bereits heute ein starker und attraktiver Standort für nationale und internationale Anbieter von digitaler Infrastruktur. Dies umfasst Anbieter von Rechenzentrumsgebäuden, Cloud-Anbieter sowie spezialisierte IT-Infrastrukturanbieter, wie zum Beispiel für Hochleistungsrechenzentren. Auch als Markt für Unternehmen, die Dienstleistungen im Bereich Digitalisierung anbieten, aber auch für Technologieunternehmen in der Digitalwirtschaft, ist Deutschland attraktiv und bietet großes Wachstumspotenzial. Zentrale Gründe dafür sind:

- Großer Nachfragemarkt: Sowohl nach digitalen Ressourcen, insbesondere aufgrund der digitalen Transformation der Wirtschaft, als auch nach RZ-Kapazität innerhalb von Deutschland, welche durch Datenschutz- und Datensicherheitsinteressen erzeugt wird,
- Hohe Qualität der leitungsgebundenen Stromversorgung, der Gebäude- und der elektrischen Infrastruktur; technisches Know-how und DE-CIX als internationaler Internetknoten,
- Politische Stabilität,
- Bildungssystem für Elektrotechnik und IT, welches Fachkräfte für den Betrieb von Rechenzentrumsgebäuden und für die Digitalbranche liefert.

Um das Wachstumspotenzial zu realisieren, sollte Deutschland jedoch seine Strategie für Digitalisierung und Digitalwirtschaft und die damit verbundene digitale Infrastruktur mit einer langfristigen Perspektive formulieren und die konkreten Wachstumschancen fördern. Deutschland braucht eine **Rechenzentrumsstrategie** für diesen zentralen Teil der digitalen Infrastruktur.

Konkrete digitalpolitische und digitalwirtschaftliche Handlungsempfehlungen für eine darauf dann aufsetzende Rechenzentrumsstrategie, die auf dieser Basis Rahmenbedingungen für den Bau und den Betrieb von Rechenzentren in Deutschland schafft, umfassen insbesondere die folgenden Aspekte:

Förderung regionaler Infrastruktur: Investitionen in regionale Netzinfrastrukturen und Ansiedlungsräume für digitale Infrastruktur, insbesondere in Regionen mit hoher erneuerbarer Energieproduktion, sind unerlässlich. Ein Fokus auf die Schaffung von Ökosystemen aus Energie, digitaler Infrastruktur, Digitalwirtschaft und IT-Dienstleistung mit regionalem Fokus ist zu empfehlen.

Europäischer Marktplatz: Die Schaffung eines gemeinsamen europäischen Marktplatzes für Rechenkapazität würde die Wettbewerbsfähigkeit Europas stärken und Synergien schaffen. Ein Marktplatz schafft neue Chancen für Import und Export mit einer neuen Differenzierungsmöglichkeit für „Rechenleistung - Made in Germany“.

Strategische Partnerschaften: Die Untersuchung strategischer Optionen, wie zum Beispiel die Stärkung der digitalen Souveränität durch gezielte Kooperationen mit internationalen Partnern, ist von großer Bedeutung. Dies gilt insbesondere, um die langfristige Nachfrage nach verschiedenen Güteklassen von Rechenkapazität zu decken, zum Beispiel für das maschinelle Lernen.

Innovation fördern: Eine verstärkte politische Unterstützung für die Entwicklung innovativer Technologien, insbesondere im Bereich der Effizienz der Gebäudetechnik, der IT-Systeme und des IT-Infrastrukturbetriebs, ermöglicht eine kosteneffizientere Herstellung von Rechenkapazität, was sowohl eine deutsche Digitalwirtschaft stärkt, als auch im internationalen Wettbewerb von Vorteil ist.

International differenzieren: Im internationalen Wettbewerb sollte sich Deutschland über Effizienz, Nachhaltigkeit und Souveränität differenzieren und sich als Standort für hochwertige digitale Infrastruktur positionieren. Hochwertige digitale Infrastrukturen stärken zudem die Wettbewerbsfähigkeit und Resilienz deutscher Schlüsselindustrien und Digitalwirtschaft.

Datenlage weiter verbessern: Die Datenlage zu Entwicklung der Rechenzentren ist immer noch unzureichend. Mit dem in Deutschland eingeführten Rechenzentrumsregister (RZReg) wurde ein wichtiger Meilenstein für mehr Transparenz geschaffen. Allerdings bestehen noch Lücken in der Datenlage und offene Fragen, zum Beispiel zur Mitwirkung von Betreiber der IT-Installationen in Colocation Rechenzentren bei Datenmeldungen. Daher sind auch im vorliegenden Gutachten vorgenommene Abschätzungen der Marktentwicklung und des Stromverbrauchs mit Unsicherheiten behaftet. Es fehlen beispielsweise noch wesentliche Informationen zum eigentlichen Nutzungsgrad und zur Effizienz des IT-Equipments in den Rechenzentren. Auch Daten zu kleineren Rechenzentren sind bislang kaum verfügbar.

Handlungsoptionen

Deutschland ist bereits heute ein starker und attraktiver Standort für nationale und internationale Anbieter von digitaler Infrastruktur und der größte Rechenzentrumsstandort in Europa. Grundsätzlich unterstützt der weitere Ausbau von Rechenzentren die Möglichkeit, das Angebot an digitalen Ressourcen in Deutschland zu erhöhen.

Infrastruktur

Die Stärken des Standorts Deutschlands liegen in der auf Nachhaltigkeit ausgerichteten Digitalpolitik, gut ausgebauten Stromnetzen und der hohen Dichte an Internetknoten. Infrastrukturelle Aspekte, wie ein unzureichender Highspeed-Internetzugang für Unternehmen in ländlichen Regionen und eine mangelnde Glasfaserabdeckung, hemmen das Wachstum der digitalen Wirtschaft und schwächen Deutschlands Wettbewerbsfähigkeit in diesem Sektor. Ein flächendeckender Glasfaser- und 5G (zukünftig 6G) Ausbau sollte Teil einer Digitalisierungsvision sein, um mehr Flexibilität und geografische Diversität bei der Ansiedlung von Rechenzenten zu ermöglichen. Eine verbesserte digitale Infrastruktur würde zudem die Attraktivität des Standorts für ausländische Investoren steigern, die insbesondere mit Blick auf Hochleistungs- und KI-Rechenzentren auf diesbezügliche Anbindungen schauen, um Kosten und Latenzzeiten zu reduzieren.

Fachkräfte

Deutschland verfügt über eine starke mittelständische IT-Dienstleisterstruktur, die als Basis für weiteres Wachstum dienen kann. Im Vergleich zu den USA, wo Digitalunternehmen jahrzehntelange Erfahrung im Aufbau großer IT-Systeme haben, fehlt in Deutschland jedoch oft das entsprechende Fachwissen und Personal mit Digitalisierungskompetenzen. Der Fachkräftemangel ist ein maßgebliches Hindernis für die Ertüchtigung und den Ausbau digitaler Infrastrukturen in Deutschland. Dies gilt dabei nicht nur für IT-Fachkräfte, sondern auch für die Fachkräfte im Bereich der Gewerke Kälte und Klima sowie Elektro. Die Förderung von Kompetenzen und die Aus- und Weiterbildung von Fachkräften in diesen Bereichen kann die Wettbewerbsfähigkeit Deutschlands erhöhen und den Rechenzentrumsstandort Deutschland weiter stärken. Insbesondere mit Blick auf den bestehenden Fachkräftemangel sollten gezielte Initiativen zur Entwicklung von IT-Kenntnissen geschaffen werden, um qualifiziertes Personal für die Digitalisierung zu sichern.

Regionale digitale Ökosysteme und Schaffung eines europäischen Marktplatzes für digitale Ressourcen

Unterstützend dazu kann die Entwicklung regionaler digitaler Ökosysteme sowie die Schaffung eines europäischen Marktplatzes für digitale Ressourcen die Wettbewerbsfähigkeit Deutschlands und Europas hinsichtlich Digitalisierung und Rechenzentrumsstandort steigern. Die konsequente Förderung von Start-ups und Technologieunternehmen sowie die Schaffung von Innovationsclustern gemeinsam mit Forschungseinrichtungen wirken positiv auf die Entwicklung regionaler digitaler Ökosysteme. Die Fokussierung auf digitale Ökosysteme maximiert die volkswirtschaftliche Wertschöpfung von digitaler Infrastruktur. Dabei sollten Rechenzentrumsgebäude, IT-Infrastrukturanbieter und IT-Dienstleister, Beratungen und Unternehmen der Digitalwirtschaft sowie die Forschungsinfrastrukturen als Ökosystem betrachtet werden, welche sich zusammen in einer Region

ansiedeln. Die Einrichtung eines Kompetenzzentrums für Energieeffizienz in Rechenzentren und die Entwicklung entsprechender disziplinenübergreifender Studiengänge könnte einen Beitrag leisten, ähnlich wie in den USA Best Practices systematisch zu dokumentieren und Schulungs- und Beratungsangebote zur Verfügung zu stellen. Die Schaffung eines europäischen Marktplatzes für digitale Ressourcen kann den Handel von digitalen Ressourcen innerhalb Europas fördern und Deutschlands digitale Souveränität erhöhen.

Regulatorik

Als Standort für digitale Infrastruktur kann sich Deutschland über Qualität und Rechtssicherheit differenzieren. Rechenzentren mit einem Standort in Deutschland stehen vielfältige Regelungen aus unterschiedlichen Rechtsbereichen gegenüber. Der Themenkomplex der Datensicherheit und der Datensouveränität ist jedoch ein wesentlicher regulatorischer Treiber für den Rechenzentrumsstandort Deutschland. Digitale Infrastrukturen, die ihren Standort in Deutschland haben und deutschem Recht unterliegen, werden nicht nur aus Deutschland nachgefragt werden. Als förderlich im Standortwettbewerb von Rechenzentren werden insbesondere beschleunigte Genehmigungsverfahren für Rechenzentren eingeordnet. Daneben werden regulatorische Anpassungen vorgeschlagen, etwa die Einführung eines deutschlandweiten Wärmenetzregisters, das mehr Transparenz über die Verfügbarkeit und Nutzung von Wärmenetzen bieten könnte. Insgesamt würden eine erhöhte Transparenz und Datenzugänglichkeit die Planung und Umsetzung des Ausbaus weiterer Rechenzentren beschleunigen.

Energieeffizienz und Flexibilisierung

Mit steigendem Stromverbrauch von Rechenzentren und zunehmendem Wettbewerb zwischen Rechenzentren und Rechenzentrumsstandorten steigt die Bedeutung der Energieeffizienz. Der Rechenzentrumsstandort Deutschland könnte in dieser Hinsicht von dem umfassenden regulatorischen Rahmen hinsichtlich Energieeffizienz durch das EnEfG profitieren. Deutschland kann sich als führendes Land im Bereich des nachhaltigen digitalen Infrastrukturausbaus positionieren, insbesondere im Hinblick auf energieeffiziente und nachhaltige Rechenzentren. Der regulatorische Rahmen weist allerdings derzeit noch Lücken auf, insbesondere im Bereich der Auslastung und dem Nutzungsgrad von IT-Infrastruktur, mit dem Ziel Ineffizienzen im Bereich von Software und IT-Infrastruktur abzubauen. Bisher gibt es keine Verpflichtung der IT-Betreiber zu Effizienzsteigerungen, obwohl der Energieverbrauch der IT ca. 80 % des Gesamtverbrauchs eines Rechenzentrums ausmacht. Auch die Ressourceneffizienz insbesondere hinsichtlich des Wassereinsatzes im Bereich der adiabaten Kühlung erfordert weitere Beobachtung und eventuell auch Maßnahmen. Grundsätzlich besteht ein Zielkonflikt zwischen der Minimierung des Wasserverbrauchs und der Energieeffizienz. Grundlage dafür bildet auch die Digitale Strategie 2025, die sehr zukunftsorientiert ist und den Ausbau der Infrastruktur – zur Überbrückung der digitalen Kluft – fokussiert, womit Deutschland anderen Ländern voraus ist.

Digitalisierung der Wirtschaft

Die digitale Zukunft von Deutschland ist primär durch den Grad der Digitalisierung in traditionellen Branchen und das Wachstum der Digitalwirtschaft geprägt. Geht man davon aus, das Deutschland seine eigene und die Digitalisierungsstrategie der EU erfüllt, ist mit einem hohen Anstieg der Nachfrage nach digitalen Ressourcen und damit auch nach digitaler Infrastruktur zu rechnen.

Insbesondere die Digitalisierung von Handel, Gesundheitswesen und dem verarbeitenden Gewerbe ist ausschlaggebend für das Nachfragewachstum an digitalen Ressourcen. Technologische Innovationen wie KI sorgen zudem für eine erhöhte Nachfrage nach Datenverarbeitung und Speicherung im Rahmen der Digitalisierung. Bei deutlichem Wachstum im Digitalsektor bei gleichzeitiger Stagnation bzw. Rückgang der Wertschöpfung im Industriesektor ergibt sich eine zunehmende Verschiebung der Wirtschaftsleistung vom Industrie- in den Dienstleistungssektor. Für die Digitalwirtschaft ist neben Fachkräften die Verfügbarkeit und der Preis von digitalen Ressourcen ein zentraler Faktor für den Erfolg der digitalen Produkte und Dienstleistungen. Dies unterscheidet die Digitalwirtschaft von anderen traditionellen Branchen, die primär andere Rohstoffe oder Handelsgüter verarbeiten, veredeln oder herstellen. Durch diese hohe Abhängigkeit von digitalen Ressourcen bedingt ein Wachstum der Digitalwirtschaft auch einen starken Anstieg der Nachfrage nach digitalen Ressourcen.

Digitalisierung der Öffentlichen Hand

Nicht zu unterschätzen ist auch die Nachfrage nach Digitalisierung durch öffentliche Einrichtungen. Viele Prozesse und Dienstleistungen für Bürger und Unternehmen müssen zukünftig effizienter und schneller erbracht werden. Aber auch die stark standardisierten und häufig gesetzlich genormten internen Bearbeitungsprozesse weisen hohe Potenziale für die Digitalisierung auf. Vielfach sind Prozesse, wie z. B. die Einreichung und Bearbeitung von Bauanträgen, noch analog organisiert. Die Nachfrage nach Digitalisierung der öffentlichen Hand kann als Treiber für die Fokussierung auf effiziente und nachhaltige digitale Infrastrukturen im Sinne grüner Leitmärkte eingesetzt werden.

Einleitung

Mit mehr als 2.000 Rechenzentren (RZ) ist Deutschland derzeit der Standort in Europa mit den meisten RZ. Als zentral gelegenes Land innerhalb Europas bietet Deutschland einen strategisch vorteilhaften Standort für den europäischen, interkontinentalen und internationalen Datenverkehr. Die höchste RZ-Dichte besteht derzeit im Raum Frankfurt am Main, wo sich neben zahlreichen Unternehmen z. B. der Finanzbranche auch der weltweit führende kommerzielle Internetknoten (DE-CIX) befindet. Ein weiterer regionaler RZ-Schwerpunkt entwickelt sich aktuell im Raum Berlin. Auch der hohe und unterbrechungsfreie Strombedarf von Rechenzentren macht Deutschland mit seinem zuverlässigen Stromnetz zu einem besonders geeigneten Standort.

Die Bundesregierung steht vor einer strategischen Herausforderung: Deutschland will in den kommenden Jahren seine digitale Infrastruktur ausbauen, um die Wettbewerbsfähigkeit der deutschen Wirtschaft zu sichern. Gleichzeitig muss dieser Ausbau im Einklang mit den Energie- und Klimazielen im Rahmen der breiten Transformation des Energiesystems erfolgen und die digitale Souveränität Deutschlands gewährleisten.

Ziel des Gutachtens:

Vor diesem Hintergrund soll das Gutachten „Stand und Entwicklung des Rechenzentrumsstandorts Deutschland“ als evidenzbasierte Grundlage für politische Entscheidungsträger zur Schaffung eines langfristigen Zukunftsbildes des Rechenzentrumsmarktes und der Ressourcennachfrage dienen. Somit soll die Entwicklung eines Ansatzes zur Stärkung der Wettbewerbsposition Deutschlands beim Ausbau digitaler Infrastrukturen und zur gezielten Förderung derjenigen Segmente des Rechenzentrumsmarktes unterstützt werden, der einen besonderen Beitrag zu wirtschaftlichem Wachstum und technologischer digitaler Souveränität leisten kann.

Herausforderungen und Hintergrund:

Die Prognosen aus diesem Gutachten zeigen, dass sich die Rechenzentrumskapazitäten bis 2030 verdoppeln könnten – unter anderem angetrieben durch Künstliche Intelligenz, Industrie 4.0 und die fortschreitende Digitalisierung der deutschen Wirtschaft. Dies stellt die Politik vor ein Trilemma:

- Der Ausbau von Rechenzentren und digitaler Infrastruktur erfordert erhebliche Stromkapazitäten in einem ohnehin angespannten Energiemarkt.
- Die Verfügbarkeit geeigneter Flächen ist besonders in wirtschaftsstarken Regionen begrenzt.
- Ein Import digitaler Ressourcen aus dem Ausland könnte die digitale Souveränität gefährden.

Um diese Herausforderungen zu meistern, muss zunächst verstanden werden, welche Kräfte die Entwicklung des Marktes für digitale Infrastruktur prägen.

Die treibende Kraft hinter dem Ausbau von Rechenzentrumsinfrastruktur in Deutschland und dem damit verbundenen Strombedarf ist die Nachfrage nach Rechenleistung, Speicherkapazität und Datenübertragung (im Folgenden „digitale Ressourcen”). Diese Nachfrage wiederum wird durch Digitalisierung der Wirtschaft, der öffentlichen Hand und der Gesellschaft, aber auch durch das Wachstum der neuen, eigenständigen Digitalwirtschaft verursacht.

Hinzu kommen Technologien, wie zum Beispiel das maschinelle Lernen und Künstliche Intelligenz, welche zunehmend Anwendung in der Digitalisierung und der Digitalwirtschaft finden und als zusätzliche Beschleuniger der Nachfrage nach digitalen Ressourcen zu betrachten sind.

Für die Untersuchung des Rechenzentrumsmarkts und dem damit verbundenen Energie- und Ressourcenverbrauch sollten die entsprechenden Kräfte entlang der Wertschöpfungskette betrachtet werden. Die Nachfrage nach digitaler Infrastruktur entsteht durch den steigenden Bedarf an digitalen Ressourcen. Der steigende Bedarf an digitalen Ressourcen entsteht durch die Anwendung von Software im Rahmen der Digitalisierung und Digitalwirtschaft.

Der digitale Ressourcenverbrauch von Software wird maßgeblich durch die eingesetzten Technologien beeinflusst. So hat zum Beispiel eine Softwareanwendung, welche maschinelles Lernen einsetzt, einen höheren Bedarf an Rechenleistung. Eine Softwareanwendung, die auf IoT-Sensorik aufbaut, sorgt für einen höheren Bedarf an Datenspeicherung und Übertragungskapazität. So lässt sich schlussfolgern, dass sowohl die Menge an eingesetzter Software im Rahmen der Digitalisierung und Digitalwirtschaft als auch die Art der Software die maßgeblichen Treiber der Nachfrage nach digitalen Ressourcen sind. Die Nachfrage nach digitalen Ressourcen wiederum treibt den Ausbau von digitaler Infrastruktur an. Der Ausbau von digitaler Infrastruktur erfordert Strom und IKT-Equipment.

Wertschöpfung entsteht in der Herstellung und Nutzung der digitalen Ressourcen durch Software.

Die primäre Wertschöpfung im digitalen Ökosystem entsteht in der Anwendung von Software für die Digitalisierung und für die Etablierung neuer Digitalunternehmen. Dieser zunehmende Einsatz von Software in allen Wirtschaftsbereichen, in Staat und Gesellschaft und für die Digitalwirtschaft führt zu einer größeren Nachfrage an digitale Ressourcen, die für den Betrieb der Software benötigt werden.

Dabei entsteht Wert durch die Veredelung der digitalen Ressourcen mit Software – unter anderem durch Algorithmen, Logik, Geschäftsmodelle und Automatisierung. Digitale Ressourcen sind kritisch für den Betrieb von Software, die Herstellung der digitalen Ressourcen hat aber im Verhältnis zur Anwendung von Software eine geringere Wertschöpfung.

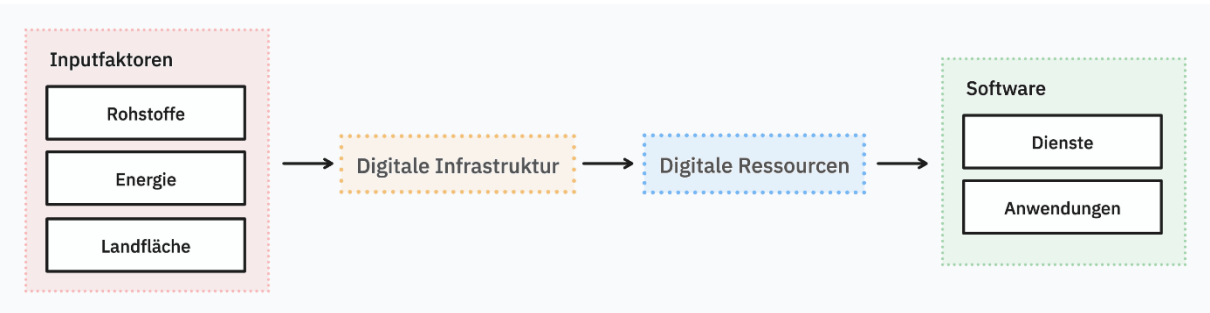


Abbildung 5: Vereinfachte Darstellung der Wertschöpfung durch digitale Infrastruktur. Quelle: Eigendarstellung SDIA

Für die Herstellung von digitalen Ressourcen wird digitale Infrastruktur benötigt. Die entsprechende Wertschöpfung in dieser Infrastruktur wurde genauer im Kapitel 3 untersucht. Mit dem Begriff „digitale Infrastruktur" werden in diesem Bericht die folgenden Komponenten beschrieben:

- Rechenzentrumsgebäude
- IKT-Equipment, wie z. B. Server-, Speicher- oder Netzwerksysteme (Datenverarbeitungssysteme")
- Glasfasernetze und Internetknoten ("Datenübertragungsnetze")

Für diesen Bericht wird unter dem Begriff "Rechenzentrum" das Gebäude und das IKT-Equipment zusammengefasst. Mit dem Begriff „digitale Infrastruktur" werden Rechenzentren und Datenübertragungsnetze zusammengefasst. In diesem Bericht werden gezielt Rechenzentren untersucht. Die Betrachtung von Datenübertragungsnetzen wird nur im weiteren Sinne als einwirkende Kraft (entweder als Barriere oder Beschleuniger der Marktentwicklung) mit einbezogen und nicht gesondert betrachtet.

Grundsätzlich ist anzumerken, dass ein Ausbau von Telekommunikationsinfrastruktur sich positiv auf die Marktentwicklung von digitaler Infrastruktur auswirkt und auch als Beschleuniger von Digitalisierung und als Wachstumstreiber für die Digitalwirtschaft verstanden werden sollte. Der Ausbau von Telekommunikationsinfrastruktur wird in diesem Gutachten nicht weiter betrachtet, da er bereits durch andere regulatorische und politische Maßnahmen abgedeckt ist.

Das Angebot an digitalen Ressourcen lässt sich durch den Ausbau von Rechenzentren oder durch den Import von digitalen Ressourcen fördern.

Im Rahmen des Gutachtens wird primär der Ausbau von Rechenzentren als Möglichkeit, den Bedarf an digitalen Ressourcen in Deutschland zu decken, betrachtet. Ein sekundärer Weg zur Bedarfsdeckung ist der Import von digitalen Ressourcen aus dem europäischen Ausland. Dieser Aspekt wird in den Kapiteln 3 und 5 betrachtet, sowohl aus der volkswirtschaftlichen Perspektive als auch aus Sicht der digitalen Souveränität.

Politische Rahmenbedingungen wirken sich auf die Nachfrage nach digitalen Ressourcen aus.

Politische Rahmenbedingungen, wie z. B. die Deutsche Digitalstrategie, Förderprogramme für Start-ups, die europäische Digitalisierungsstrategie und andere politische Maßnahmen, wirken sich auf das Wachstum der Digitalwirtschaft und auf die Geschwindigkeit der Digitalisierung aus. Die Auswirkungen dieser Rahmenbedingungen werden im Rahmen von Zukunftsszenarien (Kapitel 5) untersucht, insbesondere in Bezug auf die Auswirkungen auf die Nachfrage nach digitalen Ressourcen.

Aufbau des Gutachtens:

Kapitel 1 beschreibt die aktuellen Entwicklungen im Rechenzentrumsmarkt und ordnet die Entwicklung in Deutschland in den internationalen Kontext ein. Die Ergebnisse aus Kapitel 1 zeigen, dass vor allem die Rechenzentrumsmärkte in den USA und in China stark wachsen. Dies ist zum einen durch die hohe Konzentration von Digitalunternehmen in den USA und zum anderen durch die starke staatliche Intervention in China zu erklären. Im internationalen Kontext ist die Entwicklung der Rechenzentrumsbranche in Deutschland differenziert zu betrachten. Das Wachstum der Rechenzentren in Deutschland wird aktuell vor allem von Telekommunikationsunternehmen, Immobilien- und Infrastruktur-Fonds, welche in Colocation-Rechenzentrumsgebäude investieren, beschleunigt. Jedoch fokussiert sich das Wachstum der Rechenzentrumsbranche auf wenige

Regionen in Deutschland. Vor allem der Großraum Frankfurt/Rhein-Main mit dem Netzwerkknoten DE-CIX bleibt für die Ansiedlung von Rechenzentren attraktiv.

Kapitel 2 analysiert die Bedeutung von Rechenzentren im Energiesystem hinsichtlich Strombedarf, Energieeffizienz, Klimawirkung und Abwärmepotenziale. Aufgrund des zunehmenden Bedarfs an digitaler Infrastruktur, insbesondere durch IT-Dienste, und beschleunigt durch Künstliche Intelligenz, wird erwartet, dass der Stromverbrauch der Rechenzentren von derzeit etwa 20 TWh auf bis zu 80 TWh jährlich bis 2045 ansteigen wird. Effizienzgewinne, etwa durch Verbesserungen der PUE, können das Verbrauchswachstum nicht vollständig kompensieren. Ein Ausbau der Stromnetze ist notwendig, um die Stromversorgung zu sichern und zusätzliche Lasten durch Rechenzentren sowie erneuerbare Energiequellen zu integrieren. Da die Klimabilanz von Rechenzentren maßgeblich durch den Strommix determiniert wird, sind die Treibhausgasemissionen deutscher Rechenzentren aufgrund des derzeit noch nennenswerten Anteils fossiler Energien im deutschen Strommix höher als in Ländern mit mehr erneuerbaren Energien bzw. Atomkraft. Sind die deutschen Erneuerbaren-Ausbauziele bis 2030 erfolgreich, wird sich das aber in wenigen Jahren ändern. Rechenzentren erzeugen während ihres Betriebs erhebliche Mengen an Abwärme, die theoretisch in Wärmenetzen und für kommunale Wärmeprojekte genutzt werden könnten. Das Potenzial der aus Rechenzentren bereitgestellten Abwärme steigt durch den stark steigenden Stromverbrauch erheblich. Um größere Anteile der entstehenden Abwärme zu nutzen, müssen jedoch noch technische, organisatorische und regulatorische Herausforderungen überwunden werden.

Kapitel 3 beschreibt die Ökonomie von Rechenzentren. Dafür werden die wirtschaftliche Bedeutung digitaler Infrastruktur in Deutschland analysiert sowie Wertschöpfungspotenziale und Spillover-Effekte von Rechenzentren auf die gesamte Volkswirtschaft aufgezeigt. Während für den Bau eines Rechenzentrums temporär viele Arbeitskräfte gebraucht werden, nimmt die Zahl der permanenten Arbeitsplätze im laufenden Betrieb von Rechenzentren deutlich ab. Makroökonomische Spillover-Effekte aus Rechenzentren entstehen insbesondere durch die Veredelung digitaler Ressourcen, wie z. B. IT-Dienste, und ermöglichen Digitalisierung und Wachstum der Digitalwirtschaft. Die Nachfrage nach digitalen Ressourcen und damit auch nach Infrastruktur wird primär durch die Digitalisierung und das Wachstum der Digitalwirtschaft angetrieben. Die Ökonomie von Rechenzentren ist auch bei globaler Betrachtung durch eine geringe Wettbewerbsintensität charakterisiert: Die größten Rechenzentrumsflächenvermieter in Deutschland sind ausländische Unternehmen und internationale Anbieter dominieren das Angebot von Cloud-Diensten durch geschlossene Ökosysteme. Deutschland verfügt jedoch über eine starke mittelständische IT-Dienstleisterstruktur, die als Basis für weiteres Wachstum dienen kann.

Kapitel 4 untersucht die wesentlichen regulatorischen Rahmenbedingungen für den RZ-Standort Deutschland und entsprechende Anforderungen an die Planung und den Betrieb von Rechenzentren. Die Ergebnisse aus Kapitel 4 zeigen, dass der Themenkomplex der Datensicherheit und der Datensouveränität ein wesentlicher regulatorischer Treiber für den RZ-Standort Deutschland ist. Weitere regulative Anforderungen an Rechenzentren betreffen das Rechenzentrumsgebäude, z. B. hinsichtlich der Erfordernis einer Baugenehmigung. Eine Beschleunigung von Genehmigungsverfahren für die Errichtung von Rechenzentren könnte diesen im internationalen Vergleich existierenden Standortnachteil Deutschlands nivellieren. Ein Ergebnis der Analyse ist zudem, dass Rechenzentrumsbetreibern Mehraufwände für die Erfüllung der regulativen Anforderungen mit Blick auf Energieeffizienz entstehen. Im Zusammenhang mit den regulativen Anforderungen nach dem EnEfG wurden mögliche Zielkonflikte untersucht. Ein weiteres Ergebnis ist,

dass Kennzahlen zu Energie- und Ressourceneinsatz, wie z. B. die PUE, als Energieeffizienzparameter zwar grundsätzlich geeignet sind, jedoch für eine vollumfängliche Bewertung die Aufnahme zusätzlicher Parameter erforderlich wäre. Während PUE als Metrik für die Energieeffizienz von Rechenzentren weit verbreitet ist, so beschreibt sie die Effizienz von Rechenzentren nur unvollständig.

Kapitel 5 skizziert anhand verschiedener Digitalisierungsszenarien Optionen einer digitalen Zukunft, um die Nachfrage nach digitalen Ressourcen und die damit verbundenen Rechenzentrums- und Energiebedarfe zu verorten. Auch geht Kapitel 5 der Frage nach digitaler Souveränität nach und gibt eine Übersicht der Trends hinsichtlich Technologien im Rechenzentrumsumfeld.

Kapitel 6 beschreibt übergeordnete Handlungsempfehlungen, um die wesentlichen Aspekte eines politischen Handlungsrahmens darzustellen, die eine wettbewerbsfähige digitale Infrastruktur in Deutschland unterstützen und in einer möglichen deutschen Rechenzentrumsstrategie berücksichtigt werden sollten.

1 Entwicklungen und Trends im Rechenzentrumsmarkt

1.1 Zusammenfassung

Im Kapitel 1 wird eine umfassende Betrachtung der aktuellen Entwicklungen im Rechenzentrumsmarkt vorgenommen. Damit wird eine neutrale Bewertung der aktuellen Entwicklungen ermöglicht. Das Gutachten legt dazu ein umfassendes Begriffsverständnis des Rechenzentrumsmarktes zugrunde und betrachtet alle Marktsegmente, von kleinen IT-Installationen im mittelständischen Unternehmen bis hin zu Großrechenzentren mit Anschlussleistungen im dreistelligen Megawatt-Bereich. Der Begriff „Rechenzentren“ vereint in dem hier zugrundeliegenden Verständnis dabei sowohl die Gebäude als auch die IT-Infrastruktur und den Internetzugang.

Zunächst wird ein Überblick über die weltweite Entwicklung des Rechenzentrumsmarktes gegeben. Die Analyse zeigt, dass sich das Marktwachstum beschleunigt. Es wird davon ausgegangen, dass sich die Kapazitäten der weltweiten Rechenzentren gemessen in IT-Anschlussleistung bis 2030 auf 220 GW gegenüber heute etwa verdoppeln werden. Treiber der Entwicklung sind vor allem das starke Wachstum im Bereich Cloud Computing sowie die zunehmende Nutzung von Anwendungen im Bereich der Künstlichen Intelligenz. Durch die hohe Marktkonzentration sowohl im Bereich der digitalen Dienste aus Rechenzentren als auch in der Hardwarebereitstellung herrscht eine hohe Wettbewerbsintensität zwischen den wenigen Anbietern, die das Wachstum weiter beschleunigt.

Im weltweiten Vergleich wachsen vor allem die Rechenzentrumsmärkte in den USA und in China sehr stark, Europa liegt im Wachstum etwas zurück. Dies ist zum einen durch die hohe Konzentration von Unternehmen der Digitalbranche in den USA und zum anderen durch die starke staatliche Intervention in China zu erklären.

In Deutschland wächst der Rechenzentrumsmarkt ebenfalls dynamisch. Zwischen 2022 und 2030 ist mit einer Verdoppelung der Kapazitäten auf knapp 5 GW zu rechnen. Setzen sich die aktuellen Entwicklungen im Markt weiter fort, so wird bis 2045 ein Anstieg der Kapazitäten auf mehr als 12 GW erwartet. Getrieben ist das Wachstum auch hierzulande insbesondere durch Milliardeninvestitionen großer Hyperscale-Cloud-Anbieter, wie Amazon, Microsoft oder Google. Diese mieten bislang zumeist Gebäudeflächen von Colocation-Dienstleistern[2], wodurch der Marktanteil von Colocation-Rechenzentren in Deutschland stetig steigt. Es wird erwartet, dass im Jahr 2045 ca. 45 % der Rechenzentrumskapazitäten in Deutschland von Colocation-Anbietern bereitgestellt werden. Aufgrund der wirtschaftsstarken Industrie und insbesondere der mittelständischen Orientierung wird in Deutschland, aber auch in einer großen Zahl von Standorten die Informationstechnik “on premise”, also direkt vor Ort, betrieben. Zwar nimmt die Bedeutung dieses Bereitstellungsmodells im Vergleich zur Nutzung von Cloud- und Colocation-Angeboten weiter ab, absolut bleiben die Kapazitäten aber zumindest zwischen 2010 und 2024 mit 1 GW IT-Anschlussleistung weitgehend konstant.

Gemessen an den Kapazitäten ist Deutschland der größte Rechenzentrumsstandort in Europa. In den letzten Jahren konnte diese Position etwas ausgebaut werden. Begründet ist diese Stärke durch die

[2] für eine Übersicht der Arten und Geschäftsmodellen von Rechenzentren, siehe Kapitel 3.4

hohe inländische Nachfrage von Rechenzentrumsleistung, nicht durch den Export. Andere Staaten wie Irland, die skandinavischen Länder und die Niederlande, exportieren deutlich mehr digitale Leistungen aus Rechenzentren. Stärken weist der deutsche Rechenzentrumsstandort durch seine auf Nachhaltigkeit ausgerichtete Digitalpolitik und den guten Ausbau der Stromnetze sowie die hohe Dichte an Internetknoten auf.

Um die Vorteile des Standorts Deutschland weiter zu bündeln und zielgerichtet weiterzuentwickeln, wird empfohlen, eine nationale Rechenzentrumsstrategie zu entwickeln. Damit könnte die Wettbewerbsposition Deutschlands beim Ausbau digitaler Infrastrukturen verbessert und der Markt gezielt so gefördert werden, dass er verstärkt zu wirtschaftlichem Wachstum und digitaler Souveränität beitragen kann.

1.2 Weltweit

Die Bedeutung von Rechenzentren in der globalen Wirtschaft und Gesellschaft wächst kontinuierlich, insbesondere angetrieben durch die steigende Nachfrage nach Cloud-Diensten und Anwendungen im Bereich der Künstlicher Intelligenz. Analysten und Marktforschungsinstitute prognostizieren für die kommenden Jahre ein sehr deutliches Wachstum der weltweiten Kapazitäten und Investitionen in Rechenzentren. Dies wird durch eine Vielzahl von Studien und Berichten unterstrichen, die sowohl die physische Infrastruktur als auch den Energieverbrauch betreffen.

So zeigen die Analysen des Immobiliendienstleisters CBRE, dass die Rechenzentrumsmärkte in Nordamerika mit 24,4 %, in Asien-Pacific mit 22 % und in Europa mit knapp 20 % im vergangenen Jahr jeweils deutlich gewachsen sind (CBRE, 2024b). Diese regionalen Wachstumsraten verdeutlichen, dass der Ausbau nicht auf einzelne Märkte beschränkt bleibt, sondern ein globaler Trend ist. Andere Analysten kommen zu ähnlichen Wachstumsprognosen. Das Beratungsunternehmen Cushman & Wakefield geht davon aus, dass sich die weltweiten RZ-Kapazitäten aufgrund geplanter Projekte in den nächsten Jahren verdoppeln werden (Cushman & Wakefield, 2024). Die Analysen von Marktforschern wie Precedence Research (Precedence Research, 2024), Spherical Insights (Spherical Insights, 2023) und Fortune Business Insights (Fortune Business Insights, 2024) bestätigen die Prognosen der Verdopplung der RZ-Kapazitäten in wenigen Jahren.

Auch McKinsey prognostiziert ein sehr deutliches Wachstum im RZ-Markt. In einem mittleren Szenario könnte sich der weltweite Bedarf an Rechenzentrumskapazität zwischen 2023 und 2030 mit einer Wachstumsrate von jährlich 22 % vervierfachen. Ob dieser Bedarf auch gestillt werden kann, ist unsicher. Um ein Defizit zu vermeiden, müsste mindestens die doppelte Rechenzentrumskapazität, die seit dem Jahr 2000 gebaut wurde, in weniger als einem Viertel der Zeit errichtet werden (McKinsey, 2024). Auch für Europa sieht McKinsey hohe Wachstumsraten im Bedarf nach RZ-Kapazität von jährlich 20 % zwischen 2023 und 2030 (Granskog et al., 2024).

Die Verfügbarkeit von ausreichend Kapazitäten in der Stromerzeugung und in den Stromnetzen könnte ein begrenzender Faktor für den zukünftigen Ausbau der RZ-Infrastrukturen sein (Hintemann, Hinterholzer, & Progni, 2024). Vor allem die großen Hyperscale-Cloud-Anbieter ergreifen aktuell zunehmend Maßnahmen, sich entsprechende Stromversorgungs-Kapazitäten zu sichern, in den USA oft auch durch Verträge mit Kernkraftwerken oder auch durch Investitionen in neue Reaktortechniken (Hermann, 2024; Spiegel, 2024; Tagesschau.de, 2024).

Die Daten zum aktuell deutlichen Ausbau der RZ-Kapazitäten werden durch Prognosen zu den IT-Hardwareverkäufen gestützt. Laut Statista Market Insights wird sich das Verkaufsvolumen von Servern zwischen 2020 und 2030 nahezu verdreifachen. IDC berichtet, dass die Serververkäufe allein im Jahr 2024 um 42 % im Vergleich zum Vorjahr gestiegen sind (IDC, 2024).

Ein weiterer Indikator für das Wachstum der Rechenzentrums-Infrastrukturen ist der Energieverbrauch der Rechenzentren. Die Internationale Energieagentur (IEA) prognostiziert, dass der Stromverbrauch von Rechenzentren allein zwischen 2022 und 2026 um 75 % ansteigen wird – von 460 TWh auf über 800 TWh (Basisszenario). (IEA, 2024).

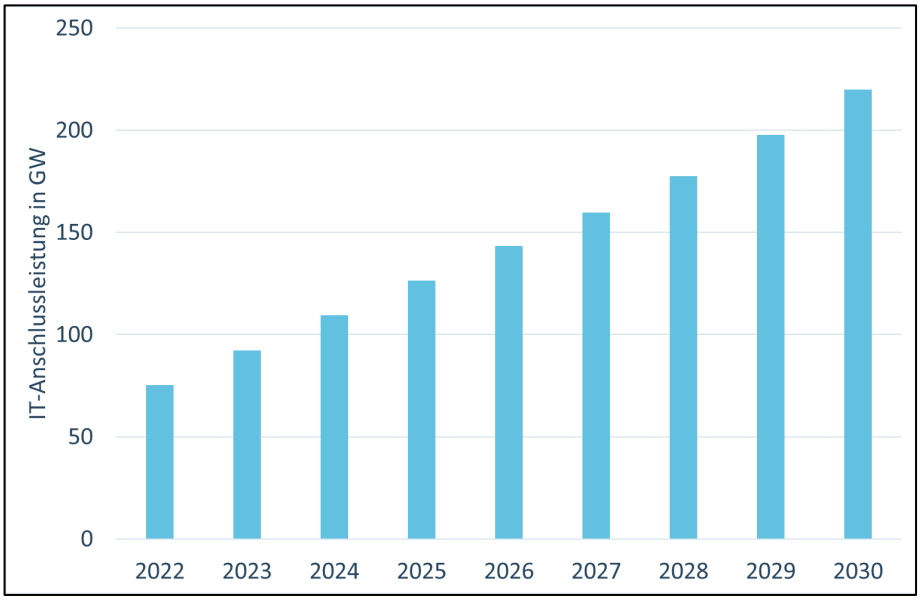


Abbildung 6: Abschätzung der Entwicklung der weltweiten Kapazitäten der Rechenzentren und kleineren IT-Installationen in den Jahren 2022 bis 2030. Quelle: Berechnungen Borderstep

Abbildung 6 zeigt eine Abschätzung der Entwicklung der weltweiten Rechenzentrumskapazitäten zwischen 2022 und 2030 auf Basis der verschiedenen Wachstumsprognosen. Danach könnten die weltweiten Rechenzentrumskapazitäten bis 2030 von aktuell 110 GW auf 220 GW ansteigen. Die Darstellung berücksichtigt nicht nur Großrechenzentren, sondern auch die Kapazitäten von kleineren IT-Installationen.

Wie gut die in Abbildung 6 gemachte Abschätzung und Prognose der Kapazitäten der weltweiten Rechenzentren die tatsächliche Situation widerspiegelt, ist nur eingeschränkt zu beurteilen. Der Abschätzung liegen insbesondere die Annahmen der Untersuchung der IEA zum stark wachsenden Stromverbrauch zwischen 2022 und 2026 zugrunde. Weiterhin werden die Prognosen, dass sich die Kapazitäten bis 2030 etwa verdoppeln könnten, in der Abschätzung berücksichtigt. Es wird nicht davon ausgegangen, dass Engpässe bei der Strombereitstellung oder andere Verzögerungen dazu führen können, dass der wachsende Bedarf nach Rechenzentrumskapazitäten nicht befriedigt werden kann.

Eine Herausforderung bei der Erstellung der Prognose liegt auch darin, dass ein systematischer Vergleich und die Zusammenfassung der unterschiedlichen Marktstudien zu Rechenzentren kaum möglich sind, da methodische und begriffliche Unterschiede oft zu Diskrepanzen in den Ergebnissen

führen. Eine wesentliche Schwierigkeit liegt in den unterschiedlichen Verständnissen und Definitionen des Begriffs „Rechenzentrum“. Manche Marktstudien decken die gesamte Bandbreite von Rechenzentren ab, von kleinen Serverräumen in Unternehmen bis hin zu großflächigen Hyperscale-Rechenzentren. Andere Studien hingegen fokussieren sich ausschließlich auf bestimmte Teilsegmente, wie etwa den Colocation- und Cloud-Markt. Insbesondere Studien, die den Immobilienmarkt betrachten, verwenden häufig eine spezifische Perspektive, die auf „vermarktbaren“ Rechenzentrumskapazitäten basiert. Diese orientiert sich – vergleichbar mit dem Büroflächenmarkt – nur auf Flächen oder Kapazitäten, die potenziell vermietet oder verkauft werden können. Diese enge Abgrenzung schließt On-Premise-Rechenzentren, Hosting- und andere Dienstleistungsrechenzentren oft aus, obwohl diese weiterhin eine bedeutende Rolle im Gesamtmarkt spielen. Auch die Art und Weise, wie Kapazitäten quantifiziert werden, variiert zwischen den verschiedenen Studien. Häufig werden Kapazitäten in Megawatt angegeben, doch diese Angabe ist nicht immer eindeutig. Es bleibt oft unklar, ob sich die genannten Zahlen auf die gesamte elektrische Anschlussleistung des Rechenzentrums beziehen oder nur auf die maximale IT-Anschlussleistung, die die tatsächlich für IT-Hardware verfügbare Kapazität beschreibt.

Das Wachstum im RZ-Markt wird maßgeblich durch die Ausweitung von Cloud-Computing-Diensten und den zunehmenden Einsatz von Künstlicher Intelligenz (KI) getrieben. Insbesondere Cloud Computing bleibt ein zentraler Wachstumstreiber für den Ausbau von Rechenzentren. Nach Angaben der Synergy Research Group stieg der globale Cloud-Markt im ersten Halbjahr 2024 um 23 % (Synergy Research, 2024b). Das weitere Wachstumspotenzial ist hoch, da die Entwicklung zu mehr Cloud Computing noch am Anfang steht. Im Jahr 2023 nutzten in der EU mit 45,2 % noch nicht einmal die Hälfte der Unternehmen überhaupt bezahlte Cloud-Dienste (Eurostat, 2024).

Das Wachstum der Cloud-Infrastruktur wird von wenigen Hyperscalern bestimmt, insbesondere von Amazon Web Services (AWS), Microsoft Azure und Google Cloud. Diese drei sind in fast allen Weltregionen Marktführer im Cloud-Markt. Ausnahme ist China, dort sind Alibaba, Tencent und China Telecom Marktführer. Zusammen haben Amazon, Microsoft und Google im zweiten Quartal 2024 nach Angaben von Synergy Research einen Marktanteil von 68 % am gesamten weltweiten Cloud-Markt (Synergy Research, 2024a). Aktuell ist festzustellen, dass sich der Ausbau der Rechenzentrumskapazitäten für Cloud Infrastruktur deutlich beschleunigt. Mit einem Wachstum von 30 % stiegen die Investitionen in Cloud-Rechenzentren im ersten Halbjahr 2024 schneller als die Umsätze mit Cloud-Services, die um 21 % stiegen (Synergy Research, 2024b). Diese Entwicklung kann als Indiz dafür gesehen werden, dass aktuell von den marktstarken Unternehmen versucht wird, über einen möglichst schnellen Ausbau von Rechenzentrumskapazitäten einen Wettbewerbsvorteil zu gewinnen.

Anwendungen im Bereich der Künstlichen Intelligenz stellen hierbei einen wesentlichen Treiber im Cloud-Markt dar. Vor allem die Nutzung von Large Language-Modellen wie GPT durch Hyperscale-Cloud-Anbieter führt zu starker Nachfrage nach Rechenleistung (Menn, 2024; online, 2024). Der globale KI-Markt wächst laut Sopra Steria dreimal so schnell wie der gesamte IT-Markt (Sopra Steria, 2024) und erfordert eine Infrastruktur, die diesen Bedarf decken kann. KI-Chips, die essenziell für das Training und die Nutzung dieser KI-Modelle sind, verzeichnen dabei ein rasantes Umsatzwachstum. Laut Gartner wird der weltweite Umsatz mit KI-Chips im Jahr 2024 um 33 % steigen (manage IT, 2024). Eine Analyse von McKinsey prognostiziert, dass die Nachfrage nach Rechenzentrumskapazitäten für KI-Anwendungen in einem mittleren Wachstumsszenario zwischen 2023 und 2030 jährlich um durchschnittlich 33 % wachsen wird (McKinsey, 2024).

Im Bereich der KI-Hardware ist eine starke Marktkonzentration festzustellen. Nvidia ist der führende Anbieter von KI-Chips und hält in diesem Marktsegment aktuell einen Marktanteil von 80 % (Investmentweek, 2024). Der Umsatz des Unternehmens hat sich innerhalb eines Jahres verdreifacht (Voigt, dpa, & AFP, 2024). Im ersten Halbjahr 2024 übertrafen die Umsätze der Rechenzentrumssparte von Nvidia die kombinierten Umsätze von Dell und Hewlett Packard Enterprise, zwei marktführender Hersteller für IKT-Hardware, im Bereich Rechenzentren bei weitem (Synergy Research, 2024b).

Im internationalen Vergleich ist festzustellen, dass insbesondere in den USA und China der Ausbau der Rechenzentrumskapazitäten sehr schnell erfolgt. In diesen beiden Ländern wird aktuell vor allem in Rechenleistung für KI-Anwendungen investiert. Europa fällt in Bezug auf den Ausbau von RZ-Kapazitäten etwas zurück.

Die Vereinigten Staaten sind führend im Markt für Rechenzentren, was auf die Dominanz großer Hyperscaler wie Amazon, Microsoft und Google zurückzuführen ist. Diese Unternehmen haben eine bedeutende Konzentration von Rechenzentren in Schlüsselregionen, wie Northern Virginia, Silicon Valley (Kalifornien), Dallas (Texas) und Phoenix (Arizona), aufgebaut. Darüber hinaus spielt die starke Digitalwirtschaft in den USA eine entscheidende Rolle. Die Nachfrage nach Cloud-Diensten und datenintensiven Technologien, wie Big Data und Künstlicher Intelligenz (KI), ist enorm. Eine zusätzliche Triebfeder ist der Bedarf an sicherer Rechenleistung in sensiblen Bereichen, wie Militär, Verteidigung und Regierungsbehörden. Diese spezifische Nachfrage unterstreicht die strategische Bedeutung von Rechenzentren und trägt maßgeblich zum Wachstum bei. Gemäß Synergy Research haben die USA einen Anteil von 44 % an den weltweiten Cloud-Umsätzen, 53 % an den weltweiten Hyperscale -Rechenzentren und 47 % am Markt der Hard- und Software-Cloud-Rechenzentren.

In China wird das Wachstum von Rechenzentren durch staatliche Strategien und Investitionen gelenkt. Die chinesische Regierung unterstützt gezielt die Entwicklung der Digitalwirtschaft und investiert massiv in Infrastrukturen zur Digitalisierung der Wirtschaft. Der Eigenbedarf an Rechenleistungen für nationale Unternehmen, Plattformen und Cloud-Anbieter, wie Alibaba Cloud und Tencent Cloud, ist dabei ein entscheidender Treiber.

Zudem verfolgt China ambitionierte Ziele im Bereich der künstlichen Intelligenz und strebt an, eine globale Führungsrolle in diesem Sektor einzunehmen. Die starke Fokussierung auf KI erfordert eine hochskalierte Datenverarbeitungskapazität, die durch den Ausbau von Rechenzentren gedeckt wird.

In Marktanalysen und -prognosen wird im Vergleich zu den USA und China in Europa ein geringeres Wachstum der RZ-Kapazitäten erwartet (CBRE, 2024b; Cushman & Wakefield, 2024; Granskog et al., 2024; Hintemann et al., 2024a; McKinsey, 2024; Synergy Research, 2024b). Im Detail unterscheiden sich die Wachstumsraten der einzelnen Staaten in Europa jedoch deutlich. Regionen mit günstigen steuerlichen, regulatorischen und klimatischen Bedingungen und niedrigen Strompreisen haben deutliche höhere Rechenzentrumskonzentrationen. Zu nennen sind hier insbesondere Irland, die Niederlande und die skandinavischen Länder.

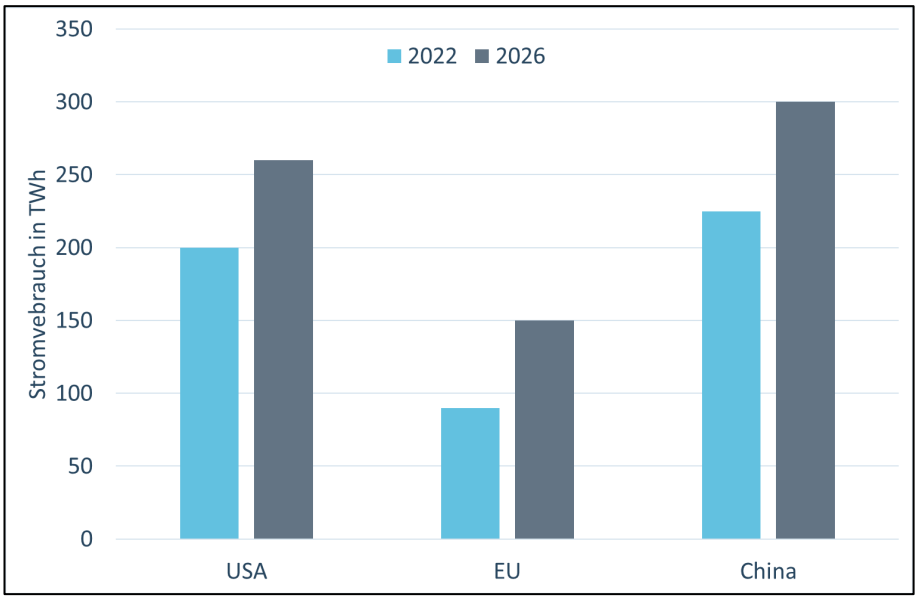


Abbildung 7: Geschätzte Stromverbräuche der Rechenzentren in den USA, der EU und China in den Jahren 2022 und 2026. Quelle: IEA (2024)

Abbildung 7 zeigt die Entwicklung des Stromverbrauchs der Rechenzentren in den USA, in der EU und in China gemäß Schätzung der IEA (2024). Auch dieser Vergleich zeigt, dass vor allem die USA und China den RZ-Markt dominieren.

Auch in anderen Weltregionen wachsen die RZ-Kapazitäten deutlich mit zweistelligen Wachstumsraten (CBRE, 2024b). Besonders stark wachsen die Kapazitäten in Indien, getrieben auch durch den Ausbau der Kapazitäten der großen Hyperscaler (Amazon, 2023a; Haring, 2023; Shankar, 2022). Es wird von einer Verdopplung der RZ-Kapazitäten in Indien zwischen 2023 und 2026 ausgegangen (Babar, 2024; www.ETCIO.com, 2024). Prognosen für ein deutliches Marktwachstum gibt es auch für Afrika, den mittleren Osten (Data Bridge, 2022; Mordor Intelligence, 2024a) und Südamerika (Latina Press, 2024; Mordor Intelligence, 2024c).

Zusammenfassend lässt sich feststellen, dass die Rechenzentrumsmärkte aktuell überall auf der Welt sehr deutlich wachsen. Entscheidend für die Ansiedlung und für die Bedeutung der Rechenzentren sind insbesondere Bevölkerungsdichte, Wirtschaftsleistung und Digitalisierungsgrad der Regionen. Allerdings gibt es auch einige andere Einflussgrößen, die dazu führen, dass sich die Kapazitäten auf einige Weltregionen besonders konzentrieren. Solche Einflussgrößen sind:

- Die Standorte der großen Digitalkonzerne und Global Player: Der internationale Cloud-Markt, die Entwicklungen bei generativer KI und auch der IT-Hardwaremarkt werden insbesondere von Unternehmen aus den USA dominiert. Das führt zu einem massiven Wachstum der RZ-Kapazitäten in den Vereinigten Staaten.
- Die Ansiedlungspolitik der Staaten zu Rechenzentren (insbesondere China, USA, Irland, Skandinavien),
- Die Verfügbarkeit von klimafreundlich produziertem Strom für Großrechenzentren,

- Infrastrukturelle Voraussetzungen, wie gut ausgebaute Telekommunikationsnetze,
- Andere Rahmenbedingungen, wie Außentemperaturen, Steuern und Regularien.

1.3 Deutschland

1.3.1 Wachstum der Rechenzentrumskapazitäten

Mit der Umsetzung digitalpolitischer Ziele aus der EU und Deutschland steigen auch die Bedarfe an digitaler Infrastruktur. Mit etwa 2.000 Rechenzentren ist Deutschland[3] bereits der größte Standort für digitale Infrastruktur in Europa. Neben großen Rechenzentren, die oft von Dienstleistern betrieben werden, gibt es in Deutschland auch ca. 50.000 kleinere IT-Installationen, die insbesondere von mittelständischen Unternehmen betrieben werden.

Die zentrale Lage, politische Stabilität, eine überdurchschnittlich zuverlässige Stromversorgung und einer der größten Internetknoten der Welt machen Deutschland zu einem strategisch vorteilhaften Standort für nationale und internationale Anbieter. Digitalisierungsvorhaben, eine wachsende Digitalwirtschaft und neue Technologien, wie zum Beispiel „Künstliche Intelligenz“, sorgen für eine steigende Nachfrage an Rechenleistung, Speicherkapazität und Datenübertragung. Diese Nachfrage lässt sich allein durch Effizienzsteigerungen bei der Informationstechnik und der Gebäudetechnik nicht decken. Um die Nachfrage zu decken, können neue Rechenzentren gebaut, der Nutzungsgrad bestehender Rechenzentren erhöht, oder Rechenkapazität aus dem europäischen Umland bezogen werden.

Der steigende Bedarf an Rechenleistung, Speicherkapazität und Datenübertragung führt zu einem kontinuierlichen Wachstum der RZ-Kapazitäten in Deutschland, das sich in den letzten Jahren sogar beschleunigt hat (Abbildung 8). Seit 2010 haben sich die Kapazitäten von Rechenzentren und kleineren IT-Installationen in Deutschland von 1.140 MW auf 2.730 MW im Jahr 2024 erhöht. Bis zum Jahr 2030 wird mit einem deutlichen weiteren Anstieg der IT-Kapazitäten auf 4.850 MW gerechnet (Bitkom 2024, ähnlich GDA 2024), was einer Verdopplung gegenüber den Kapazitäten im Jahr 2022 gleichkäme. Setzt sich dieses Wachstum auch danach weiter fort, so werden die Kapazitäten eine Größenordnung von deutlich mehr als 10.000 MW im Jahr 2045 erreichen.

[3] Da der Begriff Rechenzentrum unterschiedlich definiert wird, variiert auch die in der Literatur und Marktstudien genannte Anzahl der Rechenzentren. Die German Datacenter Association geht von 2.000 Rechenzentren in Deutschland mit mehr als 50 kW IT-Leistung aus (German Datacenter Association, 2024). Bitkom nennt 2.000 Rechenzentren mit mehr als 100 kW IT-Anschlussleistung und 3.000 Rechenzentren mit mehr als 40 kW IT-Anschlussleistung (Hintemann, Hinterholzer, & Progni, 2024). Andere Quellen berücksichtigen nur die Dienstleistungsrechenzentren und kommen daher auf geringere Zahlen. Statista meldet 522 Rechenzentren in Deutschland (Statista, 2024c), das Internetportal Datacentermap.com ermittelt am 18.11.2024 die Zahl von 399 Rechenzentren in Deutschland.

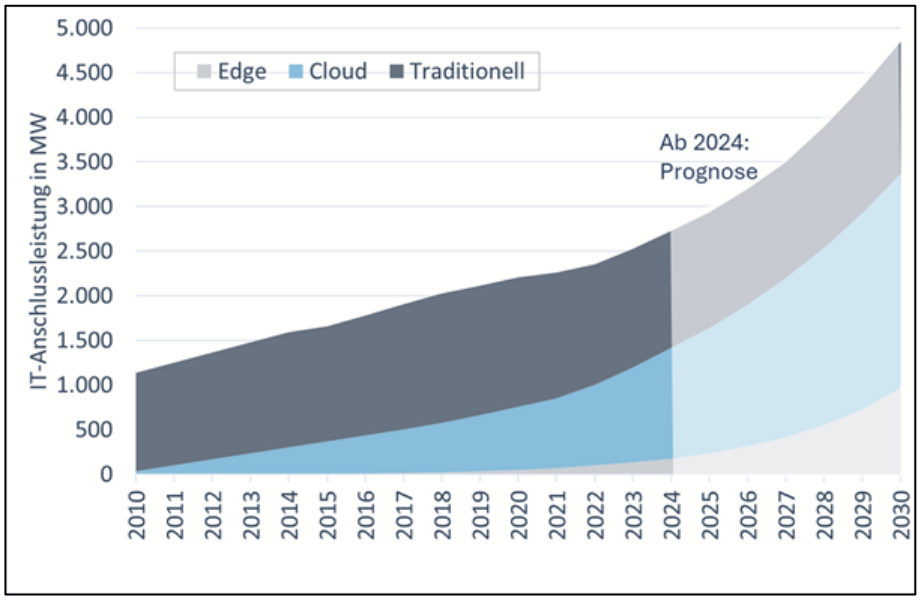


Abbildung 8: Entwicklung der Kapazitäten von Rechenzentren und kleineren IT-Installationen in Deutschland in den Jahren 2010 bis 2024 und Prognose bis 2030. Quelle: Hintemann et al. (2024a)

Ob ein solches langfristiges Wachstum der Rechenzentrumskapazitäten realisiert wird, ist von vielen Faktoren abhängig. Insbesondere folgende Aspekte können zu deutlichen Veränderungen in der Wachstumsdynamik führen:

- Langfristig können technische Änderungen und neue IT-Konzepte, wie Quanten-Computer, optoelektronische Systeme, morphologische Chips und andere, zu erheblichen Änderungen sowohl in der Nachfrage nach Rechenleistung als auch in ihrer Bereitstellung führen.
- Aufgrund der hohen Marktkonzentration im Cloud-Segment sind die Marktentwicklungen von den Entscheidungen weniger großer Digitalkonzerne abhängig. Deren strategische Entscheidungen zur Ansiedlung von Großrechenzentren können daher zu erheblichen Veränderungen der Marktstrukturen führen.
- Aktuell ist festzustellen, dass der RZ-Markt in Deutschland für Immobilien-Spekulationen interessant geworden ist. Ob die angekündigten Neu-Projekte auch wirklich realisiert werden, ist daher unsicher.
- Auch wenn aktuell viele RZ-Großprojekte in Deutschland geplant und teilweise auch bereits umgesetzt werden, ist noch unklar, wie schnell der Endausbau erfolgen wird. Bis ein Rechenzentrumscampus im dreistelligen MW-Bereich ausgebaut und vollständig ausgelastet ist, können mehr als zehn Jahre vergehen. Bei hoher Nachfrage und ohne technische oder sonstige Hemmnisse können die Projekte aber auch in sehr kurzer Zeit realisiert werden.
- Nicht zuletzt hat die verfügbare Stromanschlussleistung einen erheblichen Einfluss auf den Ausbau der Rechenzentrumskapazitäten. Ohne entsprechende Verfügbarkeiten können Groß-Rechenzentren nicht gebaut werden.

Das Wachstum der Kapazitäten findet bei den Rechenzentren mit mehr als 100 kW Anschlussleistung statt. Die Kapazitäten der kleineren IT-Installationen bleiben seit Jahren weitgehend konstant (Abbildung 9). Vor allem die Kapazitäten der Großrechenzentren mit mehr als 5 MW IT-

Anschlussleistung wachsen sehr deutlich. Zwischen 2010 und 2024 haben sich die Kapazitäten in diesem Segment fast um den Faktor fünf von 270 MW auf 1.300 MW erhöht. Aktuell gibt es etwa 100 solcher Rechenzentren in Deutschland. Auch bei den Rechenzentren mit IT-Anschlussleistungen zwischen 100 kW und 5 MW ist ein hohes Wachstum zu verzeichnen – zwischen 2010 und 2020 haben sich die Kapazitäten in dieser Größenklasse von 210 MW auf 670 MW erhöht.

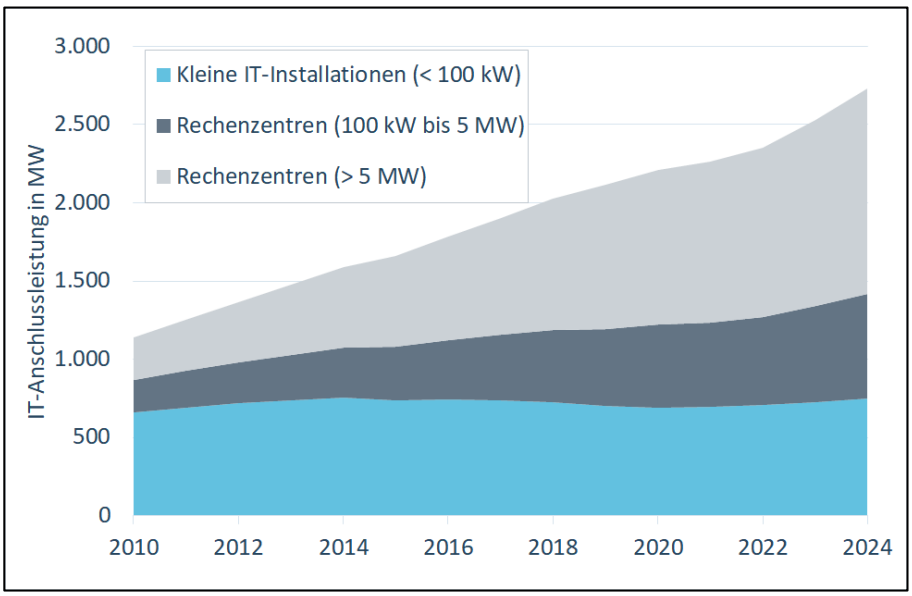


Abbildung 9:Entwicklung der Kapazitäten von Rechenzentren und kleineren IT-Installationen in Deutschland in den Jahren 2010 bis 2024 - Aufteilung in RZ-Größenklassen. Quelle: Hintemann et al. (2024a)

Das Wachstum der Rechenzentrumsbranche fokussiert sich auf wenige Regionen in Deutschland. Vor allem der Großraum Frankfurt/Rhein-Main mit dem Netzwerkknoten DE-CIX ist für die Ansiedlung von Rechenzentren weiterhin attraktiv. Allerdings stößt diese Region zunehmend an Kapazitätsgrenzen, insbesondere was die Verfügbarkeit von ausreichend Stromanschlussleistung angeht. Aktuell werden auch in der Hauptstadtregion mehrere neue Großrechenzentren geplant. Das Milliarden-Engagement von Microsoft im Rheinischen Revier und der damit verbundene Bau mehrere Großrechenzentren lassen vermuten, dass auch in dieser Region in Zukunft die Rechenzentrumskapazitäten deutlich wachsen werden. Die räumliche Konzentration von Rechenzentren und Datennetzen kann mit dem Phänomen der „data gravity“ erklärt werden. Damit ist gemeint, dass große Datenmengen eine „Anziehungskraft“ entwickeln, die weitere Daten anzieht. Begründet wird das Phänomen beispielsweise damit, dass sich große Datenmengen nur noch mit hohem Aufwand verschieben lassen und vielen Anwendungen, die auf die Daten zugreifen, geringe Latenzen erfordern. Verstärkt wird die Konzentration von Großrechenzentren in bestimmten Regionen zusätzlich dadurch, dass große Cloud-Anbieter aus Gründen der Verfügbarkeit meist drei oder mehr Rechenzentren in einer Region aufbauen. Sollte ein Rechenzentrum ausfallen, kann der Cloud-Dienst in der Region weiterhin aus den anderen Rechenzentren angeboten werden.

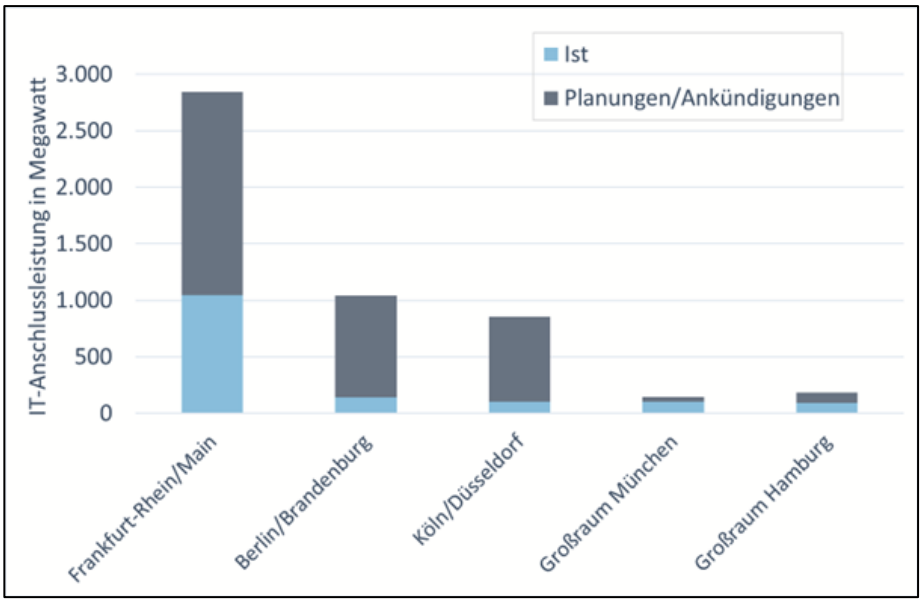


Abbildung 10: Aktuelle und in konkreter Planung befindliche IT-Anschlussleistung von Rechenzentren in verschiedenen deutschen Regionen. Quelle: Hintemann et al. (2024a)

1.3.2 Bedeutung von Cloud Computing

Wie im Rest der Welt wird auch in Deutschland das Wachstum der RZ-Kapazitäten vor allem durch die zunehmende Nutzung von Cloud-Diensten getrieben. Die Möglichkeiten der Künstlichen Intelligenz beflügeln zusätzlich die Fantasie von Branche und Spekulanten (German Datacenter Association, 2024; Goecke, Kempermann, Kestermann, Ewald, & van Baal, 2024; Hintemann et al., 2024a,b). Abbildung 8 zeigt, dass auch das Marktsegment der Edge-Rechenzentren in den letzten Jahren deutlich an Bedeutung zugenommen hat. Bei der Betrachtung dieses Marktsegments ist zu berücksichtigen, dass unter dem Begriff Edge-Rechenzentrum teilweise sehr unterschiedliche Sachverhalte verstanden werden (siehe Kasten Edge Rechenzentren).

Das zunehmende Engagement der Hyperscale-Cloud-Anbieter, wie Amazon, Microsoft oder Google, mit Milliardeninvestitionen (Amazon, 2024; Land NRW, 2024; Tagesschau, 2024a; Wirtschaftswoche, 2023) in den Aufbau von Cloud-Kapazitäten in Deutschland ist zum einen durch technische Faktoren wie Latenzen begründet, zum anderen aber auch durch den Wunsch vieler Cloud-Nutzer, dass ihre Daten im deutschen Rechtsraum oder im Rechtsraum der EU bleiben. Laut KPMG ist es für 58 Prozent der Cloud-Nutzer entscheidend, dass die Anbieter ihre Rechenzentren im Rechtsgebiet der EU betreiben. Durch aktuelle geopolitische Spannungen, Sicherheitsbedenken und Datenschutzanforderungen könnte diese Anforderung weiter an Bedeutung gewinnen (KPMG, 2023). Somit ist wahrscheinlich, dass auch nationale und regionale Cloud-Anbieter von den Anforderungen der inländischen Cloud-Nutzer an Datensicherheit und Souveränität profitieren.

Edge Rechenzentren – ganz unterschiedliche Bedeutungen

Unter Edge Computing wird das Verarbeiten von Daten nahe an ihrem Entstehungsort verstanden, beispielsweise in Kraftfahrzeugen oder in Werkzeugmaschinen. Gemäß diesem Verständnis werden als Edge Rechenzentren solche Rechenzentren bezeichnet, die gezielt aufgrund von geringen Latenzanforderungen oder zur Vermeidung der Übertragung hoher Mengen an Daten nahe am Ort der Datenentstehung aufgebaut werden. Beispiele sind Rechenzentren nahe von Produktionsanlagen in der Industrie, Rechenzentren, die für autonomes Fahren aufgebaut werden oder auch Server an 5G-Basisstationen. Mittlerweile wird der Begriff aber teilweise auch für Colocation-Rechenzentren abseits der großen Rechenzentrumsstandorte London, Frankfurt, Paris, Amsterdam oder Dublin oder regionale IT-Dienstleister verwendet. Teilweise werden auch alle On-Premise-Rechenzentren zu Edge Rechenzentren gezählt. Das Verständnis der vorliegenden Studie basiert auf der ursprünglichen Bedeutung von Edge Rechenzentren und zählt nur diese dazu, die aus technischer Sicht nicht an anderen, etwas entfernteren Standorten errichtet werden können. Gemäß dieser Definition ist das Marktsegment der Rechenzentren aktuell noch klein, weist aber hohe Wachstumsraten auf.

Das allgemeine Wachstum der Rechenzentren in Deutschland wird aktuell vor allem vom Marktsegment der Colocation-Anbieter getragen. Im Jahr 2024 befindet sich fast die Hälfte der RZ-Kapazitäten in Deutschland in Colocation-Gebäuden. Vor allem die Hyperscaler, wie Amazon, Microsoft und Google, setzen beim Aufbau ihrer Kapazitäten in Deutschland bislang zumeist auf Colocation-Angebote. Aber auch nationale Cloud-Anbieter betreiben ihre IT-Hardware auf Colocation-Flächen.

Auch wenn das Marktwachstum bei Colocation-Rechenzentren hauptsächlich von Cloud-Anbietern bestimmt ist, profitiert das Marktsegment aber ebenso von dem Trend, dass Unternehmen und Behörden zunehmend die eigene IT nicht mehr „on premise" betreiben. Dieser Trend führt dazu, dass das Marktsegment der On-Premise-Rechenzentren trotz zunehmender Digitalisierung auf einem Niveau von 1 GW weitgehend konstant bleibt und es im wachsenden Rechenzentrumsmarkt an Bedeutung verliert. Aktuell werden aufgrund der wirtschaftlichen Lage gerade bei kleineren On-Premise-Rechenzentren sogar Neu-Investitionen zurückgestellt (Hintemann et al., 2024a).

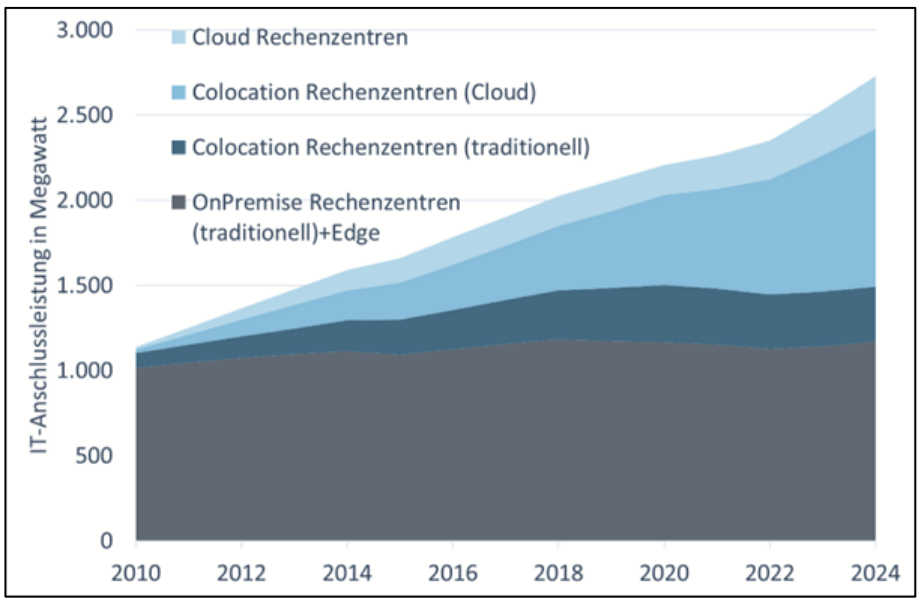


Abbildung 11: Entwicklung der Kapazitäten von Rechenzentren in Deutschland mit Anteilen von Cloud, Colocation, On Premise und Edge Rechenzentren (IT-Anschlussleistung) in den Jahren 2010 bis 2024. Quelle: Hintemann et al. (2024a)

1.4 Europäische und internationale Einordnung der deutschen Marktentwicklungen und -trends

1.4.1 Die deutsche Position in Europa

Im internationalen Kontext ist die Entwicklung der Rechenzentrumsbranche in Deutschland differenziert zu betrachten. Innerhalb von Europa hat Deutschland seine Position in den letzten Jahren stabilisiert und leicht ausgebaut. Deutschland ist gemessen an den Kapazitäten aller Rechenzentren und kleineren IT-Installationen der größte RZ-Standort in der EU27 plus Großbritannien (Abbildung 12[4]). Dies ist vor allem auf die starke Wirtschaftskraft Deutschlands und das damit verbundene Digitalisierungspotenzial zurückzuführen. Im Vergleich zu anderen Studien in der Vergangenheit kommt die aktuelle Analyse zu dem Schluss, dass die Rechenzentrumskapazitäten in Deutschland größer sind als die in Großbritannien. Dies ist insbesondere damit zu begründen, dass die vorliegende Untersuchung auch On-Premise-Rechenzentren und kleinere IT-Installationen berücksichtigt, während

[4] Basis für die Abschätzung bilden Daten zu Serverumsätze, Serverstückzahlen und Umsätzen im Colocation-Markt (z. B. Arizton, 2024; CBRE, 2023; Cushman & Wakefield, 2024; Datacentermap.com, 2024; German Datacenter Association, 2024; Hintemann, Graß, Hinterholzer, & Grothey, 2022; Hintemann & Hinterholzer, 2020; Hintemann, Hinterholzer, & Seibel, 2023; JLL, 2024; Mordor Intelligence, 2024b; Statista, 2024a, 2024b; Taylor, 2024)

viele Marktanalysten nur auf den Markt der Colocation-Kapazitäten abzielen. Aufgrund der deutlich höheren Wirtschaftsleistung in Deutschland ist es plausibel, dass die Anzahl der von Unternehmen und Behörden für eigene Zwecke betriebenen On-Premise-Rechenzentren höher liegt als in Großbritannien. Aber auch im Colocation-Segment wird der deutsche Markt im Vergleich immer stärker, wie aktuelle Marktanalysen zeigen (CBRE, 2024a; JLL, 2024). Mit 399 Rechenzentren sind auf www.datacentermap.com mehr Colocation-Rechenzentren in Deutschland erfasst als in Großbritannien (377). Die von Marktanalysten ermittelten Verkaufszahlen für Server liegen ebenfalls in Deutschland deutlich über denen in Großbritannien.

Der deutliche Anstieg der Rechenzentrumskapazitäten in Irland wird auch durch die steigenden Stromverbräuche der dortigen Rechenzentren dokumentiert. Nach Angaben des irischen Statistikamts stieg der Stromverbrauch der Rechenzentren in Irland zwischen 2015 und 2023 um mehr als den Faktor fünf und war im Jahr 2023 mit 6,33 TWh für 21 % des irischen Stromverbrauchs verantwortlich. Aufgrund dieses sehr hohen Anteils wurden vom staatlichen irischen Stromnetzbetreiber EirGrid Maßnahmen ergriffen, die künftige Ansiedlung von Rechenzentren so zu gestalten, dass sie keine negativen Auswirkungen auf die Stabilität des Stromnetzes haben. So werden Rechenzentrumsbetreiber beispielsweise dazu verpflichtet, Flexibilitäten für das Stromnetz bereitzustellen (DER STANDARD, 2022; Judge, 2023; Krempl, 2022).

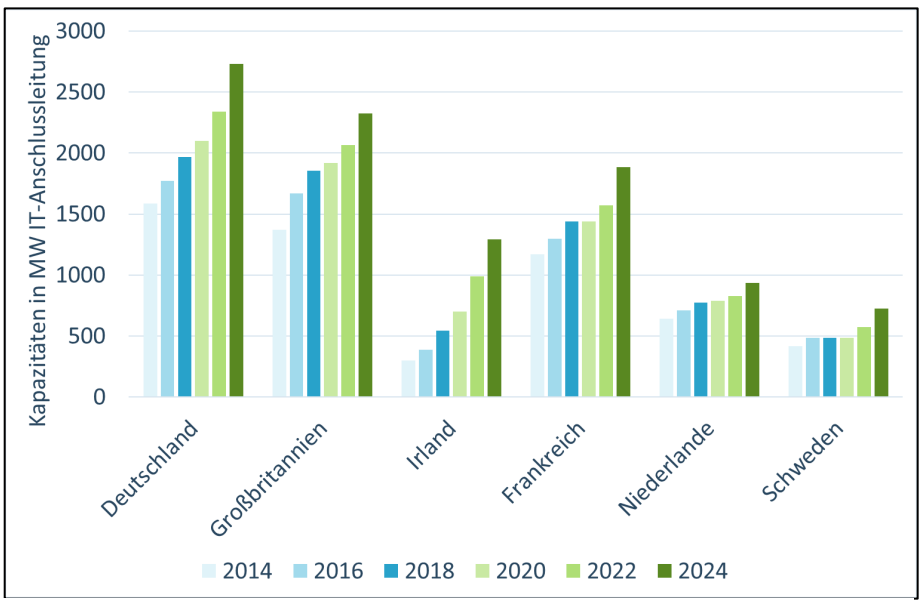


Abbildung 12: Entwicklung der Kapazitäten der Rechenzentren in den Jahren 2016 bis 2024 in ausgewählten europäischen Ländern. Quelle: Berechnungen Borderstep

Bezieht man die Rechenzentrumskapazitäten jeweils auf das Bruttoinlandsprodukt der Staaten, so liegt Deutschland bei den hier betrachteten Staaten mit 0,61 MW IT-Anschlussleistung pro Mrd. Euro BIP im Jahr 2023 im europäischen Mittelfeld. Insbesondere Irland (2,24 MW/Mrd. €) und Schweden (1,20 MW/Mrd. €) haben, bezogen auf das BIP, deutlich höhere RZ-Kapazitäten. Aber auch in den Niederlanden (0,82 MW/Mrd. €) und in Großbritannien (0,68 MW/Mrd. €) sind vergleichsweise mehr RZ-Kapazitäten vorhanden (Hintemann et al. 2024a).

1.4.2 Die deutsche und europäische Position im internationalen Vergleich

Vergleicht man die Entwicklung der RZ-Kapazitäten in Europa insgesamt mit anderen Weltregionen, so fallen Deutschland und Europa in Zukunft im Vergleich mit den USA voraussichtlich deutlich zurück (Abbildung 13). Im Vergleich zu China sind die erwarteten Wachstumsraten zwar vergleichbar, gemessen in absoluten Werten und auch in Bezug auf das BIP liegt Europa aber deutlich hinter China (2,1 MW/Mrd. €). Bereits in der Vergangenheit hat der Anteil Deutschlands an den weltweit installierten IT-Kapazitäten abgenommen. Zwischen 2015 und 2024 sank der Anteil der in den Rechenzentren in Deutschland installierten Server am Weltmarkt von 3,5 % auf 2,5 % (Hintemann et al., 2024a). Ein wesentlicher Grund für diesen Rückgang ist, dass Deutschland hinsichtlich der Nutzung von Cloud-Diensten weitestgehend ein Importland ist. Mit dem Trend zu immer mehr Anwendungen im Bereich der Künstlichen Intelligenz ist kaum davon auszugehen, dass sich die Situation ändert. Die Nachfrage nach spezieller Hardware, die für KI-Anwendungen genutzt wird, führt aktuell dazu, dass sich insbesondere in den USA und in China das Wachstum noch einmal deutlich zu beschleunigen scheint.

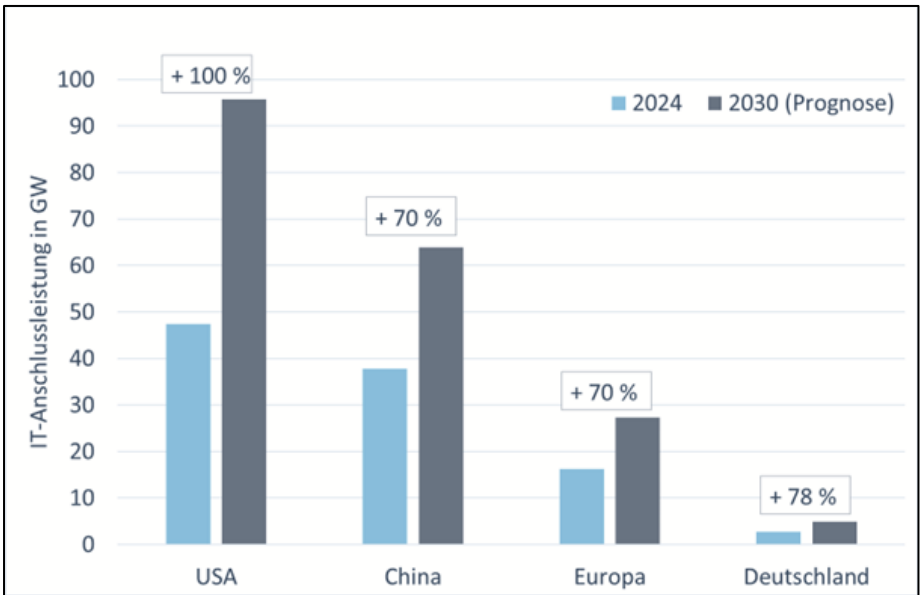


Abbildung 13: Abschätzung der Entwicklung der RZ-Kapazitäten in den USA, in China, Europa und Deutschland 2024 und 2030. Quelle: Hintemann et al. (2024a)

Als Zwischenfazit lässt sich feststellen, dass das Wachstum der Rechenzentren in Deutschland primär durch das hohe Marktpotenzial für Digitalisierung z. B. in der Industrie und im Mittelstand geprägt ist und weniger, als zum Beispiel in den USA, von einer starken Digitalwirtschaft. Aus Deutschland heraus werden nur wenig IT-Infrastruktur und -Dienstleistungen in andere Länder vermarktet – anders als z. B. in Irland, den skandinavischen Ländern und den Niederlanden. Insofern kommt einer gut ausgebauten Glasfaser-Infrastruktur in Deutschland eine besondere Bedeutung zu, um die Digitalisierung insgesamt voranzubringen und für Industrie und Mittelstand gute Voraussetzungen zu schaffen, Cloud- und IT-Dienste effizient und effektiv zu nutzen. Der Ausbau von internationalen Datenverbindungen könnte zudem zu einem verbesserten Exportpotenzial von digitalen Ressourcen aus Deutschland führen.

1.4.3 Stärken des Rechenzentrumsstandortes Deutschland

Im Rahmen einer SWOT-Analyse wurde der Rechenzentrumsstandort Deutschland detailliert mit anderen Standorten in Europa verglichen. Dabei wurde festgestellt, dass Deutschland zusammen mit den Niederlanden, Schweden und Frankreich zu den Ländern mit einer insgesamt stärkeren Digitalpolitik gehört. In diesem Sinne ist es besser positioniert als Polen, das Vereinigte Königreich oder Irland – Länder, die mit Energieabhängigkeit, regionalen Ungleichgewichten oder der Abhängigkeit von nicht-nachhaltigen Energiequellen (wie Polen) zu kämpfen haben.

Deutschlands Stärken liegen hauptsächlich in seinem umfassenden regulatorischen Rahmen und seinem starken Engagement für erneuerbare Energien, was sich auch in Gesetzen wie dem Erneuerbare-Energien-Gesetz (EEG) und dem EnEfG widerspiegelt. Dies kann Deutschland als führendes Land im Bereich des nachhaltigen digitalen Infrastrukturausbaus positionieren, insbesondere im Hinblick auf energieeffiziente und nachhaltige Rechenzentren. Auch die zuverlässige Stromversorgung und die hohe Dichte an Internetknoten sind Stärken des Standortes.

Das Engagement, mit dem Deutschland den Bereich der nachhaltigen Digitalisierung vorantreibt, zeigt sich auch in der Digitalstrategie 2025. Diese ist zukunftsorientiert und fokussiert auf den Ausbau der Infrastruktur – zur Überbrückung der digitalen Kluft, wodurch Deutschland anderen Ländern voraus ist. Was die Nachhaltigkeit angeht, haben sowohl Deutschland als auch Schweden sehr ehrgeizige Ziele für erneuerbare Energien. Schweden hat kürzlich Steueranreize für Rechenzentren abgeschafft, da das Land mit einem ungeplanten Wachstum des Krypto-Minings und weniger langfristig orientierten ausländischen Investitionen konfrontiert war. Derzeit übertrifft Deutschland die Niederlande und Irland in Bezug auf die Energieinfrastruktur. Während die Niederlande mit Problemen bei der Netzkapazität zu kämpfen haben und Irland unter der Abhängigkeit von Energieimporten leidet, könnte Deutschlands Fokus auf erneuerbare Energien und Energieeffizienz eine bessere Unterstützung für den künftigen Ausbau der digitalen Infrastruktur bieten. Insgesamt hebt sich Deutschlands ausgewogener Ansatz in Bezug auf digitale Infrastrukturen ab.

Aktuell sind die USA die Heimat großer Digitalunternehmen und -konzerne und führend in F&E-Investitionen für digitale Infrastrukturen, was auch weltweit große Investitionen in Rechenzentren anzieht. Deutschland gibt ebenfalls viel Geld für F&E aus (für erneuerbare Energien und energieeffiziente Technologien), jedoch nicht speziell, um seine Rechenzentrumslandschaft zu stärken.

1.5 Handlungsempfehlungen

Auch wenn Deutschland hinsichtlich digitalpolitischer Strategie und Rechtsrahmen innerhalb von Europa vergleichsweise gut aufgestellt ist, gibt es bislang keine konkret auf die Rechenzentrumsbranche abgestimmte Strategie. Die Entwicklung einer solchen Strategie – unter Berücksichtigung der Erkenntnisse der folgenden Kapitel des Gutachtens – wird empfohlen. Damit könnten die Wettbewerbsposition Deutschlands beim Ausbau digitaler Infrastrukturen verbessert werden und gezielt die Segmente des Rechenzentrumsmarktes gefördert werden, die besonders zu wirtschaftlichem Wachstum und digitaler Souveränität beitragen.

Trotz starker Richtlinien zu erneuerbaren Energien und Nachhaltigkeit können das Fehlen von Highspeed-Internet in einigen Regionen und der Mangel an flächendeckender Glasfaser-Konnektivität den Fortschritt und die Entwicklung verlangsamen und das Wachstum der digitalen Wirtschaft sowie Digitalisierungsbemühungen hemmen.

Die Bemühungen Irlands, die Flexibilität und Widerstandsfähigkeit des Energienetzes zu gewährleisten und zu verbessern, indem Rechenzentren verpflichtet werden, Flexibilitäten für das Stromnetz bereitzustellen, könnten ein Vorbild sein, von dem auch Deutschland profitieren könnte.

Insgesamt ist festzustellen, dass die Datenlage zur Entwicklung des Rechenzentrumsmarktes noch immer unzureichend ist. Mit dem in Deutschland eingeführten RZReg wurde ein wichtiger Meilenstein für mehr Transparenz geschaffen. Die vorliegenden Daten sind aber offensichtlich noch nicht vollständig (siehe Kapitel 2). Auch die Abschätzungen im vorliegenden Gutachten sind mit Unsicherheiten behaftet. Es fehlen immer noch wesentliche Informationen, beispielsweise zum eigentlichen Nutzungsgrad und der Effizienz des IT-Equipments in den Rechenzentren. Auch Daten zu kleineren Rechenzentren sind bislang kaum verfügbar.

2 Energiesystem und Rechenzentren

2.1 Zusammenfassung

In Kapitel 2 werden die Entwicklung des Strombedarfs und der Energieeffizienz von Rechenzentren in Deutschland analysiert und Prognosen bis 2045 formuliert. Aufgrund des zunehmenden Bedarfs an digitaler Infrastruktur, insbesondere durch Cloud-Dienste und die Anwendung von Künstlicher Intelligenz, wird erwartet, dass der Stromverbrauch der Rechenzentren weiter stark ansteigen wird – von derzeit etwa 20 TWh/a auf bis zu knapp 80 TWh/a bis 2045. Effizienzgewinne, etwa durch Verbesserungen bei der PUE, haben das Verbrauchswachstum bisher nicht kompensieren können. Der PUE-Wert deutscher Rechenzentren hat sich zwischen 2010 und 2024 zwar von 1,80 auf 1,46 verbessert, jedoch erhöht die steigende Nachfrage nach digitalen Ressourcen weiterhin den Gesamtenergiebedarf (Hintemann et al 2024a).

Darüber hinaus beschreibt dieses Kapitel die Herausforderungen bei der Integration von Rechenzentren in die Energieinfrastruktur. Um den steigenden Energiebedarf vor dem Hintergrund der Klimaziele vorrangig durch erneuerbare Energien zu decken, spielen Power Purchase Agreements (PPA) eine wesentliche Rolle. Allerdings bestehen Bedenken hinsichtlich der Verfügbarkeit ausreichender erneuerbarer Kapazitäten, da Genehmigungsverfahren für neue Anlagen oft langwierig sind.

Ein weiterer Schwerpunkt liegt auf der Nutzung der durch Rechenzentren entstehenden Abwärme, die theoretisch ein erhebliches Potenzial bietet, da der von den Rechenzentren benötigte Strom vollständig in Wärme umgewandelt wird. Praktische und technische Einschränkungen, wie die Notwendigkeit von Wärmepumpen zur Temperaturanhebung, die saisonale Schwankung in der Wärmenachfrage und die oft geringe Verfügbarkeit von Wärmenetzen in der Nähe begrenzen jedoch die tatsächliche Nutzbarkeit in vielen Fällen massiv.

Um die Abwärmenutzung zu fördern, empfiehlt sich eine verbesserte Koordination und finanzielle Unterstützung bei Infrastrukturinvestitionen für Wärmenetze. Zudem könnten staatliche Bürgschaften Investitionen in Wärmenetze erleichtern.

Daneben werden regulatorische Anpassungen vorgeschlagen, etwa die Einführung eines deutschlandweiten Wärmenetzregisters, das mehr Transparenz über die Planung, die Verfügbarkeit und die Nutzung von Wärmenetzen bieten könnte. Auch das EnEfG, das ab 2027 eine Deckung des Strombedarfs der Rechenzentren durch erneuerbare Energien vorschreibt, wird als wichtiger Schritt zur Reduktion der Emissionen betrachtet.

Abschließend identifiziert das Kapitel Herausforderungen im Netzausbau und bei der Standortwahl für Rechenzentren. Der Ausbau der Stromnetze ist notwendig, um die Stromversorgung zu sichern und zusätzliche Lasten durch Rechenzentren sowie erneuerbare Energiequellen zu integrieren. Rechenzentren, die in ihrer Standortwahl flexibel sind, können so errichtet werden, dass ihre Integration in die Energienetze Vorteile für beide Seiten bringt. Sie können lokal mit erneuerbarem Strom versorgt werden, Flexibilitäten bereitstellen oder auch große Teile ihrer Abwärme für Wärmesenken, wie Fernwärmenetze oder Industrieanwendungen, liefern und damit einen Beitrag zur Energiewende leisten.

2.2 Energieverbrauch, Energieeffizienz und Klimawirkungen von Rechenzentren

2.2.1 Strombedarf

Wie die Ausführungen in Kapitel 1 zeigen, steigen die Kapazitäten der Rechenzentren – gemessen an der IT-Anschlussleistung – signifikant. Dieses Wachstum führt zu einer deutlichen Zunahme des Stromverbrauchs, da ein Rückgang in der Auslastung der Kapazitäten nicht zu erwarten ist. Die stark wachsende Bedeutung von hoch ausgelasteten Systemen für Cloud-Dienste und Anwendungen der Künstlichen Intelligenz (KI) lässt insgesamt einen Anstieg in der Auslastung der Rechenzentren erwarten.

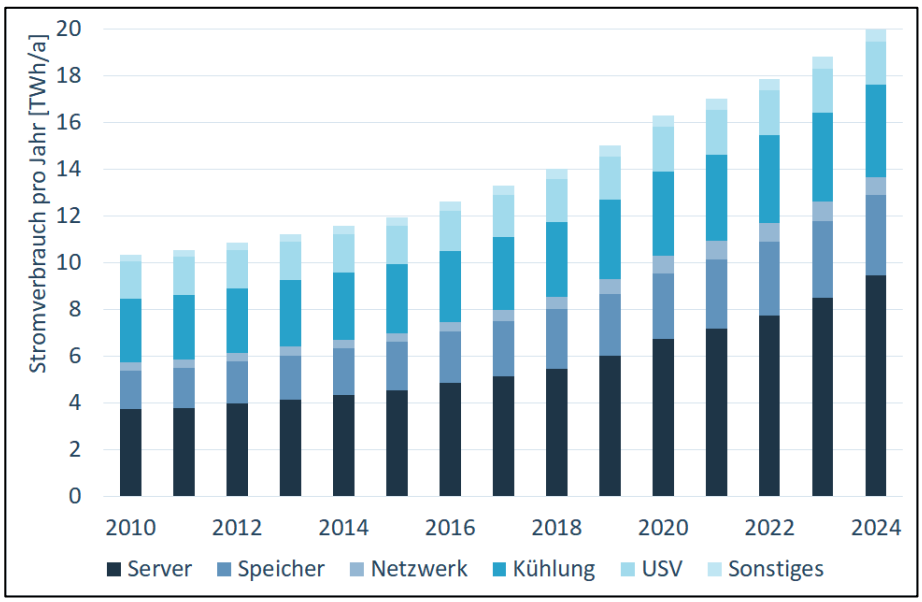


Abbildung 14: Entwicklung des Stromverbrauchs der Rechenzentren und kleineren IT-Installationen in Deutschland in den Jahren 2010 bis 2024. Quelle: Hintemann et al. (2024a)

Abbildung 14 zeigt die Entwicklung des Stromverbrauchs von Rechenzentren und kleineren IT-Installationen zwischen 2010 und 2024. Aus der Abbildung wird deutlich, dass der Stromverbrauch der Rechenzentren im Zeitraum von 2010 bis 2024 um über ca. 90 % auf 20 TWh/a angestiegen ist. Angesichts des dynamischen Wachstums der Branche ist auch in der Zukunft von einem weiteren deutlichen Anstieg des Stromverbrauchs auszugehen. Basierend auf den Prognosen für Hardwareverkäufe (Hintemann, Hinterholzer, & Konrat, 2024; Hintemann, Hinterholzer, & Progni, 2024; IDC, 2024; Statista, 2024b) und die Erweiterung von Rechenzentrumskapazitäten in Deutschland (CBRE, 2024c; German Datacenter Association, 2024; Hintemann, Hinterholzer, & Progni, 2024; JLL, 2024) dürfte der Stromverbrauch bis 2030 auf etwa 31 TWh ansteigen.

Die Berechnungen der Entwicklung des Stromverbrauchs basieren auf einem Modell der Rechenzentrumslandschaft in Deutschland.[5] Das Modell baut auf den Verkaufszahlen für IT-Hardware, wie Server, Speichersysteme und Netzwerktechnik, auf. Für unterschiedliche Gerätetypen werden jeweils typische Nutzungsdauern und Stromverbräuche ermittelt. Mit dem Modell lässt sich beispielsweise ermitteln, dass zwischen 2010 und 2024 die Gesamtzahl aller Server in Deutschland von 1,9 Millionen auf 2,4 Millionen angestiegen ist. Gleichzeitig stieg die durchschnittliche Leistungsaufnahme pro Server von 223 Watt auf 443 Watt (gemittelt über alle Servertypen). Der durchschnittliche PUE-Wert der Rechenzentren und kleineren IT-Installationen verbesserte sich im genannten Zeitraum von 1,8 auf 1,46. Die Annahmen im Modell basieren auf Befragungen, verfügbaren Marktstatistiken und technischen Daten der verkauften Hardwaresysteme und sind mit den Annahmen in anderen Studien und Statistiken vergleichbar (LBNL, 2024; SPEC, 2024; Stobbe, 2024; The Green Grid, 2024).

Die mit dem Modell abgeschätzte Höhe des Stromverbrauchs der Rechenzentren und kleinerer IT-Installationen in Deutschland ist auch beim Vergleich mit anderen Studien plausibel. Andere Untersuchungen kommen mit einem abgeschätzten Stromverbrauch der IT (ohne Gebäudetechnik) von etwa 17,5 TWh im Jahr 2024 und 25 TWh im Jahr 2030 zu etwas höheren Werten (Stobbe et al., 2024). Berücksichtigt man den Anteil von Colocation- und Hyperscale-Rechenzentren an den gesamten RZ-Kapazitäten in Deutschland (siehe dazu auch Abbildung 11), so wird der im vorliegenden Gutachten ermittelte Stromverbrauch auch durch eine Studie von BloombergNEF bestätigt, die für 2021 einen Stromverbrauch von 7,2 TWh für dieses Segment ermittelte (BloombergNEF, 2021). Gemessen am Anteil Deutschlands an der Wirtschaftsleistung der EU sind die berechneten Stromverbräuche auch im Vergleich mit entsprechenden Studien zum Stromverbrauch der Rechenzentren in der EU plausibel (Granskog et al., 2024; IEA, 2024; Kamiya & Bertoldi, 2024). Aufgrund der Vielzahl an Annahmen in dem Berechnungsmodell handelt es sich aber selbstverständlich bei den Zahlen nur um eine Abschätzung. Der tatsächliche Stromverbrauch der Rechenzentren in Deutschland kann hiervon abweichen.

Wirkungseffekt des EnEfG

Durch das EnEfG sind Vorgaben hinsichtlich der Effizienz der Gebäudetechnik in Deutschland eingeführt worden. Damit nimmt Deutschland international eine Vorreiterrolle ein. Diese Vorgaben haben eine Auswirkung auf die Entwicklung des Rechenzentrumsmarktes in Deutschland und insbesondere auf den Stromverbrauch. Diese Auswirkungen werden in Abbildung 15 abgeschätzt. Es wird erwartet, dass durch das EnEfG der Stromverbrauch der Rechenzentren und kleineren IT-Installationen im Jahr 2030 um 2,5 TWh niedriger liegt als ohne Einführung des Gesetzes.

[5] Die Berechnung der Stromverbräuche wurde mit dem in Kapitel 1 bereits vorgestellten Borderstep-Modell der Rechenzentrumslandschaft in Deutschland vorgenommen. Das Modell ist umfangreich dokumentiert (z. B. in Fichter & Hintemann 2014, Stobbe et al. 2015, Hintemann & Hinterholzer 2023)

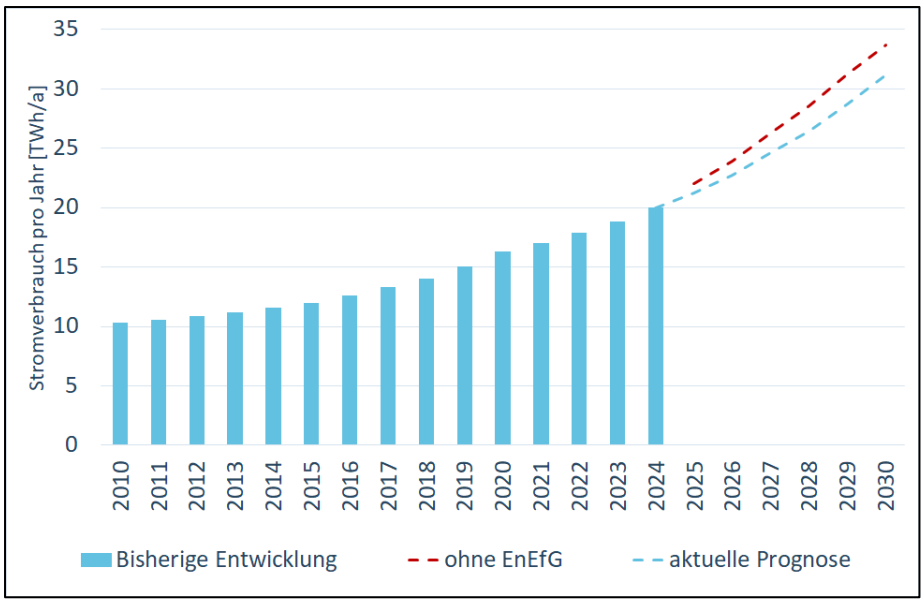


Abbildung 15: Entwicklung des Stromverbrauchs der Rechenzentren und kleineren IT-Installationen in Deutschland in den Jahren 2010 bis 2024 und Prognose bis 2030 inkl. Abschätzung der Auswirkungen des EnEfG

Der Abschätzung liegt die Annahme zugrunde, dass sich ohne EnEfG der durchschnittliche PUE-Wert in Deutschland nicht, wie in der aktuellen Prognose, von 1,46 im Jahr 2024 auf 1,25, sondern nur auf 1,35 im Jahr 2030 verbessern würde. Diese stärkere Verbesserung wird vor allem durch die Auswirkungen der EnEfG-Regelung auf neue Colocation-Rechenzentren und bestehende Rechenzentren erreicht. Rechenzentren von Hyperscale-Cloud-Anbietern erreichen international schon länger die im EnEfG geforderten Grenzwerte. Vereinfachend wird angenommen, dass die Einführung des Gesetzes keine Auswirkungen auf die Ansiedlungsabsichten für Rechenzentren und die Entwicklung der Rechenzentrumskapazitäten hat. Aus aktueller Sicht scheint das Gesetz zumindest hinsichtlich der Volumina bei der Planung und Projektierung von neuen großen Rechenzentren, insbesondere von Hyperscale-Cloud-Anbietern, keine spürbaren negativen Wirkungen zu haben. Gemäß Immobiliendienstleister CBRE wird das EnEfG von vielen Marktteilnehmern sogar als mittel- bis langfristiger Pull-Faktor für den deutschen RZ-Markt eingestuft (CBRE 2024c).

Langfrist-Prognose

Langfristige Prognosen über das Jahr 2030 hinaus sind mit erheblichen Unsicherheiten verbunden. Die raschen technologischen Fortschritte und Marktentwicklungen im Bereich der Digitalisierung und der Innovationen der Digitalwirtschaft könnten zu fundamentalen Veränderungen führen. Besondere Unsicherheitsfaktoren sind die Weiterentwicklung der KI-Technologien, Veränderungen bei den RZ-Konzepten wie Edge-Rechenzentren, Veränderungen von Investitionsentscheidungen großer Digitalkonzerne sowie technologische Innovationen wie neuartige Chiptechnologien, optoelektronische Entwicklungen und potenziell revolutionäre Systeme wie Quantencomputer. Auch die Verfügbarkeit von Stromanschlussleistungen kann die Entwicklung von Rechenzentrumskapazitäten limitieren und somit ihren Stromverbrauch begrenzen. Bereits heute sichern sich Hyperscale-Cloud-Anbieter durch langfristige Energieverträge und Investitionen in Energieinfrastruktur große Stromkapazitäten für ihre Rechenzentren. Dies ist besonders in den USA sichtbar, wo einige Konzerne in Energieverträge

bestehender Atomkraftwerke oder neue Atomenergie-Technologien investieren (Hermann, 2024; Spiegel, 2024; Swinhoe, 2024; Tagesschau, 2024b).

Trotz dieser Unsicherheiten wird auf Wunsch des Auftraggebers eine Prognose bis zum Jahr 2045 vorgenommen (Abbildung 16). Unter der Annahme, dass sich die Branche in einem vergleichbaren Rhythmus wie seit 2010 weiterentwickelt, wird der Stromverbrauch der Rechenzentren auf etwa 80 TWh im Jahr 2045 ansteigen. Trotz aller Unwägbarkeiten erscheint diese Prognose nicht unplausibel, da auch in der Vergangenheit die technologische und marktseitige Entwicklung zu einem relativ kontinuierlich steigenden Stromverbrauch geführt haben. Die Prognose berücksichtigt allerdings nicht mögliche Engpässe im Stromnetz (siehe Kapitel 2.3) oder Sättigungserscheinungen bei der Digitalisierung (siehe Kapitel 5). Solche Rahmenbedingungen könnten dazu führen, dass insbesondere ab Mitte der 2030er Jahre der Stromverbrauch nicht weiter mit steigenden Wachstumsraten zunimmt.

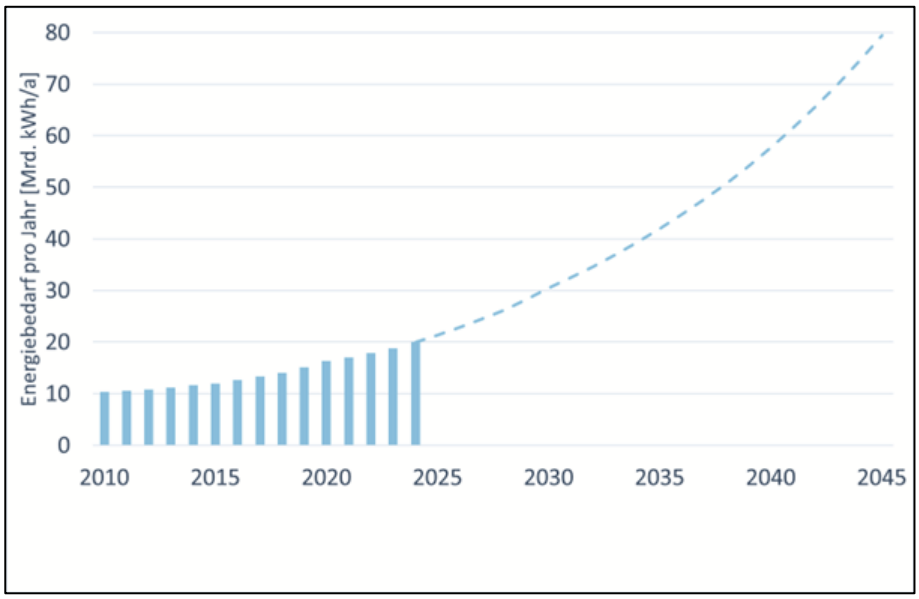


Abbildung 16: Entwicklung des Strombedarfs der Rechenzentren in Deutschland in den Jahren 2010 bis 2024 und Langfristprognose bis 2045. Quelle: Berechnungen Borderstep

2.2.2 Energieeffizienz

Innerhalb der Rechenzentrumsbranche hat sich die Kennzahl Power Usage Effectiveness (PUE) als Maß für die Effizienz der Gebäudetechnik und der Infrastruktur durchgesetzt. Sie ist innerhalb der DIN EN 50600-4-2 definiert (PUE ist in Kapitel 4.4 erläutert). Der PUE-Wert gibt das Verhältnis des Gesamtenergieverbrauches eines Rechenzentrums zum Energieverbrauch der IT an und ist per Definition immer größer als 1. Eine Aussage über die Effizienz der IT-Infrastruktur und damit des gesamten Rechenzentrums kann damit aber nicht getroffen werden, sondern lediglich Aussagen zur elektrischen- und mechanischen Effizienz der Gebäudetechnik und Hilfsinfrastruktur. Vor allem das Uptime Institut hat über die letzten Jahre Studien zur globalen Entwicklung des PUE-Werts unter Berücksichtigung verschiedener Einflussfaktoren erstellt und veröffentlicht. Diese beruhen auf Befragungen von Rechenzentrumsbetreibern. Uptime stellt Auswertungen über die zeitliche

Entwicklung, vergangene Jahre seit der Inbetriebnahme, den Einfluss der Geografie als auch die vorhandene IT-Kapazität zur Verfügung.

Von 2007 bis 2014 ist der PUE-Wert in den Befragungen im weltweiten Durchschnitt gesunken (Uptime Institute, 2022). Seitdem lag der PUE-Wert im Bereich von 1,55 bis 1,67, wobei der Trend der Abnahme von 2019 bis 2022 im Jahr 2023 unterbrochen wurde (Bizo, 2023). 2024 sank der PUE-Wert wieder minimal (Uptime Institute, 2024). Insgesamt zeigt sich, dass im globalen Durchschnitt ein PUE-Wert von weniger als 1,55 nicht erreicht wird und die Branche insgesamt hinsichtlich der Energieeffizienz der Gebäudetechnik stagniert. Die großen Effizienzgewinne im ersten Jahrzehnt sind die Folge der Ausnutzung von einfach umsetzbaren Maßnahmen, wie Kalt-/ Warmgangeinhausung oder Anhebung der Kühllufttemperaturen (siehe auch Kapitel 5).

Rechenzentren mit einer Leistung von mehr als 1 MW haben in einem Zeitraum der Inbetriebnahme von weniger als 5 Jahren im Durchschnitt den niedrigsten PUE-Wert und weisen gegenüber dem Zeitraum 6 bis 10 Jahre einen PUE-Vorteil von 0,11 und ab 11 Jahren mehr als 0,18 auf (Bizo, 2023). Dies ist die Folge der Implementierung neuer technischer Fortschritte bei Neubauten, die dadurch einen Effizienzvorteil aufweisen.

Hinsichtlich der Geografie hat Europa im Vergleich zu allen anderen Regionen einen PUE-Vorteil von 0,07 bis 0,33, je nach Vergleichsregion (Lawrence, 2020). Ursache hierfür sind zum einen die unterschiedlichen klimatischen Bedingungen und möglicherweise auch die Energiekosten. Durch das mildere Klima kann das Potenzial der freien Kühlung stärker genutzt werden, und die hohen Energiepreise liefern einen Anreiz für den effizienten Umgang mit Energie. Die entsprechende Studie hat aber die Einschränkung von teilweise sehr niedrigen zweistelligen oder gar nur einstelligen Stichprobengrößen für einzelne Regionen.

Die IT-Kapazität wiederum korreliert auch mit dem PUE-Wert. Im Regelfall gilt, je höher die IT-Kapazität eines Rechenzentrums, umso niedriger der PUE-Wert. Während im Bereich ≤ 99 kW ein gewichteter durchschnittlicher PUE von knapp 1,7 ermittelt wurde, sinkt dieser Wert bis auf 1,44 bei ≥ 30 MW (Davis, 2024).

In Deutschland hat sich der PUE-Wert, wie oben bereits erwähnt, zwischen 2010 und 2024 von 1,80 auf 1,46 verbessert (Hintemann et al. 2024a). Auch hier ist dies eine Folge der Fortschritte bei der Technik der Gebäudeinfrastruktur und vergleichbar mit dem globalen Trend. Für die Zukunft ist zu erwarten, dass sich diese Effizienzverbesserung aufgrund der Vorgaben im EnEfG weiter fortsetzt (Abbildung 17).

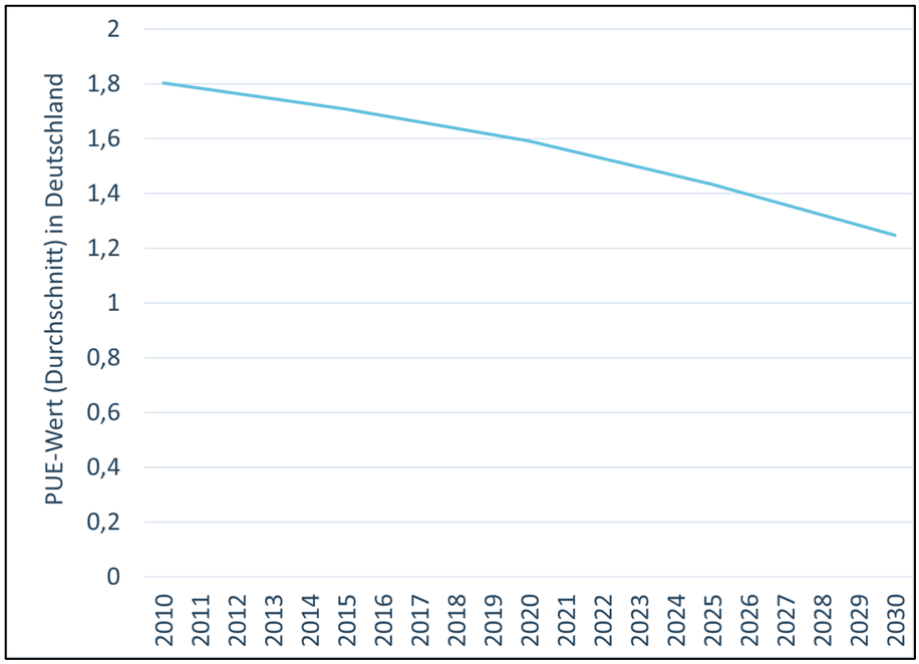


Abbildung 17: Entwicklung des PUE-Werts in Deutschland in den Jahren 2010 bis 2024 und Prognose bis 2030. Quelle: Berechnungen Borderstep

Auch wenn die Effizienz auf Basis des PUE-Werts sich in den letzten 20 Jahren positiv entwickelt hat, konnte, wie oben bereits dargestellt, ein Anstieg des Gesamtstromverbrauchs der Rechenzentren, bedingt durch die steigende Nachfrage nach digitalen Ressourcen und den Zubau an Rechenzentrumskapazität, nicht kompensiert werden. In Deutschland stieg der Stromverbrauch von knapp über 10 TWh/a (2010) auf etwa 20 TWh/a (2024) (Hintemann et al., 2024a). Weltweit stieg gleichzeitig der Energiebedarf von Rechenzentren von 200 auf 460 TWh/a (2015 - 2022) (IEA, 2024a).

Auch innerhalb der IT-Infrastruktur sorgte die technologische Entwicklung für einen Anstieg der Energieeffizienz (siehe Kapitel 5). Gleichzeitig stieg aber auch die Leistungsaufnahme der Hardware an (Mills et al., 2024; Malone et at., 2023). Gerade diese Entwicklung führt zu einem Anstieg des Gesamtstromverbrauchs der IT-Infrastruktur und bedingt zusätzliche Energieaufwendungen für sämtliche unterstützende Infrastruktur, wie z. B. die Kühlung.

Je nachdem, für welche Anwendung die IT-Infrastruktur eingesetzt wird und welche Entwicklungen hier in Zukunft auftreten, kann dies eine Auswirkung auf den Strombedarf haben. Durch ressourcensparsame Softwareentwicklung kann zum Beispiel ein Beitrag zur Effizienzsteigerung geleistet werden, während Künstliche Intelligenz oder maschinelles Lernen in Zukunft eine immer weiter steigende Bedeutung haben können (die entsprechenden Effekte wurden in Kapitel 5 berechnet).

Neben der Energieeffizienz gewinnt auch die Ressourceneffizienz zunehmend an Bedeutung. Ein Teilgebiet ist hierbei der Umgang mit Wasser, welches bei der Verdunstungskühlung eingesetzt wird. Weitere Aspekte sind der Umgang mit Rohstoffen bei der Herstellung der IT-Hardware oder bei der Errichtung des Gebäudes für das Rechenzentrum. Im Regelfall wird Trinkwasser als Wasserquelle genutzt (Zhang, 2023). In einigen Fällen wird auch Regenwasser aufgefangen und aufbereitet, um damit den Trinkwasserverbrauch zu reduzieren. Im Falle von Digital Realty sank der Trinkwasseranteil von 2017 bis 2019 von 64 auf 57 % (Digital Realty, 2019). Gleichzeitig stieg aber auch der

Gesamtwasserverbrauch. Gemäß „The World Counts“ verbrauchen Rechenzentren weltweit mehr als 4,3 Billionen Kubikmeter Wasser pro Jahr (Roundy, 2024). Der Großteil des Wasserverbrauchs eines Rechenzentrums wird indirekt durch die Stromerzeugung und Stromversorgung verursacht (Mytton, 2021). Aus diesem Grund wäre es sinnvoll, nicht nur den WUE (Water Usage Effectivness), sondern auch den WUE_{Quelle} zu betrachten (Mytton, 2021a).

Ein Rechenzentrum mit 1 MW kann bei Verwendung von Verdunstungskühlung etwa 25,5 Millionen Liter Wasser pro Jahr verwenden (Heslin, 2016). Die wenigsten Rechenzentren geben die dazugehörige Effizienzkennzahl (WUE) oder den Gesamtwasserverbrauch an. Meta ist eine der wenigen Organisationen, die den WUE für ihre Rechenzentren angibt. Microsoft und Google wiederum geben in ihren Nachhaltigkeitsberichten ihren Wasserverbrauch an (Mytton, 2021a). Amazons globaler WUE soll bei 0,19 Liter Wasser pro kWh Stromverbrauch liegen, während dieser bei Meta 2017 bei 0,24 lag, bis 2020 auf 0,30 anstieg und bis 2022 auf 0,20 sank (Zhang, 2024). Im Gegensatz dazu lag der durchschnittliche globale WUE von Microsoft 2022 bei 0,49 L/kWh (Roundy, 2024). 2022 bezogen Microsoft 23 % und Google 18 % ihres Süßwasserverbrauchs aus Wasserstressgebieten (Zhang, 2023).

In Deutschland haben 65 % aller Rechenzentren aus dem RZReg einen WUE von weniger als 0,1. Im internationalen Vergleich gehen deutsche Rechenzentrum damit sehr sparsam mit Wasser um. Es sei an dieser Stelle aber noch darauf verwiesen, dass der WUE und PUE in direkter Beziehung zueinander stehen und hier zumindest teilweise ein Trade-off besteht. Durch einen signifikanten Einsatz von Wasser bei der Rückkühlung kann der PUE durch eine Reduktion der notwendigen mechanischen Kälte reduziert werden und umgekehrt. Neben zahlreichen anderen Technologien, die den PUE (teilweise auch abseits der Kälteerzeugung) beeinflussen, steht bei der Verdunstungskälte der Strom- und Wasserverbrauch in einer reziproken Wechselwirkung. Gerade um in großen Rechenzentren Spitzenwerte unter 1,2 zu erreichen, ist die Verdunstungskühlung ein günstiger Ansatz, insbesondere an Orten, wo keine Kühlalternativen wie beispielsweise Flusswasser-/Grundwasserkühlung, zur Verfügung stehen. Die Ursache dafür, dass deutsche Rechenzentren sowohl einen im weltweiten Vergleich niedrigen PUE als auch einen WUE aufweisen, liegt vermutlich an den milden klimatischen Bedingungen und einer tendenziell moderneren IT- und Gebäudeinfrastruktur.

2.2.3 Klimawirkungen

Aufgrund der hohen Energieintensität des kontinuierlichen Betriebs der IT-Geräte (24 Stunden am Tag, 365 Tage im Jahr) sind die Klimawirkungen der Rechenzentren in Deutschland zu einem hohen Anteil (mehr als 80 %) durch den Stromverbrauch bedingt (Belkhir & Elmeligi, 2018; Bieser, Hintemann, Beucker, Schramm, & Hilty, 2020; Hintemann, Hinterholzer, & Progni, 2024). Abbildung 18 zeigt beispielhaft die Aufteilung der jährlichen Treibhausgasemissionen eines Rechenzentrums mit einer IT-Anschlussleistung von 2,5 MW.

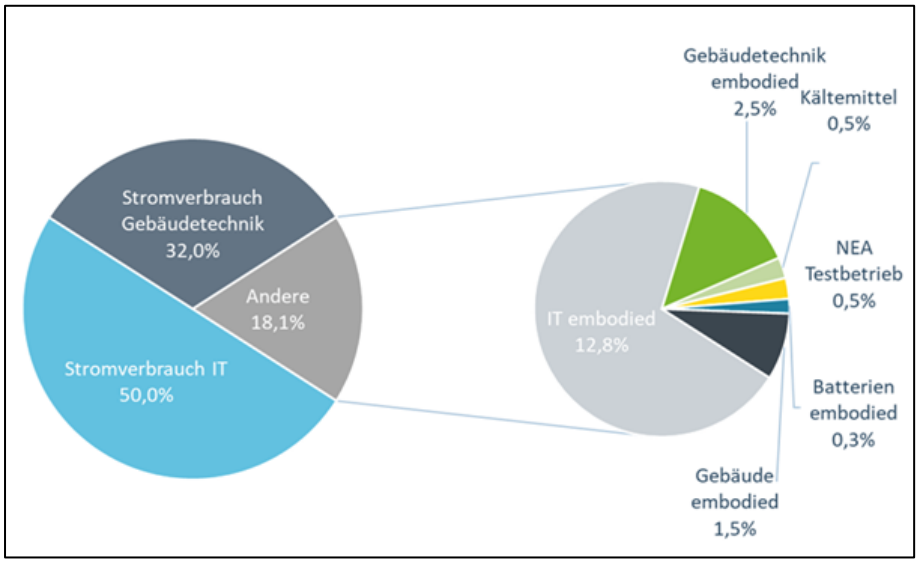


Abbildung 18: Aufteilung der Treibhausgasemissionen (in $CO_{2äq}$) eines Rechenzentrums auf verschiedene Emissionsquellen (Beispiel-RZ mit 2,5 MW IT-Anschlussleistung). Quelle: Hintemann (2024a)

Neben dem Stromverbrauch gibt es auch andere relevante Quellen für THG-Emissionen in Rechenzentren. Diese umfassen unter anderem:

- **Bau von Gebäuden und Herstellung von elektrischem, mechanischem und IKT-Equipment:** Der Bau der Gebäude sowie die Herstellung der Geräte und Anlagen, die in Rechenzentren betrieben werden, verursachen signifikante Emissionen, die zur Lebenszyklusbetrachtung eines Rechenzentrums gehören.
- **Testbetrieb von Netzersatzanlagen:** Rechenzentren sind in der Regel mit Netzersatzanlagen, wie Dieselgeneratoren, ausgestattet, um eine konstante Stromversorgung im Falle eines Stromausfalls zu gewährleisten. Diese Generatoren müssen regelmäßig getestet werden, was ebenfalls zur Emissionslast beiträgt.
- **Kältemittel in Kühlsystemen und andere klimaschädigende Gase (z. B. PFAS in Transformatoren):** Für den Betrieb von Rechenzentren ist eine effektive Kühlung notwendig, die meist mit Kältemitteln erfolgt. Weit verbreitet sind noch Kältemittel mit einer hohen klimaschädigenden Wirkung. Die Freisetzung solcher Kältemittel, beispielsweise durch Leckagen, verursacht ebenso Treibhausgasemissionen.

Die Treibhausgasemissionen von Rechenzentren durch den Stromverbrauch werden in Abbildung 19 mit dem Strommix aus Deutschland bilanziert. Da der Strommix in Deutschland aktuell noch einen nennenswerten Anteil aus fossilen Energien beinhaltet, liegen die Treibhausgasemissionen deutlich höher als in Ländern mit einem Strommix, der mehr Erneuerbare Energien bzw. Atomkraft enthält. Mit den Zielen der Bundesregierung zum Ausstieg aus der Kohleverstromung werden die

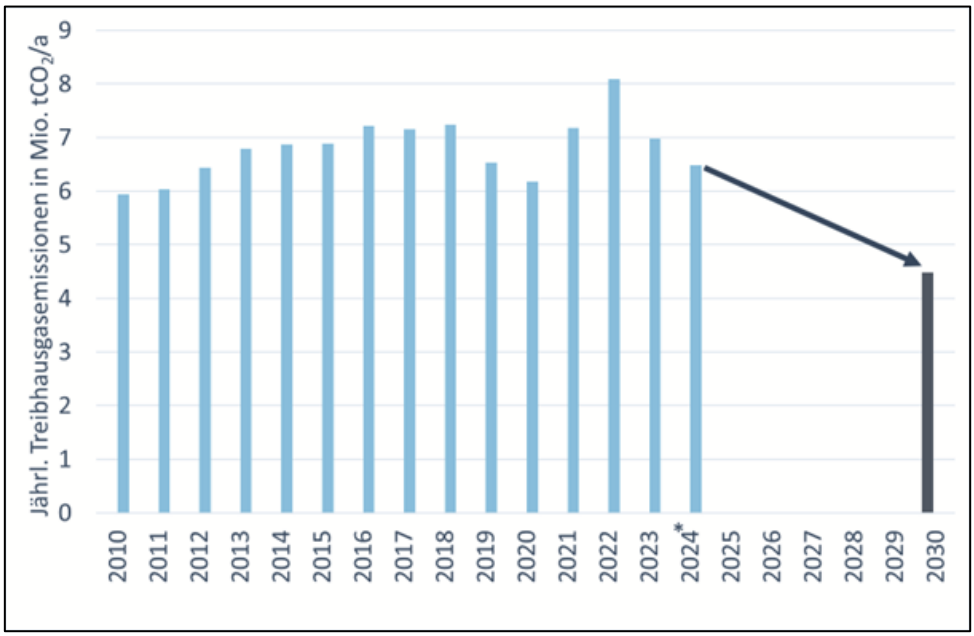


Abbildung 19: Entwicklung der durch den Stromverbrauch der Rechenzentren und kleineren IT-Installationen bedingten Treibhausgasemissionen in Deutschland in den Jahren 2010 bis 2024 und Prognose bis 2030 (* 2024 Schätzwert für den nationalen Strommix). Quelle: Hintemann et al. (2024a).

Treibhausgasemissionen durch den Stromverbrauch der Rechenzentren bis 2030 jedoch in Deutschland von 6,5 Mio. t CO_2 im Jahr 2024 auf 4,5 Mio. t CO_2 im Jahr 2030 sehr deutlich sinken.

Setzt man für die Bestimmung der Klimawirkungen der Rechenzentren nicht den nationalen Strommix, sondern die individuellen Stromlieferverträge an, verbessert sich die Klimabilanz deutlich. Viele Betreiber in Deutschland haben bereits Ökostromverträge oder Green Power Purchase Agreements abgeschlossen, was eine Reduzierung der bilanzierten Emissionen zur Folge hat. Mit der Vorgabe des EnEfG, die ab dem 1. Januar 2027 eine vollständige Deckung des Stromverbrauchs durch erneuerbare Energien vorschreibt, ist zu erwarten, dass sich die Klimabilanz der Rechenzentren in Deutschland weiter verbessern wird. Bilanziert man alle Rechenzentren mit mehr als 300 kW nichtredundanter Anschlussleistung im Jahr 2030 mit Treibhausgasemissionen von Null, so verbleiben im Jahr 2030 für die kleineren Rechenzentren maximal noch jährliche stromverbrauchsbedingte Emissionen von 1 bis 1,5 Mio. t CO_2, da auch kleinere Rechenzentren vielfach mit Ökostrom betrieben werden, ohne dass sie dazu gesetzlich verpflichtet wären.

2.3 Rechenzentren und das Stromnetz

2.3.1 Herausforderungen bei der Einbindung von Rechenzentren in das Stromnetz

Wie bereits die obigen Ausführungen zeigen, ist von einem dauerhaften Boom bei Rechenzentren auszugehen. Geht man von der Annahme aus, dass sich bis 2045 keine erheblichen Trendbrüche ergeben, kann mit einem Strombedarf der Rechenzentren in der Größenordnung 80 TWh/a ausgegangen werden (siehe Abschnitt 2.1.1). Setzt man die Strombedarfsprognosen für RZ dem für 2045 erwarteten Gesamtbruttostrombedarf von ca. 1.079 bis 1.303 TWh/a (Netzentwicklungsplan

2037/2045) gegenüber, so werden RZ für 5 bis 7 % des Strombedarfs in Deutschland verantwortlich sein. Auch in anderen Sektoren zeichnet sich laut Netzentwicklungsplan ein großer Zuwachs des Strombedarfs ab. Für Elektrolyseure gehen Prognosen von einem Strombedarf von 90 bis 150 TWh/a für 2045 aus, was wiederum 7 bis 14 % des deutschen Stromverbrauchs entspräche. Ein dritter Sektor, in dem ein starker Anstieg des Stromverbrauchs erwartet wird, ist der Verkehrssektor. Bedingt durch die dort zunehmende Elektrifizierung gehen Prognosen für 2045 von 110 bis 175 TWh/a aus, was je nach Szenario 8 bis 16 % des Gesamtstromverbrauchs entspricht (dena-Leitstudie, 2021; Agora, 2021).

Herausfordernd für die Rechenzentrumsbranche und die Energiewirtschaft wird es nicht nur sein, diese Bedarfe überhaupt zu decken, insbesondere wenn dies mit Strom aus erneuerbaren Quellen geschehen soll. Hintergrund dafür sind zum einen die gesetzlich festgelegten Klimaschutzziele, aber auch der stärker werdende Wille von Betreibern selbst, ihre Rechenzentren mit klimaneutral erzeugtem Strom zu versorgen.

Für Betreiber besteht die Möglichkeit, Strom aus erneuerbaren Quellen über Energieversorgungsunternehmen (EVU), aus der Eigenversorgung oder über Green PPA zu beziehen. Im Vergleich zum Grünstrombezug über EVU ist der Strombezug über PPA aufgrund seines klaren Beitrags zur Energiewende, günstiger Preise und der Möglichkeit, Strom auch überregional zu beziehen, vorteilhaft für Rechenzentrumsbetreiber. Fraglich ist jedoch, ob sie für den prognostizierten, starken Anstieg des Strombedarfs in ausreichendem Maße verfügbar sein werden. Dies liegt zum einen an der begrenzten Zahl an Bestandsanlagen und zum anderen an den im Vergleich zum Zubau von Rechenzentren langwierigen Genehmigungsverfahren für neue Erneuerbare Energien (EE)-Anlagen.

2.3.2 Versorgung von Rechenzentren mit erneuerbarem Strom mit Hilfe von PPAs

Eine große Rolle bei der Versorgung von Rechenzentren mit Strom aus erneuerbaren Energiequellen spielen bereits heute die PPAs. Mit rund 57 % wird aktuell mehr als die Hälfte des von Colocation-Gebäuden und Großrechenzentren benötigten Stroms über PPAs bezogen. Deutschland hat zudem nach Spanien den zweigrößten PPA-Markt Europas, und das Angebot an PPAs hat in den letzten Jahren stark zugenommen. So hat sich das Gesamtvolumen an PPAs vom Jahr 2022 im Vergleich zu 2023 um 323 % gesteigert (PPA-Marktanalyse, 2023). Die Schlüsselsegmente des starken Zuwachses liegen hierbei im Photovoltaik- und Offshore-Windbereich. Neben dem wachsenden Angebot war auch die Nachfrage nach grünen PPAs zuletzt hoch. Der deutsche Markt für PPAs zeichnet sich in diesem Zusammenhang dadurch aus, dass er zwar insgesamt sehr heterogen aufgestellt, aber auf Abnehmerseite von großen Corporate-Akteuren dominiert wird. Große Digitalkonzerne, wie Microsoft und Amazon, oder Unternehmen aus dem Energiesektor, wie Shell, gelten derzeit als die größten Abnehmer in Deutschland. Als Grund für die eher zurückhaltende Nutzung von PPAs durch KMUs und kleinere Stadtwerke werden vor allem Kreditrisiken und der Aufbau von notwendigem Know-how genannt (PPA-Marktanalyse, 2023).

Ob der voraussichtlich stark ansteigende Strombedarf von RZ maßgeblich durch Green PPA, also PPAs, die ausschließlich auf erneuerbare Energieerzeugung setzen, gedeckt werden kann, ist derzeit noch fraglich. Der Markt unterliegt einer großen Dynamik, und nach Ansicht von größeren deutschen PPA-Anbietern hat sich der PPA-Markt nach deutlich erhöhter Nachfrage in den Jahren 2022 und

2023 wieder eher zu einem Käufermarkt entwickelt. Dadurch kann zumindest aktuell mit einer guten Verfügbarkeit von regionalen PPAs gerechnet werden, auch weil beispielsweise die PPA-Nachfrage von Elektrolyseuren zuletzt hinter den Prognosen zurückblieb.

Es gibt jedoch einige Hemmnisse, die ein stärkeres Wachstum des PPA-Marktes in Deutschland begrenzen und damit auch den schnelleren Ausbau nicht besonders attraktiv machen. Dazu zählt, dass es aktuell keine Vorgabe bezüglich der PPA-Art für Stromgroßverbraucher wie RZ gibt. Statt regionale physische PPAs mit deutschen Anbietern abzuschließen, kommt es häufig zu einem Wettbewerb etwa mit spanischen PPA-Anbietern, die zwar günstigere Konditionen anbieten können, aber deren grüner Strom letztlich nie in Deutschland ankommt, sondern lediglich bilanziell durch Zertifikate von den Abnehmern ausgewiesen werden kann. Weitere Hindernisse sind die sehr langwierigen Genehmigungsverfahren für neue EE-Anlagen, auch bedingt durch Zögern einiger PPA-Interessenten, etwa aus der Stahlindustrie, die den Abschluss von Green PPAs derzeit sehr lange hinauszögern, um je nach Strategie der kommenden Bundesregierung einen womöglich ökonomisch vorteilhafteren Weg der Stromversorgung zu wählen. Zuletzt begrenzt auch die Konkurrenz zwischen der Finanzierung durch Green PPAs und der Förderung durch das EEG das Wachstum des PPA-Marktes. (PPA-Marktanalyse, 2023). Mit der Überarbeitung der EEG-Förderung ist darüber hinaus vorgesehen, dass EE-Anlagenbetreiber künftig nicht mehr monatlich zwischen der einspeisebasierten Förderung (Contracts-for-Difference) und einer Direktvermarktung wechseln können. Insbesondere für kleinere Abnehmer (bzw. kleinere Rechenzentrumsbetreiber) wird hier jedoch ein Nachteil gesehen, da diese sehr langfristige Abnahmeverträge mit PPA-Anbietern eingehen müssten (Brinkhaus & Schwartz, 2024).

2.3.3 Herausforderung durch die regionale Konzentration von Rechenzentren

Der Ausbau des Stromnetzes in Deutschland ist essenziell für die Energiewende. Insbesondere der Anschluss erneuerbarer Energiequellen, wie Wind- und Solarenergie, und die Zunahme an Großverbrauchern, wie Rechenzentren, Elektrolyseure und E-Ladeinfrastruktur, erfordert eine umfassende Netzentwicklung. Die Auswirkungen von neuen Stromgroßverbrauchern, die stark regional konzentriert auftreten, auf das Stromnetz werden auch im aktuellen Netzentwicklungsplan erwähnt und stellen die Planung vor einige Herausforderungen. So müssen die Bottom-up-Erkenntnisse zu geplanten Projekten von Stromgroßverbrauchern in der Planung berücksichtigt und mit dem Top-Down-Modellierungsansatz der Stromverbrauchsmodellierung in Einklang gebracht werden. Nur so ist ein sinnvoller Netzentwicklungsplan zu gewährleisten. Unklar ist in diesem Zusammenhang auch, inwieweit die Netzentwicklungsplanung ungewisse Bedarfe von Unternehmen berücksichtigen sollte. Einerseits könnte so die Stromnetzinfrastruktur frühzeitig auf potenzielle Großverbraucher vorbereitet werden, was schnellere Netzanschlüsse ermöglicht und Verzögerungen bei der Systemtransformation reduziert. Andererseits besteht das Risiko eines ineffizienten Netzausbaus, wenn geplante Projekte entfallen oder nicht in der ursprünglich angegebenen Größenordnung umgesetzt werden (Netzentwicklungsplan-Strom, 2024).

Abbildung 20 zeigt die Verläufe der 110 kV Hochspannungsleitungen (Bundesnetzagentur, 2023). Es existieren einige Gebiete mit einer hohen Dichte an Hochspannungsleitungen. Gleichzeitig werden aber vor allem im Norden und in einigen Knotenpunkten in Süddeutschland Engpassregionen ausgewiesen, welche an ihre Kapazitätsgrenzen stoßen. In diesen Gebieten wird sich bei ausbleibendem Netzausbau höchstwahrscheinlich kein Rechenzentrum ansiedeln können. Die

Attraktivität eines potenziellen Standorts für ein Rechenzentrum hängt aber auch von der Verfügbarkeit eines Glasfaseranschlusses ab (Vgl. Abbildung 21). Teilweise überlagern sich aber gerade die Regionen, bei denen zwar in unmittelbarer Nähe ein Glasfasernetz verläuft, gleichzeitig das Stromnetz aber an seine Kapazitätsgrenzen stößt.

2.3.4 Belastung der Stromnetze durch den Ausbau von digitaler Infrastruktur

Der beschriebene prognostizierte Zuwachs von Rechenzentren und deren Stromverbrauch stellt eine Herausforderung für die deutschen Stromnetze dar und verschärft die Netzbelastung durch kontinuierlichen Strombedarf, der besonders in Zeiten hoher Auslastung oder bei extremen Wetterbedingungen zu Engpässen führen kann.

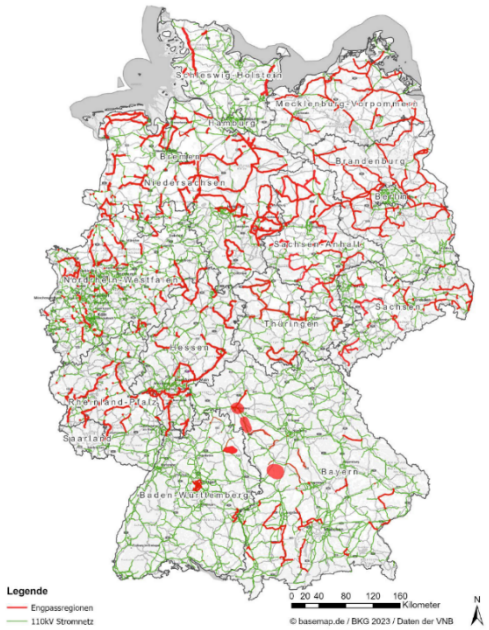


Abbildung 20: Verläufe der 110kV Hochspannungsleitungen in DE. Quelle: Bundesnetzagentur (2023)

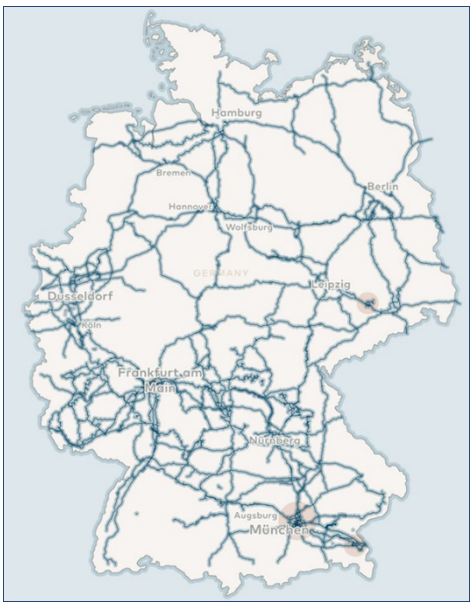


Abbildung 21: Deutsches Glasfasernetz. Quelle: NGN Fiber Network GmbH & Co. KG (2024)

Diese Spitzenlasten erfordern entweder eine Verstärkung der Netzkapazitäten oder innovative Lösungen, wie die Integration von Speichersystemen und intelligentem Lastmanagement. Selbst im Betriebsgebiet des Übertragungsnetzbetreibers (ÜNB) 50Hertz, dass sich durch eine überdurchschnittlich hohe EE-Stromproduktion und hierfür gut ausgebaute Netzinfrastruktur auszeichnet, ist die Netzkapazität nahezu aufgebraucht, und ein weiterer Zuwachs an Großstromverbrauchern wird aktuell als Risiko für die Netzstabilität eingestuft.

Darüber hinaus führt die steigende Anzahl an EE-Anlagen, aber auch an Stromgroßverbrauchern wie Rechenzentren, zu einer zunehmenden Belastung der bestehenden Netzverknüpfungspunkte (NVP). Diese Netzverknüpfungspunkte, typischerweise Umspannwerke, Umspannstationen und Schaltanlagen, sind ein essenzieller Bestandteil des modernen Stromnetzes und dienen als physische Verbindungen zwischen Stromerzeugern, Stromnetzwerken und Verbrauchern. Stand heute sind alle verfügbaren und geplanten NVP im durch 50Hertz versorgten Gebiet entweder bereits in Nutzung oder für konkrete Projekt reserviert. Die zunehmende Knappheit an verfügbaren

Netzverknüpfungspunkten hat im Kontext von Rechenzentren damit auch zur Folge, dass die Möglichkeit des Netzanschlusses einer Fläche zu einem der gewichtigsten Faktoren für die Standortwahl für Stromgroßverbraucher wie Rechenzentren geworden ist.

2.3.5 Herausforderungen für den Netzausbau

Durch die gestiegene Belastung der Stromnetze wäre ein umfassender Netzausbau die offensichtlichste Lösung. Neben den hier bekannten Herausforderungen, wie Zeitbedarf für Genehmigungsverfahren, politischer Widerstand und hohe Umsetzungskosten, wurden im Rahmen der Recherche weitere spezielle Herausforderungen hinsichtlich des Ausbaus der Stromnetze identifiziert. Hierzu zählen der massive Anstieg an Direktnetzanschlussanfragen bei den ÜNBs in den vergangenen Jahren, langfristige Planungshorizonte im Netzausbau sowie große Unsicherheiten bezüglich der zum Teil spekulativen Direktanschluss-Anfragen von Projektentwicklern.

Die Bundesnetzagentur hat 2022 den Netzentwicklungsplan 2021 - 2025 bestätigt (Bundesnetzagentur, 2022b). Darin wurden die erforderlichen Maßnahmen aufgeführt und teilweise mit einem Fertigstellungsjahr versehen (Bundesnetzagentur, 2022c). Der zweite Entwurf enthält insgesamt 6.650 km, wovon 6.350 km bestätigt und 300 km nicht bestätigt wurden. Mit dieser Maßnahme wird vor allem der windstromreiche Norden mit dem Süden von Deutschland verbunden, um damit auch die Engpässe in den Stromleitungen in Norddeutschland zu reduzieren.

Mit Blick auf die Anzahl an Direktnetzanschlussanfragen berichten ÜNBs in Deutschland von einem deutlichen Anstieg der Direktnetzanschlussanfragen in den letzten Jahren. Diese Direktanfragen unter anderem von Rechenzentrums-Projektentwicklern umfassen sowohl den Anschluss neuer Großverbraucher als auch den Ausbau bestehender Anschlüsse. Dies stellt die Netzbetreiber vor große Herausforderungen, da die Infrastruktur oft erst angepasst oder ausgebaut werden muss, um die gestiegene Nachfrage zu bewältigen. Zudem gibt es aufgrund fehlender Erfahrungswerte noch keine klare und einheitliche Vorgehensweise bei den ÜNBs hinsichtlich der Bewertung all dieser Direktanschlussanfragen.

Erschwerend hinzu kommt hierbei, dass auch nach Ansicht der konsultierten ÜNBs die Seriosität vieler Anfragen häufig in Frage gestellt werden muss, da es vermehrt zu spekulativem Sichern von Flächen durch Immobilienunternehmen und Projektentwicklern kommt, die sich große Profite zu einem späteren Zeitpunkt erhoffen. Diese Problematik besteht allerdings ebenso bei den bekannten Hyperscalern. Viele Anfragen zum Netzanschluss von Flächen durch große Cloud-Anbieter werden zunächst als äußerst dringlich adressiert und kurze Zeit später wieder verworfen. Dies macht die Auswahl von Standorten für einen Netzausbau und -anschluss für ÜNBs sehr komplex und ist mit großen Unsicherheiten behaftet. Auch eine Auswertung der Antragslage zeigt, dass Projektvorhaben und Daten einer großen Dynamik unterliegen. Im Zuge einer umfassenden Konsolidierung der Markt- und Netzbetreiberabfragen wurden zahlreiche Unterschiede und Änderungen gegenüber den ursprünglichen Datenbeständen festgestellt. Diese Unterschiede reichen von unterschiedlichen Angaben im Projektfortschritt sowie bei Leistungsdaten bis hin zu Wechseln der zu Beginn genannten Betreibergesellschaften. Einer der Gründe hierfür liegt aus Sicht der ÜNB darin, dass ein Großteil der Planungen noch nicht weit fortgeschritten ist, und eine Umsetzung oder sogar eine finale Investitionsentscheidung in der Regel noch nicht beschlossen ist. Etwa 82 % der in der Marktabfrage gemeldeten Leistung weist noch einen frühen Projektstatus der Entwurfsplanung oder

Machbarkeitsprüfung auf. Nur etwa 11 % der gemeldeten Leistung hat bereits einen fortgeschrittenen Projektstatus und durchläuft beispielsweise schon ein Genehmigungsverfahren. In unmittelbarer Umsetzung oder bereits in Betrieb sind die verbleibenden 7 % der gemeldeten Leistung. (Netzentwicklungsplan-Strom, 2024)

Auch die langfristigen Planungshorizonte und Zeitbedarfe im Netzausbau im Vergleich zu den teilweise deutlich kurzfristigeren Anschlussanfragen von Projektentwicklern stellen ein Problem dar. Bedingt durch eine szenariobasierte Planung im Netzentwicklungsplan Strom liegen Zeit- und Planungshorizonte im Netzausbau bei 10 bis 15 Jahren (dena, 2022). Diesen Planungshorizont mit dem Horizont von Rechenzentrums-Projektenwicklern zu vereinbaren, ist häufig nur schwer umsetzbar.

2.3.6 Flexibilität aus Rechenzentren

Im herkömmlichen Elektrizitätssystem stellen elektrische Verbraucher hauptsächlich eine passive Verbrauchsrolle dar und die Netzbetreiber sowie die Stromerzeuger und der Strommarkt sorgen dafür, dass die Last zu jedem Zeitpunkt und an jedem Ort versorgt wird. Großverbraucher wurden hauptsächlich über die spitzenlastbasierte Berechnung der Netzentgelte incentiviert, Lasten möglichst gleichmäßig über die Zeit zu verteilen bzw. es bestanden Anreize, abschaltbare Lasten zur Stromnetzstabilisierung für Netzbetreiber schaltbar zu machen. Im zukünftigen elektrischen Energiesystem können und sollen Verbraucher wie Rechenzentren eine wesentlich aktivere Rolle einnehmen, indem sie Flexibilität bereitstellen.

Unter Flexibilität im Energie- bzw. Stromsystem versteht man daher „die Fähigkeit von Elementen im Energiesystem, aktiv auf ein externes Signal, das die Variabilität von Stromerzeugung und Stromverbrauch widerspiegelt, mit einer Leistungsänderung zu reagieren. Elemente, die diese Fähigkeit besitzen, werden auch als Flexibilitätsoptionen bezeichnet.“ (Beucker et al., 2021)

Ob und in welchem Verhältnis die Speicherung von Energie oder die Flexibilisierung der Nachfrage geeignet ist, um auf das sich verändernde Energiesystem zu reagieren, hängt von zahlreichen Parametern (z. B. Energiemenge und -preise, Speicherdauer, Ort der Erzeugung/Speicherung) ab. Ein zentraler Indikator für die Nutzbarkeit und den Wert von Flexibilität im Energiesystem ist die Nachfrage nach dieser Leistung am Energiemarkt. Nachfrage kann z. B. auf Großhandelsmärkten für Strom (Day-Ahead, Intraday), durch die Behebung von Netzengpässen (Netzengpassmanagement) oder durch Regelenergiemärkte (Regelleistung) entstehen (Beucker et al., 2021).

Rechenzentren können sich durch flexibles Verhalten, wie auch andere industrielle und private Stromverbraucher, an Angebot und Nachfrage im Energiesystem anpassen. Dies kann prinzipiell durch folgende Maßnahmen geschehen:

- **Räumliche Verlagerung von Rechenlasten**: Z. B. durch die Verlagerung von Rechenlasten in Regionen mit geringeren Strompreisen oder emissionsärmerer Stromversorgung. Dies wird bereits für rechenintensive Prozesse (z. B. Bitcoin-Mining, Training von Large-Language-Models) betrieben, um Energiekosten und Emissionen zu reduzieren. Es ist v. a. für größere

Rechenzentrumsanbieter (z. B. Hyperscaler[6]) geeignet, die über Standorte in mehreren Regionen oder Ländern verfügen und daher Lasten zwischen Standorten verschieben können. Ungeeignet ist die Maßnahme für latenz- oder zeitkritische Anwendungen (z. B. Datenbankanwendungen oder Echtzeithandel), die in räumlicher Nähe zum Kunden bzw. Anwender angesiedelt sein müssen. Eine räumliche Verschiebung von Lasten ist aus ökonomischer und ökologischer Sicht auch nur sinnvoll, wenn die vorhandenen Systeme an den Standorten bislang nicht voll ausgelastet sind und freie Kapazitäten zur Verschiebung existieren.

- **Zeitliche Verlagerung von Rechenlasten**: Im Gegensatz zur räumlichen Verlagerung von Rechenlasten erfolgt dabei eine zeitliche Verschiebung der Abarbeitung der Rechenaufgaben. Ein Treiber kann die kurzfristige Schwankung in der Energieverfügbarkeit sein, die sich im Börsenstrompreis (z. B. Hoch- und Niedrigtarif, Intraday Handelspreise) niederschlägt. Diese Maßnahme erfordert Anwendungen, die unkritisch hinsichtlich ihrer Ausführungszeit sind, da sie zwangsläufig zur Verzögerung bzw. Verlagerung von Rechenlasten im Minuten- oder Stundenbereich (z. B. Tag/Nacht) führt. Sie wird bereits betrieben, um eine gleichmäßigere Auslastung von Rechenzentren zu erreichen. Bislang gibt es jedoch kaum ökonomische Anreize, z. B. durch flexible Strompreise oder Vorteile bei Netzentgelten, die eine zeitliche Verschiebung aus Gründen der Energiekosten, der Emissionen oder der Netzstabilität für die Betreiber attraktiv machen.
- **Flexibilitäten aus energietechnischer Infrastruktur**: Rechenzentren können mit ihrer energietechnischen Infrastruktur Netzersatzanlagen, Kühlung, etc.) an Flexibilitätsmechanismen des Energiesystems teilnehmen. Ob und in welchem Umfang dies möglich ist, hängt stark von der Art des Rechenzentrums und seiner technischen Infrastruktur ab. Hierzu gibt es sowohl erste Erfahrungen als auch konzeptionelle Überlegungen (siehe unten), die für die zukünftige Ansiedelung und die Planung von Rechenzentren relevant sind. Dieser Fall wird daher im Folgenden detaillierter behandelt.

Anlagen zur unterbrechungsfreien Stromversorgung (USV) und Netzersatzanlagen in Rechenzentren gleichen Spannungsschwankungen aus und kompensieren Stromausfälle. Größere Anlagen können als Ersatzversorgung mehrere Tage unabhängig vom Netz arbeiten, bei ausreichender Brennstoffversorgung auch unbegrenzt. Die Anforderungen an USVen und Netzersatzanlagen für Rechenzentren sind unter anderem in der Norm DIN EN 50600 beschrieben und variieren je nach Einsatzgebiet und Versorgungsdauer. Sie könnten auch zur Stabilisierung des Stromnetzes genutzt werden, da sie kurzfristig Energie speichern oder bereitstellen können (Gentsch & Tarrach, 2020).

Anlagen zur Stromversorgung von Rechenzentren bei Stromausfällen bestehen aus drei Hauptkomponenten: Erstens einer USV mit Stromrichtern und -schalter sowie Batterien bzw. Schwungradspeicher bei dynamischen USV-Systemen für kurze Stromausfälle bzw. zur kurzzeitigen Überbrückung. Zweitens Generatoren oder Brennstoffzellen (Netzersatzanlage/NEA) für die langfristige Versorgung, welche ggf. etwas Anfahrtszeit benötigen. Drittens gehören dazu Brennstoffspeicher, welche üblicherweise für mindestens 24 Stunden Vollastbetrieb Brennstoff bereithalten und auch im Betrieb befüllt werden können. Für USVen werden neben Blei-Säure-Batterien zunehmend auch Lithium-Ionen-Batterien eingesetzt, da sie z. B. Vorteile hinsichtlich

[6] Für Colocation-Rechenzentren, bei denen die Betreiber der Gebäudeinfrastruktur keinen direkten Einfluss auf die IT-Systeme haben, ist eine solche Verlagerung ohne deutliche Änderungen in den Geschäftsmodellen nicht möglich.

Flexibilität und Langlebigkeit haben (Gentsch & Tarrach, 2020). Trotz deutlich gesunkener Preise sind Lithium-Ionen-Batterien aber immer noch teurer, verursachen aber aufgrund ihrer höheren Leistungsdichte einen geringeren Platzbedarf im Rechenzentrum. Brennstoffzellen bieten einen höheren Wirkungsgrad als Motor-Generator-Systeme, sind jedoch teurer und lohnen sich nur bei intensiver Nutzung (Clausen & Beucker, 2024) und werden daher als Notstromaggregate in Rechenzentren bisher, außer in Ausnahmefällen, nicht genutzt. Als klimaneutrale Brennstoffe kommen grüner Wasserstoff oder grünes Methanol infrage, wobei Methanol aufgrund geringerer Speicheranforderungen bevorzugt wird (Umweltbundesamt, 2012, 2019; Gauto, 2023). Als direkte Alternative zum Dieselkraftstoff in klassischen Generatoren kann auch klimaneutrales hydriertes Pflanzenöl (Hydrotreated Vegetable Oil, HVO) eingesetzt werden.

Rechenzentren nutzen häufig freie Kühlung kombiniert mit Kältemaschinen. Der steigende Kühlbedarf durch leistungsstärkere IT und niedrigere geforderte Prozesstemperaturen erfordert innovative Ansätze wie Flüssigkühlung. Durch Kältespeicherung kann der Stromverbrauch zeitlich verlagert und die Netzlast reduziert werden (Friedrichsen, 2015). Insbesondere in den Sommermonaten kann der Energieverbrauch für die Kühlung dadurch deutlich reduziert werden. Zukünftig könnte Abwärme über Wärmepumpen in Wärmenetze eingespeist werden, um Flexibilität und Effizienz zu erhöhen.

Die beschriebenen Flexibilitätspotenziale aus Rechenzentren werden derzeit nur ansatzweise erprobt und genutzt. Während die räumliche und zeitliche Verlagerung von Rechenlasten von größeren Rechenzentrumsbetreibern aus betrieblichen Gründen bereits teilweise umgesetzt wird, gibt es noch wenig Anreize für einen kurzzeitigen Einsatz von Flexibilitäten aus der unterbrechungsfreien Stromversorgung oder der Kühlung von Rechenzentren. Die Gründe hierfür liegen in den auf den Geschäftsmodellen beruhenden Prioritäten der Rechenzentrumsbetreiber und unzureichenden ökonomischen Anreizen.

Für die Betreiber steht die Verfügbarkeit der Rechenzentren an erster Stelle, da ein Ausfall der Technik große wirtschaftliche Schäden zur Folge hat. Zudem dominieren vielfach die Kosten der IT-Hardware die Gesamtkosten, sodass es wirtschaftlicher ist, die vorhandene IT-Infrastruktur möglichst vollständig auszulasten. Es müssen daher entweder starke wirtschaftliche oder umweltpolitische Anreize bestehen, um von erprobten Betreibschemata abzurücken, oder es muss eine einfache Umsetzung möglich sein. Dies verdeutlichen auch die wenigen in Deutschland realisierten Beispiele aus der Praxis:

- **Arvato und RWE**: Arvato (Rechenzentrumsbetreiber) und die RWE Supply & Trading GmbH (RWEST) optimieren gemeinsam die Probeläufe von Notstromaggregaten aus einem Rechenzentrum. Durch die Verschiebung der Betriebsstunden kann Regelleistung (aFFR/mFFR) bzw. eine Spitzenlastreduktion erzielt werden. Die Flexibilisierung kann zu einer Erhöhung der direkten lokalen Emissionen führen, sofern der Notstromdiesel durch den Flexibilitätseinsatz für zusätzliche Betriebsstunden in Betrieb ist. Sie bewirkt jedoch emissionsmindernde Effekte im Energiesystem, da ggf. weniger fossile Kraftwerkskapazität für die Erbringung von Regelleistung bzw. die Spitzenlastdeckung dauerhaft vorgehalten werden muss. Grundsätzlich kann der Ansatz auch auf andere, CO_2-neutrale, Aggregate (Brennstoffzellen, Batterien, etc.) übertragen werden (Schwarz, 2024).

- **Googles 24/7 Clean Energy Ansatz**: Google hat sich zum Ziel gesetzt, seine Rechenzentren (d. h. auch die Standorte in Deutschland) bis zum Jahr 2030 zu 100 % mit Ökostrom zu versorgen. Dabei sollen stündlich bzw. halbstündlich erstellte Nachweise der EnergyTag Initiative sicherstellen, dass die Rechenzentren auch vollständig (24/7) mit Ökostrom betrieben werden. Damit soll erreicht werden, dass
    - die Ökostromerzeugung sowie der Verbrauch zeitgleich erfolgen,
    - möglichst auf lokale Stromerzeugung (Strommarkt-Gebotszone) zurückgegriffen wird,
    - möglichst neue Erzeuger bzw. Erzeugungskapazitäten genutzt werden (24/7 Clean Energy – Data Centers – Google, o. D.).

Der Ansatz zielt zwar nicht primär auf die Flexibilisierung des Betriebs ab, allerdings ist die flexible Verlagerung von Rechenlasten einer der Mechanismen, mit denen Google das Ziel erreichen will (Koningstein, 2021).

Für eine stärkere Mobilisierung der Flexibilitätspotenziale sind deutliche zusätzliche Anreize für einen netzdienlichen und emissionsärmeren Betrieb notwendig. Dies können z. B. sein:

- Variable Bestandteile des Strompreises oder der Netzentgelte, die sich stärker an der System- bzw. Netzdienlichkeit orientieren. Die Preisunterschiede müssen so hoch sein, dass sie die Anpassung von Betriebsweisen oder die Nutzung energietechnischer Infrastrukturen zur Flexibilitätsbereitstellung rechtfertigen.
- Kompensationsmechanismen (z. B. Zertifikatsgutschriften für CO_2-Emissionen), die eine an dem Einsatz erneuerbarer Energien orientierte Betriebsweise oder die Nutzung energietechnischer Infrastrukturen rechtfertigen. Durch die Gutschriften, die aus der Flexibilitätsbereitstellung resultieren, könnte dann die Gesamtbilanz (z. B. für CO_2-Emissionen) des Betriebs eines Rechenzentrums verbessert und dadurch auch ein wirtschaftlicherer Betrieb gesichert werden.

Zu beachten ist, dass derzeit noch eine Netzentgeltreduzierung bei einer hohen Volllaststundenzahl von 7.000 h/a gewährt wird. Diese Regelung soll aber ersetzt werden und Netzentgelte sollen künftig Anreize für einen netzdienlichen Betrieb liefern. Ein weiterer indirekter Anreiz könnte zukünftig aus der Nutzung der Abwärme aus Rechenzentren entstehen. Hierfür sind jedoch neue Betreiber- und Finanzierungsmodelle, z. B. für Investitionen in Wärmepumpen, gefragt, die gemeinsam mit Contractoren oder Energiedienstleistern Lösungen für spezifische Standorte entwickeln und diese in Wärmenetze oder Speicherkonzepte integrieren.

Eine unter dem Gesichtspunkt von Flexibilitäten und Klimaschutz optimierte Notstromversorgung eines Rechenzentrums könnte folgende Elemente bzw. Aggregate enthalten:

- Eine flexibilisierbare Batterie-USV (z. B. aus Lithium-Eisen-Phosphat-Batterien o. ä.), die auf hohe Lade- und Entladezyklen ausgelegt ist. Ggf. kann die Kapazität der Batterie auch über dem Bedarf des Rechenzentrums liegen, um einerseits genügend Reserve für die Notstromversorgung des Standortes und andererseits genügend handelbare Leistung bereitzustellen.

- Ein Stromerzeugungsaggregat mit Verbrennungsmotor oder einer Brennstoffzelle, das möglichst mit CO_2-armen Brennstoffen (z. B. grünem Methanol/E-Fuel) versorgt wird. Solche Stromerzeugungsaggregate könnten netzdienlich eingesetzt werden, indem das Rechenzentrum bei Stromengpässen durch sie versorgt wird und der Netzanschluss nur als „Redundanz" genutzt wird, sollte es zu Ausfällen bei den Stromerzeugungsaggregaten kommen.
- Durch Verlagerung von Kühlleistung und die Nutzung der Abwärme über Wärmepumpen in einem Niedertemperatur-Wärmenetz oder durch Einspeisung in ein vorhandenes, größeres Wärmenetz können weitere Flexibilisierungspotenziale erschlossen werden.

Abschätzung: Bereitstellung von Flexibilität aus Rechenzentren durch USV und Netzersatzanlagen

Größere kommerzielle Rechenzentren sind in der Regel mit einer USV und einer NEA ausgerüstet. Die Kapazitäten dieser Anlagen werden im Folgenden abgeschätzt. Außerdem wird eine erste Diskussion hinsichtlich der Möglichkeiten ihrer Nutzung für die Bereitstellung von Flexibilitäten (z. B. für Systemdienstleistungen, Lastmanagement oder Kapazitätsreserve) vorgenommen. Diese erste Abschätzung und Analyse zeigt, dass es sinnvoll wäre, die Potenziale einer solchen Verwendung dieser Anlagen in einer weiterführenden Untersuchung genauer zu bestimmen und insbesondere die erforderlichen Anpassungen bei der eingesetzten Technik, den Betriebskonzepten und in der Regulierung zu ermitteln und zu bewerten. In großen Rechenzentren wird üblicherweise eine Kombination aus USV (Batterie, Schalter, Umrichter; teilw. auch Schwungradspeicher) zur kurzfristigen und störungsfreien Überbrückung von Netzausfällen verwendet (sowie als Netzfilter etc.) und zusätzlich ein NEA (häufig Dieselgenerator, seltener Gasturbinen oder Brennstoffzellen) zur längerfristigen Überbrückung. Neben der Installation in Rechenzentren finden sich solche Anlagen unter anderen auch in folgenden Bereichen: Krankenhäuser, Müllverbrennungsanlagen, Kläranlagen, energieintensive Industrien, Banken, Telekommunikationsanlagen.

> Sowohl im Sprachgebrauch als auch in Normen und Vorschriften werden USV, NEA, Notstromaggregat/-diesel etc. nicht einheitlich verwendet. In diesem Dokument wird nach dem Kriterium Kurzzeitversorgung (=USV), Langzeitversorgung (Netzersatzanlage/NEA) unterschiede, Als Überbegriff für USV und NEA wird „Ersatzstromversorgungssystem" verwendet.

Basierend auf den oben ermittelten Kapazitäten der Rechenzentren in Deutschland und ihrer Verteilung auf die verschiedenen Größenklassen kann davon ausgegangen werden, das aktuell grob geschätzt etwa 1.500 bis 2.000 MW elektrischer Nennleistung in Form von Notstrom-Generatorleistung in den Rechenzentren existieren. Eingesetzt werden dafür heute fast ausschließlich Dieselgeneratoren. Dieser Abschätzung liegt die Annahme zugrunde, dass kleinere Rechenzentren unter 100 kW, einige Forschungsrechenzentren oder Rechenzentren mit anderweitiger Absicherung oft nicht über Generatoren verfügen. Außerdem wird davon ausgegangen, dass aus Redundanzgründen zwar mehr Generatoren als notwendig vorhanden sind, zumeist aber „nur" eine n+1-Redundanz genutzt wird. Bis zum Jahr 2030 dürfte die Leistung der Generatoren auf etwa 3.500 MW anwachsen. Wächst der Rechenzentrumsmarkt weiterhin sehr stark, wie in der Langfristprognose unterstellt, so könnten im Jahr 2045 etwa 10.000 MW Generatorleistung in den Rechenzentren installiert sein.

Ob sich die Generator-Kapazitäten im Bereich Rechenzentren tatsächlich so wie erwartet entwickeln, ist von einer Reihe von Faktoren abhängig:

- Wie oben bereits dargestellt, ist die Langfristprognose mit großen Unsicherheiten behaftet, die sich insbesondere aus der Markt- und Technologieentwicklung ergeben, aber auch aus den verfügbaren Stromnetzkapazitäten.
- Möglicherweise verändern sich Redundanzkonzepte für Rechenzentren (Standortredundanz anstatt Stromredundanz). Anstatt von redundanter Stromversorgung wird die hohe Verfügbarkeit durch dynamische Verlagerung von Services auf andere Rechenzentrumsgebäude (ggf. an anderen Orten) erreicht (sog. Standortredundanz; ggf. auch Georedundanz bei >200 km Entfernung). Dies erfordert jedoch, dass jederzeit sichergestellt ist, dass an anderen Standorten ausreichend freie IT-Kapazitäten entsprechend den Redundanz-Anforderungen (z. B. N+1, 2N…) vorgehalten werden.
- Für neue Spezialrechenzentren AI/High-Performance-Computing könnte auf Notstromgeneratoren verzichtet werden, da ein Stromausfall des Rechenzentrums geringere Auswirkungen hätte. Solche Konzepte werden heute bereits bei HPC-Rechenzentren im Forschungsbereich praktiziert (Beispiel Green Cube des GSI Helmholtz-Zentrum, HLRS Stuttgart).

Ob sich diese Generatorkapazitäten für die Bereitstellung von Flexibilitäten nutzen lassen, ist von einer Reihe von Faktoren abhängig, die eingehender untersucht werden sollten:

- Der primäre Zweck von Ersatz-Stromversorgungssystemen ist die Sicherstellung der Stromversorgung von Rechenzentren bei Ausfall der Stromversorgung aus dem Netz, da größere Rechenzentren zu den kritischen Infrastrukturen gehören. Ein Konzept zur Bereitstellung von Flexibilitäten darf diese Sicherheitsfunktion nicht beeinträchtigen. Aus aktueller Sicht scheint es technisch insbesondere realisierbar zu sein, dass Rechenzentren auf Anfrage vom Stromnetzbetreiber den Bezug von Strom aus dem Netz vollständig unterbrechen und die Stromversorgung aus den Netzersatzanlagen im Inselbetrieb realisieren. Das Netz kann dabei ggf. als Redundanz dienen. Eine solche Verfahrensweise (z. B. Black Building Test) wird heute bereits regelmäßig zur Überprüfung der Abläufe bei Stromausfällen und der Funktionsweise von Netzersatzanlagen (teilweise im Zusammenhang mit TIER 3/4 Zertifizierung, KritisV und EN 50600 Zertifizierung) durchgeführt.
- Eine Bereitstellung von Flexibilitäten aus Notstromgeneratoren erfordert neben der technischen Anpassung der Anlagen auch eine Änderung in den Betriebskonzepten und den organisatorischen Abläufen im Rechenzentrum. Es stellt sich auch die Frage, wer der eigentliche Anbieter der Leistung ist, wie die jeweilige Konstellation aus Sicht der Regulierung des Energiemarktes zu bewerten ist und wie entsprechende Leistungen zu vergüten sind.
- Eine flexible Bereitstellung von Stromleistung aus Netzersatzanlagen erfordert wahrscheinlich größere Anpassungen in der technischen Ausstattung. Es ist anzunehmen, dass die bestehenden Anlagen (insb. die Leistungselektronik) nicht dafür ausgelegt wurden. Damit können die Anlagen zwar in das Rechenzentrum einspeisen, nicht zwangsläufig aber auch in das Stromnetz. Dies ist derzeit auch regulatorisch nicht vorgesehen, eine Rückspeisung muss stets sicher vermieden werden, auch beim Testbetrieb der Netzersatzanlagen (insb. EnWG, VDE TAR1).

- Für Dieselgeneratoren existieren Ausnahmeregelungen für Notbetrieb in der 44. Bundes-Immissionsschutz-Verordnung (BImSchV), wodurch ggf. höhere Abgase und Lärmemissionen erzeugt werden dürfen als bei allgemeinen Verbrennungsmotoranlagen zur Stromerzeugung.
- In der Praxis gibt es schon Lösungen, bei denen die notwendigen Testläufe der Notstromaggregate zeitlich so geplant werden, dass sie netzdienlich (insb. für Regelleistung) genutzt werden können. Damit ist allerdings das technische Flexibilisierungspotenzial bei Weitem nicht ausgeschöpft.
- Neue technischen Lösungen, wie Brennstoffzellen oder Gasturbinen, könnten in Zukunft eine Alternative zu Dieselgeneratoren darstellen. Dann wären aber auch Anpassungen in den Betriebskonzepten von Netzersatzanlagen notwendig.
- Mit einer systemischen Verbindung von Großrechenzentren, Großbatteriespeichern und Elektrolyseuren könnten Synergien erreicht und damit Vorteile, sowohl für die jeweiligen Anlagenbetreiber als auch für die Stromnetzbetreiber, generiert werden.

Die Nutzung von Batteriekapazitäten für kurzfristige Maßnahmen zur Netzstabilisierung ist prinzipiell mit ähnlichen Herausforderungen verbunden wie die Nutzung der Notstromgeneratoren. Hier besteht allerdings noch ein größerer Bedarf für detailliertere Untersuchungen. Dieser zusätzliche Bedarf ist durch folgende Aspekte begründet:

- Die Abschätzung der Größenordnung der Batteriekapazitäten in Rechenzentren ist schwieriger als die Abschätzung der Kapazitäten der Generatoren. Dies ist zum einen dadurch begründet, dass es neben Batterien auch Schwungradspeicher gibt. Bisher sind allerdings kaum Daten zu den Marktanteilen der verschiedenen Technologien begründet. Schwungradspeicher dienen meist nur für sehr kurze Zeiten zur Überbrückung von Stromausfällen, sie sind also für Flexibilitäten nicht geeignet. Zum zweiten ist für die Bestimmung der Kapazitäten von Bedeutung, für welche Überbrückungszeit die Batterien ausgelegt sind. Auch hierzu liegen kaum verlässliche Marktdaten vor.
- Ganz grob geschätzt kann davon ausgegangen werden, dass eine Leistung von etwa 1.500 bis 2.000 MW durch Batteriespeicher in Rechenzentren bereitgestellt wird. Wenn sich die Technologie und die Redundanzkonzepte nicht fundamental ändern, ist zukünftig eine Entwicklung vergleichbar zu den Generatoren wahrscheinlich. Diese Leistung steht aber aktuell nicht für Flexibilitäten zur Verfügung, da die Batterien so dimensioniert wurden, dass ihre Kapazitäten für eine jeweils vorgegebene Zeit ausreichen, die Rechenzentren mit Strom zu versorgen. Eine Nutzung der Batterien für Flexibilitäten würde damit die Verfügbarkeit/Reichweite (also Überbrückungsdauer) einer USV einschränken.
- Sollen Batterien in Rechenzentren für Flexibilitäten genutzt werden, müssten ihre Kapazitäten, die Betriebskonzepte und die organisatorischen Abläufe entsprechend angepasst werden.
- Aktuell werden, wie bereits beschrieben, überwiegend Bleibatterien in Rechenzentren eingesetzt. Bleibatterien eignen sich aber aufgrund ihrer geringeren Zyklenfestigkeit kaum für die Bereitstellung von Flexibilitäten.

Zusammenfassend kann festgehalten werden, dass bisher sowohl die Netzersatzanlagen (meist Dieselgeneratoren) als auch die in Rechenzentren vorhandenen Batteriekapazitäten primär als Notversorgung für einen Ausfall des Stromnetzes konzipiert wurden. Inwiefern diese Kapazitäten für Stromnetz oder Strommarkt für Systemdienstleistungen oder Kapazitätsreserven bereitgestellt werden

können, hängt von zahlreichen technischen, ökonomischen, organisatorischen wie auch regulativen Faktoren ab.

2.4 Rechenzentren und die Abwärmenutzung

2.4.1 Die Entstehung von Abwärme in Rechenzentren und Bedingungen der Abwärmenutzung

Wie in vielen anderen technischen Prozessen entsteht auch in Rechenzentren durch die Umwandlung von elektrischem Strom Wärmeenergie. Gemäß dem Energieerhaltungsgesetz kann angenommen werden, dass die verbrauchte elektrische Energie vollständig in Wärmeenergie umgewandelt wird. Demnach entspricht die in Rechenzentren theoretisch entstehende Wärme etwa dem gesamten Stromverbrauch der Rechenzentren, welcher zuvor beschrieben wurde. Es existieren jedoch diverse technische Einschränkungen in Rechenzentren, welche die vollständige Nutzung dieser Abwärme verhindern:

- **Strukturelle Barrieren in älteren Rechenzentren:** In älteren Rechenzentren ist die technologische Infrastruktur häufig nicht darauf ausgelegt, Wärme effizient auszukoppeln und nutzbar zu machen. Eine Studie der Universität Kassel zeigt, dass es in vielen älteren Rechenzentren erhebliche strukturelle Anpassungen bräuchte, um die Abwärme auszukoppeln (Orozaliev, 2023). Gerade bei heterogenen Kühlkonzepten mit vielen dezentralen Kühlkreisläufen (besonders relevant bei CRAC) ist eine zentrale Erfassung und Auskopplung mit enormen technischen Herausforderungen verbunden.
- **Begrenzte Nutzbarkeit der Wärme:** Nicht die gesamte durch den Stromverbrauch der Rechenzentren erzeugte Wärme lässt sich technisch sinnvoll nutzen. Aufgrund von technischen Restriktionen und Wirkungsgradverlusten z. B. bei der Wärmeübertragung kann nur ein bestimmter Anteil dieser Wärme weiterverwendet werden. Grob abschätzen kann man, dass die Wärme, die direkt durch den Stromverbrauch der IT-Geräte entsteht und über die Kühlanlagen abgeführt wird, weitgehend genutzt werden kann. Die Wärme, die durch den Stromverbrach der Gebäudetechnik entsteht, wird zum großen Teil kaum wiederverwendet werden können.
- **Abhängigkeit von Kühltechnologie und Temperaturniveau**: Die Effizienz der Abwärmenutzung ist stark abhängig von der eingesetzten Kühltechnologie und dem Temperaturniveau der erzeugten Abwärme. Neuere Kühlmethoden, wie das Direct-Liquid-Cooling oder die Immersionskühlung, bieten hierbei erhebliche Vorteile, da sie die Abwärme in einem höheren Temperaturniveau bereitstellen können (siehe auch Kapitel 5 – Technologische Trends). Die tatsächliche Qualität der Abwärme hängt dann aber auch von der gewählten Eintrittstemperatur des Kühlmediums ab. Gemäß ASHRAE sind zwar Temperaturen von über 45 °C erlaubt, aber in der Realität kann mit niedrigeren Kühlwassertemperaturen gefahren werden. Das Open Compute Project (OCP) schlägt diesbezüglich eine einheitliche Kühlmitteltemperatur von 30 °C vor (Mills et al., 2024).
- **Verfügbarkeit von Wärmenetzen**: Die externe Nutzbarkeit der Abwärme ist davon abhängig, ob in der Nähe eines Rechenzentrums ein vorhandenes Fern- oder Nahwärmenetz existiert, ein neues Wärmenetz geplant ist oder aufgrund der Ansiedlung des Rechenzentrums geplant

werden kann. Hierbei kann auch relevant sein, ob in der kommunalen Wärmeplanung (KWP) Eignungsgebiete für Wärmenetze ausgewiesen werden bzw. eine Nahwärmeversorgung von z. B. Industriebetrieben vorgesehen ist. Aus den kommunalen Wärmeplänen lassen sich Eignungsgebiete für die Fernwärmeversorgung erkennen. Sie werden nach verschiedenen Kriterien ermittelt, die in die Bestands- und Potenzialanalyse eingeflossen sind. Dies kann also ein erster Anhaltspunkt sein, um zu prüfen, ob das Potenzial für ein Fernwärmenetz besteht. Es ist allerdings zu beachten, dass es sich hierbei nur um die Ausweisung von Eignungsgebieten handelt und nicht mit Sicherheit gesagt werden kann, ob das Wärmenetz tatsächlich gebaut wird oder nicht. Auch kann nicht ausgeschlossen werden, dass sich in den nächsten Jahren weitere Eignungsgebiete finden lassen könnten, wenn die Wärmebedarfsdichte dort ausreichend hoch ist für eine zentrale Versorgung.

Als Zwischenfazit kann festgehalten werden, dass das Potenzial der aus Rechenzentren bereitgesellten Abwärme erheblich ist, strukturelle und technische Faktoren die tatsächliche Nutzbarkeit aber einschränken. Wird die Abwärmenutzung schon beim Bau der Rechenzentren vorgesehen, wie dies im EnEfG ab 1. Juli 2026 vorgeschrieben ist, so kann rein technisch am Rechenzentrum ein großer Teil der durch den IT-Betrieb verursachten Wärme zur Verfügung gestellt werden.

2.4.2 Herausforderungen bei der Nutzung von Abwärme aus Rechenzentren auf Seiten des Wärmenetzes und der Wärmesenken

Durch die KWP und das EnEfG wurden bedeutende Fortschritte erzielt, um die Nutzung von Abwärme aus Rechenzentren zu befördern und damit zur Dekarbonisierung der Wärmenetze beizutragen.

So werden im Rahmen der KWP lokale Potenziale für erneuerbare Energien und Abwärmequellen identifiziert. Hinzu kommt der regulatorische Rahmen in Form des EnEfG, das verbindliche Energieeffizienzstandards und die Nutzung von Abwärme aus Rechenzentren festlegt. Trotz dieser Fortschritte sowie bereits existierender Praxisbeispiele für die Abwärmenutzung aus Rechenzentren bestehen technische, organisatorische, wirtschaftliche und regulatorische Herausforderungen, die im Folgenden beschrieben werden.

Die **technischen Herausforderungen** reichen dabei von der Qualität der Abwärme über die notwendige Infrastruktur bis hin zu den dynamischen Anforderungen von Fernwärmesystemen.

Wie bereits beschrieben ergibt sich eine große Herausforderung aus der Temperatur der Abwärme. Die Abwärme aus Rechenzentren ist oft von niedriger Qualität, mit Temperaturen unter 50°C, typischerweise 30 bis 35°C.Diese niedrigen Temperaturen sind für bestehende Hochtemperatur-Fernwärmesysteme, welche Temperaturen von bis zu 120°C benötigen, oft unzureichend. Um diese Diskrepanz zu überwinden, müssen Wärmepumpen eingesetzt werden, die die Temperatur auf ein nutzbares Niveau anheben. Dies erfordert zusätzliche Investitionen und erhöht den Energiebedarf des Systems. Hierbei muss aber auch erwähnt werden, dass die Abwärme in einem konstanten Temperaturbereich zur Verfügung steht. Das bedeutet, dass im Gegensatz zu Großwärmepumpen, welche Umgebungstemperaturen (Außenluft, Flusswasser, Seewasser etc.) nutzen, vor allem im Winter ein höherer COP erzielt werden kann. Eine weitere Möglichkeit bietet die Einspeisung der Abwärme in den Rücklauf des Wärmenetzes, wodurch die benötigte Energie zur Erreichung der

Vorlauftemperatur reduziert werden kann. Hierbei müssen weniger hohe Temperaturen erreicht werden, wodurch je nach Netztemperatur auf eine zusätzliche Wärmepumpe verzichtet werden kann.

Für die Integration der Abwärme ist eine geeignete physische Verbindung zwischen dem Rechenzentrum und dem Fernwärmenetz erforderlich. Dies umfasst die Installation von Rohrleitungen, Wärmetauschern und gegebenenfalls Wärmepumpen. Die Auslegung dieser Infrastruktur ist komplex, da die Temperatur, der Druck und die Strömung des Mediums optimiert werden müssen, um Verluste zu minimieren und eine effiziente Wärmeübertragung zu gewährleisten. Hierbei muss auch auf eine optimale Leitungsführung und auf eine thermische Isolation der Rohre geachtet werden. Besonders in Niedertemperatur- oder Ultratiefnetzen sind präzise Anpassungen notwendig, um die Abwärme optimal in das Netz einzuspeisen.

Eine weitere Herausforderung ist die zeitliche und räumliche Diskrepanz. Rechenzentren produzieren Abwärme konstant über das gesamte Jahr, während die Wärmenachfrage in Fernwärmenetzen saisonal schwankt und im Sommer deutlich geringer ist. Diese zeitliche Diskrepanz kann dazu führen, dass ein Großteil der Abwärme im Sommer ungenutzt bleibt. Eine Lösung bieten thermische Energiespeicher, wie z. B. große Erdbeckenspeicher oder geothermische Speicher, die überschüssige Wärme für späteren Gebrauch speichern. Allerdings erhöhen diese Systeme die Komplexität und die Investitionskosten und sind besonders bei großen Netzwerken mit hohen Anforderungen an die Speicherleistung verbunden.

Die auch im Tagesverlauf dynamische Natur der Wärmeproduktion und -nachfrage stellt eine weitere Herausforderung dar. Rechenzentren liefern eine konstante Wärmemenge, während Fernwärmenetze mit stark schwankenden Lasten umgehen müssen. Abhängig von der Größe des Wärmenetzes fungiert dieses bis zu einem gewissen Maße als Puffer. Je nachdem wie steuerbar die restlichen Erzeugungsanlagen sind, kann allerdings ein Steuerungssystem notwendig sein. Fortgeschrittene Regelungs- und Steuerungssysteme sind notwendig, um diese Diskrepanzen auszugleichen und die Wärmeversorgung effizient zu gestalten. Die Systeme müssen in der Lage sein, die Einspeisung der Abwärme in Echtzeit an die aktuelle Nachfrage anzupassen und gleichzeitig den Betrieb von Wärmepumpen und Speichersystemen zu koordinieren. Herausfordernd ist hierbei zusätzlich, dass Rechenzentren typischerweise schrittweise den Betrieb aufnehmen. Das bedeutet, dass nicht immer bekannt ist, wie viel Wärme ab welchem Zeitpunkt zur Verfügung stehen wird.

Die Integration von Abwärme in ein bestehendes Fernwärmenetz kann die thermodynamischen und hydraulischen Eigenschaften des Systems beeinflussen. Beispielsweise kann das Einspeisen von Wärme mit niedriger Temperatur das Rücklaufwasser im Netz beeinflussen, was sich auf die Effizienz anderer Wärmequellen und auf den Gesamtwirkungsgrad des Systems auswirken kann. Darüber hinaus müssen die Strömungsdynamik und der Druck im Netz so angepasst werden, dass es nicht zu unerwünschten Störungen kommt.

Ab einer gewissen Größe des Rechenzentrums ist es außerdem schwierig, einen großen Teil der Abwärme zu nutzen, wenn nicht genug Wärmenachfrage besteht oder das Wärmenetz nicht genügend Kapazität hat. Es existieren mittlerweile mehrere Projekte für Rechenzentren-Campi oder auch KI-Rechenzentren im hohen zweistelligen oder sogar dreistelligen Megawattbereich, was gerade in ländlicher Umgebung kaum ein Wärmenetz aufnehmen kann.

Die Ergebnisse der vorliegenden Studie zeigen (Hintemann et al., 2024a), dass für die Standortwahl der Rechenzentren vor allem die Stromnetz- und die Breitbandanbindung die wichtigsten Aspekte sind. Zusätzlich ist das Vorhandensein von erneuerbarem Strom ein entscheidender Faktor und damit auch das Vorhandensein von Potenzialflächen. Diese Flächen finden sich oft in dünn besiedelten Gebieten, was wiederum häufig eine geringere Wärmenachfrage zur Folge hat. Im Rahmen einer Strategie für digitale Infrastruktur sollte die gezielte Ansiedlung von Rechenzentren in Regionen mit Wärmebedarf untersucht werden. Das bedeutet, dass sich Rechenzentren auch an Orten ansiedeln können, in welchen nur eine geringe Wärmenachfrage besteht. Durch die thermischen Verluste bei längerem Transport von Wärme kann Abwärme (im Gegensatz zu Strom) nicht wirtschaftlich über längere Distanzen transportiert werden.

Die zentrale Fragestellung mit Blick auf die **organisatorischen Herausforderungen** bei der Abwärmenutzung von Rechenzentren besteht darin, wie Wärme(netz)nutzer und Rechenzentren als Wärmeerzeuger und potenzielle Einspeiser zusammenfinden.

Durch die Plattform für Abwärme ist bereits ein erster wichtiger Schritt in die Umsetzung gekommen, um die Abwärme von Rechenzentren systematisch zu erfassen und deren Nutzung in Wärmenetzen zu stärken. Gemäß EnEfG sind Unternehmen mit einem jährlichen Gesamtendenergieverbrauch von mehr als 2,5 GWh verpflichtet, ihre Abwärmedaten auf der „Plattform für Abwärme“ zu melden. Dies umfasst zwar auch die Betreiber von Rechenzentren, jedoch werden für die Wärmeabnahme auch kleinere Wärmemengen wichtig, also auch jene Rechenzentren, deren Gesamtenergieverbrauch unter dem im EnEfG festgelegten Schwellenwert liegt. Die Summe der kleinen Wärmemengen bildet ein großes Potenzial, das durch das Instrument des Abwärmeregisters derzeit nicht erfasst, sondern nur im Rahmen der KWP identifiziert und ggf. gehoben werden kann.

Mit Blick auf die verfügbaren Informationen zu Wärmenetzen bzw. zur Fernwärmeversorgung ist die derzeitige Datenlage deutlich schlechter als bei Strom und Gas. Da Fernwärme- und Verteilungsunternehmen (FVU) nicht unter die Vorgaben des Energiewirtschaftsgesetzes (EnWG) fallen, sind sie zu deutlich weniger Datenlieferungen gegenüber staatlichen Stellen verpflichtet. Eine solide Datengrundlage über vorhandene und geplante Wärmenetze ist jedoch Voraussetzung für die politische Steuerung und das Monitoring der Wärmewende. Die dahingehende Transparenz sollte daher dringend verbessert und den zuständigen Planungsverantwortlichen zugänglich gemacht werden.

Im Kontext der im Rahmen der KWP zu erstellenden Wärmepläne ist festzuhalten, dass laut dem Bundesinstitut für Bau-, Stadt- und Raumforschung (BBSR) bisher nur wenige fertige Wärmepläne vorliegen. Diese Zahl differiert stark nach Bundesland. In Baden-Württemberg gab es bereits vor der im Wärmeplanungsgesetz (WPG) verankerten Verpflichtung eine landesspezifische Verpflichtung zur Erstellung von kommunalen Wärmeplänen, die daher dort für große Kommunen auch durchweg bereits vorliegen.

Hinzu kommt jedoch, dass Wärmepläne keiner bundesweiten Veröffentlichungspflicht unterliegen, und dies auf den Länderebenen geregelt werden muss. Ferner zeigt der aktuelle Stand fertiger Wärmepläne, dass diese kurzfristig keine ausreichende Grundlage für die Lokalisierung verfügbarer Wärmenetze und weitere Informationen zu geplanten Wärmenetzen liefern können. Während sich Kenntnisse über bestehende und vorgesehene Wärmenetze grundsätzlich nur lokal ableiten lassen, werden weitere wesentliche Informationen nicht dokumentiert bzw. zugänglich gemacht. Technische

und energiewirtschaftliche Details sind alleinige Sache der Wärmenetzbetreiber, die dafür zu keiner Transparenz verpflichtet sind. Gemäß §32 WPG sind alle Wärmenetzbetreiber dazu verpflichtet, bis Ende 2026 einen Wärmenetzausbau- und Dekarbonisierungsfahrplan zu erstellen und auf der Internetseite des Betreibers des Wärmenetzes zu veröffentlichen. Daten nach § 11 Absatz 4 können jedoch durch den Betreiber des Wärmenetzes von der Veröffentlichung ausgenommen werden.

Auch auf Daten, die durch Fernwärmeunternehmen laut EnStatG § 5 zu übermitteln sind, gibt es nur eingeschränkten Zugriff. In der durch das Forschungsprojekt DEKADE-F-Wärme aufbereiteten Übersicht werden große Wärmenetze abgebildet, während viele weitere in Mittel- und Kleinstädten bzw. auf Quartiersebenen nicht dargestellt sind. Einen noch unvollständigen Überblick für die Ebene der Bundesländer liefert der Datenkompass des Kompetenzzentrums Kommunale Wärmewende (KWW)[7].

Eine weitere Herausforderung bilden die unterschiedlichen Planungshorizonte beim Ausbau von Wärmenetzen und bei der Planung von Rechenzentren. Ferner fehlt es an Abstimmungen zentraler Akteure sowohl auf regionaler als auch auf kommunaler Ebene. Letzteres betrifft die planungsverantwortlichen Stellen der KWP, lokale Wärmenetzbetreiber und Vertreter der kommunalen oder regionalen Wirtschaftsförderung.

Hervorzuheben sind die Herausforderungen, aber zugleich auch die strategischen Potenziale von Kohlestandorten, bei denen Fernwärmeleitungen bereits überwiegend vorhanden sind, Internetknoten jedoch nicht (z. B. Lausitz). Die Ansiedlung von Rechenzentren an diesen Standorten kann möglicherweise zum erforderlichen Infrastrukturwandel in diesen Regionen beitragen (siehe auch Kapitel 3), abhängig von der Wirtschaftlichkeit der Abwärmenutzung. Zu prüfen sind dafür die Entfernung zum versorgten Gebiet, der Wärmebedarf sowie die Verfügbarkeit von Fördermitteln. Die Braunkohleregionen profitieren vom 2020 verabschiedeten Investitionsgesetz Kohleregionen (InvKG) in den Bundesländern Brandenburg, Nordrhein-Westfalen, Sachsen und Sachsen-Anhalt, zudem von zahlreichen Maßnahmen des Bundes und der Länder, unter anderem durch Innovationsprogramme und die Europäischen Strukturfonds. Da es sich bei den Braunkohleregionen überwiegend um strukturschwache Regionen handelt, stehen ihnen zudem die Fördermöglichkeiten der Bund-Länder-Gemeinschaftsaufgabe „Verbesserung der regionalen Wirtschaftsstruktur“ (GRW) zur Verfügung.

2.4.3 Herausforderungen der Wirtschaftlichkeit der Abwärmenutzung

Neben den technischen und organisatorischen Herausforderungen gibt es bei der externen Nutzung von Abwärme aber vor allem auch **ökonomische Herausforderungen**, die eine umfassendere Nutzung erschweren. Als erster Punkt zu nennen sind hierbei vor allem die hohen Anfangsinvestitionen. Die Implementierung von Abwärmenutzungsprojekten erfordert erhebliche Investitionen. Dazu zählen die Installation von Wärmetauschern, die Anpassung bestehender Wärmenetze und die Verlegung von Infrastruktur für die Wärmeübertragung. Die anfänglichen Kapitalkosten sind besonders hoch, wenn bestehende Netze nicht für den Betrieb mit Niedertemperaturwärme ausgelegt sind, da dann zusätzlich eine Wärmepumpe und gegebenenfalls weitere Komponenten installiert werden müssen.

[7] KWW-Datenkompass zur Kommunalen Wärmeplanung - Kompetenzzentrum Kommunale Wärmewände

Die Amortisationszeit für solche Projekte beträgt häufig mehr als 10 Jahre, insbesondere wenn Wärmespeicher oder teure Verbindungsleitungen gebaut werden müssen. Dieser Amortisationszeitraum deckt sich nicht mit den Amortisationszeiten, mit denen Unternehmen normalerweise rechnen, was eine Implementierung deutlich erschwert (Li et al., 2021; He et al., 2018). Diese lange Amortisationszeit basiert auf der begrenzten Rentabilität der Abwärmenutzung. Niedrige Wärmepreise, wie sie in vielen europäischen Märkten üblich sind, erschweren es, aus Abwärme eine wirtschaftlich tragfähige Einnahmequelle zu machen.

In Deutschland ist hierbei aktuell vor allem noch die Konkurrenz zu den niedrigen Gaspreisen schwierig. In Zukunft wird der Wärmepreis, verglichen mit anderen erneuerbaren Erzeugungstechnologien, allerdings konkurrenzfähiger sein. Kostentreiber sind im Falle der Abwärmenutzung nicht die Abwärme selbst, sondern die Investitionen und Betriebskosten der Wärmeauskopplung.

Die physische Distanz zwischen Rechenzentren und potenziellen Wärmenutzern erhöht die Investitionskosten erheblich. Lange Verbindungsleitungen bedeuten höhere Material- und Baukosten sowie Energieverluste bei der Wärmeübertragung. Diese Kosten machen kleinere Abwärmequellen oder entfernte Rechenzentren weniger attraktiv für die Integration in Wärmenetze. In Abbildung 22 sind die Einbindungskosten für unterschiedliche RZ-Größen und Entfernungen zum Anschlusspunkt dargestellt. Für die Berechnung wurden nur die Kosten des Wärmetauschers und der Rohrleitungen bis zum Anschlusspunkt inkl. Planungs- und Tiefbaukosten sowie die Betriebskosten des Wärmetauschers und der Anbindungsleitung berücksichtigt. Eine Wärmepumpe sowie die Betriebskosten zum Betrieb der Wärmepumpe wurden nicht berücksichtigt.

2.4.4 Einschätzung wirtschaftlich nutzbarer Abwärmemengen aus Rechenzentren

Aktuell wird in Rechenzentren in Deutschland verhältnismäßig wenig Abwärme genutzt, wobei die Zahl der Rechenzentren mit Abwämenutzung steigt und mittlerweile zumindest mehr als die Hälfte der in Hintemann et al., (2024a) befragten Unternehmen Teile ihrer Abwärme nutzen. Ausgehend von dem in Kapitel 1 beschriebenen Zuwachs an Rechenzentren bis 2045 sowie den technischen und ökonomischen Hürden auf Seiten der Rechenzentren, Wärmenetze und Wärmesenken sowie dem regulatorischen Rahmen kann eine grobe Prognose für in Zukunft genutzte Abwärme getroffen werden.

Geht man davon aus, dass die neuen gesetzlichen Vorgaben des EnEfG von den neu errichteten Rechenzentren eingehalten werden, so könnte im Jahr 2030 etwa 1 TWh Wärme aus Rechenzentren beim Endverbraucher genutzt werden. Im Jahr 2035 ständen bereits mehr als 3 TWh Wärme aus Rechenzentren zur Verfügung. Bis zum Jahr 2045 würde sich die nutzbare Abwärmemenge auf etwa 10 TWh pro Jahr erhöhen[8]. Dieser starke relative Anstieg (im Vergleich zu der Entwicklung der Gesamtkapazität von Rechenzentren) lässt sich damit begründen, dass die Vorgaben zur Abwärmenutzung im EnEfG nur für neue Rechenzentren ab der Inbetriebnahme zum 1. Juli 2026 gelten.

[8] Es ist davon auszugehen, dass ohne die Einführung des EnEfG nur geringe Mengen an Abwärme genutzt würden.

Diese prognostizierten Mengen können sich, abhängig von der Entwicklung des Rechenzentrumsmarktes sowie dem regulatorischen Rahmen sowohl auf Seiten der Rechenzentren als auch bei der Netz- und Abnehmerseite (z. B. KWP), noch stark nach unten oder oben verändern. Auch technologische Trends wie die breite Einführung von Niedertemperaturwärmenetzen der 4. und 5. Generation sowie Distributed Computing oder Edge-Rechenzentren könnten die Entwicklung noch stark beeinflussen.

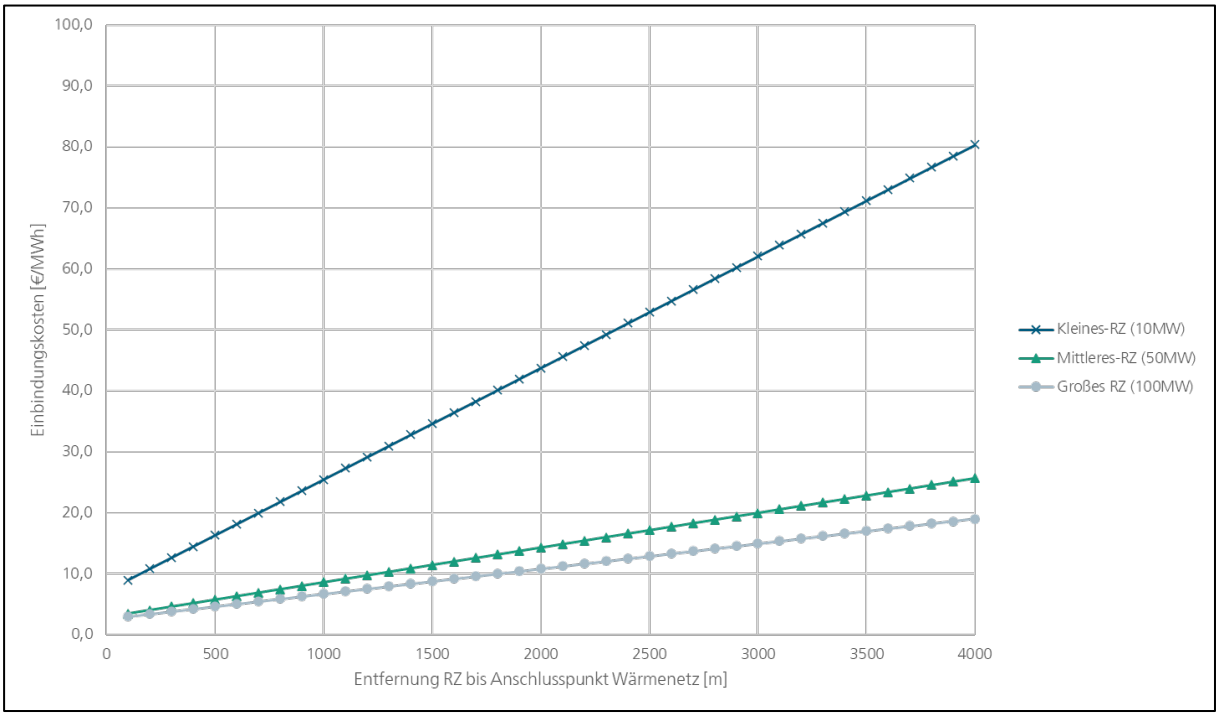


Abbildung 22: Einbindungskosten von Abwärme aus Rechenzentren in € pro MWh in Abhängigkeit der Entfernung (in Meter), Quelle: Berechnung Fraunhofer ISI

Zusätzlich ergeben sich aus den bereits beschriebenen technischen Herausforderungen auch ökonomische Herausforderungen. Die Aufbereitung durch Wärmepumpen oder andere Technologien führt zu zusätzlichen Energiekosten, was die Betriebskosten signifikant erhöht. Dadurch wird die wirtschaftliche Attraktivität solcher Projekte weiter gemindert. Auch die Notwendigkeit von Speicherlösungen, um die zeitliche Verschiebung der Wärmenachfrage trotz gleichbleibendem Abwärmeangebot zu decken, kann die wirtschaftliche Amortisation verlängern. Während große Rechenzentren potenziell genug Abwärme erzeugen, um wirtschaftlich tragfähige Projekte zu ermöglichen, sind kleinere RZ häufig nicht in der Lage, die erforderlichen Skaleneffekte für einen Anschluss an ein Fernwärmenetz zu erzielen. Die Fixkosten für die Infrastruktur und Technologien übersteigen in diesen Fällen die wirtschaftlich nutzbaren Erträge. Aufgrund der niedrigen Marktpreise für Wärme und fehlender Anreizmechanismen ist es für Betreiber von Wärmenetzen oft nicht attraktiv, Abwärme aus RZ in ihre Systeme zu integrieren. Hinzu kommt, dass insbesondere die Kostenverantwortlichkeit für die zur Abwärmenutzung erforderliche Infrastruktur, Anlagen und Maßnahmen nicht klar definiert und abgegrenzt werden kann. Hinzu kommt, dass gemäß Umsatzsteuergesetzgebung auch bei unentgeltlicher Abgabe der Abwärme die Umsatzsteuer anfällt.

Ferner sind im Bereich der Fernwärme auf lokaler Ebene seit Jahrzehnten monopolartige Strukturen zu verzeichnen. Entsprechend hoch sind – lokal sehr unterschiedlich und abhängig vom jeweiligen

lokalen FVU – die Widerstände für die Verbesserung der Datentransparenz für Anspruchsberechtigte. Das Monopol im Fernwärmemarkt ist auch hinderlich vor dem Hintergrund eines fehlenden wettbewerblichen Anreizes für die Abnahme bzw. Einspeisung von Abwärme durch Wärmenetzbetreiber, die auch – wie oben ausgeführt – mangels gesetzlicher Verpflichtung eine Einspeisung ablehnen können. Denn die Einspeisung von Abwärme (aller Art) basiert immer auf privatrechtlichen Verträgen, die nur bei Interesse des FVU zustande kommen, unabhängig der gesamtökonomischen und vor allem ökologischen Vorteile.

In Deutschland ist die Fernwärme heute noch zu 86 % von fossilen Kraft-Wärme-Kopplungs-Anlagen geprägt. Dagegen ist der Beitrag der klimaneutralen Wärmeerzeugung mit 19 % vergleichsweise gering. Bis 2030 soll deren Anteil bei mindestens 50 % liegen. Um das zu erreichen, soll künftig für Fernwärme und erneuerbare Wärme ein überragendes öffentliches Interesse gelten. Das gegenwärtig noch vorherrschende Geschäftsmodell über Kraft-Wärme-Kopplung behindert direkt und indirekt (z. B. über die Berechnungsformel zum Primärenergiefaktor) eine schnellere Dekarbonisierung (dena, 2023).

2.5 Handlungsempfehlungen

Im Folgenden werden aus den Erkenntnissen dieses Kapitels Handlungsempfehlungen abgeleitet, um den Rechenzentrumsstandort Deutschland zu stärken und gleichzeitig eine nachhaltige Integration der Rechenzentren ins Energiesystem zu unterstützen.

Flexibilisierung und Integration von Rechenzentren in den Strommarkt sowie Anreize für geographisch sinnvolle Ansiedlung

- Neben bereits existierenden Anreizen am Strommarkt, den Stromverbrauch in Zeiten von hohem Dargebot an Erneuerbaren Energien zu verlagern, könnten durch **neue Ökostrom-Bilanzierungsmethoden** wie die von energytag[9] oder CO_2-Gutschriften weitere Flexibilitäten angeregt werden.
- Der Bau eines Rechenzentrums setzt heute einen sehr umfangreichen Genehmigungsprozess voraus. Um sowohl Bürokratie zu reduzieren als auch einen Anreiz für eine flexible Ansiedlung zu bieten, könnten Kommunen, unterstützt durch die Länder/den Bund lokal verschiedene Flächenprüfungen vorwegnehmen, eng verzahnt mit der KWP. Ein Standort, der sowohl die Stromnetzkapazität bietet als auch für hohe Abwärmenutzungsanteile (über denen im EnEfG vorgegebenen) geeignet ist, könnte vorab auf weitere Kriterien (Umweltverträglichkeit, TA-Lärm, etc.) geprüft werden und dann mit einem vereinfachten Genehmigungsverfahren von Rechenzentren bebaut werden.
- **Prüfung und Anpassung der Netzanschlussanträge bei besonders großen Rechenzentren**: Die Anträge für neue Netzanschlüsse sollten angepasst werden, um die Netzverträglichkeit und Flexibilitätspotenziale neuer Rechenzentren strukturiert zu bewerten und gerade bei neuen Hyperscale-Rechenzentren (>30 MW) eine Minimalanforderung an die Demand-Response Fähigkeiten stellen.

[9] https://energytag.org/

- Im Bereich der Flexibilität aus Rechenzentren **fehlen derzeit noch fundierte Erkenntnisse zu den tatsächlichen Potenzialen** sowie der technischen, organisatorischen und regulatorischen Umsetzbarkeit. Hier sollten weitere wissenschaftliche Untersuchungen stattfinden sowie Pilotprojekte/Leuchttürme mit hoher Strahlkraft unterstützt werden.

Stärkung der Abwärmenutzung und -integration in Wärmenetze

- Kommunen sollten zusätzlich gefördert werden, wenn sie Abwärme aus Rechenzentren integrieren. Die Nutzung von industrieller Abwärme (z. B. aus Rechenzentren), die ansonsten in die Umwelt abgeführt wird, sollte immer als klimaneutral gelten, während für fossile Kraft-Wärme-Kopplung ehrliche Emissionsfaktoren angelegt werden sollten. Hierzu sollte auch das etablierte Bewertungsverfahren der Fernwärme (insb. die Berechnung anhand eines Verdrängungsstrommix) kritisch betrachtet werden, welches der fossilen KWK bis heute große Vorteile ermöglicht.
- Unabhängig von Rechenzentrumsabwärme sind Wärmenetze ein Schlüssel zur Dekarbonisierung, auch für andere Abwärme, Großwärmepumpen oder Biomasse. Da bei Investoren wenige Erfahrungswerte im Hinblick auf Wärmenetze mit Rechenzentren vorliegen, ist es häufig jedoch schwierig, eine Finanzierung von über 10 Jahren zu erhalten, da die Angst besteht, ein Rechenzentrum könnte nach einigen Jahren obsolet sein. Bei Wärmenetzen handelt es sich jedoch um eine Infrastrukturinvestition, welche unabhängig von einem konkreten Rechenzentrum für die Wärmewende notwendig ist. Deshalb wäre **ein staatlicher Fonds/eine Bürgschaft sinnvoll, welche die Investitionen in Wärmenetze absichern**.
- Die KWP sollte den Bedarf und die **Nutzbarkeit von Abwärme aus Rechenzentren regionenübergreifend analysieren**, um eine wirtschaftliche Nutzung auch über die Kommunengrenzen hinweg zu ermöglichen. Hierbei könnten der Bund oder die Länder eine Schlüsselfunktion einnehmen, indem sie die Wärmebedarfe aus der KWP zusammentragen und potenzielle Hotspotregionen, welche sich besonders für Rechenzentrumsabwärme eignen, identifizieren.

Transparenz und Datenzugänglichkeit für Planung und Umsetzung

- Ein deutschlandweites Wärmenetzregister sollte aufgebaut werden, das zentrale Kenndaten zur Verfügbarkeit und Nutzung von Wärmenetzen beinhaltet. Die Teilnahme für Wärmeversorgungsunternehmen (FVU) sollte verpflichtend sein, um die Datentransparenz für relevante Planungsakteure zu gewährleisten und damit auch die Möglichkeiten von Rechenzentren zu stärken, ihre Wärme in bestehende oder geplante Wärmenetze einzuspeisen.
- Mit dem **EnEfG werden systematisch Daten** von Rechenzentren zu deren Anschlussleistung und Energieverbrauch erhoben. Diese Daten (ggf. grobe Leistungsklasse der jährlich möglichen Wärmeauskopplung) sollten den Kommunen sowie Stadtwerken für ihre **lokale Wärmeplanung** zugänglich gemacht werden, sofern dies durch die Plattform für Abwärme der Bundesstelle für Energieeffizienz nicht ohnehin vorgesehen ist. Damit können die Kommunen ggf. auf bestehende Rechenzentren zugehen und mögliche Kooperationen für zukünftige Abwärmenutzung ausloten.

Technologieförderung und Anreize zur Effizienzsteigerung und Abwärmenutzung kombinieren

- Die öffentliche Beschaffung und Förderung spezifischer Kühltechnologien (z. B. Flüssigkühlung) für Rechenzentren sollte erweitert werden, um deren Energieverbrauch zu reduzieren und die Abwärmequalität für Fernwärmenetze zu erhöhen.
- Incentives, wie CO_2-Gutschriften oder etwas mehr Spielraum bei den Effizienzanforderungen (z. B. PUE-Werte des EnEfG), sollten Rechenzentren angeboten werden, um für eine umfangreichere Nutzung von Abwärme zu motivieren (bei neuen RZ über den vorgeschriebenen Werten und bei bestehenden RZ generell bei Abwärmenutzung).

3 Ökonomie der Rechenzentren

3.1 Zusammenfassung

Das dritte Kapitel analysiert die wirtschaftliche Bedeutung von digitaler Infrastruktur in Deutschland und zeigt deren Wertschöpfungspotenziale sowie mögliche Spillover-Effekte auf. Die Analyse der Beschäftigungseffekte macht deutlich: Für den Bau eines Rechenzentrums mit 100 MW werden durchschnittlich etwa 6,9 Arbeitskräfte pro MW und Baujahr über einen phasenweisen Gesamtbauzeitraum von 10 Jahren benötigt (Copenhagen Economics, 2020). Nachdem die Bauphase abgeschlossen ist, werden für den Betrieb permanente Arbeitsplätze geschaffen. Im Schnitt generieren deutsche Rechenzentrumsanbieter 9 Arbeitsplätze für den Betrieb eines 1 MW Rechenzentrumsgebäudes, während internationale Betreiber nur etwa 3 Arbeitsplätze pro Megawatt in Deutschland schaffen. Dies ist darauf zurückzuführen, dass ein Teil der Arbeitsplätze im Ausland entsteht, z.B. in zentralen Konzernfunktionen.

Die Wertschöpfungskette bei Rechenzentren gliedert sich in drei Veredelungsstufen: Rechenzentrumsgebäude, IT-Infrastruktur und IT-Dienste. Dabei steigt die Wertschöpfung mit jeder Veredelungsstufe. Während Rechenzentrumsgebäude primär Strom und Fläche in hochverfügbare Infrastruktur umwandeln, schaffen IT-Infrastrukturbetreiber, welche das IT-Equipment in RZ-Gebäuden betreiben, mit 3 bis 25 Mitarbeitern pro Megawatt bereits deutlich mehr Arbeitsplätze. Die höchste Wertschöpfung entsteht jedoch bei IT-Dienstleistern, die für den Betrieb eines Megawatts IT-Dienste zwischen 35 und 140 Mitarbeiter beschäftigen.

Der Markt wird aktuell stark von ausländischen Investoren geprägt. Die größten Rechenzentrumsflächen-Vermieter in Deutschland sind mehrheitlich in ausländischer, hauptsächlich US-amerikanischer, Hand. Auch bei den Cloud-Marktplätzen dominieren geschlossene Systeme internationaler Anbieter. Deutschland verfügt jedoch über eine starke mittelständische IT-Dienstleisterstruktur, die zukünftig als Basis für weiteres Wachstum dienen kann.

Eine besondere Herausforderung stellt der Fachkräftemangel dar, insbesondere für den Betrieb großer IT-Infrastrukturen. Anders als die USA, wo Digitalunternehmen jahrzehntelange Erfahrung im Aufbau großer IT-Systeme haben, fehlt in Deutschland oft das entsprechende Know-how.

Um diese Potenziale zu heben, empfiehlt die vorliegende Analyse die Entwicklung regionaler digitaler Ökosysteme sowie die Schaffung eines europäischen Marktplatzes für digitale Ressourcen. Weitere zentrale Handlungsfelder sind die gezielte Fachkräfteausbildung, die Verbesserung des Kapitalzugangs für den IT-Mittelstand sowie die Stärkung des Wettbewerbs durch die Öffnung geschlossener Märkte.

3.2 Wirtschaftliche Bedeutung von digitaler Infrastruktur in Deutschland

Um die wirtschaftliche Bedeutung von Rechenzentren in Deutschland zu untersuchen und zu verstehen, muss die gesamte Wertschöpfungskette mit einbezogen werden. Für diese Betrachtung definieren wir die Digitalwirtschaft als einen eigenen Sektor. Unter diesem Sektor verstehen wir die Infrastruktur und die verschiedenen Veredelungsschritte als eigene Wirtschaftszweige. Zur

Digitalwirtschaft zählen die Hard- und Softwareindustrie, die Telekommunikationsbranche (als Teil der Infrastruktur), die (digitale) Medienbranche und natürlich auch die IT-Dienstleistungsbranche.

Rechenzentren (Gebäude mit IKT-Equipment) bilden in diesem Sektor einen Teil der Infrastruktur, die digitale Ressourcen bereitstellt, die von anderen Wirtschaftszweigen veredelt werden. Eine visuelle Darstellung des so skizzierten Digitalsektors findet sich in Abbildung 23.

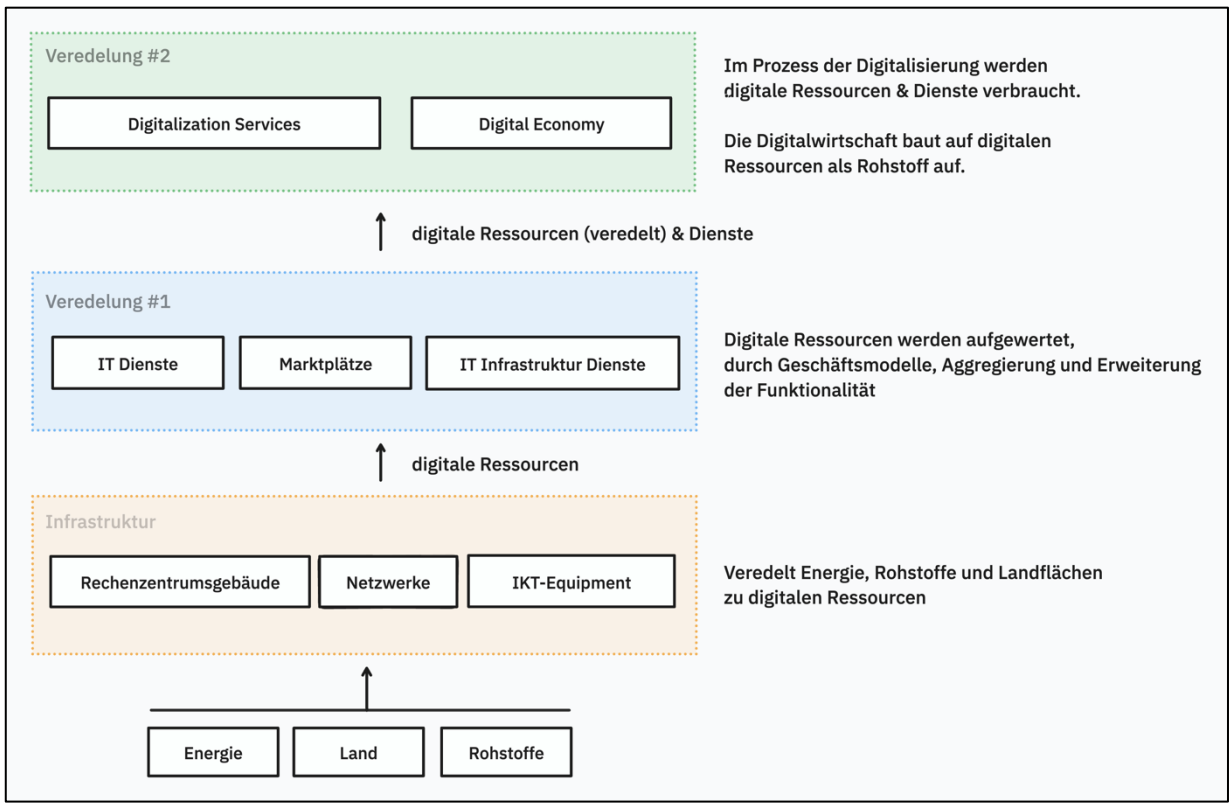


Abbildung 23: Vereinfachte Darstellung des Digitalsektors. Quelle: Eigendarstellung SDIA

Die Nachfrage nach digitalen Ressourcen und damit auch nach Infrastruktur wird primär durch die Digitalisierung und das Wachstum der Digitalwirtschaft angetrieben. Mit Digitalisierung ist hierbei der Prozess der Transformation von nicht-digitalen Sektoren und Branchen mit Hilfe von Software gemeint. Dieser Prozess wird primär durch IT-, Digital- und Unternehmensberatungen oder eigenständige, dafür zuständige Unternehmensbereiche implementiert. Dabei werden Software und digitale Ressourcen (Rechen-, Speicher- und Datenübertragungskapazität) für die Digitalisierung von Geschäftsprozessen, Arbeitsplätzen sowie Produkten und Dienstleistungen implementiert. Szenarios zum Digitalisierungsfortschritt in Deutschland finden sich in Kapitel 5, auch mit einer Ableitung der dafür notwendigen digitalen Ressourcen.

Im Rahmen der Digitalisierung kommen auch Produkte und Lösungen aus der Digitalwirtschaft zum Einsatz. Die Digitalwirtschaft kann als Branche im Digitalsektor verstanden werden und hat als Alleinstellungsmerkmal, dass die Unternehmen dieser Branche komplett digitale Produkte und Dienstleistungen, meist primär über den Internetmarktplatz, anbieten. In dieser Branche werden sowohl Produkte als auch Dienstleistungen für Endverbraucher entwickelt, wie zum Beispiel Online-Shopping-Angebote oder Reisebuchungsplattformen, aber auch Unternehmenslösungen angeboten, wie zum Beispiel digitale Lösungen zur Prozessautomatisierung, Textverarbeitung oder andere branchenspezifische Softwarelösungen.

Aus volkswirtschaftlicher Betrachtung entsteht durch Digitalisierung, bzw. die damit verbundenen Dienstleistungen und durch die Digitalwirtschaft, die höchste Wertschöpfung für die Gesellschaft.

3.2.1 IT-Dienstleister und -Beratungen sind neben der Digitalwirtschaft für den Großteil der Wertschöpfung verantwortlich

IT-Beratungen sind maßgebliche Anbieter von Digitalisierungsdienstleistungen in Deutschland. Laut Berichten von Marktanalysten (IBISWorld, 2024) hat dieser Bereich ein Marktvolumen von 66,2 Milliarden Euro und schafft mehr als 260.000 Arbeitsplätze in rund 36.000 Unternehmen. Die Branche wächst um ca. 8 % pro Jahr.

Laut dem IKT-Branchenbild der Bundesregierung, welches von ZEW erstellt wurde, werden die Dienstleistungen der Digitalisierung und der Digitalwirtschaft gemeinsam betrachtet. Für das Jahr 2020 wurde in diesen Branchen ein Umsatz von 188 Milliarden Euro erwirtschaftet[10] (siehe Abbildung 24). Laut den Daten des Statistischen Bundesamtes gibt es im Jahr 2022 in Deutschland 1,35 Mio. Beschäftigte in der IKT-Branche (Statistisches Bundesamt, 2024).

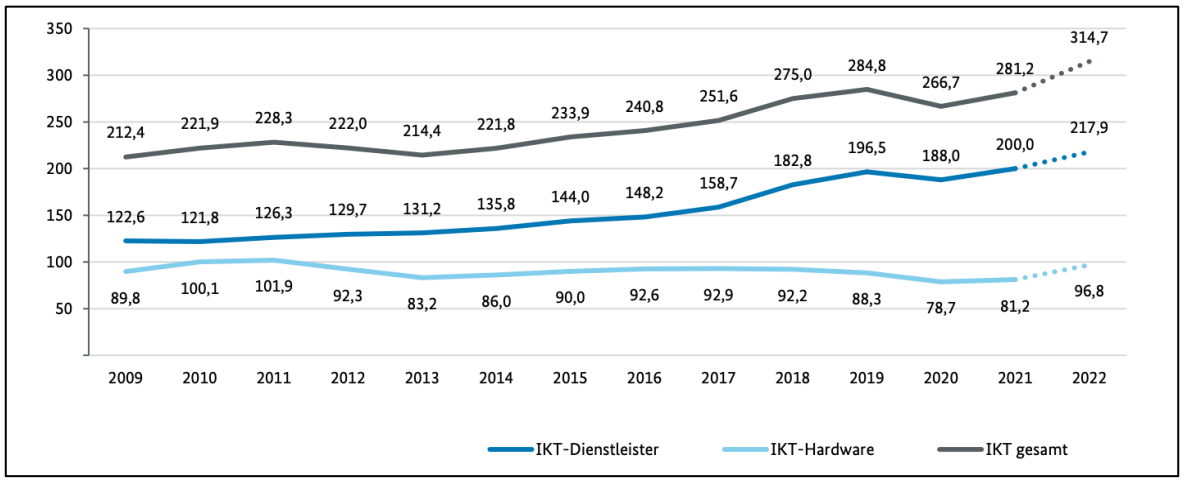


Abbildung 24: Entwicklung des Umsatzes in der IKT-Branche in Deutschland in den Jahren 2009 bis 2022 (in Milliarden Euro). Quelle: Berechnungen des ZEW, 2023. Werte am aktuellen Rand approximiert.

Im Folgenden betrachten wir die gesamtwirtschaftliche und regionale Bedeutung sowie den volkswirtschaftlichen Beitrag der digitalen Infrastruktur und der Veredelungsschritte, welche die digitalen Ressourcen für die Digitalisierungs-Dienstleistungsbranche und die Digitalwirtschaft bereitstellen und somit die Wertschöpfung ermöglichen.

[10] https://de.digital/DIGITAL/Navigation/DE/Lagebild/IKT-Branchenbild/ikt-branchenbild.html

3.3 Digitale Infrastruktur veredelt Energie und Rohstoff

Digitale Infrastruktur veredelt primär Energie und Rohstoffe in die Kapazität, Daten zu speichern, zu transportieren und zu verarbeiten („digitale Ressourcen"). Der Prozess der Veredelung lässt sich in die folgenden Schritte unterteilen.

Digitale Infrastruktur/Rechenzentrum:

1. **Gebäude***:* RZ-Gebäude werden aus Baustoffen wie Stahl und Beton gebaut, das mechanische und elektrische Equipment erfordert zudem weitere Rohstoffe, Metalle und seltene Erden. Mit den Gebäuden wird primär Strom aus dem Stromnetz aufgewertet und zu Strom mit Verfügbarkeitsgarantie (üblicherweise 99.999 %) und mit festen Qualitätskriterien (Redundanz, unterbrechungsfreie Stromversorgung, keine Spannungsschwankungen) veredelt. Zudem wird auch Kühlleistung aus Strom bereitgestellt. Die Fläche wird mit Sicherheitsanlagen, gesicherten Räumen, komplexem Brandschutz und einem Internetanschluss veredelt.
2. **IKT-Equipment***:* Mit hohem Energieaufwand werden Rohstoffe in Chips, IT-Equipment, Netzwerksysteme und Speichersysteme veredelt und in das Rechenzentrumsgebäude verbaut. Die vom Gebäude bereitgestellte Strom-, Kühl- und Internet-Übertragungsleistung wird vom IT- und Netzwerkequipment in digitale Ressourcen verwandelt. Die erzeugten digitalen Ressourcen sind nur nutzbar, wenn das IKT-Equipment mit Strom versorgt wird. Zudem kann Rechenkapazität nicht gespeichert werden, sondern muss direkt verbraucht werden. Solange ein IT-System läuft, wird Strom konstant in digitale Ressourcen gewandelt. Dieser Veredelungsprozess erzeugt Wärme, die durch die Kühlleistung aus dem Gebäude abgeführt wird.

IT-Infrastruktur, Marktplätze und Dienste:

Im zweiten Veredelungsschritt werden die digitalen Ressourcen in verschiedenen Varianten aufgewertet.

1. **IT-Infrastruktur***:* Durch Virtualisierung, Orchestrierung, Automatisierung oder manuelle Bereitstellung werden bedarfsgerecht digitale Ressourcen, z. B. mit hoher Verfügbarkeit (Redundanz), kosteneffizient oder mit hoher Qualität (z. B. geringe Latenz) zur Miete angeboten.

2. **Marktplätze („Cloud" bzw. Cloud-Infrastruktur)***:* Digitale Ressourcen können über Marktplätze bedarfsbasiert bezogen werden und zwar On-Demand (zum sofortigen Verbrauch), Spot (zum Verbrauch, wenn verfügbar, z. B. mit Preisgrenze) oder als Future (garantierte Ressourcen für einen Zeitraum mit fixiertem Preis).
3. **IT-Dienste***:* Digitale Ressourcen werden mit Software-Anwendungen kombiniert und als Dienstleistungen angeboten. Software veredelt in diesem Falle die digitalen Ressourcen und damit indirekt den Strom und die Rohstoffe, welche in den vorhergegangenen Veredelungsstufen eingebracht wurden. Zwei Beispiele für einen IT-Dienst:

   3.1. Ein Dienst, der eine Datenbank für die strukturierte Speicherung von Daten bereitstellt,

   3.2. ein Dienst, der die Möglichkeit bietet, Videos zu komprimieren.

Die Wertschöpfungskette lässt sich noch um einen Schritt erweitern. Digitale Produkte und Dienstleistungen, die aus verschiedenen Kombinationen von IT-Diensten und eigener proprietärer oder open-source-Software softwarebasierte Produkte bereitstellen. Die Hersteller der digitalen Produkte sind die primären Akteure der Digitalwirtschaft. Im Prozess der Digitalisierung werden oft digitale Produkte aus der Digitalwirtschaft in anderen Wirtschaftsbereichen für die Prozess- oder Produktdigitalisierung eingesetzt.

Am Beispiel einer Video-Konferenz-Lösung (digitales Produkt, z. B. Zoom, WebEx, Microsoft Teams) lässt sich der Sachverhalt veranschaulichen:

Eine Video-Konferenz-Lösung liefert eigene Software für die Anwender aus, z. B. eine WebEx-Desktop-Anwendung oder eine App für Smartphones.

Diese Applikation verbindet sich mit verschiedenen IT-Diensten, um die Funktionalitäten, wie z. B. eine Videokonferenz oder eine Gesprächsaufzeichnung bereitzustellen. Dies ist insbesondere der Fall, wenn das Endgerät nicht über ausreichend eigene digitale Ressourcen verfügt.

Die IT-Dienste laufen auf IT-Infrastruktur (Server-, Netzwerk- und Speichersysteme) und nutzen ggf. zusätzlich andere IT-Dienste für die Speicherung und Verarbeitung von Daten, z. B. für die Kompression von Video-Aufzeichnungen. In moderner Software-Architektur werden IT-Dienste zunehmend auf eine Funktion reduziert und rufen sich konstant gegenseitig auf (Stichwort “Microservices”).

Die IT-Infrastruktur steht in einem Rechenzentrumsgebäude, welches den Strom und die Kühlleistung bereitstellt und den gesicherten Betrieb der IT-Infrastruktur sicherstellt.

3.4 Zentrale Geschäftsmodelle

Rechenzentrumsgebäude-Betreiber, IT-Infrastruktur-Anbieter und IT-Dienstleister differenzieren ihre angebotenen Produkte mit verschiedenen Geschäftsmodellen. Bevor die einzelnen Geschäftsmodelle betrachtet werden, im Folgenden eine Übersicht der Produkte und Merkmale in den Veredelungsstufen:

Veredelungsstufe	Produkt	Merkmale
Rechenzentrumsgebäude-Vermieter	Quadratmeter Stellfläche für IT-Equipment	hochverfügbarer Strom, Kühlleistung, Sicherheitsgarantien, Brandschutz und Internet-Anbindung
IT-Infrastruktur-Anbieter	Kapazität zur Verarbeitung, Speicherung und Übertragung von Daten („digitale Ressourcen”)	Verfügbarkeitsgarantien, Sicherheitsgarantien, Internet-Anbindung und Datenverkehr
IT-Dienste-Anbieter	Die Verfügbarkeit eines IT-Dienstes	Leistungskapazität, Verfügbarkeit, Qualität, Erreichbarkeit und Sicherheitseigenschaften

Eine Übersicht der verschiedenen Geschäftsmodelle der Anbieter folgt. Die Übersicht hat nicht den Anspruch, alle Geschäftsmodelle und Variationen darzustellen, sondern die primären, marktüblichen Modelle hervorzuheben.

Veredelungsstufe	Geschäftsmodell	Merkmale
Rechenzentrumsgebäude	Co-Location	Ein Gebäude wird von mehreren Mietern geteilt („Bürogebäude mit verschiedenen Unternehmen pro Etage”).
	Single Tenant	Ein Gebäude wird für einen einzelnen Mieter gebaut und betrieben („Das Unternehmen mietet das gesamte Rechenzentrumsgebäude für sich allein”).
	Campus	Eine Landfläche, die mit mehreren Gebäuden (z. B. für verschiedene Single Tenants oder gemischt mit Co-Location-Gebäuden) bebaut wird, die sich einen Stromanschluss teilen.
IT-Infrastruktur	Dediziert („dedicated”, „Bare Metal”)	In diesem Modell werden Pakete von digitalen Ressourcen angeboten (Anmietung eines dedizierten, kompletten Server-Systems mit Rechen-, Speicher- und Netzwerk-Kapazität). Die Ressourcen stehen dem Käufer dediziert zur Verfügung und werden nicht mit anderen Kunden geteilt. Der Käufer kann selbst bestimmen, wie stark bzw. wann die digitalen Ressourcen genutzt werden - der Anbieter stellt die Ressourcen lediglich bereit und erhält dafür üblicherweise eine monatliche Mietgebühr.
	Geteilt oder virtualisiert („Shared”)	In diesem Modell werden Pakete von digitalen Ressourcen angeboten (Rechen-, Speicher- und Netzwerk-Kapazität), jedoch nicht dediziert. Es kann vorkommen das zwar Kapazität angemietet wurde, aber durch die Nutzung eines anderen Kunden nicht zur Verfügung steht. Die Kapazität wird von IT-Systemen bereitgestellt, die von mehreren Kunden geteilt wird. Für die Bereitstellung von Kapazität entfällt üblicherweise eine monatliche Mietgebühr. Oft sind diese „Shared” Pakete jedoch wesentlich günstiger, da der Anbieter die Kapazität oft mehrfach vermieten kann, wenn z. B. ein Kunde zwar Kapazität reserviert hat, aber nicht aktiv nutzt.
	Cloud (optional auf „hyperskaliert” Infrastruktur)	Unterscheidet sich durch ein nutzungsbasiertes Abrechnungsmodell für den Verbrauch von digitalen Ressourcen und die Möglichkeit, einzelne Klassen von digitalen Ressourcen einzeln zu nutzen (z. B. Speicherplatz, Rechenleistung und Netzwerkübertragung).

Veredelungsstufe	Geschäftsmodell	Merkmale
		Hyperskaliert („Hyperscale") bezieht sich auf den Aufbau der IT-Infrastruktur (und die Größe der dafür notwendigen Gebäude). In einer Hyperscale-Infrastruktur wird das gesamte IT-Equipment in einem Gebäude als ein einzelnes System verwaltet (anstatt jedes oder einzelne Gruppen von Server-, Netzwerk- oder Speichersystemen zu verwalten), was den Verwaltungs- bzw. Betriebsaufwand reduziert und insgesamt eine bessere Auslastung der Anlagen ermöglicht. Der Nachteil dieser Infrastruktur ist die Nachvollziehbarkeit. Es ist nur schwer möglich zu bestimmen, von welchem physischen System eine digitale Ressource erzeugt wurde, lediglich in welchem Gebäude (bzw. in welcher „Availability Zone") diese erzeugt wurde, kann ermittelt werden.
IT-Dienste	Managed	In diesem Modell werden IT-Dienste von einem Dienstleister für den Kunden betrieben. Dabei garantiert der Anbieter oft eine Verfügbarkeit und kümmert sich um den Einkauf der notwendigen digitalen Ressourcen und den Betrieb der für den Dienst notwendigen IT-Infrastruktur. Die IT-Dienste selbst bestehen entweder aus proprietären Software-Komponenten oder sind Open-Source-Software-Dienste, die für den Anbieter ohne Kosten genutzt und als IT-Dienst für Kunden angeboten werden können. Einige Anbieter von „Managed"-IT-Diensten haben eigene IT-Infrastruktur, produzieren also die notwendigen digitalen Ressourcen selbst.
	Cloud („Cloud Services", „Platform as a Service")	Wie auch bei Cloud-Infrastruktur werden cloudbasierte IT-Dienste über ein nutzungsbasiertes Abrechnungsmodell eingekauft. Die für den Dienst notwendigen digitalen Ressourcen werden üblicherweise aus der Cloud-Infrastruktur des Anbieters (dem eigenen geschlossen Marktplatz) bezogen. Dies führt zu „Bundling"-Effekten, da die IT-Dienste fest mit dem eigenen Marktplatz verbunden sind – die Dienste können nicht mit digitalen Ressourcen von anderen Anbietern betrieben werden. Wie auch bei „Managed"-Modellen basieren die IT-Dienste der Cloud-Anbieter meistens auf Open-Source-Software die kostenlos zur Verfügung steht. Jedoch werden diese oft so angepasst, dass sie nicht interoperabel sind, das heißt, das IT-Anwendungen, die auf den Diensten aufgebaut werden, nicht mit dem Dienst eines anderen Anbieters kompatibel

Veredelungsstufe	Geschäftsmodell	Merkmale
		sind (auch wenn dieser auf derselben Open-Source-Technologie aufbaut). So entsteht ein Lock-In Effekt. Zusammen mit dem Bundling der digitalen Ressourcen und dem Lock-In-Effekt entsteht ein profitables Geschäftsmodell, welches sowohl den Verkauf von Cloud-Infrastruktur fördert (wo der Großteil der Marge erwirtschaftet wird) als auch über die IT-Dienste zusätzliche Einnahmen generiert.
	Hosting	Ein Hosting Anbieter bietet üblicherweise spezielle IT-Dienste, die für den Betrieb von Internet-Präsenzen notwendig sind. Dies umfasst Onlineshops, Webseiten, aber auch Online-Speicherplatz. Hosting-Dienste werden oft als Nischen-Modell betrachtet und werden teilweise durch cloudbasierte IT-Dienste ersetzt. Jedoch bieten sie häufig eine einfache Möglichkeit für Unternehmen und Endverbraucher, eine Internet-Präsenz zu erstellen, mit hoher Verfügbarkeit und ohne operatives Betriebs-Know-how zur Verfügung zu stellen.

3.5 Volkswirtschaftliche Effekte

Eine Herausforderung bei der Analyse der volkswirtschaftlichen Bedeutung ist die begrenzte Transparenz über die tatsächlichen steuerlichen Einnahmen und die geschaffenen Arbeitsplätze, die von Anbietern digitaler Infrastrukturen generiert werden.

Viele internationale Immobilienfonds und Infrastrukturbetreiber veröffentlichen nur selektiv Informationen über Kapazitäten, Auslastung, Mitarbeiter und Finanzdaten oder machen diese über komplexe Unternehmenskonstrukte schwer zu erheben. Dies erschwert eine genaue Einschätzung. Die in diesem Bericht verwendeten Zahlen basieren daher auf

- öffentlich verfügbaren Unternehmensberichten,
- Branchenumfragen im Rahmen dieser Studie,
- Expertengesprächen.

Diese Quellen ermöglichen eine fundierte Schätzung der Wertschöpfung, können aber von den tatsächlichen Werten abweichen. In der folgenden Abbildung sind die Mitarbeiterschätzungen, jeweils pro MW-äquivalent (z. B. Betrieb von 1 MW RZ-Gebäude, 1 MW IT-Infrastruktur, 1 MW IT-Dienste) dargestellt. Eine Übersicht der Kernkennzahlen findet sich in Abbildung 25.

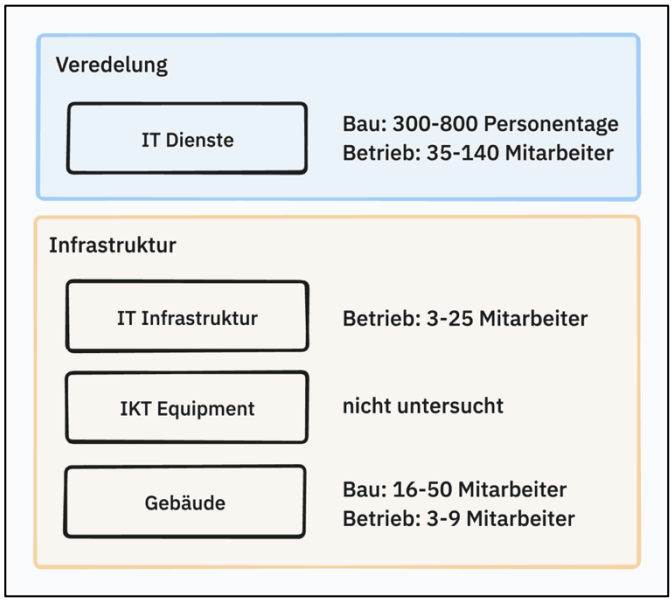


Abbildung 25: Übersicht der Kennzahlen zur volkswirtschaftlichen Wertschöpfung in Mitarbeitenden in Rechenzentren. Quelle: Eigendarstellung SDIA

Rechenzentrumsgebäude

Arbeitsplätze in der Planung und im Bau von Gebäuden

Für den Bau von Rechenzentrumsgebäuden werden primär Arbeitskräfte aus der Baubranche benötigt. Zudem werden spezialisierte Ingenieurbüros für die Planung eingesetzt. Für den Bau von kleineren Gebäuden im Bereich von 1 bis 10 MW werden 50 bis 160 Arbeitskräfte benötigt. Für größere Gebäude kann mit im Schnitt 16 Arbeitskräften pro MW gerechnet werden. So werden für den

phasenweisen Bau eines am Ende 100-MW-Rechenzentrumsgebäudes, laut einer Studie für ein Rechenzentrum von Google in Finnland, im Schnitt ca. 690 Vollzeitarbeitskräfte über einen Bauzeitraum von 10 Jahren benötigt (Oxford Economics, 2020).

Je nach Bauzeitraum werden diese Arbeitskräfte für 3 bis zu maximal 10 Jahre beschäftigt, dies variiert jedoch je nach Bauphase. Ist die Hauptstruktur bzw. das Hauptgebäude mit einer höheren Zahl an Arbeitskräften errichtet, werden weniger Arbeitskräfte für den Ausbau der Innenflächen benötigt.

Arbeitsplätze im Betrieb der Gebäude

Für den Betrieb der Gebäude wird primär Sicherheitspersonal und Wartungspersonal für die elektrischen und mechanischen Anlagen benötigt.

Auf Basis von veröffentlichten Finanzdaten von Gebäudebetreibern in Deutschland lassen sich zwei Beobachtungen hinsichtlich der Mitarbeiterzahlen treffen. Internationale Immobilienfonds, welche RZ-Gebäude in Deutschland betreiben, schaffen pro MW RZ-Gebäude im Schnitt ca. 3 Arbeitsplätze mit einem durchschnittlichen Jahresbruttogehalt von ca. 90.000 EUR. Betreiber, die ausschließlich in Deutschland ihren Hauptsitz haben und RZ-Flächen betreiben, schaffen im Schnitt ca. 9 Arbeitsplätze pro MW RZ-Gebäude mit einem durchschnittlichen Jahresbruttogehalt von ca. 75.000 EUR.

Die Unterschiede in der Zahl der Mitarbeiter ist darauf zurückzuführen, dass internationale Immobilienfonds viele betriebsrelevante Funktionen zentral von ihrem Hauptsitz aus betreiben, zum Beispiel aus den USA, Japan oder den Niederlanden.

Dabei ist wichtig zu beachten, dass die Arbeitsplatzzahl pro Megawatt nicht linear skaliert. Insbesondere bei großen und hocheffizienten Gebäuden in der Größenklasse 20 MW oder mehr kann maximal mit 0,3 bis 0,8 Arbeitsplätzen pro MW gerechnet werden. Neben der Energie- und Flächeneffizienz dieser Gebäude ist ein wichtiger Faktor der geringe Personalaufwand, der für den Betrieb der Gebäude notwendig ist. Da die Energie- und Personalkosten die Basis der operativen Betriebsausgaben bilden, sind Effizienzgewinne in beiden Bereichen für die Wettbewerbsfähigkeit des Anbieters ausschlaggebend.

Steuereinnahmen

Beim Kauf der notwendigen Flächen für den Bau von Rechenzentren wird Grunderwerbssteuer an das jeweilige Bundesland abgeführt. Zudem fällt für jedes Gebäude die Grundsteuer an, die an die jeweilige Gemeinde abgeführt werden muss. Alle Gebäudebetreiber müssen zudem Körperschaftssteuer abführen.

Bei der Erhebung der Gewerbesteuer müssen in der Betrachtung Betreiber, welche Gebäude an mehreren Standorten in verschiedenen Gemeinden betreiben, und solche, die nur an einem einzelnen Standort sowohl ihren Hauptsitz haben als auch ein Gebäude betreiben, differenziert werden. Im ersten Fall wird die Gewerbesteuer am Hauptsitz abgeführt.

Für Betreiber mit verschiedenen Standorten wird oft eine einzelne Betriebsstätte in einer Gemeinde etabliert (in manchen Fällen bewusst in Gemeinden mit niedrigerem Hebesatz). Die zum Betreiber gehörenden Gebäude in anderen Gemeinden werden lediglich als Objektgesellschaft (welche lediglich die Immobilie hält) etabliert und führen häufig keine Gewerbesteuer an die Gemeinde ab, da weder

Arbeitsplätze noch Geschäftsaktivitäten dort verbucht werden. Die Arbeitskräfte werden in der Hauptgesellschaft (Betriebsstätte) eingestellt, auch wenn die eigentliche Arbeit ggf. an einem anderen Standort ausgeführt wird. Auch kommen häufig externe Dienstleister zum Einsatz, welche das entsprechende Personal für die Gebäude an anderen Standorten als Arbeitnehmerüberlassung anbieten. Die Rechnungslegung für die Miete von Rack- und Serverstellplätzen wird häufig auch über die Hauptgesellschaft abgewickelt, auch wenn der eigentliche Stromverbrauch und die Landnutzung in der Gemeinde geschieht, in der keine Gewerbesteuer abgeführt wird. Für Gemeinden besteht die Möglichkeit, Vereinbarungen über Garantiezahlungen mit den Betreibern zu vereinbaren (Ministerium für Wirtschaft, Industrie, Klimaschutz und Energie des Landes Nordrhein-Westfalen, 2024).

Für die Vermietungsaktivitäten der Gebäudebetreiber fällt zudem ein Umsatzsteuersatz von 19 % an.

Investitionen in Infrastruktur

Betreiber von Gebäuden investieren in die digitale Infrastruktur von Deutschland. Dabei entstehen Investitionen in die Gebäude selbst als auch in die mechanische elektrische Infrastruktur. Einige größere Gebäude investieren zudem in eigene Trafo-Stationen und schließen sich direkt an die Übertragungsnetze an. Das notwendige Equipment kann unter anderem von deutschen Herstellern bezogen werden, welches sich positiv auf die Wertschöpfung auswirken kann.

Zudem bemühen sich große Gebäudebetreiber, Strom direkt von erneuerbaren Energieerzeugungsanlagen einzukaufen (PPAs, siehe auch Kapitel 2), was wiederum eine positive Wirkung auf die Entwicklung von erneuerbaren Energien in Deutschland haben kann. Es ist anzumerken, dass diese Art von Stromeinkauf vielen kleineren und mittelständischen Gebäudebetreibern nicht zur Verfügung steht, da hohe Kapitalauflagen für langfristige Stromabnahmeverträge notwendig sind, die diese Unternehmen nicht leisten können.

Wissenstransfer und Spillover-Effekte

Der Wissenstransfer und Spillover-Effekte aus den Gebäuden sind begrenzt. Denn die Gebäude bieten lediglich sogenannten „White-Space", also Fläche für IKT-Equipment, an. Diese Flächen werden erst durch IT-Infrastrukturbetreiber durch den Betrieb von IKT-Equipment nutzbar gemacht (erst dann entstehen digitale Ressourcen, die für IT-Dienste, Digitalisierung und Digitalwirtschaft nutzbar sind).

Wissenstransfer entsteht primär im Bereich der elektrischen Infrastruktur und auch im Betrieb von großen Kühlsystemen, da die RZ-Gebäude hohe Maßstäbe an redundante Kühlsysteme sowie komplexe und hochwertige elektrische und mechanische Infrastruktur setzen.

IKT-Equipment

Herstellung von Equipment

In der Herstellung von IKT-Equipment entstehen grundsätzlich viele Arbeitsplätze – dies umfasst sowohl die Herstellung von Chips und Speichern als auch den Zusammenbau der IKT-Systeme. Prominente Unternehmen in diesem Umfeld sind TSMC (Chipherstellung, ca. 80.000 Mitarbeiter) und Foxconn (Montage, 200.000 Mitarbeiter). Der Großteil der Unternehmen, die in den

Herstellungsprozess von IKT-Equipment involviert sind, befindet sich nicht in Deutschland, weshalb dieser Aspekt volkswirtschaftlich nicht weiter betrachtet wird.

Es ist anzumerken, dass das Wachstum von digitaler Infrastruktur und Rechenzentren in Deutschland abhängig von im Ausland gefertigtem IKT-Equipment ist und damit eine Abhängigkeit darstellt (welche im Rahmen der digitalen Souveränität im Kapitel 5 betrachtet wird).

Des Weiteren ist anzumerken, dass es kleine und mittelständische Unternehmen in Deutschland gibt, die die Montage von Serversystemen von IT-Systemen auch lokal durchführen, Beispiele dafür sind die Thomas Krenn AG[11] (167 Mitarbeiter, Bundesanzeiger 2021) oder die Wortmann AG (650 Mitarbeitende bzw. 2.000 Mitarbeitende mit Beteiligungen, Bundesanzeiger, 2021).

Distribution von Equipment

Neben der Herstellung von IKT-Equipment ist die Distribution bzw. die Systemintegration der Systeme ein weiterer Schritt in der Wertschöpfung. Die Unternehmen in diesem Bereich sind regional und national in Deutschland zu finden. Ein prominentes Beispiel ist die Bechtle AG (15.000 Mitarbeiter, Eigenangabe).

Viele Systemhäuser und Systemintegratoren bieten jedoch mittlerweile auch IT-Dienste und Beratungsleistungen an, weshalb eine eigenständige Betrachtung der Distribution und deren volkswirtschaftlichen Leistungen nicht weiterverfolgt wurde.

Steuereinnahmen, Spillover-Effekte und Investitionen in Infrastruktur

Die weiteren volkswirtschaftlichen Effekte von IKT-Equipment wurden im Rahmen dieser Studie nicht weiter untersucht, da der Anteil der Wertschöpfung in Deutschland nicht signifikant genug ist, um eine detaillierte Analyse zu ermöglichen.

IT-Infrastruktur und Dienste

Betrieb von IT-Infrastruktur

Im Betrieb von IT-Infrastruktur muss in zwei Kategorien unterschieden werden. Die erste Kategorie kann als virtualisierte IT-Infrastruktur beschrieben werden und die zweite Kategorie als dedizierte IT-Infrastruktur. Virtualisierte Infrastruktur benötigt für den Betrieb weniger Personal, denn das gesamte IKT-Equipment wird gebündelt und als eine geteilte Infrastruktur betrieben. Im Fall von dedizierter Infrastruktur muss für jeden Kunden bzw. jeden IKT-System-Verbund ein eigenständiger Betrieb sichergestellt werden.

Für den Betrieb von virtualisierter IT-Infrastruktur sind im Schnitt 3 Mitarbeitende pro MW IT-Infrastruktur notwendig. Im Bereich dedizierter IT-Infrastruktur ist eine hohe Varianz der Mitarbeiterzahlen von 5 bis 25 Mitarbeitenden pro MW zu beobachten. Der Aspekt des geringeren Personalaufwands beschleunigt den Trend hin zu virtualisierten IT-Infrastrukturen, zu denen unter anderem auch Cloud-Infrastrukturen gehören.

[11] http://krenn.ag/

Im Falle von virtualisierten IT-Infrastrukturen kommen jedoch noch hohe initiale Investitionskosten hinzu, die sich primär in Personalaufwand ausdrücken. Schätzungen im Rahmen dieser Studie zufolge beläuft sich dieser Aufwand auf ungefähr 90 bis 110 Personenjahre.

In beiden Kategorien ist zu beachten, dass der Betrieb der IT-Infrastruktur zum Großteil aus der Ferne, zentralisiert und global durchgeführt werden kann. Insbesondere bei Unternehmen, die ihren primären Hauptsitz im Ausland haben und dort auch meist die Betriebs- und Entwicklungsteams beschäftigen, werden keine direkten Arbeitsplätze in Deutschland geschaffen. Bei Unternehmen, die ihren primären Hauptsitz in Deutschland haben und IT-Infrastruktur betreiben, wird ein Großteil der Arbeitsplätze laut Umfragen auch in der Region des Unternehmens geschaffen. Aufgrund von Fachkräftemangel werden jedoch zunehmend Optionen für Near-Shoring und Off-Shoring von deutschen Unternehmen betrachtet.

Entwicklung und Bereitstellung von IT-Diensten

Für die Entwicklung von IT-Diensten ist im ersten Schritt ein einmaliger Entwicklungsaufwand in Höhe von circa 300 bis 800 Personentagen pro Dienst notwendig. Dies variiert je nach Komplexität des Dienstes und der Erfahrung des Herstellers in der Entwicklung von IT-Diensten.

Für den Betrieb eines Äquivalents von einem MW an IT-Diensten ist mit einem Personalaufwand von 35 bis 140 Mitarbeitern zu rechnen. Der Personalaufwand für einen einzelnen IT-Dienst lässt sich nur schwer berechnen, da die meisten Anbieter viele verschiedene Dienste mit einem geteilten Pool von Mitarbeitenden betreiben. Zudem gibt es Skalierungseffekte, da die Menge der IT-Dienste, die ein Betriebsteam betreiben kann, je nach Kompetenz und Erfahrung des Teams und je nach Komplexität der angebotenen Dienste stark variiert.

Auch bei der Entwicklung und Bereitstellung von IT-Diensten ist zu beachten, dass die damit verbundenen Arbeitsplätze meist am Hauptsitz oder in Entwicklungszentren von globalen Unternehmen geschaffen werden. Damit führen sie nur bedingt zu einem Wachstum von Arbeitsplätzen in Deutschland. Auch hier ist der Unterschied zu Unternehmen, die in Deutschland ihren Hauptsitz haben, zu beachten. Diese schaffen wiederum Arbeitsplätze für die Entwicklung und den Betrieb von IT-Diensten in Deutschland, solange die notwendigen Fachkräfte verfügbar sind.

Steuereinnahmen

Bei den IT-Infrastrukturbetreibern und Anbietern von IT-Diensten (im folgenden „Anbieter") ist, genauso wie bei den RZ-Gebäuden, zwischen regionalen bzw. nationalen Betreibern mit einem Hauptsitz in Deutschland und internationalen Unternehmen, die einen Hauptsitz im Ausland haben, zu unterscheiden.

Anbieter, die ihren Hauptsitz in Deutschland haben und ihre Dienste sowie die Infrastruktur primär auf dem deutschen Markt verkaufen, unterliegen der Umsatzsteuer, der Gewerbesteuer an ihrem Hauptsitz und der Körperschaftsteuer.

Bei Anbietern, die ihren Sitz im Ausland haben, ist die Situation komplexer. Da es sich bei digitalen Ressourcen und IT-Diensten um virtuelle Güter handelt, die global verkauft werden können, ist die Nachvollziehbarkeit, in welchem Land sie produziert wurden, oft erschwert. Einige Anbieter konsolidieren alle ihre Umsätze aus dem Verkauf von digitalen Ressourcen und IT-Diensten über eine

zentrale europäische Entität, wie zum Beispiel in Irland, wo entsprechende Umsatzsteuer und Unternehmenssteuern abgeführt werden. Dies kann dazu führen, dass ein Anbieter, der eigene Rechenzentrumsgebäude in Deutschland betreibt und diese mit IKT-Equipment ausrüstet, um damit digitale Ressourcen zu produzieren, mitunter die Umsätze aus dem Verkauf nicht in Deutschland versteuert. Damit entgeht Deutschland aus volkswirtschaftlicher Sicht ein erheblicher Teil der Wertschöpfung, der durch die IT-Infrastruktur und -Dienste entsteht.

Investitionen in Infrastruktur

In Fällen, in denen ein Anbieter Rechenzentrumsflächen lediglich anmietet, investiert dieser primär in das IKT-Equipment und Software, die für den Betrieb der Infrastruktur notwendig ist. Das IKT-Equipment ist primär ein Importgut und ist daher in der volkswirtschaftlichen Betrachtung für Deutschland weniger relevant.

Einige IT-Infrastrukturanbieter sind jedoch gleichzeitig auch Gebäudeeigentümer und somit Investoren in sowohl Gebäude als auch IT-Infrastruktur.

Wissenstransfer und Spillover-Effekte

Wie bereits im vorigen Abschnitt erklärt, können IT-Infrastruktur und -Dienste auch aus der Ferne betrieben werden. In solchen Fällen, wo der Betrieb der Infrastruktur aus dem Ausland geschieht, ist der Wissenstransfer eher gering.

Wenn der Betrieb jedoch in Deutschland durchgeführt wird, entsteht wichtiges neues Wissen im Bereich des Betriebs von großen IT-Infrastrukturen und -Diensten, welches langfristig für das Wachstum der digitalen Infrastruktur in Deutschland notwendig ist.

Die Spillover-Effekte auf andere Branchen sind hoch, da IT-Infrastruktur die digitalen Ressourcen und IT-Dienste produziert, welche für die Digitalisierung anderer Branchen und das Wachstum der Digitalwirtschaft notwendig sind.

3.6 Mechanismen der Spillover-Effekte

Digitale Infrastruktur produziert digitale Ressourcen. Diese digitalen Ressourcen lassen sich von IT-Beratungen und IT-Dienstleistern in Produkte und Dienstleistungen veredeln, welche im Rahmen der Digitalisierung von anderen Branchen eingesetzt werden können. Dies kann als ein Spillover-Effekt von digitaler Infrastruktur angesehen werden.

Besonders wichtig bei dieser Betrachtung ist der Aspekt der qualifizierten Arbeitskräfte. Auch wenn genug digitale Ressourcen bereitgestellt werden, so erfordert deren Veredelung in digitale Produkte und Dienstleistungen spezialisierte IT-Fachkräfte, die diese Veredelung durch Programmierung, Beratung, Architektur und Betrieb realisieren. Allein die Verfügbarkeit von digitalen Ressourcen sorgt nicht für eine beschleunigte Digitalisierung.

Gleiches gilt für die Digitalwirtschaft. Die Digitalwirtschaft benötigt zwingend digitale Ressourcen für den Bau und Betrieb ihrer Produkte und Dienstleistungen. Jedoch erfordert es auch hier qualifizierte Fachkräfte, die mithilfe von Software-Programmierung aus digitalen Ressourcen Produkte herstellen.

Um die möglichen Spillover-Effekte aus digitaler Infrastruktur zu maximieren, sollte diese als Ökosystem gedacht werden. Wird zum Beispiel in einer Gemeinde ein Rechenzentrumsgebäude errichtet, so können die Spillover-Effekte maximiert werden, indem auch IT-Infrastrukturanbieter sowie IT-Beratung und Dienstleister gleichermaßen angesiedelt werden. Mit der Ansiedlung der Anbieter werden qualifizierte IT-Fachkräfte auch in die Region gezogen bzw. es können Weiterbildungsangebote geschaffen werden, die wiederum mit Dienstleistungen und Produkten die Digitalisierung von regionalen Unternehmen befähigen können.

Grundsätzlich ist festzustellen, dass die Verfügbarkeit von digitalen Ressourcen allein nicht ausreichend ist. Nur in Kombination mit entsprechenden Fachkräften, die in der Lage sind, diese digitalen Ressourcen zu veredeln, können Spillover-Effekt entstehen, welche sich direkt auf die deutsche Gesamtwirtschaft auswirken.

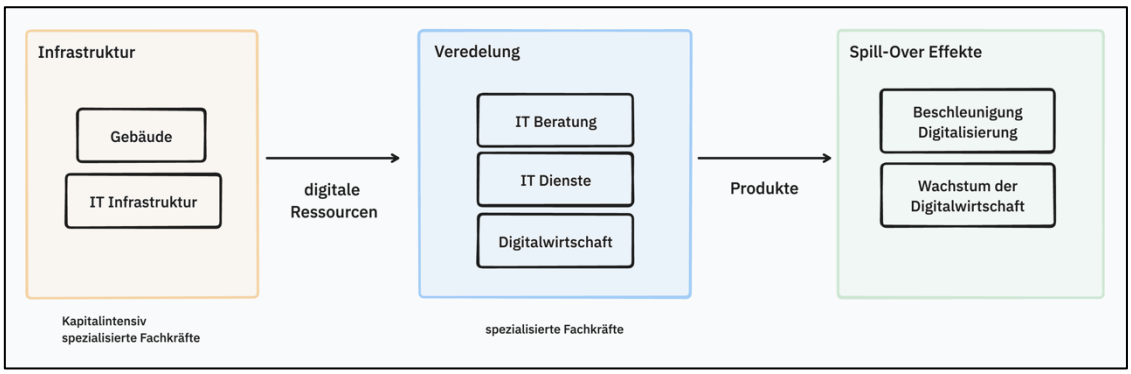


Abbildung 26: Wertschöpfungskette und Spillover-Effekte der Digitalwirtschaft: Von der Infrastruktur über die Veredelung durch digitale Ressourcen bis hin zu den resultierenden Produkten und deren Auswirkungen auf die Gesamtwirtschaft. Quelle: Eigendarstellung SDIA

Von einer möglichen Bewertung von Produktivitätseffekten durch Digitalisierung oder durch digitale Produkte wird in dieser Studie abgesehen.

Anhand von Finanzzahlen großer Digitalunternehmen lässt sich der Zusammenhang aus digitalen Ressourcen und Fachkräften darstellen. In Abbildung 27 ist der prozentuale Anteil der Kosten für digitale Ressourcen im Verhältnis zum Umsatz zu sehen. Betrachtet werden hier fünf verschiedene Anbieter von digitalen Produkten, welche digitale Ressourcen von IT-Infrastrukturanbietern beziehen (in diesem Falle Cloud-IT-Infrastrukturanbietern). Im Schnitt sind die digitalen Ressourcen für 50 % der operativen Kosten dieser Unternehmen verantwortlich. Die anderen 50 % werden zum Großteil für Fachkräfte aufgebracht.

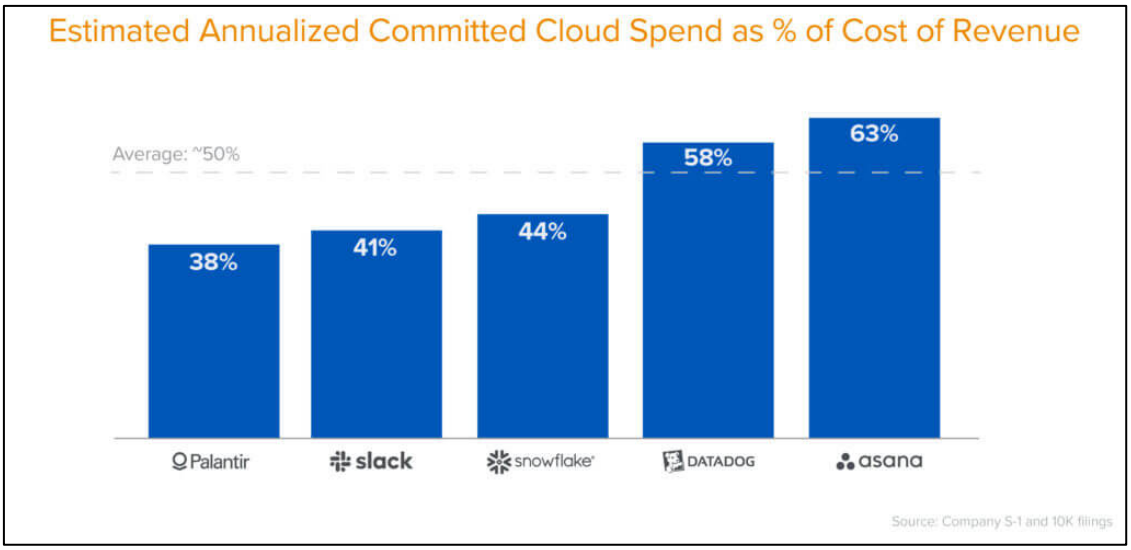


Abbildung 27: Prozentualer Anteil der Kosten für digitale Ressourcen im Verhältnis zum Umsatz großer Digitalunternehmen. Quelle: Wang, Casado (2021)

3.7 Beschränkter Marktzugang als Hemmnis

IT-Infrastrukturanbieter, die mit einem Cloud-Geschäftsmodell operieren, bieten Kunden einen geschlossenen Marktplatz für digitale Ressourcen an. In diesem Marktplatz können digitale Ressourcen entweder nach Bedarf („On Demand“), nach Preissignal („Spot“) oder im Voraus reserviert werden („Futures“). Abbildung 28 zeigt beispielhaft die Produktauswahl beim Cloud-Anbieter Amazon. Innerhalb dieser Märkte können Kunden den Einkauf ihrer digitalen Ressourcen, die sie für ihr eigenes Geschäft benötigen, durch die verfügbaren Einkaufsoptionen optimieren.

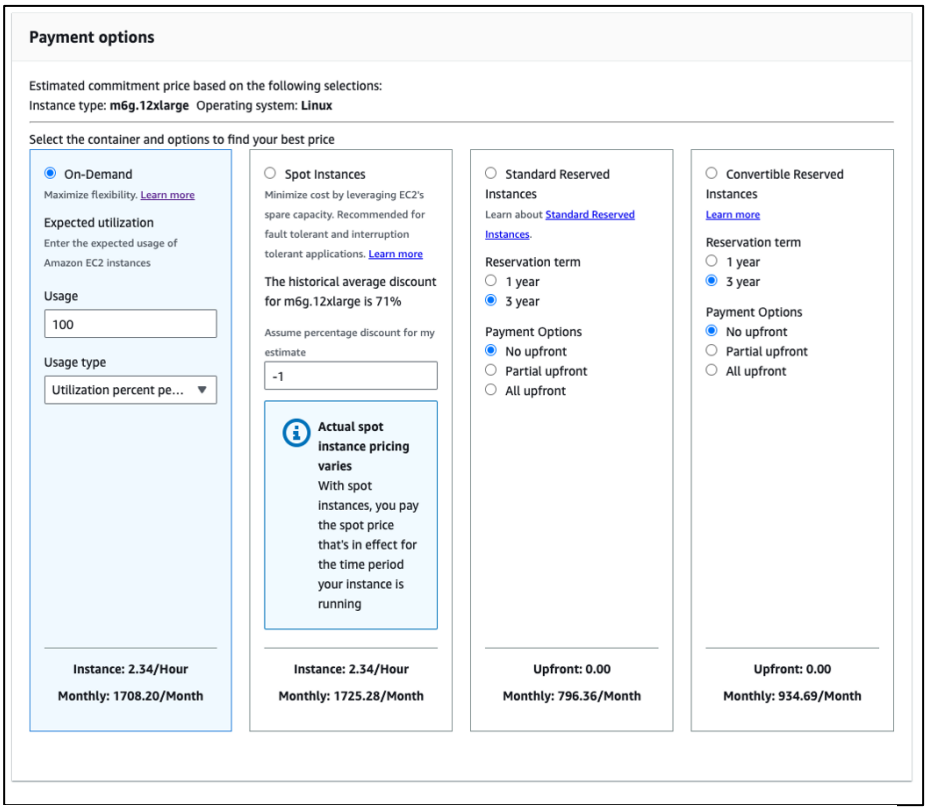
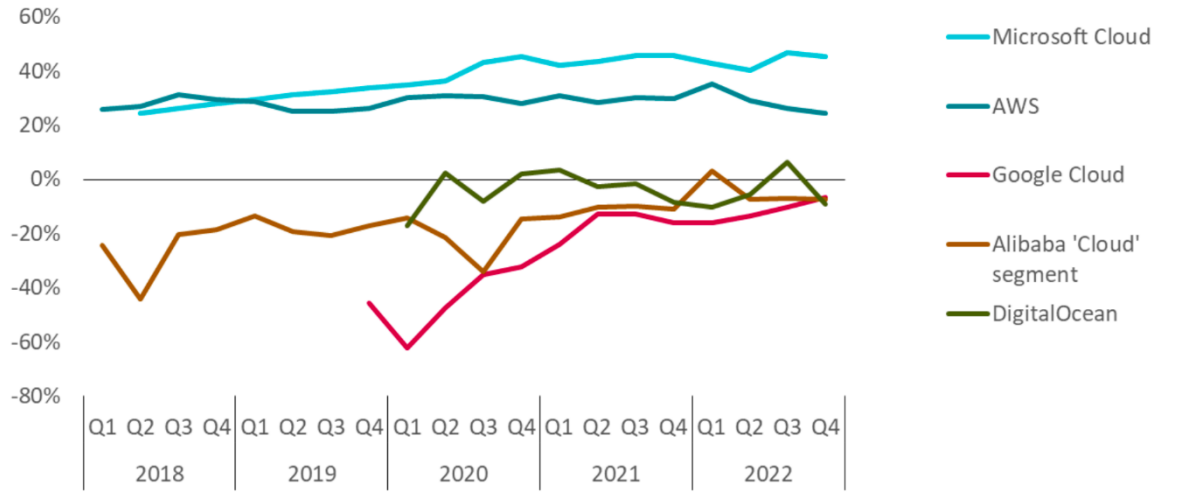


Abbildung 28: Screenshot Amazon mit Auswahloption für die Miete einer EC2-Instanz (Rechenleistung mit optionalem Datenspeicher) aus dem Kostenkalkulator von Amazon. Dargestellt sind die Zahlungsmodelle (On-Demand, Spot, Reserved). Quelle: Amazon (2024)

Die Renditen innerhalb der Cloud Marktplätze, insbesondere im Zusammenspiel mit eigener IT-Infrastruktur und Rechenzentren, sind hoch. Die britische Regulierungsbehörde Ofcom hat 2023 eine Studie, die den Cloud-Markt untersucht, veröffentlicht (Ofcom, 2023). Die Abbildung 29 stammt aus dem Bericht und zeigt die weltweite EBIT-Marge des Cloud-Geschäfts von Google, Microsoft, Amazon, Alibaba und Digitalocean. Zum Zeitpunkt der Studie befand sich Google noch im Aufbau und in der Investitionsphase für das Cloud-Geschäft, weshalb die EBIT-Marge noch sehr niedrig ist (jedoch ist bereits der stark steigende Verlauf zu erkennen).



Source: Ofcom analysis of cloud providers' financial data reported by S&P Capital IQ, information provided by Microsoft in response to our information requests[55] and Ofcom assumptions. We have used the latest filings (incorporating restatements).

Abbildung 29: Quartalsentwicklung in Prozent der führenden Cloud-Anbieter in den Jahren 2018 bis 2022. Quelle: Ofcom-Analyse

Jeder Anbieter stellt dabei seinen eigenen geschlossenen Markt bereit, in dem nur digitale Ressourcen gekauft werden können, die von ihm selbst erzeugt wurden. Um die Abhängigkeit von einem einzelnen Marktplatz zu reduzieren, haben viele große Einkäufer von digitalen Ressourcen sogenannte Multi-Cloud-Strategien implementiert. Dabei verteilen sie ihre Einkäufe über mehrere Märkte hinweg („derisking").

Einige IT-Infrastrukturanbieter bieten zudem in denselben Marktplätzen auch IT-Dienste an, die mit denselben Einkaufsoptionen (On-Demand, Spot und Futures) bereitgestellt werden. In diesen IT-Diensten sind die digitalen Ressourcen, die für den Betrieb des Dienstes notwendig sind, bereits inkludiert („bundling"). Baut ein Unternehmen seine eigenen digitalen Produkte oder seine Digitalisierungsmaßnahmen auf diesen IT-Diensten auf, ist es nur bedingt möglich, eine Multi-Cloud-Strategie zu implementieren, um das Risiko der Abhängigkeit von einem einzelnen Marktplatz zu reduzieren. Die Interoperabilität von IT-Diensten zwischen den verschiedenen Märkten ist nicht gegeben. Zudem ist es nicht möglich, die IT-Dienste einzelner Anbieter mit digitalen Ressourcen eines anderen Anbieters zu betreiben.

Viele IT-Beratungen und IT-Dienstleistungsanbieter werden von IT-Infrastrukturanbietern incentiviert Digitalisierungsmaßnahmen mit Hilfe von IT-Diensten aus ihrem Markt umzusetzen. Dies wird als schnellere Alternative zum Aufbau eigener IT-Dienste oder zur Nutzung von Open-Source-Diensten angeboten. Die Nutzung der IT-Dienste aus einem einzelnen Marktplatz führt jedoch mitunter zu einer hohen Abhängigkeit von einem einzelnen Anbieter. Insbesondere, wenn ein Unternehmen die Kosten für die digitalen Ressourcen, welche für die Dienste notwendig sind, reduzieren möchte, z. B. indem diese von einem anderen Marktplatz eingekauft werden, ist dies nicht mehr möglich. Daher sind die IT-Dienste, die von Marktplatzanbietern angeboten werden, als Beschleuniger für das Wachstum der Marktplätze zu verstehen, deren Umsatz primär aus dem Verkauf der dafür notwendigen digitalen Ressourcen entsteht.

3.8 KI als Treiber der Nachfrage

IT-Dienste und digitale Produkte, die mit KI-Funktionen angereichert werden, haben einen erhöhten Verbrauch von digitalen Ressourcen. Dabei können diese KI-Funktionen für eine verbesserte Funktion der Produkte oder eine bessere Nutzbarkeit der Dienstleistungen führen. Genaue Daten dazu sind zum Zeitpunkt der Studie nicht vorhanden.

Schätzungen von Experten zufolge steigert das Hinzufügen einer KI-Funktionalität zu einem bestehenden digitalen Produkt, wie z. B. einer Google-Suche, den digitalen Ressourcenaufwand um das Vierfache. Durch diesen steigenden Aufwand entsteht auch ein erhöhter Bedarf an IKT-Equipment, Gebäudeflächen und Strom. Auch mit Effizienzgewinnen im IKT-Equipment führt das Hinzufügen von KI-Funktionalitäten zu einem massiven Anstieg des benötigten Equipments (siehe auch Zukunftsszenarien in Kapitel 5).

Hinzu kommt, dass die bestehende Gebäudeinfrastruktur möglicherweise für die hohe Dichte von IKT-Equipment, die für KI-Anwendungen notwendig ist, umgebaut werden muss, und die Kühlkapazitäten erhöht werden müssen.

Zudem sorgen Engpässe in der Herstellung von KI-Equipment, wie zum Beispiel spezialisierten Chips, zu erheblichen Kostensteigerungen für den Betrieb von KI-spezifischer IT-Infrastruktur, die für die Herstellung von digitalen Ressourcen für KI-Anwendungen notwendig ist.

3.9 Fachkräftemangel als Hemmnis

Durch die im Rahmen der Studie durchgeführten Umfragen wurde ein Mangel an spezialisierten Fachkräften im Bereich IT-Infrastruktur deutlich.

Diese Fachkräfte unterscheiden sich von klassischen IT-Fachkräften dadurch, dass sie ihren Schwerpunkt nicht auf der Entwicklung von Softwareanwendungen haben, sondern das Know-how für den Betrieb von IT-Infrastrukturen, insbesondere großen Server- und Speicheranlagen, besitzen. Dieses Know-how ist insbesondere für den Betrieb von skalierbaren Cloud- und IT-Infrastrukturen notwendig.

Neben einer spezifischen Ausbildung brauchen diese Fachkräfte praktische Erfahrung im Betrieb von hochskalierten IT-Infrastrukturen, von denen es in Deutschland bis heute nur wenige gibt. Hier haben Länder wie die USA einen Wettbewerbsvorteil, denn dort haben viele Digitalunternehmen ihren Hauptsitz und bauen seit Jahrzehnten große IT-Infrastrukturen für die eigenen Produkte auf. Dieses Wissen und die entsprechenden Fachkräfte setzen diese Unternehmen nun ein, um eigene IT-Infrastrukturgeschäfte aufzubauen, wie sie z. B. im Cloud-Bereich bekannt sind.

Ein weiteres Hemmnis für die Produktion von digitalen Ressourcen in Deutschland ist die Strom- und Flächenverfügbarkeit für die Gebäude, die benötigt werden, um das IKT-Equipment zu betreiben. Dieser Aspekt wird in Kapitel 2 und Kapitel 5 weiter betrachtet. Hier wird auch die Frage betrachtet, inwieweit Deutschland seine Nachfrage an digitalen Ressourcen durch Importe aus dem europäischen Ausland langfristig decken könnte, ohne dabei eigene Flächen und Strommengen bereitstellen zu müssen.

3.10 Handlungsempfehlungen

Eine strategische Entwicklung von digitaler Infrastruktur in Deutschland ist erforderlich.

Um die volkswirtschaftliche Wertschöpfung von digitaler Infrastruktur zu maximieren, sollte ein Fokus auf digitale Ökosysteme gelegt werden. Dabei sollten Rechenzentrumsgebäude, IT-Infrastrukturanbieter und IT-Dienstleister, Beratungen und Unternehmen der Digitalwirtschaft als Ökosystem betrachtet werden, welche sich zusammen in einer Region ansiedeln.

Um diese regionale Wirtschaftsentwicklung zu befähigen, sollte eine nationale Rechenzentrumsstrategie entwickelt werden, welche insbesondere Empfehlungen formuliert, wo solche Ökosysteme angesiedelt werden können (unter Einbezug von Strom- und Flächenverfügbarkeit), und eine Analyse beinhaltet, welche regionalen Mehrwerte daraus geschöpft werden können. Zu diesen Mehrwerten sollte z. B. eine Beschleunigung der Digitalisierung in bestimmten Wirtschaftsbereichen, die in der Region besonders stark vertreten sind, gehören.

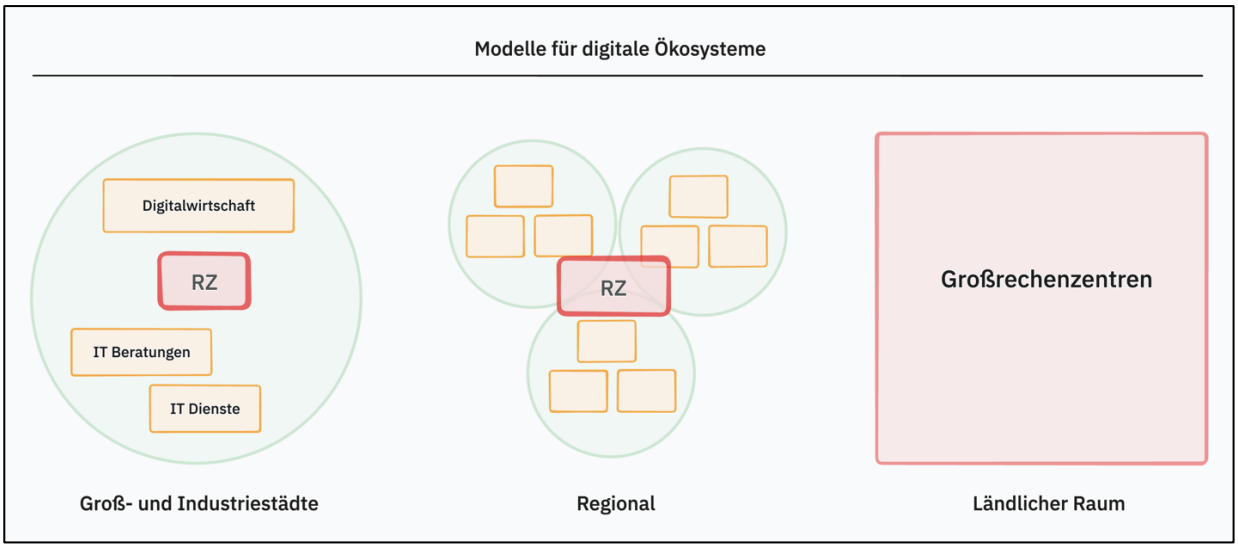


Abbildung 30: Modelle für digitale Ökosysteme: Vergleichende Darstellung der Entwicklungskonzepte für Groß- und Industriestädte, regionale Zentren und ländlichen Raum mit unterschiedlicher Ausprägung der IT-Infrastruktur und Vernetzung. Quelle: Eigendarstellung SDIA

Teil dieser Strategie sollten mindestens drei verschiedene Entwicklungskonzepte sein: für Groß- und Industriestädte (1), mittelgroße Gemeinden (2) und den ländlichen Raum (3), welche in Abbildung 30 beispielhaft dargestellt sind.

1. Das Entwicklungskonzept für Großstädte und Industriestädte sollte sich auf kleinere Rechenzentren und IT-Infrastruktur fokussieren, aus denen die Abwärme in der Stadt genutzt und diese nahtlos in das Energiesystem integriert werden kann. Zudem sollte diese IT-Infrastruktur im Ökosystem mit Digitalwirtschaft und IT-Dienstleistungen entwickelt werden, mit einem Schwerpunkt auf digitalen Produkten und Dienstleistungen, die eine niedrige Latenz (nähe zum Rechenzentrum) benötigen.

2. In mittelgroßen Gemeinden sollte ein überregionales Rechenzentrum geplant und gefördert werden und zusammen mit einem Ökosystem aus IT-Beratungen, -Dienstleistern und einer regionalen Digitalwirtschaft aufgebaut werden.

3. Im ländlichen Raum, in dem eine hohe Flächenverfügbarkeit und EE-Stromverfügbarkeit herrscht, sollten Großrechenzentren zusammen mit IT-Infrastruktur-Unternehmen (Arbeitsplätze und Steuereinnahmen für die ländliche Region) entwickelt werden, die digitale Ressourcen mit weniger Latenzsensitivität für andere Gemeinden bereitstellen (z. B. für Datenanalysen, KI-Training, oder als Backup/Redundanzsystem).

Durch diese Modelle können für die jeweilige Region proportional signifikante Steuereinnahmen und Arbeitsplätze geschaffen werden.

Zugang zu Kapital für den Ausbau regionaler digitaler Infrastruktur für kleine und mittelständische IT-Infrastruktur-Betriebe erleichtern.

Der Bau von Rechenzentrumsgebäuden lässt sich zwar über Kredite finanzieren, viele Hausbanken und Kreditinstitute sind jedoch nicht entsprechend aufgestellt, um entsprechende Investitionsvorhaben einzuschätzen. Damit besteht die Gefahr, dass sie regionalen Projekten eine hohe Risikoklasse zuweisen, was zu hohen Bürgschafts- und Eigenkapitalanforderungen führt, welche kleine und mittelständische regionale IT-Infrastruktur-Anbieter oft nicht bewältigen können. Dies hindert eine regionale Infrastruktur- und Marktentwicklung. Zudem haben kleine und mittelständische Unternehmen nicht gleichermaßen Zugang zum Kapitalmarkt wie internationale Wettbewerber, sind nicht vertikal integriert und damit nicht in der Lage, mit anderen Geschäftsbereichen den Ausbau der Infrastruktur initial zu subventionieren (wie z. B. große Cloud-Anbieter es mit Werbegeschäften, Onlineshopping oder einem etablierten Software-Lizenz-Geschäft tun). Es wird empfohlen Bürgschaften oder andere Maßnahmen zu ergreifen, um KMUs im Bereich digitaler Infrastruktur auf regionaler Ebene zu stärken.

Ein weiterer Aspekt ist der Kauf bzw. das Leasing von IKT-Equipment, was bei einem 1 MW Standard-Rechenzentrum bereits im Bereich von 30 bis 60 Millionen Euro liegen kann. Auch hier ist der Kapitalmarktzugang ein Hindernis. KMUs verfügen nicht über ausreichend Kapital, um notwendige Investitionen tätigen zu können, die die Marktteilnahme am IT-Infrastruktur- oder Cloud-Markt ermöglichen.

Gezielte Investitionen in die Ausbildung von Fachkräften mit Spezialisierung auf den Betrieb von großen und skalierbaren digitalen Infrastrukturen sollten gefördert werden.

Um die IT-Infrastrukturbranche in Deutschland (insbesondere im Kontext der digitalen Souveränität, siehe Kapitel 5) weiter zu stärken, sollte gezielt in die Ausbildung von spezialisierten Fachkräften investiert werden.

Hierfür können verschiedene Ansätze verfolgt werden:

- Die Entwicklung eigener hochskalierter, öffentlicher IT-Infrastruktur für Know-how-Aufbau und eine öffentliche Trainingsumgebung, die von Bildungseinrichtungen mitgenutzt werden kann.

- Gezielt spezialisierte Ausbildungsangebote an Universitäten und technischen Hochschulen schaffen für IT-Infrastruktur-Expertise.
- Internationale Zusammenarbeit: Austauschprogramme mit anderen Ländern schaffen, die bereits hochskalierte IT-Infrastruktur betreiben und bei denen ein Wissenstransfer möglich ist.
- Weiterbildungsprogramme für bestehende IT-Fachkräfte, um sich im Bereich hochskalierte Infrastruktur zu schulen.

Regionalität als Kriterium in den Einkauf von IT-Diensten und -Infrastruktur auf kommunaler und Bundesebene mit einbeziehen.

Eine weitere Option, die regionale und nationale Entwicklung von digitaler Infrastruktur zu fördern, ist es, regionale Ökosysteme und Anbieter gezielt in Einkaufsbedingungen einzubeziehen und Wettbewerb rund um die lokale Wertschöpfung zu schaffen. Dazu können auch Beiträge zum Energiesystem positiv gewertet werden (siehe Kapitel 2, Abwärme und Flexibilität), ebenso wie Beiträge zu lokalen Arbeitsplätzen, zur Ansiedlung von Niederlassungen, Investitionen in die lokale Bildung von Fachkräften o. ä.

Aus Sicht der Nachhaltigkeit und Resilienz von Infrastruktur ist der regionale Einkauf zu fördern, um eine diversifizierte Marktentwicklung in Deutschland zu unterstützen.

Wettbewerb im IT-Infrastruktur-Markt sicherstellen und geschlossene Märkte für alle Marktteilnehmer öffnen

Die geschlossenen Cloud-Marktplätze von internationalen Unternehmen sollten auf dem Europäischen Markt über regulatorische Maßnahmen geöffnet werden, um einen fairen Wettbewerb zu ermöglichen. Eine Öffnung der Märkte würde deutschen und europäischen Produzenten von digitalen Ressourcen die Möglichkeit geben, ihre Ressourcen auch über die Cloud-Märkte der großen Cloud-Anbieter zu vertreiben. Auch könnten dann IT- und Cloud-Dienste, welche in den geschlossenen Märkten angeboten werden, auf digitalen Ressourcen von Drittanbietern aufbauen und damit “Lock-In Effekten” entgegenwirken. Verbraucher von digitalen Ressourcen, insbesondere europäische Unternehmen, hätten eine bessere Entscheidungsfreiheit und könnten über die Cloud-Marktplätze regional produzierte digitale Ressourcen einkaufen und mit den IT-Diensten der Marktplatzbetreiber frei kombinieren. So kann ein Teil der Wertschöpfung der Marktplätze in die europäische Volkswirtschaft geleitet werden.

Etablierung eines europäischen Marktplatzes, um den Handel von digitalen Ressourcen innerhalb von Europa zu fördern.

Ein weiteres Modell, um den IT-Infrastrukturmarkt in Deutschland zu fördern und gleichermaßen Resilienz und Nachhaltigkeit in der Beschaffung von digitalen Ressourcen sicherzustellen, ist die Etablierung eines europäischen Marktplatzes für digitale Ressourcen.

In diesem Marktplatz können in erster Linie Über- und Unterkapazitäten ausgeglichen werden, um damit den Nutzungsgrad der bestehenden Rechenzentren zu verbessern. Gleichzeitig kann in einem offenen Markt ein Wettbewerb geführt werden, wie er in den bisher üblichen geschlossenen Märkten nur bedingt möglich ist. So können Wettbewerber sich über Preis, Qualität und Nachhaltigkeitsaspekte

differenzieren. Hat eine IT-Anwendung z. B. eine Latenzanforderung, so können regionale Anbieter im Wettbewerb besser dastehen. Hat eine Anwendung hohe Nachhaltigkeitsanforderung, so können sich Wettbewerber in Regionen mit hohem erneuerbaren Energieanteil oder IT-Infrastrukturanbieter mit wiederaufbereitetem IKT-Equipment differenzieren.

Für die öffentliche Hand, sowohl auf Bundes- als auch auf EU-Ebene, wird zudem der Einkauf von digitalen Ressourcen erheblich vereinfacht. Diese Ressourcen können über den Marktplatz nach entsprechenden Kriterien eingekauft werden (z. B. souverän, innerhalb der Landesgrenzen, nachhaltig, kosteneffizient etc.). Die Ausgaben für IT-Infrastruktur können über viele kleine, mittelständische und Großunternehmen in Europa verteilt werden, die allein zu klein, aber gemeinsam ausreichend IT-Infrastruktur bereitstellen können, um den Ressourcenbedarf von Regierungen, Industrie und anderen Wirtschaftssektoren decken zu können. Dies würde auch das Argument der Skalierbarkeit, welches von Anbietern großer geschlossener Marktplätze oft angeführt wird, mit einem dezentralen, europäischen Ansatz widerlegen.

Auslastung & Nutzungsgrad von IT-Infrastruktur erhöhen

Ein weiterer Aspekt, der im Rahmen der Studie durch Interviews und Umfragen deutlich wurde, ist, dass die IT-Infrastruktur, insbesondere das IKT-Equipment, nicht optimal genutzt wird. Im Falle eines der größten Anbieter für IT-Dienste und -Infrastruktur in Deutschland lag der effektive, jährliche Nutzungsgrad der Infrastruktur gerade einmal bei 15 %.[12]

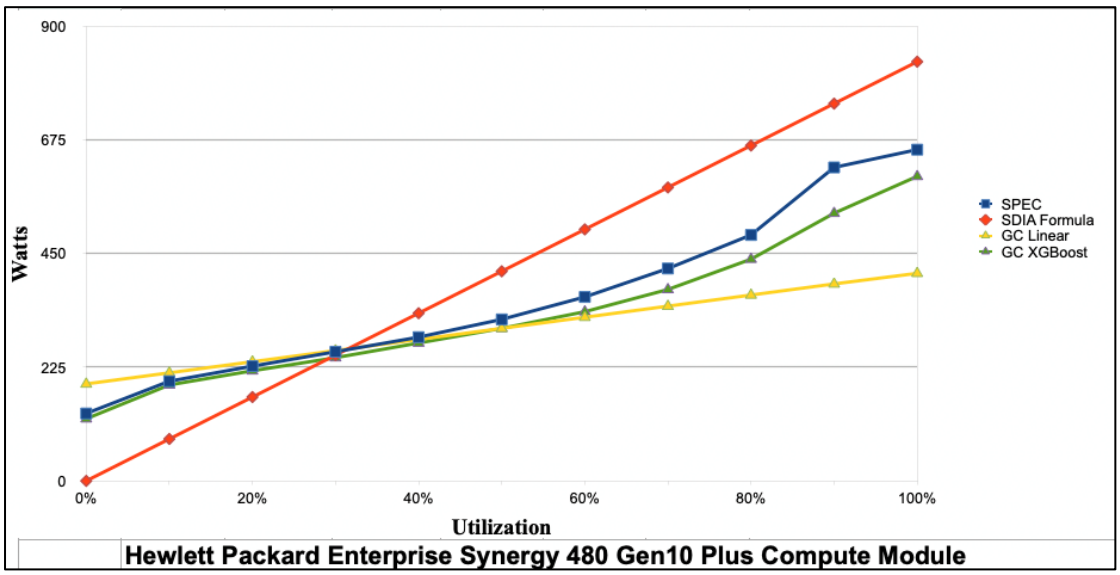


Abbildung 31: Server-Stromverbrauch in Watt in Abhängigkeit vom Nutzungsgrad in Prozent. Die Grafik verdeutlicht den Zusammenhang zwischen Serverauslastung und Energieverbrauch eines HPE Synergy 480 Gen10 Plus Servers, basierend auf Daten der SPEC Organisation. Quelle: Green Coding Solution, Cloud Energy Modell (2022)

[12] https://github.com/green-coding-solutions/cloud-energy

Einige Serverhersteller veröffentlichen den Stromverbrauch der Geräte bei verschiedenen Nutzungsgraden in einer Datenbank der SPEC-Organisation (SPEC, 2024). Mit diesen Daten lässt sich abschätzen, dass ein System zwischen 0 und 40 % Nutzungsgrad, einen nahezu gleichbleibenden Stromverbrauch hat (im Falle des dargestellten HPE Synergy 480 Gen10 Plus, ca. 225 W). Wird ein System nur 10 % ausgelastet, entsteht ein Verlust von 168W, oder 75 %. Jedes Server-System hat dabei, wie auf der Abbildung 31 ersichtlich, einen „Optimalen Nutzungsbereich“. Bei über 80 % wird das System in den Hochgeschwindigkeitsmodus versetzt, was zu einem signifikanten Anstieg des Stromverbrauchs führt. Im Falle des gezeigten Systems sollte der Nutzungsgrad bei 40 bis 65 % liegen, um den Strom mit maximaler Effizienz zu nutzen.

Der oft niedrige Nutzungsgrad hat verschiedene Ursachen, unter anderem die hohen Redundanzanforderungen, die für viele IT-Anwendungen gesetzt werden (in denen 1-2 redundante Systeme parallel laufen müssen, im Falle des Ausfalls des Primärsystems). Weitere Ursachen sind die Architektur der Infrastruktur und fehlende Anreize, um die Infrastruktur zu optimieren. Der ineffiziente Nutzungsgrad wird oft an die Kunden in Form von höheren Kosten weitergegeben.

Regulatorisch gibt es in diesem Bereich de facto keine Auflagen. Dies haben auch unsere Experten-Gespräche bestätigt. Es bieten sich einige Handlungsoptionen:

- Standards für die einheitliche Messung des Nutzungsgrads in IT-Infrastrukturen zu etablieren
- Verpflichtung für IT-Infrastrukturanbieter, den durchschnittlichen Nutzungsgrad in der eigenen Infrastruktur zu veröffentlichen oder in das bestehende RZReg zu melden (in dem bereits IT-Infrastruktur neben Rechenzentrumsgebäuden erfasst werden).
- Richtwerte für den Nutzungsgrad im Rahmen des EnEfG, zusätzlich zu den Auflagen für Rechenzentrumsgebäuden, zu etablieren.
- Regulatorische Rahmenbedingungen für IT-Anwendungen dahingehend anpassen, dass diese nicht zu unnötigen Redundanzanforderungen führen, bzw. alternative Optionen für die Hochverfügbarkeit ermöglichen (z. B. IT-Grundschutz).
- In Forschung zu investieren, um effektivere Virtualisierungs-Technologien und Technologien zur Erhöhung des Nutzungsgrads in Deutschland zu entwickeln.

4 Rechenzentren und Regulatorik

4.1 Zusammenfassung

Im Kapitel 4 werden die wesentlichen regulatorischen Rahmenbedingungen für den Standort Deutschland und entsprechende Anforderungen an die Planung und den Betrieb von Rechenzentren untersucht. Diese Untersuchung zeigt auf, dass für Rechenzentren mit einem Standort in Deutschland vielfältige Regelungen aus unterschiedlichen Rechtsbereichen gelten und dass der Themenkomplex der Datensicherheit und der Datensouveränität ein wesentlicher regulatorischer Treiber für den Rechenzentrumsstandort Deutschland ist. Dies führt dazu, dass Rechenzentren, die ihren Standort in Deutschland haben und deutschem Recht unterliegen, am Markt nachgefragt sind.

Regulatorische Anforderungen an Rechenzentren bestehen zunächst in dem Erfordernis einer Baugenehmigung (betreffend das Rechenzentrumsgebäude) sowie unter bestimmten Voraussetzungen in dem Erfordernis einer Genehmigung auf Grundlage des Bundes-Immissionsschutzgesetz (BImSchG). Aufgrund der jeweiligen gesetzlichen Konzentrationswirkung dieser Genehmigungen werden von diesen die Zustimmungen und das Einvernehmen anderer Fachbehörden einbezogen hinsichtlich der Einhaltung weiterer öffentlich-rechtlicher Vorschriften. Dies betrifft etwa eine Kühlwasserversorgung des Rechenzentrums und wasserrechtliche Anforderungen für entsprechende Erlaubnisse und Befreiungen, eine wasserrechtliche Eignungsfeststellung mit Blick auf die Errichtung von Kraftstoff-Lagertanks für den Betrieb der USV-Anlage oder die Einhaltung der Anforderungen der AwSV (Verordnung über Anlagen zum Umgang mit wassergefährdenden Stoffen), die Gewährleistung abfallrechtlicher Anforderungen, arbeitsschutzrechtlicher Anforderungen betreffend den umweltbezogenen Gesundheitsschutz, ggf. zudem noch natur- und artenschutzrechtliche Befreiungen, Waldumwandlungsgenehmigungen, straßenrechtliche Ausnahmen und Zustimmungen, ggf. auch Sondernutzungserlaubnisse, ggf. luftverkehrsrechtliche Zustimmungen. Eine Erkenntnis aus den im Rahmen der Studie durchgeführten Stakeholder-Workshops ist, dass die Genehmigungsprozesse mit Blick auf erforderliche Baugenehmigungen sowie immissionsschutzrechtliche Genehmigungen aus Sicht der Projektentwickler und Betreiber zu viel Zeit in Anspruch nehmen und Aufwände im Rahmen der Antragstellung und der einzureichenden Nachweise zu hoch sind. Hinzu kommen im Rahmen der (baurechtlichen und immissionsschutzrechtlichen) Genehmigungsprozesse Abstimmungsprobleme zwischen Genehmigungsbehörden und weiteren Fachbehörden und damit einhergehende Verzögerungen.

Die Analyse des Rechtsrahmens hat weiterhin gezeigt, dass Rechenzentrumsbetreiber Adressaten verschiedener Regelungen sind, die im Kern eine Steigerung der Energieeffizienz zum Ziel haben. Dabei wurde ein Fokus auf die Vorgaben des EnEfG gelegt. In diesem Zusammenhang konnten insbesondere in Bezug auf die Abwärmenutzung bzw. die Nutzbarmachung der Abwärme gesetzgeberische Regelungspotenziale in bestimmten Bereichen identifiziert werden:

- Regelungspotenziale mit Blick auf Angebot und Nachfrage von Abwärmemengen und -temperaturen
- keine wechselseitigen (Informations-) Pflichten zwischen Rechenzentrumsbetreibern und Verantwortlichen der KWP

- Unklarheiten bei der Kostenverantwortung hinsichtlich Abwärmenutzungsangeboten zu Gestehungskosten sowie der Möglichkeit und Implikationen unentgeltlicher Wertabgaben.

Die Abhängigkeit von der Anbindung an notwendige Infrastrukturen wie Stromnetze oder Wärmenetze schränkt die Standortwahl innerhalb Deutschlands ein. Als mögliches Hemmnis für den Rechenzentrumsstandort sind dabei eine fehlende integrierende und übergreifende Planung für die Ausweisung entsprechend geeigneter Flächen für die Netzinfrastruktur (insbesondere für Strom und Wärme) und Rechenzentren sowie fehlende strukturierte Kommunikationsformate zwischen Rechenzentrumsbetreibern und Verantwortlichen der KWP einzuordnen (siehe auch Kapitel 2).

Im Zusammenhang mit der Abwärmenutzung von Rechenzentren und einer entsprechenden Anbindung an ein Wärmenetz ist zwar die verpflichtende KWP förderlich. Jedoch stellt eine im Ergebnis fehlende Transparenz für Rechenzentrumsbetreiber aufgrund von Regelungen, die nur eingeschränkt Zugriff auf entsprechende Datenbestände für bestehende und geplante Wärmenetze ermöglichen, eine Herausforderung dar. Insbesondere fehlen Regelungen für eine bundesweite Übersichtskarte bzw. ein bundeseinheitliches Wärmekataster für die Identifizierung geeigneter Flächen für Rechenzentren. Auch könnten kleinere Wärmenetze bei Aufnahme sehr großer Abwärmemengen aus Rechenzentren an Kapazitätsgrenzen stoßen.

Im internationalen Vergleich werden Best Practices identifiziert und hierfür Ableitungen aus den jeweils geltenden regulatorischen Rahmenbedingungen getroffen. Als förderlich im Standortwettbewerb von Rechenzentren werden in diesem Rahmen insbesondere beschleunigte Genehmigungsverfahren für Rechenzentren eingeordnet. Weiterhin ergibt die Analyse, dass die PUE als Energieeffizienzparameter, mit deren Wert weitreichende regulatorische Anforderungen beim Betrieb von Rechenzentren einhergehen, ihre eigenen Grenzen hat.

4.2 Dimensionen politisch-rechtlicher Rahmenbedingungen

Die Analyse der regulatorischen Rahmenbedingungen, die im Rahmen dieser Studie durchgeführt wurde, hat eine Vielzahl von Rechtsbereichen identifiziert, die Einfluss auf die Entscheidung zur Ansiedlung von Rechenzentren in Deutschland haben.

Der wesentliche regulatorische Treiber für den Rechenzentrumsstandort Deutschland ist der Themenkomplex der Datensicherheit und der Datensouveränität. Beide führen dazu, dass die Rechenzentren, die ihren Standort in Deutschland haben und deutschem Recht unterliegen, am deutschen Markt nachgefragt sind.

Die maßgeblichen regulatorischen Faktoren, die die Standortwahl (ggf. negativ) beeinflussen, sind unter Berücksichtigung der Erkenntnisse aus den durchgeführten Stakeholder-Workshops neben dem Themenkomplex der Datensicherheit und Datensouveränität öffentlich-rechtliche Genehmigungserfordernisse sowie entsprechende Prozesse und Abläufe. In diesem Zusammenhang steht auch die kommunale Bauleitplanung mit Blick auf die Herstellung von Baurecht für geeignete Standorte für Rechenzentren. Weiterhin ist die notwendige Anbindung an Infrastrukturen ein zentraler Standortfaktor, wobei entsprechende gesetzliche Rahmenbedingungen mangels kombinierender und integraler Ansätze für Anschlussansprüche potenziell zu Investitions- und Planungsunsicherheiten führen können.

4.2.1 Datensicherheit und Datensouveränität

Die DSGVO der Europäischen Union als umfassendes Regelwerk gewährleistet den Schutz personenbezogener Daten innerhalb der EU und den Transfer solcher Daten in Drittländer. Die DSGVO beinhaltet zwar keine speziell für Rechenzentren anwendbare Vorschriften. Gleichwohl hat die DSGVO erhebliche Auswirkungen auf deren Betrieb, insbesondere mit Blick auf Datenspeicherung und -verarbeitung und das hierbei zu gewährleistende Schutzniveau, das konkrete Anforderungen an die Verfügbarkeit, Verlässlichkeit und an die Sicherheit in Rechenzentren umfasst. Weiterhin werden bestehende datenschutzrechtliche Vorschriften aus dem Telekommunikationsgesetz (TKG) und dem Telemediengesetz (TMG) im Telekommunikation-Telemedien-Datenschutz-Gesetz (TTDSG) zusammengefasst.

Weitere IT-sicherheitsrelevante Anforderungen ergeben sich etwa aus dem IT-Sicherheitsgesetz 1.0 bzw. 2.0 und dieses ergänzend aus der KRITIS-Verordnung. Hierfür gibt es Kriterien über die Auslastung und die Verfügbarkeit aller Ressourcen im Rechenzentrum. Entscheidend ist dabei, dass nicht alle Systeme unter Volllast laufen, sondern dass sie ggf. gemeinsam den Ausfall einer Einheit kompensieren können, und dass neben dem Datennetz auch das Stromnetz umfassend gesichert sein muss. Ebenfalls betrifft dies den Brandschutz, den Einbruchschutz, die Cybersicherheit oder den Schutz gegen Wasserschäden. Nach der KRITIS-Verordnung werden u. a. Anforderungen an Informationssicherheits-Managementsysteme für Betreiber kritischer Infrastrukturen aufgestellt. Dabei sind Rechenzentren und Anbieter von IT-Infrastruktur bei einer vertraglich vereinbarten IT-Leistung ab 3,5 MW nach § 5 KRITIS-Verordnung als kritische Infrastruktur einzuordnen. KRITIS-Betreiber müssen die Umsetzung von Maßnahmen der Cybersicherheit und des Datenschutzes je nach Schutzbedarf regelmäßig durch Prüfer belegen lassen. Die Prüfungen orientieren sich dabei am geltenden Stand der Technik. Je nach Schutzbedarf gibt es hierfür unterschiedliche Anforderungen nicht nur an die IT-Infrastruktur, sondern auch an das Rechenzentrumsgebäude und dessen Umgebung, unter anderem an Aspekte der Baukonstruktion, des Brandschutzes oder der Sicherheitssysteme. Hierzu stellt etwa das BSI auf branchenspezifische Sicherheitsstandards ab und hat hierzu Standortkriterien (z. B. bestimmte Abstandsgebote) entwickelt. Diese sollten im Rahmen der Risiko- und Schutzbedarfsanalyse für die Planung neuer Rechenzentren herangezogen werden. Sie gelten ggf. auch im Rahmen bestimmter branchenspezifischer Zertifizierungen. Diese spezifischen Standortkriterien und Anforderungen an das Rechenzentrumsumfeld beeinflussen die Eignung bestimmter Flächen für die Ansiedlung eines Rechenzentrums.

Weiterhin wird in Umsetzung der CER-Richtlinie das KRITIS-Dachgesetz Betreiber kritischer Infrastrukturen verpflichten, diese in besonderer Weise zu schützen, um Störungen der Verfügbarkeit, Integrität, Authentizität und Vertraulichkeit zu vermeiden. Weiterhin wird das noch nicht in Kraft getretene NIS2-Umsetzungsgesetz zur Umsetzung der EU-Richtlinie NIS-2 weitere umfassende Maßnahmen zur Stärkung der Cybersicherheit umfassen. Durch den am 09.12.2024 in Kraft getretenen Cyber Resilience Act der EU (Verordnung) werden zudem weitere sicherheitsrelevante Anforderungen an digitale Produkte gestellt. So soll etwa gewährleistet werden, jederzeit Security-Updates aufzuspielen.

Die spezielle Regulatorik für KI ist eigenen Aussagen der Betreiber von Rechenzentren aus den Stakeholder-Workshops zufolge für eine Standortwahl in Deutschland unmittelbar nicht relevant, weil die Betreiber nicht Adressaten der Regulierung sind und insoweit nur die Infrastruktur nach den Bedürfnissen ihrer Kunden bereithalten. Mittelbar mag diese spezielle Regulierung die Nachfrage

nach Rechenleistung in Deutschland beeinflussen sowie den für die Standortwahl relevanten Energiebedarf, aber direkten Einfluss auf den Betrieb von Rechenzentren und auf die Standortwahl hat sie nicht. Hier ist zudem zu beachten, dass die KI-Regulierung EU-weit harmonisiert ist.

4.2.2 Energieeffizienz und Abwärmenutzung

Der Bereich der Energieeffizienz und der Abwärmenutzung ist der einzige Bereich, für den bereits jetzt ein spezifisches „Rechenzentrumsrecht“ existiert, geregelt im EnEfG. Nach dem EnEfG (vgl. § 11) müssen neue Rechenzentren, die ab dem 1. Juli 2026 den Betrieb aufnehmen, eine Energieverbrauchseffektivität von kleiner oder gleich 1,2 erreichen. Diese Anforderung ist spätestens zwei Jahre nach Inbetriebnahme im Jahresdurchschnitt dauerhaft zu erreichen. Rechenzentren, die vor dem 1. Juli 2026 den Betrieb aufgenommen haben, müssen ab dem 1. Juli 2027 eine Energieverbrauchseffektivität von kleiner oder gleich 1,5 und ab dem 1. Juli 2030 eine Energieverbrauchseffektivität von kleiner oder gleich 1,3 im Jahresdurchschnitt erreichen. Zudem müssen Rechenzentrumsbetreiber den Stromverbrauch in ihren Rechenzentren bilanziell ab dem 1. Januar 2024 zu 50 % und ab dem 1. Januar 2027 zu 100 % durch Strom aus erneuerbaren Energien decken.

Weiterhin umfasst das EnEfG in Bezug auf Rechenzentren bestimmte Meldepflichten sowie die Einrichtung von Energie- oder Umweltmanagementsystemen. In diesem Zusammenhang sind für die Meldung entsprechender Kennzahlen kontinuierlich Messungen bei den wesentlichen Komponenten eines Rechenzentrums sowie Maßnahmen zur Verbesserung der Energieeffizienz vorzunehmen.

Die Vorgaben des EnEfG zur Verwertbarkeit der unvermeidbaren Abwärme werden ergänzt um die Vorgaben des WPG. Das WPG schafft die Grundlagen für die Planung und die Errichtung eines Wärmenetzes, um eine Umstellung der Erzeugung und der Versorgung mit Raumwärme (u. a.) auf erneuerbare Energien zu schaffen. Insoweit setzt das WPG unvermeidbare Abwärme auf eine Stufe mit Wärme aus erneuerbaren Energien (vgl. § 1 WPG). Das WPG adressiert Rechenzentren zwar nicht ausdrücklich, allerdings werden Rechenzentren mittelbar in den Anwendungsbereich einbezogen. So definiert § 3 Abs. 1 Nr. 13 WPG unvermeidbare Abwärme als Wärme, die als unvermeidbares Nebenprodukt in einer Industrieanlage, einer Stromerzeugungsanlage oder im tertiären Sektor anfällt und ohne den Zugang zu einem Wärmenetz ungenutzt in die Luft oder in das Wasser abgeleitet werden würde. Abwärme gilt als unvermeidbar, soweit sie aus wirtschaftlichen, sicherheitstechnischen oder sonstigen Gründen im Produktionsprozess nicht nutzbar ist und nicht mit vertretbarem Aufwand verringert werden kann. Zwar ist die Definition nicht deckungsgleich mit der Definition in § 3 Nr. 27 EnEfG zur „technisch unvermeidbaren Abwärme“. Danach ist technisch unvermeidbare Abwärme der Teil der Abwärme, der aufgrund thermodynamischer Gesetzmäßigkeiten entsteht und nicht durch Anwendung des Standes der Technik mit vertretbarem Aufwand vermieden oder reduziert werden kann. Allerdings wird die unvermeidbare Abwärme aus Rechenzentren ohne Weiteres dem tertiären Sektor zugerechnet werden können (unter Umständen gilt sie bereits als unvermeidbares Nebenprodukt aus einer Industrieanlage). Auch an anderer Stelle werden Rechenzentren bzw. deren Betreiber als Produzenten unvermeidbarer Abwärme adressiert (hierauf wird nachfolgend noch eingegangen).

Insoweit kann das Zusammenspiel zwischen den konkreten Regelungen des EnEfG für Rechenzentren und den Vorgaben des WPG durchaus als eine atypische Sektorenkopplung

betrachtet werden. Nichtsdestotrotz gilt es, auch hier das bereits bestehende Regelungsgefüge durch gezielte Änderungen an neu gewonnene Erkenntnisse anzupassen, um bestehende Potenziale vollständig auszuschöpfen.

4.2.3 Regelungspotenziale im Bereich Abwärmenutzung

Der Regelungsrahmen aus EnEfG und KWP – wie zuvor erläutert – trifft auf bestehende Hemmnisse, welche dazu führen, dass Abwärme aus Rechenzentren bislang vielfach ungenutzt bleibt. Eine gezielte Datenerhebung und Informationsbereitstellung sowie Anpassungen der Regulatorik können die Abwärmenutzung jedoch vereinfachen und dazu beitragen, entsprechende Ausbaupotenziale zu heben.

Mit Blick auf die **Nutzbarkeit bzw. die Abnahme der in Rechenzentren** erzeugten unvermeidbaren Abwärme besteht ein generelles Informationsdefizit. So werden oftmals Daten zu Abwärmemengen nicht erhoben oder nicht in ausreichendem Maße bereitgestellt. Eine nicht ausreichende Transparenz der lokalen Wärmeversorgung, insbesondere fehlende Informationen zu vorhandener oder geplanter Wärmeinfrastruktur (Wärmeerzeuger als auch Wärmenetztrassen) verschärfen die Situation. Während in § 14 EnEfG ausdrücklich die Errichtung eines Energieeffizienzregisters für Rechenzentren geregelt und mittlerweile bereits durch das RZReg umgesetzt ist, besteht mit Blick auf die Wärmeversorger kein Äquivalent. Zwar sind die Weichen für die Nutzbarmachung der Abwärme durch das EnEfG als auch durch die Ausbauziele für klimaneutrale Wärmeversorgung im WPG gestellt – die Umsetzung ist vielerorts jedoch noch an Herausforderungen gebunden. In der Praxis müsste der zukünftige Betreiber eines neuen Rechenzentrums für die Standortwahl zunächst mögliche Kommunen kontaktieren um Informationen über den Bestand und die Planung einer Wärmeinfrastruktur ersuchen. Dies kann zu einer Verzögerung im Planungsprozess eines neuen Rechenzentrums führen. In Anbetracht der Tatsache, dass der Rechenzentrumsbetreiber die notwendige Infrastruktur zur Abgabe der Wärme bereithalten muss, um nach § 11 Abs. 3 Nr. 3 EnEfG von den Anforderungen des § 11 Abs. 2 Satz 1 Nr. 2 EnEfG befreit zu sein, kann dies zudem zu Mehrkosten führen.
Das WPG hat mit der verbindlichen KWP die Grundlage geschaffen, Potenziale für klimaneutrale Wärmequellen als auch den Verlauf bestehender sowie potenzieller Wärmenetze sichtbar zu machen. Damit werden zwar grundsätzlich Potenziale für eine Nutzung der Abwärme von Rechenzentren aufgezeigt. Um die Einspeisung von Abwärme in Wärmenetzen zu erleichtern, wäre es empfehlenswert, verfügbare Informationen in einer zentralen Plattform – wie beispielsweise einem für relevante Akteure zugänglichen Wärmeinfrastrukturatlas – zu bündeln. Damit würde die Planung neuer Rechenzentren auf Seiten der Betreiber erleichtert.

Um die Erschließung bestehender Abwärmepotenziale weiter zu forcieren, könnten Regelungen geschaffen werden, die eine Koordination lokaler Planungen zwischen Rechenzentrumsbetreibern und Wärmenetzbetreibern regeln (ähnlich wie dies bereits in § 143 TKG für die Eigentümer und Betreiber öffentlicher Versorgungsnetze und die Eigentümer und Betreiber öffentlicher Telekommunikationsnetze vorgesehen ist). Gerade in Anbetracht der Tatsache, dass Rechenzentrumsbetreiber bei der Standortauswahl ebenfalls reguliert sein können, bietet es sich an, diese im Wege einer Transparenzsymmetrie und durch weitere Koordinationsmöglichkeiten bei der Planung und der Errichtung neuer Rechenzentren zu unterstützen (siehe hierzu auch Unterabschnitt

„Planung“). Gleichsam würden hiervon die Wärmeversorger profitieren, die auf ein hinreichend großes Angebot von Wärme angewiesen sind.

Ansatzpunkte für weitere Potenziale sind:

Ein potenzielles Leerlaufen der Verpflichtung zur Abwärmenutzung ist im Rahmen der Ausnahmeregelungen hinsichtlich der Vorgabe der Abwärmenutzung für den Fall möglich, dass der Betreiber eines in der Umgebung befindlichen Wärmenetzes ein Angebot zur Nutzung wiederverwendeter Energie zu Gestehungskosten nicht innerhalb von sechs Monaten annimmt. Voraussetzung ist, dass der Betreiber des Rechenzentrums die notwendige Infrastruktur zur Bereitstellung der Wärme, insbesondere in Form einer Wärmeübergabestation, bereithält. So sieht das Gesetz weder eine regelmäßige Prüfung oder Wiedervorlage des Angebots noch eine alternative Verpflichtung zur Abwärmenutzung vor.

Hinzukommt, dass mit der Ausnahme von der Vorgabe einer Abwärmenutzung für den Fall, dass der Anteil an wiederverwendeter Abwärme nach Inbetriebnahme durch nachträgliche Ereignisse und ohne Verschulden des Betreibers nicht mehr den Anforderungen entspricht, nicht durch die Aufnahme von Regelbeispielen eingegrenzt ist, welche nachträglichen Ereignisse hiervon umfasst sein könnten.

Ein weiteres mögliches tatsächliches Hemmnis ist die **Abwärmetemperatur.** Der Umstand, dass die Abwärme in bestimmten Konstellationen nicht der zur weiteren Wärmenutzung erforderlichen Vorlauftemperatur entspricht und die Temperatur daher erhöht werden muss, kann die Gesamteffizienz der Abwärmenutzung senken. Für die Betreiber von Rechenzentren wirkt sich in diesem Zusammenhang allerdings positiv aus, dass der entsprechende Strombedarf von Wärmepumpen nicht in die Berechnung des PUE-Werts einbezogen wird. Entsprechend kann die Abwärmenutzung in den Sommermonaten zu einer Verbesserung der PUE-Werte beitragen.

Zudem zeigt auch der Blick auf Meldepflichten von Daten mit Nachhaltigkeitsbezug auf, dass der damit einhergehende Aufwand als hemmendes Kriterium eingeordnet werden könnte. Denn Meldepflichten – sogar unterschiedlichen Umfangs – bestehen nicht nur auf Grundlage des EnEfG, sondern auch auf Grundlage der im Juni 2024 in Kraft getretenen Delegierten Verordnung (EU) 2024/1364 zur Energieeffizienzrichtlinie (nachfolgend Delegierte Verordnung).

Dabei gilt es zu beachten, dass das EnEfG in § 18 Möglichkeiten zu Ausnahmen und Befreiungen vorsieht und die damit einhergehende Anreizwirkung anscheinend nicht zugleich auf die nach der Delegierten Verordnung umfassten Meldepflichten durchgreift. Hinzu kommt, dass nach der Delegierten Verordnung bestimmte Kennzahlen jeweils bei Kunden von Co-Location-Betreibern zu erfassen und für diese jeweils separat auszuweisen sind. Eine gesetzlich verpflichtende Handhabe im Verhältnis des Rechenzentrumsbetreibers zu seinen Kunden gibt es hierbei jedoch nicht. Umgekehrt gehen an anderer Stelle die geforderten Kennzahlen über die Forderungen der Delegierten Verordnung hinaus. So sollen nach § 15 EnEfG Co-Location-Rechenzentrumsbetreiber die direkt ihren jeweiligen Kunden zuzuordnenden Energieverbräuche pro Jahr abbilden. Zu konstatieren ist, dass generell Aufwände für die Branche, die deutlich über die EU-harmonisierten Verpflichtungen hinausgehen, jedenfalls im europäischen Vergleich einen Standortnachteil bedeuten können. Andererseits kann dies insbesondere für die Kunden von Rechenzentren vorteilhaft sein, und dies könnte wiederum den Rechenzentrumsstandort Deutschland stärken.

Weiterhin gibt es – außerhalb des EnEfG bzw. der Energieeffizienzrichtlinie und der Delegierten Verordnung – weitere Berichtspflichten auf Grundlage der CSRD der Europäischen Union als Teil des Europäischen Green Deals. Diese adressiert für Rechenzentrumsbetreiber Pflichten bei der Nachhaltigkeitsberichterstattung zwar nicht als spezifisches „Rechenzentrumsrecht“, gleichwohl sind hiernach Rechenzentren als Teil der energieintensiven Branche regelmäßig betroffen. Für einen künftigen Gleichlauf entsprechender Rechte und Pflichten für Rechenzentrumsbetreiber müsste jedenfalls der deutsche Gesetzgeber tätig werden.

Aus Stellungnahmen der Branche ergeben sich Kritikpunkte dergestalt, dass die Verantwortlichkeiten der Beteiligten noch nicht im erforderlichen Maße definiert sind, und zwar sowohl hinsichtlich der Frage der Kostentragung als auch mit Blick auf die Frage der Wärmeplanung.

Im Hinblick auf die **Kommunikation** der Beteiligten im Kontext der Abwärmenutzung besteht nach Aussage der Branche der Bedarf nach spezifischen Kommunikationsformaten und -routinen. Es wäre zudem zu klären, wer diese Kommunikation organisiert, insbesondere für künftige Standorte für Rechenzentren. Hier enthält das EnEfG lediglich die Verpflichtung für den Rechenzentrumsbetreiber, die Abwärme „anzubieten“. Aus regulatorischer Sicht gilt es in diesem Zusammenhang zu berücksichtigen, dass auch das WPG hierzu bisher keine entsprechenden Regelungen enthält. Gerade mit Blick auf die Planung künftiger Standorte könnten die zuvor skizzierten Möglichkeiten eines Wärmeinfrastrukturatlas oder der Koordination von Baumaßnahmen dazu beitragen, die Potenziale zur Abwärmenutzung weitgehender zu heben.

Überdies kann die **Kostenverantwortung** einen Unsicherheitsfaktor darstellen (siehe auch Kapitel 2.4). So bleibt unklar, was aus der Verpflichtung, ein Angebot zur Abwärmenutzung zu „Gestehungskosten“ gemäß § 11 EnEfG zu machen, konkret folgt. Nach Auskünften Beteiligter in den Stakeholder-Workshops wäre es eine Erleichterung, wenn die Möglichkeit bestehen würde, die Abwärme weiterzugeben, ohne hierfür eine Gegenleistung in Form der Gestehungskosten zu verlangen. Hintergrund hierfür ist, dass die Nutzung der Abwärme aus Sicht der Rechenzentrumsbetreiber als Erfüllung einer gesetzlichen Pflicht angesehen wird und nicht als Bestandteil des Business Case, sodass betriebswirtschaftlich auch etwaige Einnahmen verzichtbar sind. Insoweit könnte der Gesetzgeber die rechtliche und steuerliche Zulässigkeit eines unentgeltlichen Angebots zur Abwärmenutzung vorsehen.

4.2.4 Planung

Aus Sicht von Branchenvertretenden wird teilweise die einseitige und zu ambitionierte Fokussierung auf die PUE als Hemmnis benannt. Es wird u. a. kritisiert, dass die ambitionierten zeitlichen Vorgaben des EnEfG nicht der technischen Entwicklung und auch nicht der Marktnachfrage entsprächen. Eine Herausforderung bestehe darin, dass das EnEfG einen möglichst geringen PUE-Wert erfordert. Allerdings verschlechtern nach Aussage der Branchenvertretenden Anforderungen an Verfügbarkeiten und Redundanz von Ressourcen im Rechenzentrum sowie gängige Geschäftsmodelle den PUE-Wert.

In diesem Zusammenhang wird von Branchenvertretenden zur Diskussion gestellt, ob Defizite bei der Energieeffizienz ggf. an anderer Stelle ausgeglichen werden könnten. Beispielsweise durch die vorzeitige Erfüllung von Dekarbonisierungs-Vorgaben bei der Energiebeschaffung (beispielsweise durch eine vollständige Versorgung mit Grünstrom mittels Onsite PPAs).

Ausweislich der Gesetzesbegründung des EnEfG zielen die Regelungen rund um die Nutzung von Abwärme darauf ab, dass sich Rechenzentrumsbetreiber bei der Planung bereits auf die Auskopplung eines bestimmten Anteils der im Rechenzentrum entstehenden Abwärme einstellen und entsprechende Investitionen vornehmen. Auch gilt es, aus rechtlicher Sicht die Vorgaben AVBFernwärmeV für die Laufzeit von entsprechenden Versorgungsverträgen von maximal zehn Jahren zu beachten.

Die dargestellten Potenziale zeigen auf, dass die bestehende Regulatorik einerseits und die entsprechenden Geschäftsmodelle der Branche mit Blick auf die zugrunde zu legenden Parameter für die Abwärmenutzung von Rechenzentren andererseits teilweise zu Herausforderungen führen können. Um Regulatorik verstärkt als Enabler und als für den Standort Deutschland sprechendes Kriterium zu nutzen, könnte sie entsprechend weiterentwickelt und als integraler Bestandteil der entsprechenden Geschäftsmodelle und der Planung der erforderlichen Infrastruktur (z. B. Wärmenetze) einbezogen werden. Die oben dargestellten Instrumente zum Wärmeinfrastrukturatlas bzw. zu einer Regelung ähnlich der Koordination von Baumaßnahmen in § 143 TKG könnten insoweit einen ersten Ansatz darstellen.

4.2.5 Bauplanungsrecht und Außenbereichsprivilegierung

Die Errichtung von Rechenzentren unterliegt baurechtlichen Genehmigungserfordernissen. Zu konstatieren ist, dass es für die Errichtung von Rechenzentren im Bereich des Bauplanungsrechts ein spezielles „Rechenzentrumsrecht“ nicht gibt. Mit Blick auf eine Beschleunigung bzw. Vereinfachung der Baugenehmigungsprozesse ist zu fragen, ob es sachgerecht wäre, Rechenzentren in den Kanon bauplanungsrechtlich privilegierter Außenbereichsvorhaben nach § 35 Abs. 1 BauGB aufzunehmen. Dies hätte zur Folge, dass eine gemeindliche Bauleitplanung für entsprechende Vorhaben zur Ansiedlung von Rechenzentren entbehrlich würde, was den Prozess beschleunigen dürfte. Denn die kommunale Bauleitplanung nimmt erfahrungsgemäß viel Zeit in Anspruch, da hierfür öffentliche und private Belange, mithin Aspekte des Umweltschutzes und etwa Belange der Anwohner einzubeziehen sind und auch die Erschließung der Flächen gesichert sein muss. Eine übermäßige Einschränkung der Rechte der Standortgemeinde wäre damit nicht verbunden, weil das gemeindliche Einvernehmen nach § 36 BauGB weiterhin erforderlich ist und weil es der Gemeinde unbenommen ist, Standorte anderweitig zu überplanen.

Aus kommunaler Perspektive ist an dieser Stelle auf die Bauleitplanung Bezug zu nehmen. Gegenstand der Bauleitplanung ist die Schaffung von „Baurecht“ als Prüfmaßstab für Baugenehmigungen. Im Rahmen des kraft Gesetzes vorgesehenen Beitrags für eine nachhaltige städtebauliche Entwicklung mitsamt entsprechenden umweltschützenden Anforderungen (§ 1 BauGB) steht derzeit als Steuerungsinstrument etwa die Möglichkeit eines **städtebaulichen Entwicklungskonzeptes** zur Verfügung. Mit diesem Steuerungsinstrument kann die Ansiedlung von Rechenzentren auf entsprechend geeigneten Flächen innerhalb der jeweiligen Kommune koordiniert werden. Das städtebauliche Entwicklungskonzept kann mithin auch eine Rechenzentrumsstrategie umfassen, § 1 Abs. 6 Nr. 11 BauGB. Auch dies würde einen Beitrag zur Prozessbeschleunigung leisten können, setzt aber aufgrund der kommunalen Selbstverwaltungsgarantie eine entsprechende Willensbildung und -betätigung der einzelnen Kommunen voraus.

Unter dem Gesichtspunkt der Abwärmenutzung von Rechenzentren besteht als weitere kommunale Steuerungsmöglichkeit die KWP gemäß Gesetz für die Wärmeplanung und zur Dekarbonisierung der Wärmenetze. Daneben steht aber auch die Möglichkeit der Regelung energieversorgungsrelevanter Themen mittels Vereinbarung **städtebaulicher Verträge**. Hiervon erfasst werden können insbesondere Verträge für die Planung und Umsetzung von Wärmepumpen und Wärmenetzen sowie die Verpflichtung zur Nutzung entsprechender Wärme (auf Abnehmerseite). Auch erfasst werden können Planungen und Verträge für den entsprechenden Ausbau von Einrichtungen zur dezentralen und zentralen Erzeugung, Verteilung, Nutzung oder Speicherung von Strom. Daneben besteht, ungeachtet entsprechender städtebaulicher Verträge, die Möglichkeit eines Anschluss- und Benutzungszwangs (auf Abnehmerseite). Die Kommunen können so zur Inangriffnahme einer klimafreundlichen Wärmeplanung beitragen, wobei städtebauliche Verträge sogar neben einem Anschluss- und Benutzungszwang (durch Satzung) zum Einsatz kommen können.

Insgesamt ist festzuhalten, dass die Herstellung von Baurecht im Rahmen der Bauleitplanung ein politischer Prozess ist, dessen Ausgang nicht immer vorhersehbar ist. Die Aufstellung und Änderung von Bebauungsplänen bedeuten eine deutliche Verzögerung für Rechenzentrumsentwickler. Ein hohes Potenzial ist hingegen für die Ansiedlung von Rechenzentren zu verzeichnen, in denen über die Bauleitplanung entsprechende Flächen für Rechenzentren bereits ausgewiesen sind.

Hilfreich im Rahmen der Bauleitplanung aber auch im Genehmigungsverfahren wäre es, wenn zunehmend für mehr Transparenz und eine koordinierte Infrastruktur **integrierte Prozesse/Verfahren aufgesetzt, digitalisiert, vereinfacht und zeitlich verkürzt bzw. gestrafft würden**. In diesem Zusammenhang ist etwa das am 7. Juli in Kraft getretene Gesetz zur Stärkung der Digitalisierung im Bauleitplanverfahren sowie die konkret beschriebenen Vorhaben der Bundesregierung im Rahmen ihrer beschlossenen Wachstumsinitiative vom 5. Juli 2024 positiv einzuordnen. Auch gibt es mittlerweile bereits im Rahmen der Umsetzung des Onlinezugangsgesetzes jedenfalls partiell die Möglichkeit digitaler Bauanträge und digitaler Baugenehmigungen.

4.2.6 Genehmigungserfordernisse

Bei einem Rechenzentrum handelt es sich um eine bauliche Anlage, die im Regelfall nach den bauordnungsrechtlichen Bestimmungen der Länder einer Baugenehmigung bedarf. Sind eine Notstromanlage und/oder eine sonstige Erzeugungsanlage zur netzunabhängigen Energieversorgung (z. B. Standortkraftwerk) Bestandteil des Rechenzentrums, so kann zusätzlich und unabhängig davon eine Genehmigungserfordernis nach Bundes-Immissionsschutzgesetz (BImSchG) bestehen. Dies richtet sich danach, ob die Energieanlage eine bestimmte Leistungsschwelle erreicht. Gemäß § 4 BImSchG in Verbindung mit Anhang 1 Nr. 1 zur 4. Bundesimmissionsschutzverordnung (BImSchV) ist eine immissionsschutzrechtliche Genehmigung erforderlich, wenn die Anlage eine Feuerungswärmeleistung von 20 MW oder mehr aufweist. Liegt die Feuerungswärmeleistung bei 50 MW oder mehr, so ist sogar ein Genehmigungsverfahren mit Öffentlichkeitsbeteiligung durchzuführen.

Für beide Genehmigungsverfahren ergibt sich ein mehr oder weniger identischer Prüfungsmaßstab mit entsprechenden Restriktionen:

- Bauplanungsrecht: Die Errichtung des Rechenzentrums (gegebenenfalls mit Energieanlage) muss den Festsetzungen eines vorhandenen Bebauungsplans entsprechen (§ 30 Baugesetzbuch – BauGB) oder es muss sich, sofern ein Bebauungsplan nicht vorhanden ist,

nach Art und Maß der baulichen Nutzung in die nähere Umgebung einfügen (§ 34 BauGB). Eine erhebliche Erschwernis kann sich dabei in den Fällen ergeben, in denen die Energieanlage eine Leistungsschwelle erreicht, die ein immissionsschutzrechtliches Genehmigungserfordernis auslöst. § 15 Abs. 3 der Baunutzungsverordnung (BauNVO) stellt klar, dass die Zulässigkeit von Anlagen in den Baugebieten nicht allein nach den verfahrensrechtlichen Einordnungen des BImSchG und der auf seiner Grundlage erlassenen Verordnungen zu beurteilen ist. Trotzdem ist anerkannt, dass ein immissionsschutzrechtliches Genehmigungserfordernis einen gewichtigen Anhaltspunkt dafür bietet, dass die Anlage nicht zu den nicht-störenden Gewerbebetrieben zählt und daher nur in einem Industriegebiet GI (§ 9 BauNVO) zulässig ist, nicht aber in einfachen Gewerbegebieten GE (§ 8 BauNVO) (vgl. Ernst/Zinkahn/Bielenberg/Krautzberger/Söfker BauNVO § 15 Rn. 34). Flächen in bauplanungsrechtlich ausgewiesenen Industriegebieten sind aber nur begrenzt verfügbar. Naturgemäß würde eine vorrangige Verortung von Rechenzentren in Industriegebieten zudem die Nutzung der in ihnen entstehenden Abwärme erschweren, weil Industriegebiete typischerweise eine größere Entfernung zu potenziellen Nutzern der Wärme aufweisen.

- Bauordnungsrecht: Darüber hinaus müssen Rechenzentrum und Energieanlage den bauordnungsrechtlichen Vorgaben im jeweiligen Bundesland entsprechen. Das heißt, sie müssen den dortigen Anforderungen in Bezug auf Standsicherheit, Brandschutz, gesunde Wohn- und Arbeitsverhältnisse etc. genügen. Rechenzentren werden dabei nach vielen landesrechtlichen Vorschriften als sogenannte Sonderbauten einzustufen sein, für die insoweit verschärfte Anforderungen gelten.
- Immissionsschutzrecht: Etwa mit dem Rechenzentrum verbundene Energieanlagen (Notstromanlage, Standortkraftwerk o. ä.) unterliegen den Anforderungen an mittelgroße Feuerungsanlagen gemäß der 44. BImSchV. Dies gilt unabhängig davon, ob für die Anlage ein gesondertes immissionsschutzrechtliches Genehmigungserfordernis besteht. Die Verordnung sieht in erster Linie Emissionsgrenzwerte für unterschiedliche Parameter, wie Ammoniak, Stickoxide, Schwefeldioxid, Quecksilber etc., vor. Sie geht dabei insoweit über die Vorgaben der ihr zugrunde liegenden sogenannten europäischen MCP-Richtlinie (Medium Combustion Plant = mittlere Verbrennungsanlage; RL (EU) 2015/2193) hinaus, als sie nur eingeschränkt von der dort in Art. 6 Abs. 3 vorgesehenen Möglichkeit Gebrauch macht, Anlagen mit einer Betriebsdauer von bis 500 Stunden pro Jahr von wesentlichen Teilen der Anforderungen auszunehmen. Die 44. BImSchV sieht entsprechende Erleichterungen nur für eine Betriebsdauer bis zu 300 Stunden pro Jahr und in reduziertem Umfang vor (vgl. etwa § 16 der 44. BImSchV für Notstromanlagen in Form von Verbrennungsmotoranlagen). Auch insoweit besteht somit ein gesteigertes genehmigungsrechtliches Hemmnis.
- Wasserrecht: Spezifische Anforderungen können sich schließlich aus dem Regelungsregime über Anlagen zum Umgang mit wassergefährdenden Stoffen (§§ 62 ff. WHG in Verbindung mit der Verordnung über Anlagen zum Umgang mit wassergefährdenden Stoffen (AwSV)) ergeben. Sofern Notstrom- oder sonstige Energieanlagen Teil des Rechenzentrums sind, werden diese regelmäßig über Kraftstofftanks mit einem ausreichenden Volumen verfügen, welche in den Anwendungsbereich dieser Vorschriften fallen. Danach sind Vorgaben z. B. zu Auffang- und Rückhaltevorrichtungen sowie strikte Überwachungs- und Prüfpflichten zu beachten.
- Straßenrecht, sonstige Rechtsgebiete: Unter bestimmten Voraussetzungen bedarf es zudem straßenrechtlicher Ausnahmen und Zustimmungen (z. B. nach § 9 FStrG) und ggf. auch

Sondernutzungserlaubnisse (ggf. bei Querung von Straßen mit Leitungen oder für das Aufstellen von Gerüsten und Containern, wobei wiederum für letztere ggf. auch straßenverkehrsrechtliche Anordnungen erforderlich sein können). Weiterhin können luftverkehrsrechtliche Zustimmungen erforderlich werden, etwa wegen Bauschutzbereichen von Flugplätzen oder für Bauwerke von mehr als 100 m Höhe (relevant für Schornsteine der USV-Anlagen) nach Luftverkehrsgesetz. Ebenfalls stehen ggf. denkmalschutzrechtliche Erlaubnisse in Rede.

Im Zusammenhang mit dem baurechtlichen- bzw. immissionsschutzrechtlichen Genehmigungsprozess identifizierten Teilnehmende des dritten Workshops im Rahmen dieses Auftrags die Umständlichkeit bzw. den Aufwand der Antragstellung aufgrund vorzulegender Nachweise und der Nachreichung von Unterlagen/Nachweisen an die Genehmigungsbehörde. Dies lasse sich auf Abstimmungsprobleme und fehlende Kommunikationsstrukturen zwischen Genehmigungsbehörde und Fachbehörden zurückführen, was wiederum im Zusammenhang zur bau- und immissionsschutzrechtlichen Konzentrationswirkung steht. Hierzu ist in § 54 Abs. 3 S. 3 der Landesbauordnung geregelt, dass wenn es für die Erteilung der Baugenehmigung des Einvernehmens oder der Zustimmung einer anderen Stelle bedarf, so gilt diese als erteilt, wenn sie nicht innerhalb eines Monats nach Eingang des Ersuchens unter Angabe der Gründe verweigert wird. Mögen solche Zustimmungsfiktionen zwar ein wichtiges Signal zur Beschleunigung von Genehmigungsprozessen senden, so räumen sich Genehmigungsbehörden und Fachbehörden nach Aussagen von Workshopteilnehmenden eine gewisse Flexibilisierung in Gestalt einer Verlängerung dieser Frist ein, indem für den Fristbeginn nicht auf den Eingang des Ersuchens (Bauantrag) abgestellt wird, sondern auf den Zeitpunkt von Nachreichungen von Dokumenten für die Antragstellung.

4.2.7 Anschluss/Nutzung Infrastrukturen

Ein zentraler Themenkomplex, der die Standortwahl für Rechenzentren beeinflusst und der zumindest in zweiter Linie auch regulatorischer Art ist, ist die Frage des Anschlusses an die notwendigen Netzinfrastrukturen und die insoweit zur Verfügung stehenden Kapazitäten.

In Bezug auf den **Stromnetzanschluss:** Hier stellt sich in regulatorischer Hinsicht die Frage, inwieweit die „allgemeinen“ Netzanschluss- und Netznutzungsansprüche aus § 17 EnWG und § 20 EnWG (jeweils auch mit Einschränkungen und Kostentragungspflichten) die Interessen von Rechenzentrumsbetreibern einerseits und Stromnetzbetreibern andererseits ausreichend abbilden. Die Herausforderung liegt dabei in **enormen Leistungsbedarfen** nicht nur von Rechenzentren, sondern in Kombination mit weiteren Großverbrauchern in räumlicher Nähe und in einer auf Kapazitätsmängeln beruhenden Anschlusskonkurrenz. Diese verschärft sich noch mit der nach dem EnEfG erforderlichen Abwärmenutzung. Denn faktisch sind Wärmenetze in räumlicher Nähe von Rechenzentren tatsächlich eher in (Groß-) Städten vorhanden. Damit wird aber die Ansiedlung in lastgeprägten Netzgebieten verstärkt. Auch die für die Abwärmenutzung erforderliche Temperaturerhöhung durch Wärmepumpen führt zu weiteren Lasten. Netzbetreiber sind mit Blick auf den Stromnetzanschluss und dessen Nutzung verpflichtet, einen **diskriminierungsfreien Netzanschluss und -zugang zu gewähren. Konkrete Kriterien** sind dabei gesetzlich nicht etabliert. Netzbetreiber können solche Netzanschlussbegehren bspw. chronologisch bearbeiten. Daneben greifen Netzbetreiber gerade mit Blick auf Anschlussbegehren von Betreibern besonders energieintensiver Anlagen auch auf das sog. Repartierungsverfahren zurück, das die verfügbaren

Kapazitäten ermittelt und anteilig und gleichberechtigt auf die Netzanschlussbegehrenden verteilt. Einer rechtssicheren Privilegierung von Anschlussbegehren durch Rechenzentren könnte am Ehesten damit begegnet werden, indem der Gesetzgeber konkrete Auswahlkriterien oder Privilegierungstatbestände für Rechenzentren bei Anschlusskonkurrenzen aufnimmt, die rechtliche Zulässigkeit einer Privilegierung von Rechenzentren und eine entsprechende energiepolitische Abwägung zugunsten von Rechenzentren vorausgesetzt.

Als für Anschlussbegehren positiv zu werten ist die ausstehende Umsetzung der Regelung in Art. 6a der EU-Strombinnenmarktrichtlinie über flexible Netzanschlussverträge. Denkbar ist hier, den Netzanschluss für Rechenzentren einem verbindlichen Reservierungsverfahren zu unterziehen, wie dies für konventionelle Erzeuger in § 4 KraftNAV bereits geltendes Recht ist und für EE-Anlagen derzeit geplant ist (§ 8a EEG-Entwurf).

4.2.8 Erfordernis einer überörtlichen Bedarfsplanung

Die vorstehenden Punkte führen zu der Frage, ob nicht (zusätzlich zur oben beschriebenen Möglichkeit der örtlichen Planung) eine überörtliche Bedarfsplanung und eine entsprechende Flächenausweisung für Rechenzentren probate Mittel wären, um die Standortwahl zu steuern und Planungssicherheit für alle Akteure zu erzeugen. Im Ausgangspunkt ist festzuhalten, dass Rechenzentren in mehrfacher Hinsicht von Netzinfrastrukturen abhängen, die sich ganz unmittelbar auf die Standorteignung auswirken. Dies betrifft den schon angesprochenen Stromnetzanschluss genauso wie den Anschluss und Zugang zum Telekommunikationsnetz sowie das Vorhandensein von Wärmenetzen, um die Abwärme einer sinnvollen Nutzung zuzuführen.

Insoweit wird teilweise die Forderung erhoben, diese – und andere Standortfaktoren – in eine überörtliche und integrale Bedarfsplanung und eine **Potenzialanalyse** einfließen zu lassen. In diesem Zusammenhang kann beispielsweise auf die Potenzialanalyse im Bereich des Gigabit-Ausbaus verwiesen werden.

4.3 Erörterung der Kennzahlen zum Energie- und Ressourceneinsatz

In diesem Unterkapitel werden die wichtigsten Energie- und Ressourceneffizienz-Kennzahlen für Rechenzentren auf Basis der Normenfamilie DIN EN 50600 dargestellt. Hierbei spielt die detaillierte Analyse des Effizienzparameters PUE eine zentrale Rolle, da es sich hierbei um **eine der häufigsten verwendeten Metriken zur Bewertung der Energieeffizienz** von Rechenzentren handelt. Zudem werden im Rahmen des **§ 11 EnEfG** weitreichende **regulatorischen Anforderungen** zum PUE formuliert. Daher werden im Anschluss zur Definition der Kennzahlen sowohl die historische Entwicklung des PUE genauer betrachtet als auch die Vor- und Nachteile dieser Kennzahl diskutiert.

4.3.1 Definition der Kennwerte

In diesem Abschnitt werden die drei wichtigsten Effizienzkennzahlen zum Strom-, Wasser- und Kohlenstoffeinsatz auf Basis der DIN EN 50600 aufgeführt.

Kennzahl zur eingesetzten Energie / Power Usage Effectiveness (PUE) nach DIN EN 50600-4-2

Die PUE misst das **Verhältnis zwischen der gesamten im Rechenzentrum verbrauchten Energie und der Energie, die für die IT-Geräte** wie Server, Speicher und Netzwerkausrüstung verwendet wird. Ursprünglich wurde der PUE-Wert von der Organisation The Green Grid im Jahr 2007 vorgeschlagen (The Green Grid, 2007), um Fortschritte in der Energieeffizienz von Rechenzentren im Laufe der Zeit zu messen. Inzwischen wurde PUE auch in internationalen Standards wie EN 50600-4:2 und ISO/IEC 30134-2 formalisiert und ist somit eine international anerkannte Metrik.

Der PUE-Wert von 1 stellt das ideale Ziel dar, bei dem die gesamte Energie eines Rechenzentrums ausschließlich für die IT-Geräte genutzt wird. Ein PUE größer als 1 bedeutet, dass zusätzliche Energie für andere Zwecke, wie Kühlung, Beleuchtung oder andere nicht IT-bezogene Systeme, verbraucht wird. Ein sehr effizientes Rechenzentrum hat typischerweise einen PUE-Wert von 1,2 bis 1,4, während der Wert bei durchschnittlichen Rechenzentren zwischen 1,4 und 2,0 liegt. Weniger effiziente Rechenzentren können PUE-Werte über 2,0 aufweisen.

Rechenzentren können ihre PUE durch verschiedene Maßnahmen verbessern. Dazu gehören die Implementierung effizienter Kühltechnologien, Freikühlung oder Warm- und Kaltgangeinhausung. Ein weiterer Ansatz ist die Optimierung der Stromverteilung, beispielsweise durch die Reduktion von Energieverlusten in der Stromversorgung oder die Nutzung energieeffizienter, unterbrechungsfreier Stromversorgungssysteme. Zudem kann durch die kontinuierliche Überwachung und das Management des Energieverbrauchs mithilfe von Energiemanagement-Strategien der Energieverbrauch optimiert werden. Andererseits führt eine räumliche oder zeitliche Teilauslastung von Rechenzentren zu einer Verschlechterung des PUEs durch die vorhandenen Überkapazitäten bei der Strom- und Kälteversorgung. Eine Gegenmaßnahme kann die bessere Ausnutzung der vorhandenen Kapazitäten für Server und Racks sein. Dies kann allerdings zu Zielkonflikten mit den oben beschriebenen Ansätzen führen, Flexibilitäten aus Rechenzentren zur Verfügung zu stellen.

Die Kennzahl zur eingesetzten Energie wird folgendermaßen bestimmt:

$$PUE = \frac{E_{RZ}}{E_{IT}}$$

E_{RZ} der Gesamtenergieverbrauch des Rechenzentrums (jährlich) in kWh

E_{IT} der Energieverbrauch der IT-Ausstattung (jährlich) in kWh.

Der PUE wird auf jährlicher Basis berechnet, wofür eine Messung über einen Zeitraum von 12 Monaten erforderlich ist. Hierbei ist der PUE immer größer als 1. Der Idealwert ist 1. Die anzusetzende Häufigkeit der Messungen ist in der DIN EN 50600 allerdings nicht festgeschrieben.

Zu den Haupteinflussfaktoren auf den PUE zählen insbesondere (i) die Verfügbarkeitsklassen und die Auslastung; (ii) die geografische Lage und die klimatischen Bedingungen sowie (iii) die eingesetzten technischen Lösungen zu Klimatisierung und Kühlung, Stromversorgung, Brandschutz und Sicherheitstechnik.

Es gibt drei Kategorien des PUE, welche von einem einfachen (Kategorie 1) bis zum fortgeschrittenen Niveau (Kategorie 3) eine erhöhte Auflösung bieten, diese unterscheiden sich basierend auf dem Ort

der Energieverbrauchsmessung der IT-Ausstattung. Kategorie 1 misst den Energieverbrauch der IT-Ausstattung an dem USV-Ausgang und schließt somit veränderliche IT- und Kühlungslasten mit ein. Kategorie 2 misst am PDU-Ausgang und schließt somit den Einfluss von mit PDU-Transformatoren und statischen Schaltern zusammenhängenden Verlusten aus. Kategorie 3 misst am Eingang der IT-Ausstattung und schließt somit den Einfluss von mit elektrischen Verteilerbaugruppen und nicht IT-bezogenen Geräten zusammenhängenden Verlusten aus.

Die höheren Kategorien bieten zunehmend eine genauere Messung des Energieeinsatzes, weil die Messungen näher an den Energieverbrauchern durchgeführt werden. Gleichzeitig bietet es mehr Möglichkeiten zur Verbesserung der Energieeffizienz. Die Wahl der Messgenauigkeit ist dabei vom PUE-Wert abhängig. Bei einem PUE < 1,5 sollte es mindestens die Kategorie 2 sein.

Effektivität der Vermeidung von CO_2-Emissionen / Carbon Usage Effectiveness (CUE) nach DIN EN 50600-4-8

Die Effektivität der Vermeidung von CO_2-Emissionen (CUE) ist wie folgt definiert:

$$CUE = \frac{RZ_{CO_2}}{E_{IT}}$$

RZ_{CO_2} die CO_2-Emissionen des Rechenzentrums in kg

E_{IT} der Energieverbrauch der IT-Ausstattung (jährlich) in kWh.

Der CUE kann Werte von 0 bis unendlich annehmen, wobei der Idealwert 0 ist. Die Berechnung der CUE erfordert die Erfassung und Dokumentation der RZ-CO_2-Werte und von E_{IT} für eine übereinstimmende Zeitspanne von 12 Monaten. Es erfolgt zusätzlich eine Differenzierung nach Kategorie 1 (grundlegend) und 2 (mittel).

Auch die CUE kann in zwei Kategorien betrachtet werden. Bei der Kategorie 1 werden die CO_2-Emissionen der gesamten internen und externen Stromversorgung betrachtet. Zudem wird bei der Kategorie 1 ausschließlich elektrische Energie betrachtet. Die CUE der Kategorie 2 deckt sowohl die CO_2-Emissionen als auch die CO_2e für die gesamte externe und interne Stromversorgung, für sämtliche zusätzliche Energieversorgung sowie für sämtliche weiteren CO_2e-relevanten Emissionsquellen ab.

Effektivität der Wasserverwendung / Water Usage Effectiveness (WUE) nach DIN EN 50600-4-9

Die Effektivität der Wasserverwendung (WUE) ist wie folgt festgelegt:

$$WUE = \frac{W_U}{E_{IT}}$$

E_{IT} der Energieverbrauch der IT-Ausstattung (jährlich) in MWh

W_U die Wasserverwendung (jährlich) eines Rechenzentrums, gemessen als Gesamtvolumen in m^3.

Auch der WUE kann in mehreren Kategorien (1 bis 3) betrachtet werden und unterscheidet sich hinsichtlich der Genauigkeitsstufen, des betrachteten Wasserzugangs, der betrachteten Wasserabgabe und hinsichtlich dessen, ob das regionale Wasserniveau und der Flächenverbrauch

mit berichtet werden. Zusätzlich kann über den Faktor der Wiederverwendung von Wasser (WRF) die Wasserwiederverwendung eines Rechenzentrums berechnet werden.

Zusätzlich zu den im vorherigen Abschnitt erläuterten Kennzahlen gibt es eine **Vielzahl an weiteren Kenngrößen**. Hierzu gehören unter anderem der Anteil der Nutzung von erneuerbaren Energien (DIN EN 50600-4-3), die Kühlungseffizienz (DIN EN 50600-4-7), der energieeffiziente Einsatz von IT-Geräten sowie Leerlaufanteile von Servern und des Rechenzentrums (IEA 4e, 2021).

Betrachtung historischer Effizienzsteigerungen hinsichtlich Energieverbrauch

Um die Dynamik und das Potenzial der Effizienzverbesserungen, gemessen am PUE, besser einordnen zu können, ist ein historischer Blick sowie die Betrachtung der PUE-Werte unterschiedlicher Größen von Rechenzentren aufschlussreich.

Das Uptime Institut hat im Zuge seiner jährlich stattfindenden globalen Rechenzentrumsbefragung umfassende Daten zu PUE-Werten gesammelt. Die nachfolgende Grafik stellt dazu den Verlauf des durchschnittlichen PUE des größten Rechenzentrums der Befragten (n = 526) von 2007 bis 2024 dar. Von 2007 bis 2014 sank der PUE um 34 %. In den 7 darauffolgenden Jahren wiederum hat der PUE keine vergleichbare Reduktion erfahren. Vielmehr stagniert der Wert im Bereich von 1,56 (für die Einordnung PUE und EnEfG siehe auch Kapitel 2.2. S. 42 sowie Abbildung 17). Der Grund für die schnelle Abnahme zu Beginn ist das Resultat der Ausnutzung von kostengünstigen Effizienzmaßnahmen, wie die Kalt- und Warmgangeinhausung, optimierte Kühlungssteuerung und eine Erhöhung der Zulufttemperatur. Zusätzliche Energieeinsparungen, insbesondere in Bestandsrechenzentren, sind mit einem erhöhten technischen und ökonomischen Aufwand verbunden.

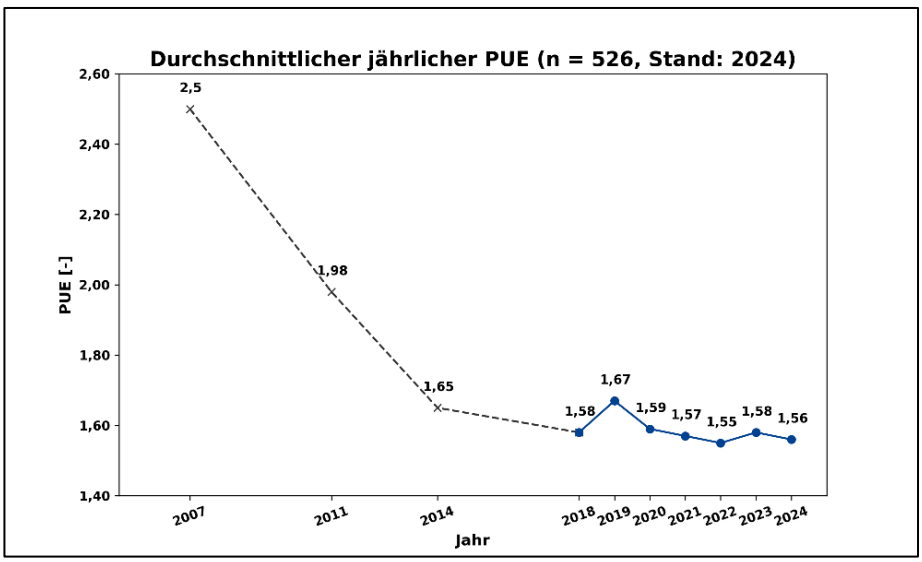


Abbildung 32: Weltweite durchschnittliche jährliche PUE-Werte in den Jahren 2007 bis 2024 (Uptime Institute, 2024)

Im Jahr 2023 (n = 558) differenzierte das Uptime Institut den PUE zusätzlich nach der IT-Kapazität der Rechenzentren. Mit zunehmender Leistungsgröße sinkt der PUE bis zu einem Grenzwert in Höhe von 1,44. Prognosen zur Entwicklung des PUE in Zukunft sind in Kapitel 2.2 (Abbildung 17) dargestellt. Ab einer Leistung von 20 MW zeigten sich keine zusätzlichen Energieeffizienzgewinne, da nicht mehr die Größe der Anlagen, sondern deren Anzahl erhöht wird, es treten also keine weiteren Skaleneffekte auf. Die ermittelten Unterschiede zwischen den einzelnen Leistungsklassen liegen im Bereich von 1 bis 4 %. Größere Rechenzentren sind in der Regel effizienter. Sie sind meistens neu und verwenden modernste Geräte mit effizienteren Kühlkonzepten und optimierten Steuerungen. Bei der Modernisierung kleinerer Anlagen ist es dagegen weniger wahrscheinlich, dass sich die Investition in Energieeinsparungen auszahlt. Daneben sind in der Regel größere technische Anlagen (z. B. Kältemaschine, Pumpen etc.) auch effizienter, da sie im Vergleich zu kleineren technischen Anlagen oftmals sorgfältiger optimiert sind.

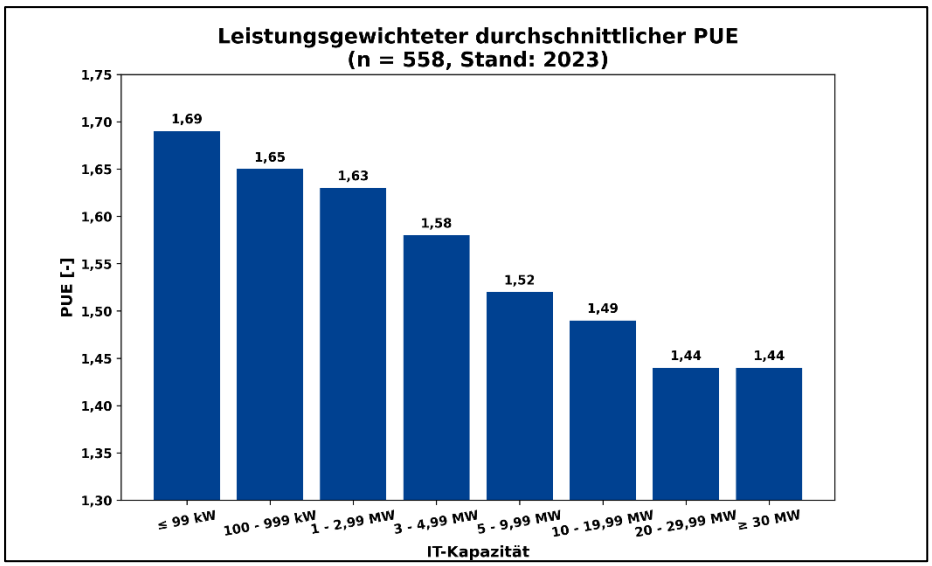


Abbildung 33: Weltweite gewichtete durchschnittliche PUE-Werte nach IT-Kapazität in kW/ MW (Uptime Institute, 2024)

Problematik der Messung von Energieeffizienz in Rechenzentren

Während PUE als Metrik für die Energieeffizienz von Rechenzentren weit verbreitet ist, stößt sie in der Praxis auf Herausforderungen. Der **Begriff der Energieeffizienz, der klassisch als Verhältnis von nutzbarem Output zu Energieinput definiert ist**, lässt sich in Rechenzentren schwer anwenden. Während der Energieinput gut messbar ist, ist der nutzbare **Output** – wie z. B. die Rechenleistung, der Speicherplatz oder der Netzwerkdurchsatz – deutlich **schwerer zu definieren und zu messen**. Rechenzentren erbringen eine Vielzahl von Diensten, die von der Speicherung und Verarbeitung von Daten über das Hosten von Websites bis hin zu Cloud-Computing-Diensten reichen. Zudem haben unterschiedliche Anwendungen unterschiedliche Anforderungen an Leistung, Latenz, Speicherbedarf und Ausfallsicherheit. Ein Rechenzentrum, das für hochsensible Finanztransaktionen optimiert ist, hat einen völlig anderen „Output" als eines, das für das Hosting von Websites ausgelegt ist.

Vor- und Nachteile des PUE

Obwohl PUE eine der am häufigsten verwendeten Metriken ist, misst es nicht die Gesamtenergieeffizienz im ingenieurwissenschaftlichen Sinne. Vielmehr hat sich PUE als einfacher und praktischer Ansatz etabliert, um zumindest die **Effizienz der Energieverteilung** im Rechenzentrum zu bewerten, anstatt zu versuchen, eine allgemeine Energieeffizienz zu definieren.

Die **Einfachheit und Verständlichkeit** des PUE ist einer seiner größten Vorteile. Die Berechnung erfordert nur zwei Werte: den Gesamtenergieverbrauch des Rechenzentrums und den Energieverbrauch der IT-Geräte. Diese Einfachheit hat dazu beigetragen, dass PUE sich als **international anerkannte Metrik** etabliert hat. Zudem ermöglicht der PUE Benchmarking innerhalb eines Rechenzentrums über die Zeit, was Einblicke in die Effizienzverbesserungen liefert.

Trotz seiner Vorteile weist PUE auch signifikante Schwächen auf. Der PUE **berücksichtigt nicht die Effizienz der IT-Hardware**. Ein Rechenzentrum kann eine gute PUE aufweisen, aber ineffiziente IT-Geräte betreiben, die viel Energie verbrauchen, ohne entsprechend viel Rechenleistung zu liefern. Der Fortschritt bei der IT-Energieeffizienz, insbesondere durch Moore’s Law, das ungefähr eine Verdopplung der Leistung pro Watt alle zwei Jahre verspricht, wird durch den PUE nicht direkt erfasst. Im Gegenteil, verringert sich der Stromverbrauch durch Effizienzverbesserungen in der IT, durch Nutzung von Flexibilitäten oder durch Abschalten von IT-Systemen in Zeiten niedriger Last, so verschlechtert sich der PUE. Ein entsprechender Nachteil ist, dass PUE **keinen umfassenden Vergleich der Energieeffizienz ermöglicht**. Die Metrik spiegelt nicht die Produktivität oder den tatsächlichen Output des Rechenzentrums wider, wie z. B. die Rechenleistung oder die Anzahl der verarbeiteten Daten. Zudem ist PUE **nicht immer einheitlich anwendbar**, da Unterschiede in Klima, Design und Betrieb von Rechenzentren zu verschiedenen Interpretationen führen können. Die Norm ISO/IEC 30134-2 legt fest, dass PUE nicht zum Vergleich von Rechenzentren untereinander verwendet werden sollte, was die Aussagekraft des Vergleichs zwischen verschiedenen Rechenzentren einschränkt.

Weitere Aspekte, die PUE nicht berücksichtigt, sind die IT-Auslastung, die Effizienz der Vor-Ort-Stromerzeugung, die Nutzung erneuerbarer Energien oder die Abwärmenutzung. Diese Faktoren spielen jedoch eine wichtige Rolle bei der Gesamtbewertung der Energieeffizienz eines Rechenzentrums.

4.3.2 Alternative Kennwerte

Im Laufe der Zeit wurden mehrere alternative Metriken entwickelt, um die Schwächen des PUE auszugleichen. Hierzu gehören die oben definierten Metriken der **Carbon Usage Effectiveness** (CUE) sowie der **Water Usage Effectiveness** (WUE). Diese Metriken liefern jeweils zentrale weitere Aspekte zur Betrachtung der Nachhaltigkeit von Rechenzentren. Weitere Alternativen, wie **Energy Reuse Effectiveness** (ERE) und **Green Energy Coefficient** (GEC), fördern die Wiederverwendung von Energie und den Einsatz erneuerbarer Energien. Diese Metriken haben ihre eigenen Vor- und Nachteile und ergänzen die PUE, bieten aber ebenso keine umfassende Lösung für die Bewertung der Energieeffizienz.

4.3.3 Fazit für zukünftige Kennwerte

PUE bietet einen praktikablen Ansatz, die Effizienz der Gebäudetechnik von Rechenzentren zu messen, ohne die schwer messbaren Details des Outputs berücksichtigen zu müssen. Allerdings bleibt das **zentrale Problem bestehen, eine robuste und praktische Metrik für die Produktivität oder den „Output" eines Rechenzentrums zu finden**.

4.4 Internationale gute Beispiele politisch-rechtlicher Ansätze

4.4.1 Überörtliche Bedarfsplanung und Standortwahl

In den Niederlanden wurde 2019 die „Ruimtelijke Strategie Datacenters" (REOS, 2019) etabliert, die über die Raumplanung „Ruimtelijke Ordening" eine koordinierte Standortwahl für Rechenzentren sicherstellt. Dabei werden Ansiedlungsprojekte gezielt in Regionen mit ausreichender Netz- und Energieinfrastruktur gefördert und weniger geeignete Standorte bewusst ausgeschlossen. Ziel ist es, eine effiziente Flächennutzung zu erreichen, Nutzungskonflikte zu vermeiden und energetische Synergien zu nutzen. Ein ähnlicher Ansatz wird in Schweden verfolgt, wo in enger Zusammenarbeit zwischen staatlichen Stellen und EVU geeignete Standorte mit Zugang zu kostengünstiger und CO_2-neutraler Energie, insbesondere aus Wasserkraft, identifiziert und entwickelt werden. Beispielsweise hat Microsoft in Schweden Rechenzentren in Gävle, Sandviken und Staffanstorp errichtet (REOS, 2019), die vollständig mit erneuerbarer Energie betrieben werden. Diese Projekte werden von lokalen und nationalen Behörden unterstützt, um die Infrastruktur und die Rahmenbedingungen für Rechenzentren zu optimieren.

Eine überörtliche Bedarfsplanung, z. B. auf Länderebene, wäre prinzipiell auch für Deutschland denkbar. Durch zentral koordinierte Planungsstellen könnten die Standortwahl effizienter gesteuert und Rechenzentren gezielt in Regionen mit vorhandener Infrastruktur gelenkt werden. Allerdings erfordert eine solche Planung erhebliche Abstimmungsprozesse zwischen kommunalen und überregionalen Akteuren – schon auf Länderebene, erst recht auf Bundesebene. Auch ist eine enge Zusammenarbeit mit Netz- und Energiebetreibern erforderlich. Dies wäre zunächst erstmal mit Aufwand verbunden. Auch sollten dadurch keine Verzögerungen für die aktuelle Ansiedlung von Rechenzentren entstehen.

4.4.2 Finanzielle und fiskalische Anreize für nachhaltige Rechenzentren

Gezielte finanzielle und steuerliche Anreize könnten wichtige Hebel sein, um den Rechenzentrumsstandort Deutschland zu stärken, indem Rechenzentren zu energieeffizienteren Technologien motiviert werden und die strategische Nutzung von erneuerbaren Energien und Abwärme intensiviert wird. In Deutschland sind mehrere in Rechenzentren eingesetzte Technologien im Rahmen der Bundesförderung für Energie- und Ressourceneffizienz in der Wirtschaft (EEW) förderfähig.

In Irland bietet das „Accelerated Capital Allowances Program“[13] Rechenzentren steuerliche Anreize für Investitionen in energieeffiziente Technologien. Einen ähnlichen Ansatz verfolgt Großbritannien mit dem "Full Expensing"[14]-Programm, das es Unternehmen ermöglicht, 100 % der Investitionsausgaben für qualifizierte effiziente Anlagen und Maschinen im Jahr der Anschaffung abzuschreiben. Darüber hinaus gibt es das britische „Climate Change Agreement“ (CCA)[15], das Rechenzentren reduzierte Energiesteuern bietet, wenn sie konkrete Energie- und CO_2-Reduktionsziele erreichen.

4.4.3 Beschleunigte Genehmigungsverfahren

Die langen Planungs- und Genehmigungszeiten in Deutschland werden häufig als Investitionshemmnis betrachtet. Derzeit erstrecken sich Genehmigungsverfahren häufig über mehrere Jahre, da komplexe Umweltprüfungen, Brandschutzauflagen und Abstimmungen zwischen verschiedenen Behörden erforderlich sind. Insbesondere die Umweltverträglichkeitsprüfung, die Bauleitplanung und die Abstimmung mit EVU und Netzbetreibern führen häufig zu erheblichen Verzögerungen.

Irland hat mit dem „Strategic Infrastructure Act“[16] einen Rahmen geschaffen, der Rechenzentren als strategische Infrastruktur anerkennt. Dieses Gesetz ermöglicht es, Anträge für bestimmte strategische Infrastrukturprojekte direkt bei An Bord Pleanála, der irischen Planungsbehörde, einzureichen, anstatt den üblichen Weg über die lokalen Planungsbehörden zu gehen. Ziel ist es, die Genehmigungsverfahren für Projekte von strategischer Bedeutung zu beschleunigen und zu vereinfachen. Sie durchlaufen einen gestrafften Genehmigungsprozess mit festen Bearbeitungsfristen. Zentrale Anlaufstellen bündeln alle Schritte des Genehmigungsverfahrens, einschließlich der Umwelt- und Bauprüfungen. Daraus ergeben sich klare Fristen und eine zentrale Koordination, die das Verfahren für die Unternehmen vereinfacht und verkürzt. In den Niederlanden optimiert die „Ruimtelijke Ordening“[17] (Raumordnung) das Verfahren, indem große Infrastrukturprojekte eine Sondergenehmigung erhalten und auf überörtlicher Ebene in der Raumplanung priorisiert werden. Die Genehmigungsdauer verkürzt sich durch eine verbindliche Abstimmung auf nationaler Ebene und vorgefertigte Umwelt- und Planungsstudien, die eine schnelle Entscheidungsfindung ermöglichen.

Für Deutschland könnte ein vergleichbarer Ansatz in Form eines ‚Fast-Track‘ umgesetzt werden. Dieser könnte Rechenzentren als kritische Infrastruktur klassifizieren, die zentrale Genehmigungsverfahren durchlaufen und standardisierte Prüfungen nutzen. Ein zentrales Verfahren, bei dem Umweltverträglichkeitsprüfung und Baugenehmigung parallel und nicht sequenziell geprüft werden, könnte Zeit sparen. Zudem könnten vorgefertigte Gutachten, ähnlich den niederländischen Standards, den Prüfungsumfang reduzieren und eine sofortige Anwendung ermöglichen. Nachteile könnten in der Einschränkung der kommunalen Planungsautonomie und im zusätzlichen Ressourcenbedarf für eine zentrale Anlaufstelle liegen.

[13] https://www.neh.gov.ie/business-supports/accelerated-capital-allowances
[14] https://www.gov.uk/government/publications/full-expensing/spring-budget-2023-full-expensing
[15] https://www.gov.uk/guidance/climate-change-agreements--2
[16] https://www.irishstatutebook.ie/eli/2024/act/16/enacted/en/html
[17] https://www.rijksoverheid.nl/onderwerpen/ruimtelijke-ordening-en-gebiedsontwikkeling

4.4.4 Offizielle Empfehlungen und systematische Erfassung bewährter Praktiken

Die systematische Erfassung und Veröffentlichung von bewährten Praktiken und Empfehlungen zur Energieeffizienz bzw. Nachhaltigkeit von Rechenzentren kann eine wichtige Rolle spielen, um Betreiber für nachhaltige und wirtschaftlich sinnvolle Betriebsweisen zu sensibilisieren. In der EU zielt der Code of Conduct for Data Centre Energy Efficiency, eine freiwillige Initiative der Europäischen Kommission, darauf ab, den Energieverbrauch und die Energieeffizienz von Rechenzentren zu verbessern. Der Code bietet Empfehlungen, Best Practices und Benchmarks, um Rechenzentren energieeffizienter zu gestalten, ohne die Funktionalität oder Leistung zu beeinträchtigen.

Ein internationales Beispiel ist das Center of Expertise for Energy Efficiency in Data Centers[18] des Lawrence Berkeley National Laboratory in den USA. Dieses vom US-Energieministerium geförderte Forschungszentrum stellt Betreibern von Rechenzentren umfassende Ressourcen zur Verfügung, um die Energieeffizienz ihrer Einrichtungen zu verbessern. Zu den Angeboten gehören detaillierte Technologiekataloge[19], die verschiedene Energieeffizienzmaßnahmen systematisch erfassen und beschreiben, sowie ein „PUE Estimator“[20], mit dem Betreiber den Energieverbrauch in ihren Rechenzentren ermitteln und mögliche Einsparpotenziale identifizieren können. Der PUE-Rechner bietet zudem eine Kosten- und Einsparungsbewertung verschiedener Technologien und kann Betreibern damit helfen, gezielte Investitionsentscheidungen zu treffen.

Für Deutschland wäre die Einrichtung eines nationalen Kompetenzzentrums unter Federführung einer Bundesbehörde oder einer anderen fachlich ausgestatteten, aber neutral agierenden Institutionen sinnvoll. Ein solches Zentrum könnte Best-Practice-Beispiele zur Energieeffizienz systematisch dokumentieren und durch Schulungs- und Beratungsangebote zugänglich machen. Ein möglicher Nachteil einer zentralen Anlaufstelle liegt in den Initialkosten und dem Bedarf an Fachpersonal für die kontinuierliche Aktualisierung und Pflege der Angebote des Kompetenzzentrums.

4.4.5 Energieeffizienzanforderungen für Rechenzentrumsinfrastruktur

In Deutschland gelten für die Infrastruktur von Rechenzentren mit den §§ 11 ff. EnEfG verbindliche Vorgaben zur Verbesserung der Energieeffizienz. Rechenzentren, die vor dem 1. Juli 2026 in Betrieb genommen werden, müssen bis Juli 2027 einen PUE von maximal 1,5 erreichen und diesen bis Juli 2030 auf maximal 1,3 senken. Für neue Rechenzentren, die nach dem 1. Juli 2026 in Betrieb genommen werden, gelten noch strengere Anforderungen mit einem maximalen PUE von 1,2 und der Verpflichtung, mindestens 10 % der Abwärme mit Staffelung bis 20 % bis 2028 wiederzuverwenden.

Internationale Regelungen zeigen ähnliche Ansätze zur Festlegung von Mindesteffizienzstandards. In China legt der Three-Year Action Plan on New Data Centres (2021 - 2023) (Brocklehurst, 2024) strenge PUE-Vorgaben für große Rechenzentren mit einer Anschlussleistung von 7,5 MW_{el} oder mehr fest und fordert für neue Einrichtungen einen PUE von maximal 1,3. Darüber hinaus wurde mit dem Standard GB 40879-2021 (Liu, 2023) eine Vorgabe für die öffentliche Verwaltung eingeführt, die vorschreibt, dass Rechenzentren staatlicher Organisationen ab Juni 2023 einen PUE von maximal 1,4 erreichen müssen, der bis 2025 auf 1,3 gesenkt wird. China hat kürzlich den „Sonderaktionsplan für

[18] https://datacenters.lbl.gov/
[19] https://datacenters.lbl.gov/sites/default/files/DCProMasterList02112016.pdf
[20] https://datacenters.lbl.gov/dcpro

die grüne und kohlenstoffarme Entwicklung von Rechenzentren“ veröffentlicht. Dieser setzt klare Ziele für die kohlenstoffarme Weiterentwicklung von Rechenzentren. Der Plan sieht vor, dass bis 2025 die PUE von neu gebauten und erweiterten großen und sehr großen Rechenzentren auf 1,25 oder weniger gesenkt werden soll und dass die PUE von nationalen Rechenzentrumsprojekten 1,2 nicht überschreiten soll.

In Japan gibt es keine Mindeststandards, aber das Energieeinsparungsgesetz (Brocklehurst, 2024) verlangt von Betreibern, über PUE-Werte zu berichten und einen Zielwert von 1,4 anzustreben. Bei Nichteinhaltung folgen Beratungen, Inspektionen und möglicherweise öffentliche Sanktionen.

In den Niederlanden gibt es anstelle von PUE-Vorgaben eine Energieeinsparverpflichtung (energiebesparingsplicht)[21], die Rechenzentren mit einem Jahresverbrauch von mehr als 50 MWh_{el} zur Umsetzung von Energieeffizienzmaßnahmen verpflichtet, sofern diese eine Amortisationszeit von fünf Jahren oder weniger aufweisen und auf einer offiziellen Liste des Wirtschaftsministeriums stehen. Diese Regelung sieht auch eine Berichtspflicht alle vier Jahre vor, um die umgesetzten Maßnahmen zu dokumentieren (Informatieplicht energiebesparing)[22].

Trotz der Vielfalt internationaler Standards gibt es Kritik an der PUE-Metrik als Mindesteffizienzanforderung, da der PUE-Wert lediglich das Verhältnis zwischen Infrastruktur- und IT-Energieverbrauch widerspiegelt, aber keine Aussage über die technische Energieeffizienz oder den Nutzungsgrad und die Effizienz der IT-Infrastruktur trifft. Dieser Kritikpunkt wird sowohl in China als auch in Deutschland intensiv diskutiert, da die festgelegten PUE-Grenzwerte die Marktrealitäten und technologischen Möglichkeiten nicht immer adäquat abbilden und oftmals zu hoch angesetzt erscheinen. Auch die niederländische Energieeinsparverpflichtung zeigt Herausforderungen auf, indem sie auf pauschale Einsparmaßnahmen fokussiert und die spezifischen betrieblichen Anforderungen und Betriebsprofile von Rechenzentren nur bedingt berücksichtigen kann. Dies kann dazu führen, dass Effizienzmaßnahmen umgesetzt werden müssen, die für bestimmte Standorte oder Anwendungen technisch oder wirtschaftlich nicht praktikabel sind.

4.4.6 Energieeffizienzanforderungen für IT- und Netzwerkequipment

Server, Datenspeicher und die Netzwerkstromausrüstung zählen zu den wesentlichen Energieverbrauchern in Rechenzentren. Energieeffizienzstandards haben sich als wirksames Instrument erwiesen, um Betreiber zum Umstieg auf energieeffizientere Geräte zu bewegen und damit sowohl Energieverbrauch als auch Betriebskosten zu senken. Die Europäische Union hat bereits umfassende Anforderungen an die Energieeffizienz von Servern und Datenspeicherprodukten festgelegt, die auch für Deutschland verbindlich sind. Die im Rahmen der Ökodesign-Richtlinie entwickelte Verordnung (EU) 2019/424 (Europäische Kommission, 2019) legt seit März 2020 spezifische Anforderungen an die Effizienz von Netzteilen, die Leerlaufleistung und die Effizienz im aktiven Zustand von Servern und Speichergeräten fest.

Ein Bereich ohne spezifische EU-Regelung ist die USV von Rechenzentren, obwohl 2014 hierzu eine Vorstudie der Europäischen Kommission (Boulos at al, 2014).) veröffentlicht wurde. Konkrete Vorschläge für eine Ökodesign-Verordnung liegen jedoch noch nicht vor. Ähnlich sieht es bei anderen

[21] https://english.rvo.nl/topics/energy-saving-obligation/what-energy-saving-obligation
[22] https://english.rvo.nl/topics/energy-saving-obligation/energy-saving-notification-obligation

Komponenten, wie Netzwerk-Switches und -Routern, Power Distribution Units (PDUs) und Batteriesystemen zur Notstromversorgung, aus. Regulierung könnte hier effektiv zu Effizienzsteigerungen führen. Gründe dafür, dass diese Komponenten bisher nicht in das Ökodesign einbezogen wurden, sind die technische Komplexität und die unterschiedlichen Betriebsanforderungen, die eine Standardisierung auf EU-Ebene erschweren.

Da nationale Anforderungen aufgrund des EU-Binnenmarktes wenig sinnvoll wären, liegt die Verantwortung für eine weitergehende Regulierung bei der Europäischen Kommission und ihrem Ökodesign- und Labelling-Arbeitsprogramm. Eine umfassendere Betrachtung weiterer IT-Komponenten könnte das bestehende Programm sinnvoll ergänzen und zusätzliche Effizienzpotenziale in Rechenzentren freisetzen. Im internationalen Vergleich gibt es nur wenige Benchmarks für eine solche Regulierung der Energieeffizienz in Rechenzentren. Japan hat zwar begonnen, Mindesteffizienzstandards für die Energieeffizienz von Servern und Computern (METI, 2019) zu entwickeln, aktuelle Informationen über den Stand dieser Initiative waren den Bearbeitenden des vorliegenden Gutachtens nicht zugänglich.

4.4.7 Abwärmenutzung und -integration in Wärmenetze

In der Praxis gibt es in Deutschland zahlreiche Hemmnisse, die eine effektive Integration von Abwärme in Wärmenetze erschweren. Dazu zählen die hohen Investitionsausgaben für Wärmepumpen, die notwendig sind, um die oft niedrigen Abwärmetemperaturen auf ein nutzbares Niveau anzuheben, sowie die Kosten für den Ausbau der Fernwärmeinfrastruktur, die zwischen Rechenzentrums- und Wärmenetzbetreibern aufgeteilt werden müssen.

Dänemark hat 2022 eine Gesetzgebung eingeführt, die zwei zentrale Mechanismen enthält, um die Nutzung von Abwärme aus Rechenzentren zu stärken (Jerez Monsalves et al, 2022). Zum einen gibt es eine regulierte Preisobergrenze für Abwärme, die in das Fernwärmenetz eingespeist wird. Diese Preisobergrenze wird von der nationalen Regulierungsbehörde Forsyningstilsynet festgelegt und basiert auf den Produktionskosten der günstigsten alternativen Wärmequelle, z. B. Biomasse oder Solarthermie. Dieser Ansatz soll die Wärmenetzbetreiber motivieren, Abwärme als attraktive und kosteneffiziente Wärmequelle zu betrachten: Ist der ausgehandelte Preis geringer als die Preisobergrenze, wird die Abwärme für das Wärmenetz wettbewerbsfähig. Gleichzeitig stellt die Preisgrenze sicher, dass Abwärme nur dann genutzt wird, wenn sie im Vergleich zu anderen Energiequellen wirtschaftlich konkurrenzfähig bleibt. Der zweite Anreizmechanismus ist eine Steuerbefreiung für Abwärme aus Stromnutzung, die wiederum konkrete finanzielle Anreize für die Bereitstellung von Abwärme schafft.

Die dänischen Maßnahmen bieten eine vielversprechende Perspektive für Deutschland. Ein weiteres Hemmnis für die Abwärmenutzung sind die hohen Transaktionskosten für Verhandlungen und Abstimmungen zwischen Rechenzentrumsbetreibern und Fernwärmeversorgern. Hier zeigt das Beispiel Dänemark, wie gezielte Maßnahmen zur Stärkung des Marktes für Intermediäre, also spezialisierte Drittanbieter, die Abwärme einkaufen und an Wärmenetzbetreiber weiterverkaufen, Abhilfe schaffen können.

4.4.8 Flexibilitätsbereitstellung durch Rechenzentren

Das Thema nachfrageseitige Flexibilität (Demand Side Flexibility) ist für Rechenzentren aus zwei zentralen Gründen relevant. Erstens können Rechenzentren durch Lastverschiebung ihre Energieverbräuche an Zeiten hoher erneuerbarer Energieeinspeisung oder geringer Stromnetzlast anpassen. Prozesse, die keine Echtzeitbearbeitung erfordern, wie Batch-Verarbeitungen oder Datensicherungen, können flexibel geplant werden, um Netzspitzen zu vermeiden und die Nutzung erneuerbarer Energien zu optimieren. Zweitens können Rechenzentren aktiv Systemdienstleistungen bereitstellen, wie Frequenzregelung oder Spannungsstabilisierung. Über ihre Energiesysteme, etwa Batteriespeicher (z. B. aus USV), könnten Rechenzentren kurzfristig Energie ins Netz einspeisen oder aufnehmen. Darüber hinaus sind auch ihre Kühl- und Lüftungssysteme potenziell geeignet, positive Regelenergie bereitzustellen.

Im europäischen Vergleich entwickelt sich Deutschland zu einem wichtigen Markt für nachfrageseitige Flexibilität (Murley, Ferris, 2024). Im Bereich der Systemdienstleistungen sind regulatorische Fortschritte, wie die Zulassung von Aggregatoren und Gebotsgrößen unter 1 MW, mit Vorreitern wie Frankreich, den Niederlanden und Schweden vergleichbar. Jedoch besteht im Bereich der Verteilnetzflexibilität weiterhin Handlungsbedarf. Dies betrifft beispielsweise dynamische Netzentgelte (indirektes Demand Response) oder Ausschreibungen für feste Flexibilität (explizites Demand Response).

Ein ambitioniertes, aber kontroverses Beispiel liefert der irische Netzbetreiber EirGrid: Neue Rechenzentren in der Region Dublin müssen in der Lage sein, ihre Stromaufnahme auf Signal des Netzbetreibers zu reduzieren (EirGrid, 2020). In der Praxis setzen Betreiber hierfür häufig auf Gasturbinen vor Ort, die bei Bedarf hochgefahren werden.

4.5 Handlungsempfehlungen

Ein möglicher Ansatz zur Förderung des Ziels des EnEfG, von Rechenzentren produzierte Abwärme zu nutzen und damit einen Beitrag für Klimaneutralität zu leisten, wäre, das Delta zwischen Abwärmemengen von Rechenzentren und der Menge der Abnahme an Abwärme zu verkleinern. Während auf der Abgabeseite, also bei den Rechenzentren, unter bestimmten Voraussetzungen die Pflicht zur Abwärmenutzung und damit auch zur Abgabe der Wärme besteht und diese Seite daher reguliert ist, bestehen auf der Abnahmeseite noch Potenziale, um die Einbindung in Wärmenetze zu erleichtern. Deren Fehlen kann auch potenziell durch das teilweise lokale Vorhandensein monopolartiger Strukturen in Bezug auf Wärmenetze und eine entsprechende Einspeisung (vgl. auch Kapitel 2) nicht allein durch (wettbewerbliche) Anreize ausgeglichen werden. Um zu verhindern, dass die Verpflichtung aus dem EnEfG zur Abwärmenutzung durch Rechenzentren nicht ins Leere läuft, wäre es sinnvoll, eine regelmäßige Wiedervorlage eines vorher abgelehnten Angebots zur Nutzung von Abwärme gegenüber Betreibern eines in der Umgebung befindlichen Wärmenetzes vorzusehen und damit den Ausnahmefall in Bezug auf die Pflicht zur Abwärmenutzung nicht zu perpetuieren. Weiterhin wäre es sinnvoll, die Kostenverantwortung im Bereich der Abwärmenutzung klarer zu definieren und hier die Kostenlast von Rechenzentrumsbetreibern einerseits und etwa Wärmenetzbetreibern andererseits klar voneinander abzugrenzen. Allerdings wird die wirtschaftliche Kalkulation von Rechenzentren überwiegend als Betriebsgeheimnis behandelt. Insbesondere bieten

sich normkonkretisierende Verwaltungsvorschriften an – eine entsprechende Anpassung der bereits bestehenden Verordnungsermächtigungen im EnEfG vorausgesetzt.

Mit Blick auf die kommunale Bauleitplanung und die langwierige „Schaffung von Baurecht“ in den jeweiligen Kommunen würde die Aufnahme von Rechenzentren in den Kanon bauplanungsrechtlich privilegierter Außenbereichsvorhaben nach § 35 Abs. 1 BauGB den Prozess beschleunigen. Weiterhin wäre mit Blick auf die Anschlusskonkurrenz für das Stromnetz die Etablierung eines verbindlichen Reservierungsverfahrens für den Netzanschluss von Rechenzentren begrüßenswert, vergleichbar mit der Regelung in § 4 KraftNAV, um hier mehr Planungssicherheit für Rechenzentrumsbetreiber zu gewährleisten.

Internationale politisch-regulatorische Ansätze bieten potenzielle Erfolgsmodelle für den deutschen Rechenzentrumsmarkt. Um die Energieeffizienz von Rechenzentren zu fördern, könnten auf EU-Ebene Ökodesign-Standards für bisher nicht regulierte Komponenten, wie Netzwerk-Switches und Batteriesysteme, eingeführt werden. Alternativ zu PUE-Grenzwerten könnte die Einführung von Energieeinsparverpflichtungen nach niederländischem Vorbild Rechenzentren verpflichten, Maßnahmen mit einer Amortisationszeit von beispielsweise maximal fünf Jahren umzusetzen und zu melden. Parallel dazu könnte eine zentrale Anlaufstelle für beschleunigte Genehmigungsverfahren eingerichtet werden, die Planungsprozesse standardisiert und durch Fristen strafft. Steuerliche bzw. finanzielle Anreize, wie in Irland und Großbritannien, könnten die Betreiber bei der Umsetzung energieeffizienter Technologien und der Abwärmenutzung unterstützen. Um speziell die Abwärmenutzung zu erleichtern, wäre wie in Dänemark die Förderung von Marktintermediären sinnvoll, die Transaktionskosten senken und Betreiber bei der Infrastrukturplanung unterstützen. Schließlich könnte die Einrichtung eines Kompetenzzentrums für Energieeffizienz in Rechenzentren, ähnlich wie in den USA, Best Practices systematisch dokumentieren und Schulungs- und Beratungsangebote zur Verfügung stellen.

5 Blick in die Zukunft der Rechenzentren in Deutschland

5.1 Zusammenfassung

Im Rahmen des Gutachtens wird primär der Ausbau von Rechenzentren als Möglichkeit, das Angebot an digitalen Ressourcen in Deutschland zu erhöhen, betrachtet. Die Nachfrageseite wird in Kapitel 5 im Rahmen von Zukunftsszenarien modelliert und für das Guthaben als Vorhersagewerkzeug für möglichen Bedarf an digitalen Ressourcen und den damit verbundenen Rechenzentrums- und Energiebedarf genutzt. Ziel des Gutachtens ist es dabei nicht, über die bereits bestehenden politischen Rahmenbedingungen hinweg neue Annahmen zum Grad der Digitalisierung in Deutschland oder dem Wachstum der Digitalwirtschaft zu treffen, sondern die bestehenden Trends in eine Vorhersage zum Bedarf an digitalen Ressourcen in Deutschland umzurechnen.

Die Vorhersagen im Gutachten sind über die Nachfrage der digitalen Ressourcen modelliert. Für die Herstellung von digitalen Ressourcen werden wiederum IKT-Systeme (Server-, Speicher- und Netzwerksysteme) benötigt. Diese IKT-Systeme sind der primäre Faktor bei der Vorhersage von Strom- und Flächenbedarf.

Die Nachfrage nach digitalen Ressourcen lässt sich mindestens über zwei Wege decken. Der erste Weg ist der Ausbau von digitaler Infrastruktur in Deutschland. Damit einhergehend müssen dafür Flächen und Strominfrastruktur bereitgestellt werden. Dies ist insbesondere im Kontext der Souveränität bzw. strategischer Autonomie von Bedeutung und wird im folgenden Kapitel genauer betrachtet.

Der zweite Weg ist der Import von digitalen Ressourcen aus dem europäischen Ausland. Einige europäische Länder, wie Schweden, Norwegen und Finnland, haben für den Wettbewerb im Markt der digitalen Ressourcen teilweise bessere Ausgangsbedingungen, z. B. verfügbare Flächen und erneuerbare Energien. Die konkrete Wettbewerbssituation wurde im Kapitel 4 genauer untersucht. Ein in Summe erheblicher Import von digitalen Ressourcen aus Märkten außerhalb von Europa ist aus technologischer Sicht bis 2030 unwahrscheinlich, da die Distanzen in den Daten-Übertragungsnetzen zu Latenzen führen, welche sich negativ auf einen Großteil von Softwareanwendungen auswirken würden.

Die Bedeutung von Rechenzentren in der globalen Wirtschaft und Gesellschaft wächst kontinuierlich. Aufgrund der fortschreitenden Digitalisierung, und angetrieben durch die steigende Nachfrage nach Cloud-Diensten und Anwendungen im Bereich der Künstlichen Intelligenz, wird für die kommenden Jahre ein sehr deutliches Wachstum der weltweiten Kapazitäten und Investitionen in Rechenzentren prognostiziert.

5.2 Digitale Dimensionen bei Wirtschaftswachstum und Energiebedarf

Geht man davon aus, das Deutschland seine eigene und die Digitalisierungsstrategie der EU verfolgt, ist mit einem hohen Anstieg der Nachfrage nach digitalen Ressourcen und damit auch nach digitaler Infrastruktur zu rechnen.

Insbesondere die Digitalisierung von Handel, Gesundheitswesen und dem verarbeitenden Gewerbe ist ausschlaggebend für das Nachfragewachstum an digitalen Ressourcen. Technologische Innovationen wie KI sorgen zudem für eine erhöhte Nachfrage nach Datenverarbeitung und Speicherung im Kontext der Digitalisierung.

Im Digitalsektor ist zudem durch stagnierende oder rückläufige Trends in einigen Industriebranchen zunehmend mehr Wachstum zu erwarten, indem Wirtschaftsleistung in diesen neuen Sektor verschoben wird.

Für die Digitalwirtschaft sind neben Fachkräften Verfügbarkeit und Preis von digitalen Ressourcen zentrale Faktoren für den Erfolg der digitalen Produkte und Dienstleistungen. Dies unterscheidet die Digitalwirtschaft von anderen, traditionellen Branchen, die primär andere Rohstoffe oder Handelsgüter verarbeiten, veredeln bzw. herstellen. Durch die hohe Abhängigkeit von digitalen Ressourcen bedingt ein Wachstum der Digitalwirtschaft auch einen starken Anstieg der Nachfrage nach digitalen Ressourcen.

Um den Bedarf an digitalen Ressourcen bis zum Jahr 2045 abschätzen zu können, wurden im Rahmen der Studie eine Modellierung der Digitalisierungsintensität in Deutschland vorgenommen und das voraussichtliche Wachstum der Digitalwirtschaft abgeschätzt. Im Folgenden finden sich die Ergebnisse dieser Modellierung und relevante Hintergrundinformationen zu den untersuchten Szenarien.

5.2.1 Digitalisierungsgrad der Gesamtwirtschaft

Bis 2045 könnte Deutschland einen durchschnittlichen Digitalisierungsgrad der Gesamtwirtschaft von 42 % erreichen (2024 lag er schätzungsweise bei 20 %).

Modellierungsansatz

Die Grundlage der Modellierung sind die folgenden Datensätze, welche vom Statistischen Bundesamt bezogen wurden:

- Datensatz „Statistisches Unternehmensregister - Rechtliche Einheiten, Beschäftigte und Umsatz im Berichtsjahr 2022” vom Statistischen Bundesamt (Statistisches Bundesamt, 2024d), aus dem die Beschäftigungszahlen und Anzahl an Unternehmen in den verschiedenen Wirtschaftsabschnitten übernommen werden.
- Die Anzahl der Unternehmen wird in die Gruppen Klein- und Kleinstunternehmen (Klein), mittlere Unternehmen (Mittel) und Großunternehmen (Groß) unterteilt. Dies wird, basierend auf dem Datensatz des Statistischen Bundesamtes (Statistisches Bundesamt, 2024b), welcher nur die Gesamtzahl der Unternehmen enthält, vorgenommen, um eine Verteilung zu berechnen. Bei der Verteilung auf Klein-, Mittel- und Großunternehmen wird für die Wirtschaftsbereiche Bergbau und Gewinnung von Steinen und Erden, Verarbeitendes Gewerbe, Energieversorgung, Wasserversorgung, Entsorgung, Beseitigung von Umweltverschmutzung, Erbringung von Finanz- und Versicherungsleistungen, Gesundheits- und Sozialwesen ein höherer Anteil von Großunternehmen auf Basis von eigenen Recherchen angenommen.

- Die Anzahl der Beschäftigten wurde auf die Unternehmensgrößen verteilt (siehe Statistisches Bundesamt[23]).

Auf Basis dieser Grundsatzdaten wurde der Digitalisierungsgrad der Unternehmen in jedem Wirtschaftsabschnitt bestimmt. Der Digitalisierungsgrad wurde über 3 Faktoren bestimmt:

- **Digitalisierung der Arbeitsplätze:** Anteil der Beschäftigten, die über einen digitalen Arbeitsplatz bzw. digitale Arbeitswerkzeuge (Internet-basierte Software) verfügen.
- **Digitalisierung der Unternehmensprozesse:** Anteil der Prozesse, die mit Hilfe von digitalen Prozess-Werkzeugen, Plattformen oder sonstiger Software digitalisiert wurden. Die Menge der Prozesse pro Unternehmen wurde abhängig von der Größenklasse des Unternehmens modelliert.
- **Digitalisierung der Produkte und Dienstleistungen:** Anteil der Endprodukte und Dienstleistungen, die vom Unternehmen angeboten werden und teilweise oder vollständig digital ausgeliefert werden (Beispiel: Onlinebanking, virtuelle Beratung, etc.).

Der Grad der Prozess-, Arbeitsplatz- und Produkt-Digitalisierung wurde für jeden Wirtschaftsabschnitt und jedes Jahr bis 2045 modelliert. Die Ausgangslage wurde auf Basis des Digitalisierungsindizes der Bundesregierung modelliert[24] und darauf basierend eine gesamtwirtschaftliche Durchschnittsdigitalisierung von 20 % für 2024 angesetzt. Dies wurde durch Umfragen validiert und deckt sich proportional mit dem Stromverbrauch aus Rechenzentren, die in Kapitel 1 erhoben wurden.

Annahmen und Abgrenzung

- Für die Modellierung des Digitalisierungsfortschritts der verschiedenen Wirtschaftsabschnitte wurden qualitative Umfragen im Rahmen der Studie durchgeführt und quantitative Vorhersagen aus Branchenanalysen einbezogen. Die Datenlage zu branchenspezifischen Digitalisierungsstrategien, Zielen und Fortschrittsdaten ist allerdings allgemein schlecht.
- Bei der Modellierung wurden Endgeräte, insbesondere für digitale Arbeitsplätze (Laptops, Desktop-Computer und Smartphones) nicht betrachtet, sondern lediglich die für die Arbeitsplätze notwendige Rechen- und Speicherkapazität („digitale Ressourcen") berechnet, da diese für die Abschätzung von Flächen- und Strombedarf von Rechenzentren ausschlaggebend sind.
- Netzwerkverkehr und die damit verbundene Infrastruktur und der damit verbundene Stromverbrauch wurden nicht betrachtet.
- Für das Modell wurde eine Standardserverkonfiguration („Standard-Server-Äquivalent") mit 128 GB Arbeitsspeicher, 32 TB Festplattenspeicher und 500 Watt Leistung verwendet (Durchschnittswerte aus den Einkaufsdaten von Server-Systemen in Deutschland aus der Modellierung in Kapitel 1)
- Des Weiteren wird angenommen, dass die für die Digitalisierung erforderlichen Fachkräfte verfügbar sind und dass sich das regulatorische Umfeld weitestgehend nicht verändert bzw.

[23] https://www.destatis.de/DE/Themen/Branchen-Unternehmen/Unternehmen/Kleine-Unternehmen-Mittlere-Unternehmen/aktuell-beschaeftigte.html
[24] https://www.de.digital/DIGITAL/Navigation/DE/Lagebild/Digitalisierungsindex/digitalisierungsindex.html

sich nicht auf die Geschwindigkeit der Digitalisierung in den verschiedenen Wirtschaftsabschnitten auswirkt.

- Für das Wirtschaftswachstum über alle Wirtschaftsabschnitte hinweg wurde im Durchschnitt mit 1,3 % Wachstum pro Jahr bis 2045 kalkuliert (ifo Institut, 2024).

Vergleich mit anderen Studien

Die bestehenden Studien im Markt wurden hinsichtlich ihrer Validität betrachtet (u. a. Bitkom). Jedoch gibt es keine langfristigen Vorhersagen für die Digitalisierung in Deutschland oder in relevanten anderen Ländern, die für einen Vergleich herangezogen werden konnten.

Entwicklung der Digitalisierung in den Wirtschaftsabschnitten - Energie, Gesundheitswesen, IKT und Finanzwirtschaft als Treiber

Im Rahmen der Analyse wurden die Digitalisierungsverläufe der verschiedenen Wirtschaftsabschnitte modelliert. Dabei wird davon ausgegangen, dass sich der Energiesektor bezogen auf die Energiewende schneller digitalisiert als andere Industriezweige. Auch kann davon ausgegangen werden, dass das Gesundheitswesen und die Finanzwirtschaft von einem hohen Grad der Digitalisierung profitieren. Daher ist anzunehmen, dass bis 2045 auch diese beiden Bereiche einen hohen Grad der Digitalisierung aufweisen. Wie auch im Digitalisierungsindex der Bundesregierung dargestellt, ist es realistisch anzunehmen, dass der IKT-Sektor selbst einen hohen Digitalisierungsgrad aufweisen wird.

In der folgenden Abbildung findet sich ein Überblick der verschiedenen Digitalisierungsverläufe der Wirtschaftsabschnitte vom Jahr 2024 bis hin zum Jahr 2045.

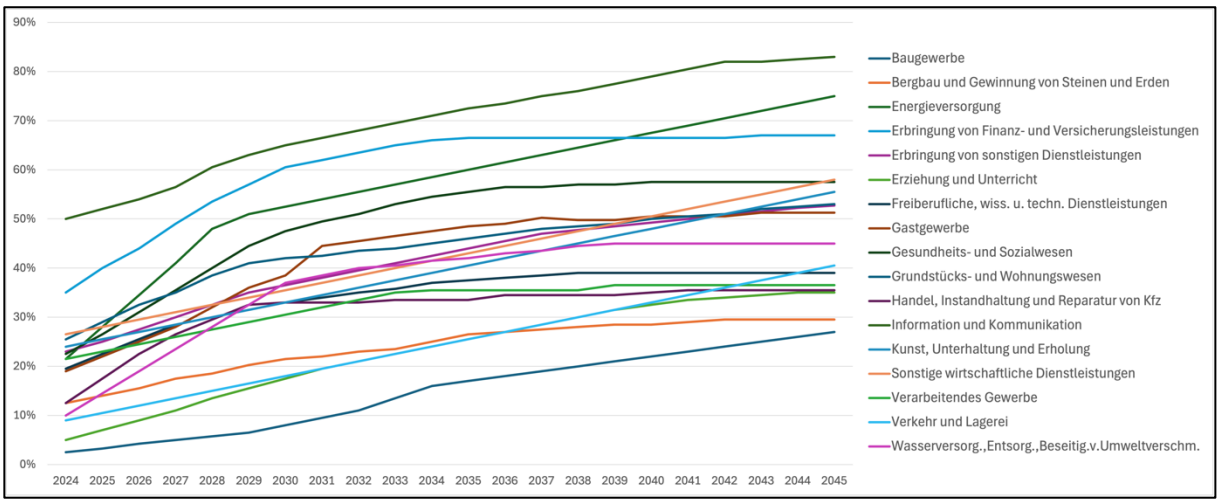


Abbildung 34: Verlauf des Digitalisierungsgrades der verschiedenen Wirtschaftsabschnitte in Deutschland von 2024 bis 2045. Quelle: Berechnungen der SDIA

Besonders hervorzuheben ist eine gewisse Sättigung ungefähr ab dem Jahr 2040, da viele Wirtschaftsabschnitte in diesem Zeitraum das realistisch anzunehmende Digitalisierungsmaximum der Arbeitsplätze, Prozesse und Produkte erreichen werden. Im Baugewerbe oder im Bergbau kann nicht davon ausgegangen werden, dass der Digitalisierungsgrad dieser Unternehmen mehr als 35 % überschreitet. Selbst im Falle einer hohen Digitalisierung der Arbeitsumgebungen sind die

fundamentalen Betriebsprozesse zum Großteil nicht digitalisierungsfähig. Selbst in Fällen, in denen roboterbasierte Automatisierung möglicherweise das Produkt- oder Dienstleistungsangebot eines Unternehmens digitalisieren kann, führt dies nur bedingt auch gleichzeitig zu einer Reduzierung von Arbeitsplätzen und hat damit somit nur einen begrenzten Einfluss auf den Gesamt-Digitalisierungsgrad.

Für die Gesamtwirtschaft wurde bis 2045 ein Digitalisierungsgrad von 39 % errechnet. Dabei wurden die digitalen Produkte und Dienstleistungen aus der Digitalwirtschaft nicht einbezogen. Sie werden in diesem Kapitel besonders betrachtet. Die Arbeitsplätze und Prozesse der Digitalwirtschaft wurden im Rahmen des Wirtschaftsabschnitts IKT mit einbezogen. Die möglichen Beschleunigungseffekte durch KI-Technologien wurden bis 2030 einbezogen.

Um den Digitalisierungsgrad von 39 % zu erreichen, benötigt Deutschland digitale Ressourcen, die 19 Millionen Standard-Server-Äquivalenten entsprechen.

Um die notwendigen digitalen Ressourcen bereitzustellen, müssen zwischen 2024 und 2045 zusätzlich rund 19 Millionen Standard-Server-Äquivalente im Wert von circa 70 bis 100 Milliarden Euro bereitgestellt werden. Dies ist in Abbildung 36 dargestellt. Geht man davon aus, dass 5 % der installierten Serverkapazität bis 2045 aus KI-Systemen besteht, ist mit einem zusätzlichen Investitionsaufwand von 300 Milliarden Euro zu rechnen (auf Basis aktueller Marktpreise).

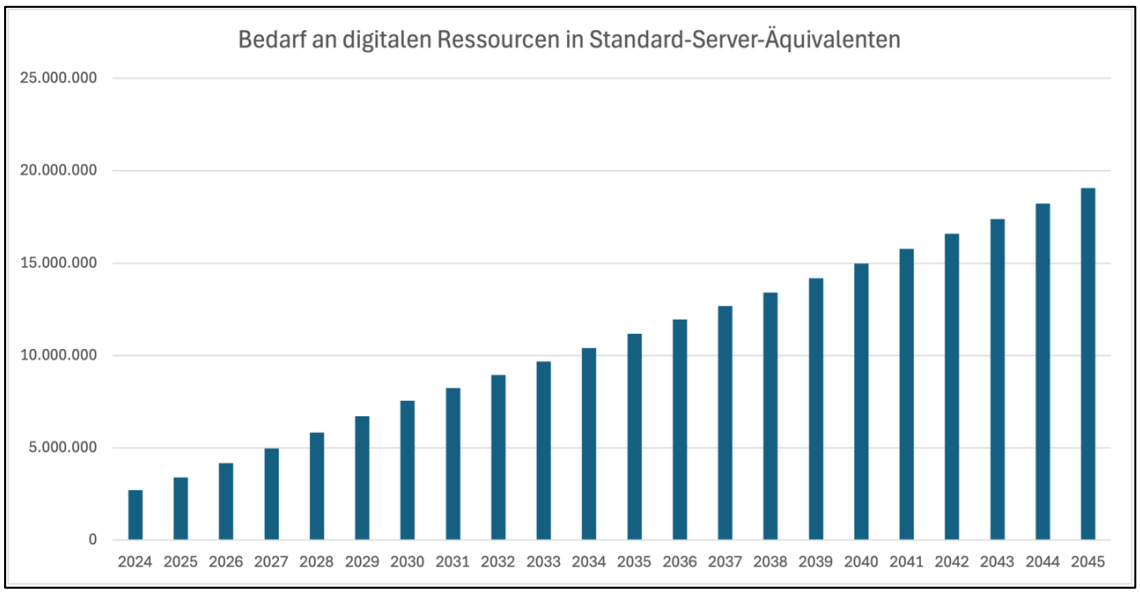


Abbildung 35: Berechneter Bedarf an digitalen Ressourcen (Standard-Server-Äquivalente) auf Basis der Schätzung des Digitalisierungsgrades der Wirtschaftsabschnitte. Quelle: Berechnungen SDIA

Es kann jedoch nicht davon ausgegangen werden, dass alle digitalen Ressourcen für die Digitalisierung auch aus Rechenzentren bzw. IT-Infrastruktur in Deutschland bezogen werden.

Geht man davon aus, dass auf 1 m^2 Rechenzentrumsfläche im Schnitt 30 Standard-Server-Äquivalente Platz finden (mit einer Leistungsdichte von 5 - 15 kW pro Rack), so wären im Jahr 2045 ungefähr 650.000 m^2 zusätzliche Rechenzentrumsfläche notwendig. Dies entspricht der Fläche von 91 Fußballfeldern.

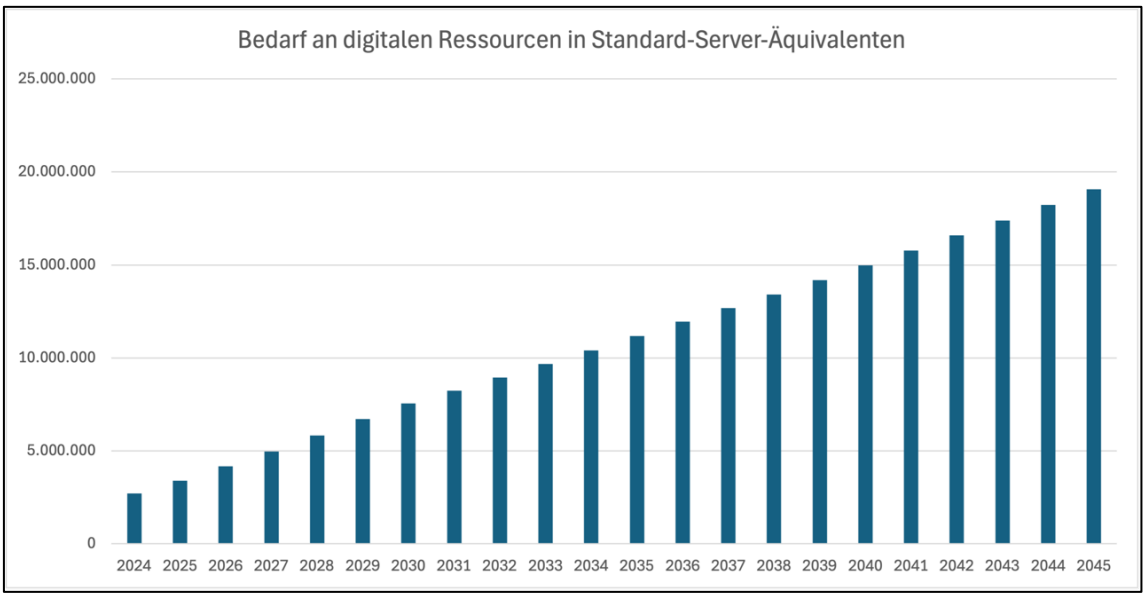


Abbildung 36: Prognose der Entwicklung des Bedarfs an digitalen Ressourcen (Standard-Server-Äquivalente) in den Jahren 2024 bis 2045 auf Basis der Schätzung des Digitalisierungsgrades der Wirtschaftsabschnitte. Quelle: Berechnungen SDIA

5.2.2 Einflüsse auf den Energiebedarf

Mithilfe des Modells konnten verschiedene Energieeffizienzmaßnahmen genauer betrachtet werden:

- **Baseline Szenario**: Ineffizienzen in der Software und der IT-Infrastruktur führen zu einer Bedarfssteigerung von 5 % Standard-Server-Äquivalenten pro Jahr.
- **Szenario 1:** Verbesserungen der Software und der IT-Infrastruktur führen dazu, dass der Bedarf an Standard-Server-Äquivalenten bei steigender Digitalisierung pro Jahr nur um 2 % steigt (die Software und IT-Infrastruktur wird weniger ineffizient).
- **Szenario 2**: Verbesserungen der Software und der IT-Infrastruktur führen dazu, dass der Bedarf an Standard-Server-Äquivalenten bei steigender Digitalisierung pro Jahr gleichbleibt.
- **Szenario 3**: Es wird angenommen, das die Software und IT-Infrastruktur von Jahr zu Jahr 5 % effizienter wird und damit ein geringerer Verbrauch von digitalen Ressourcen pro Anwendung entsteht.

In Abbildung 37 sind die verschiedenen Szenarien dargestellt, mit dem geschätzten Bedarf an Standard-Server-Äquivalenten pro Jahr. Hier wird deutlich, dass sich durch eine kontinuierliche Verbesserung von Software-Anwendungen hinsichtlich der Ressourceneffizienz ein Teil des Bedarfs an digitalen Ressourcen decken lässt. Insbesondere wenn die Digitalisierung der Wirtschaftsbereiche sich ab ca. 2040 sättigt, wirkt sich die Steigerung der Ressourceneffizienz von Software und IT-Infrastruktur positiv aus und kann den weiter steigenden Bedarf teilweise abfangen.

5.2.3 Entwicklung der Digitalwirtschaft

Wächst die Digitalwirtschaft bis 2045 mit durchschnittlich 10 % pro Jahr, so entstehen 8 % der Nachfrage nach digitalen Ressourcen aus diesem Wirtschaftsbereich.

Für die Modellierung der Nachfrage nach digitalen Ressourcen aus der Digitalwirtschaft fehlen Informationen zum Umsatz der Digitalwirtschaft sowie Informationen zu detaillierten Beschäftigungszahlen. Daher wurde für eine grobe Schätzung das investierte Risikokapital als Indikator verwendet.

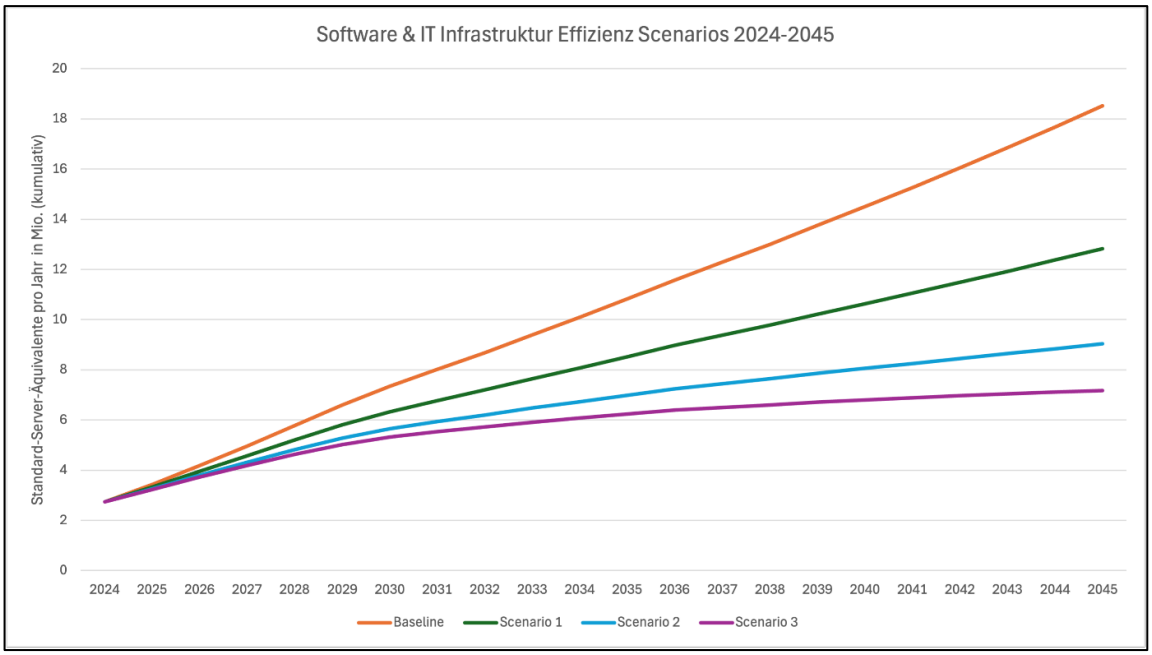


Abbildung 37: Prognose der Entwicklung der Unterschiede des Bedarfs an Standard-Server-Äquivalenten durch verschiedene Effizienzmaßnahmen in den Jahren 2024 bis 2045. Quelle: Berechnungen SDIA

Grundlage für die Zahlen zu Kapitalinvestitionen in Start-ups in Deutschland bietet das EY-Startup-Barometer vom Januar 2024[25]. Für die Berechnung wurde angenommen, dass die Digitalwirtschaft im Durchschnitt 10 % mehr Investorenkapital erhält. Weiterhin wurde angenommen, dass aus diesem Kapital 15 % in IT-Infrastruktur bzw. in den Einkauf von digitalen Ressourcen fließen.

Mit Hilfe der Kennzahl zu den potenziellen Ausgaben für IT-Infrastruktur wurde die Menge der Standard-Server-Äquivalente berechnet. Das Ergebnis ist eine parallele Entwicklung zur Digitalisierung der anderen Wirtschaftszweige, bei der die Digitalwirtschaft ca. 7 bis 9 % des digitalen Ressourcenbedarfs **zusätzlich** verursacht. Um den Ressourcenbedarf der Digitalwirtschaft im Jahr 2045 zu decken, werden Standard-Server-Äquivalente, berechnet auf aktuellen Marktpreisen, im Wert von 7 bis 8 Milliarden Euro benötigt.

[25] https://startup.ey.com/en/news/ey-startup-barometer-germany-january-2024

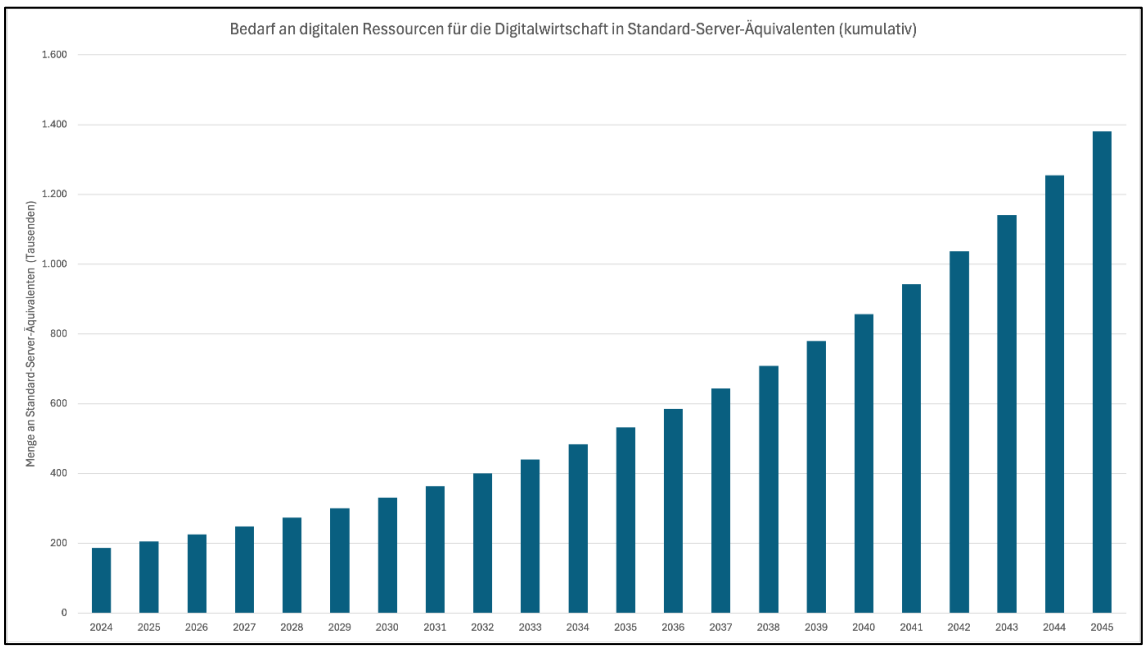


Abbildung 38: Prognose der Entwicklung der Nachfrage an Standard-Server-Äquivalenten in den Jahren 2024 bis 2045, die aus dem vorhergesagten Wachstum der Digitalwirtschaft bis 2045 entstehen könnten. Quelle: Berechnungen SDIA

Es ist anzumerken, dass ein Großteil der Digitalwirtschaft digitale Ressourcen und IT-Infrastruktur aus geschlossenen Cloud-Märkten bezieht. Diese könnten möglicherweise die Bereitstellung auch aus dem Ausland oder Europa realisieren.

5.3 Digitale Souveränität

Die digitale Souveränität Deutschlands gewinnt im Kontext der fortschreitenden Digitalisierung und der zunehmenden Bedeutung von Daten, Rechenleistung, Speicher- und Datenübertragungskapazität als strategische Ressource immer mehr an Bedeutung. Rechenzentren, IT-Infrastruktur und IT-Dienste bilden dabei die Grundlage der digitalen Infrastruktur und spielen eine zentrale Rolle bei der Verarbeitung, Speicherung und Verwaltung großer Datenmengen. Ihre Entwicklung beeinflusst maßgeblich die Fähigkeit eines Landes, Kontrolle über seine digitale Wertschöpfung zu behalten und seine strategischen Interessen zu wahren. **In diesem Zusammenhang versteht man unter digitaler Souveränität die Fähigkeit eines Staates, seiner Bürger und Unternehmen, die Kontrolle über digitale Technologien, Datenflüsse und die damit verbundenen Infrastrukturen zu behalten, ohne von externen, insbesondere ausländischen Einflüssen, Technologien oder Anbietern abhängig zu sein.** Dabei geht es jedoch nicht um Autarkie, die in der heutigen, globalen und durch vielschichtige Interdependenzen gezeichneten Welt beinahe unmöglich ist, sondern um die Möglichkeit der bewussten Auswahl unvermeidbarer Abhängigkeiten innerhalb persönlicher, geschäftlicher und internationaler Verflechtungen.

Deutschland steht in Bezug auf die digitale Souveränität vor verschiedenen Herausforderungen, darunter der zunehmende Wettbewerb durch internationale Cloud-Anbieter, die Abhängigkeit von globalen Immobilienfonds und Technologieanbietern und der steigenden Anforderungen an strategische Autonomie, Sicherheit, Datenschutz und Nachhaltigkeit. Die Entwicklung hin zu einer

resilienten, sicheren und eigenständigen digitalen Infrastruktur kann erheblich zur Stärkung der digitalen Souveränität beitragen. Dies erfordert jedoch eine differenzierte Betrachtung und Bewertung, wie gut die Infrastruktur aufgestellt ist, um den erwähnten Herausforderungen der digitalen Welt zu begegnen.

Im Folgenden werden mögliche Bewertungskriterien für die digitale Souveränität definiert, die bei der strategischen Entscheidungsfindung miteinbezogen werden sollten. Im Anschluss wird ein kurzer Marktüberblick aus der Perspektive der Souveränität aufgezeigt, welcher bei der Priorisierung der strategischen Fragestellungen hilfreich sein kann

5.3.1 Einflussfaktoren

Um die Auswirkungen der Entwicklung der digitalen Infrastruktur auf die digitale Souveränität Deutschlands sinnvoll bewerten zu können, wurden verschiedene Bewertungskriterien identifiziert. Diese Kriterien werden den drei Oberkategorien Datensouveränität, operative Souveränität und Infrastruktursouveränität zugeordnet.

Datensouveränität

Datensouveränität oder Datenhoheit bezeichnet die Kontrolle über die eigenen Daten und die Fähigkeit, deren Nutzung, Verarbeitung und Speicherung zu steuern. Sie ist entscheidend für den Schutz sensibler Informationen und für die Sicherstellung der Einhaltung lokaler Gesetze und Vorschriften (siehe auch Kapitel 4.2. zu den gesetzlichen Rahmenbedingungen).

- **Lokale Speicherung**: Die Speicherung von Daten und Metadaten innerhalb der Landesgrenzen oder eines Gebietes mit Souveränitätsbestrebungen ist essenziell, um sicherzustellen, dass sie den nationalen Rechtsvorschriften unterliegen. Dies reduziert die Risiken, die durch ausländische Zugriffsmöglichkeiten oder rechtliche Anfragen entstehen können.
- **Rechtliche Kontrolle**: Die vollständige Kontrolle über die rechtliche Zuständigkeit für gespeicherte Daten bedeutet, dass keine ausländischen Gesetze oder Gerichtsbeschlüsse Einfluss auf die Daten nehmen können. Dies stellt sicher, dass nur nationale Gesetze Anwendung finden.
- **Externe Verschlüsselung:** Die Verschlüsselung von Daten nach dem BYOK-Prinzip (Bring Your Own Key) mit Hilfe externer Schlüssel, die vom Kunden selbst verwaltet werden, gewährleistet, dass selbst der Betreiber des Rechenzentrums keinen Zugriff auf die Daten hat. Dies erhöht die Datensicherheit und stärkt die Souveränität der Datenverwaltung.

Operative Souveränität

Operative Souveränität bezieht sich auf die Fähigkeit, die eigenen operativen Prozesse und die Verwaltung der IT-Infrastruktur unabhängig zu steuern. Diese Kontrolle ist entscheidend, um Abhängigkeiten von ausländischen Anbietern und deren operativen Einflüssen zu minimieren.

- **Lokale Unternehmen**: Rechenzentren sollten von lokalen Unternehmen betrieben werden, die den nationalen Gesetzen und Regulierungen unterliegen. Dies verringert das Risiko von ausländischer Einflussnahme und sorgt für eine bessere Abstimmung auf nationale Interessen.
- **Lokaler Betrieb**: Die Durchführung aller operativen Tätigkeiten vor Ort ermöglicht eine direkte Kontrolle über den Betrieb und die Wartung der Rechenzentren. Dies ist wichtig, um schnelle Reaktionszeiten auf Vorfälle zu gewährleisten und den Einfluss externer Akteure zu minimieren.
- **Resilienz:** Die Verfügbarkeit von vielen geografisch getrennten Rechenzentrumsstandorten sowie die Nutzung von Offline-Archiven erhöhen die Resilienz der Dateninfrastruktur. Dadurch kann der Betrieb auch im Falle eines Ausfalls eines RZ aufrechterhalten werden, was die Widerstandsfähigkeit und operative Unabhängigkeit stärkt.

Infrastruktursouveränität

Infrastruktursouveränität umfasst die Kontrolle über die physische und technische Infrastruktur, die für den Betrieb von IT-Infrastruktur und -Diensten erforderlich ist. Diese Art der Souveränität stellt sicher, dass die Infrastruktur sicher und unabhängig betrieben werden kann.

- **Anbieterunabhängigkeit**: Die Möglichkeit, alle Daten und Prozesse ohne Abhängigkeiten von bestimmten Anbietern zurückzuholen oder zu migrieren, ist entscheidend, um Anbieterabhängigkeiten zu vermeiden. Dies stellt sicher, dass ein Wechsel des Anbieters oder des Marktplatzes jederzeit möglich ist, ohne dass es zu operativen Einschränkungen kommt.
- **Sicherheitszertifizierung**: Die Zertifizierung der Sicherheit der IT-Infrastruktur nach anerkannten Standards gewährleistet, dass die höchsten Sicherheitsanforderungen erfüllt werden. Dies schafft Vertrauen in die Integrität und den Schutz der Infrastruktur.
- **Netzwerksicherung**: Durch das sogenannte Zero-Trust-Modell wird jede Interaktion unabhängig überprüft und eine strikte logische und Netzwerksegmentierung vorgenommen, die das Risiko von Sicherheitsverletzungen minimieren kann. Dadurch wird die Kontrolle über die gesamte Infrastruktur verstärkt.
- **Isolation kritischer Systeme**: Die physische Trennung und Isolierung von kritischen Systemen, beispielsweise durch luftabgeschottete Umgebungen (Air Gaps), verhindert unbefugte Zugriffe und stellt sicher, dass sensible Daten und Prozesse vor äußeren Bedrohungen geschützt sind.

5.3.2 Wertschöpfungsstufen

Im Folgenden wird ein weiterer strategischer Aspekt dargelegt, aus der Perspektive der Wertschöpfung und Volkswirtschaft. Die dargelegten strategischen Fragestellungen können immer sowohl aus einer europäischen oder nationalen Sicht gleichermaßen betrachtet werden.

Die erste strategische Frage, die in diesem Zusammenhang gestellt werden muss: Ab welcher Stufe der Wertschöpfungskette sollte Deutschland souverän sein, um sowohl die volkswirtschaftliche Wertschöpfung zu maximieren als auch dem Souveränitätsanspruch gerecht zu werden?

- Rechenzentrumsgebäude - souveräne Gebäudefläche?
- IKT-Equipment – Herstellung der Hardware in Deutschland?
- IT-Infrastruktur - souveräne digitale Ressourcen?
- IT-Dienste - souveräne digitale- und IT-Dienste für Digitalisierung und Digitalwirtschaft?

Im Folgenden eine kurze Analyse dieser verschiedenen Ebenen und eine Darstellung der Ausgangslage.

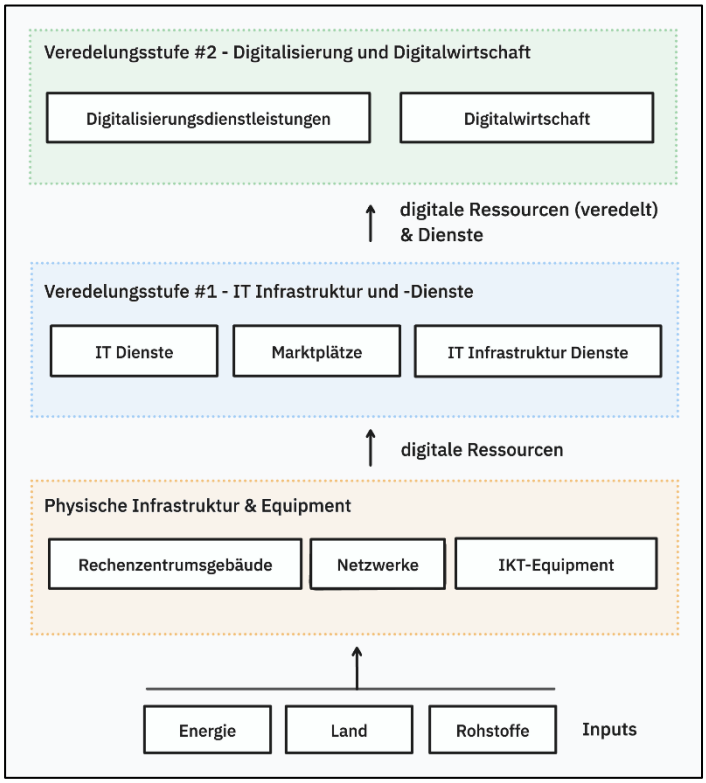


Abbildung 39: Wertschöpfungsstufen der digitalen Transformation: Von grundlegenden Inputs über physische Infrastruktur zu IT-Diensten bis hin zur Digitalwirtschaft. Die Darstellung zeigt den mehrstufigen Veredelungsprozess digitaler Ressourcen und deren Transformation in Digitalisierungsdienstleistungen und digitale Wirtschaftsaktivitäten.

Rechenzentrumsgebäude

Ein Großteil der deutschen RZ-Gebäudeinfrastruktur befindet sich in den Händen von ausländischen Investoren und Eigentümern. Ein Auszug der größeren Gebäudevermieter im Raum Frankfurt und Berlin, wie folgt:

UNTERNEHMEN	STRUKTUR	STANDORTE
EQUINIX	börsennotierter US-Immobilieninvestmentfond (REIT)	u. a. Frankfurt, München, Hamburg, Düsseldorf
NTT	börsennotierter japanischer Telekommunikationskonzern	u. a. Frankfurt, Berlin, Hamburg
VANTAGE	Privatunternehmen im Besitz eines US Private-Equity Fonds (Digital Bridge)	u. a. Berlin
MAINCUBES	Privatunternehmen mit Art.Invest und DTCP als Investoren, jeweils deutsche Investmentfonds	u. a. Frankfurt, Berlin
DIGITAL REALITY	börsennotierter US-Immobilieninvestmentfond (REIT)	u. a. Frankfurt
CYRUSONE	Privatunternehmen in den USA mit KKR (Private Equity) und GIP (Infrastruktur-Investor) als Hauptinvestoren	u. a. Frankfurt
TELEHOUSE	Privatunternehmen in den USA	u. a. Frankfurt
DATA CASTLE	deutscher RZ-Gebäudeentwickler mit Angelo Gordon, einem US-Infrastrukturfonds, als Investor	u. a. Berlin
NORIS NETWORKS	Privatunternehmen, Betreiber deutscher RZ-Gebäude und IT-Infrastruktur	u. a. Nürnberg, München

IT-Infrastruktur

Auf der Wertschöpfungsebene der IT-Infrastruktur kann in IT-Equipment und IT-Infrastrukturbetrieb unterschieden werden. Wie aus den Daten vom Statistischen Bundesamt (Statistisches Bundesamt, 2024a) ersichtlich ist, wird ein Großteil des IT-Equipments für Deutschland aus dem Ausland importiert.

Entwicklung des Außenhandels mit IKT-Gütern in Deutschland nach wichtigsten Importländern in %

Land	2020	2021	2022	2023
Volksrepublik China	41,3	45,7	44,7	42,6
Taiwan	5,0	5,7	7,1	8,4
Malaysia	4,0	4,1	4,7	5,0
Vereinigte Staaten von Amerika	4,3	3,9	4,2	4,1
Vietnam	3,9	3,4	3,4	3,6
Niederlande	5,1	4,0	3,6	3,5
Japan	3,4	3,2	3,0	3,2
Ungarn	3,4	3,0	3,0	3,2
Thailand	2,1	2,6	3,0	3,1
Tschechien	4,9	3,9	3,9	3,0

Stand 23. Februar 2024

Abbildung 40: Entwicklung des Außenhandels mit IKT-Gütern in Deutschland mit ausgewählten Ländern in den Jahren 2020 bis 2023. Quelle: Statistisches Bundesamt

Beim Betrieb von IT-Infrastruktur gibt es in Deutschland viele Großunternehmen, die keine eigenen Rechenzentrumsflächen betreiben, sondern diese von Gebäudeanbietern mieten. Ein Auszug einiger deutscher IT-Infrastrukturanbieter im Folgenden:

UNTERNEHMEN	STRUKTUR
STACKIT	Verbundunternehmen der Schwarz-Gruppe, IT-Infrastruktur und Cloud-Anbieter
IONOS	Börsennotierter IT-Infrastruktur- und Cloud-Anbieter aus Deutschland
HETZNER	Privatunternehmen, einer der größten IT-Infrastruktur-Anbieter in Europa
SCALE UP TECHNOLOGIES	Privatunternehmen im IT-Infrastruktur-Bereich
PLUSSERVER	ehemals deutscher Anbieter, akquiriert vom US-Anbieter “Go Daddy”
HOSTINGER	Privatunternehmen im IT-Infrastruktur-Bereich

UNTERNEHMEN	STRUKTUR
MITTWALD	Privatunternehmen im IT-Infrastruktur-Bereich
DOGADO	Privatunternehmen im IT-Infrastruktur-Bereich
NORIS NETWORKS	Privatunternehmen, deutscher RZ-Gebäude- und IT-Infrastruktur-Betreiber

IT-Dienste

Insbesondere für IT-Dienstleistungen und IT-Beratungen, unter anderem Managed Services, gibt es in Deutschland viele Großunternehmen, die sowohl international, national als auch regional ihre Leistungen anbieten.

UNTERNEHMEN	STRUKTUR
SVA SYSTEM VERTRIEB ALEXANDER GMBH	Privatunternehmen im Bereich IT-Dienste und IT-Beratung
BECHTLE AG	Börsennotiertes IT-Systemhaus, welches IT-Dienste und -Beratung anbietet
T-SYSTEMS INTERNATIONAL GMBH	Verbundunternehmen der Deutschen Telekom AG mit IT-Diensten, IT-Beratung, IT-Infrastruktur und eigenen Rechenzentrumsgebäuden
AKQUINET GMBH	Privatunternehmen mit Angebot im Bereich IT-Beratung, IT-Diensten, IT-Infrastruktur und eigenen Rechenzentrumsgebäuden; verkauft an Private-Equity-Fond “Deutsche Beteiligungs AG”
NETGO GROUP GMBH	Privatunternehmen, Anbieter von IT-Diensten und -Beratung
ARVATO SYSTEMS GMBH	Verbundunternehmen der Bertelsmann AG mit IT-Diensten, IT-Beratung, IT-Infrastruktur und eigenen Rechenzentrumsgebäuden
CONTROLWARE GMBH	Privatunternehmen mit Angeboten zu IT-Diensten und IT-Beratung
MR DATENTECHNIK VERTRIEBS- UND SERVICE GMBH	Privatunternehmen mit Angeboten zu IT-Diensten und IT-Beratung
OPERATIONAL SERVICES GMBH & CO.KG	Privatunternehmen mit Angeboten zu IT-Diensten und IT-Beratung

UNTERNEHMEN	STRUKTUR
NETPLANS GMBH	Privatunternehmen mit Angeboten zu IT-Diensten und IT-Beratung
KUTZSCHBACH GMBH	Privatunternehmen mit Angeboten zu IT-Diensten und IT-Beratung

Risiken bei der Abhängigkeit von Gebäuden, IKT-Equipment und geschlossenen Cloud-Marktplätzen

Aus Sicht der Souveränität kann diese Abhängigkeit bei Rechenzentrumsgebäuden und IKT-Equipment ein Risiko darstellen. Zudem kann es als Risiko gesehen werden, dass deutsche Unternehmen im Rahmen der Digitalisierung, aber auch Technologieunternehmen aus der Digitalwirtschaft, zunehmend ihren digitalen Ressourcenbedarf über geschlossene Marktplätze von ausländischen Cloud-Anbietern decken.

5.3.3 Regionale Ebenen

Zuletzt muss sich die strategische Frage gestellt werden, auf welcher Ebene sich Souveränität in Bezug auf digitale Infrastruktur wirklich erreichen lässt.

Insbesondere bei den notwendigen Kapitalinvestitionen und dem Strombedarf für digitale Ressourcen muss die Frage geklärt werden, ob eine europäische Souveränität nicht sinnvoller erscheint, insbesondere im Kontext der Nachhaltigkeit. Innerhalb der EU gibt es viele Mitgliedstaaten, die über Stromüberkapazitäten und die notwendigen Landflächen verfügen, um großflächige digitale Infrastruktur zu schaffen. Diese kann in Form von digitalen Ressourcen über den Internetmarkt nach ganz Europa exportiert werden.

Zudem wird eine strategische Abhängigkeit bezüglich des IKT-Equipments auch mittelfristig unumgänglich sein. Auch Investitionen in die RZ-Gebäudeinfrastruktur von ausländischen Investoren, deren Gelder oft aus Pensionskassen kommen, sind mit geringerem Risiko verbunden, da diese im Ausnahmefall enteignet werden können.

Hingegen ist eine strategische Abhängigkeit von geschlossenen Cloud-Marktplätzen nicht unbedingt notwendig und mit Risiken der Steuerbarkeit und Rechtssicherheit verbunden. Es wird empfohlen, offene Märkte zu entwickeln, an denen sowohl europäische IT-Infrastruktur-Akteure als auch ausländische Anbieter teilnehmen können, indem sie digitale Ressourcen zum Kauf anbieten (siehe auch Handlungsempfehlungen, Kapitel 5.6)

5.3.4 Schlussfolgerungen und Ansätze für mehr digitale Souveränität

Digitale Infrastruktur ist ein unverzichtbarer Bestandteil der digitalen Souveränität, da sie die physische Grundlage für die Speicherung, Übertragung und Verarbeitung von Daten bildet und eine zentrale Rolle in der Wertschöpfung spielt. Ohne eine starke nationale digitale Infrastruktur ist Deutschland in den Bereichen der Digitalisierung und Digitalwirtschaft abhängig von ausländischen

Anbietern, was die Souveränität gefährden könnte. Allerdings ist der Betrieb von eigener digitaler Infrastruktur allein nicht ausreichend. Es sollte eine umfassende Strategie, die auch Aspekte wie Cybersecurity, eigene Hard- und Softwarelösungen und eine robuste rechtliche Grundlage umfasst, entwickelt werden.

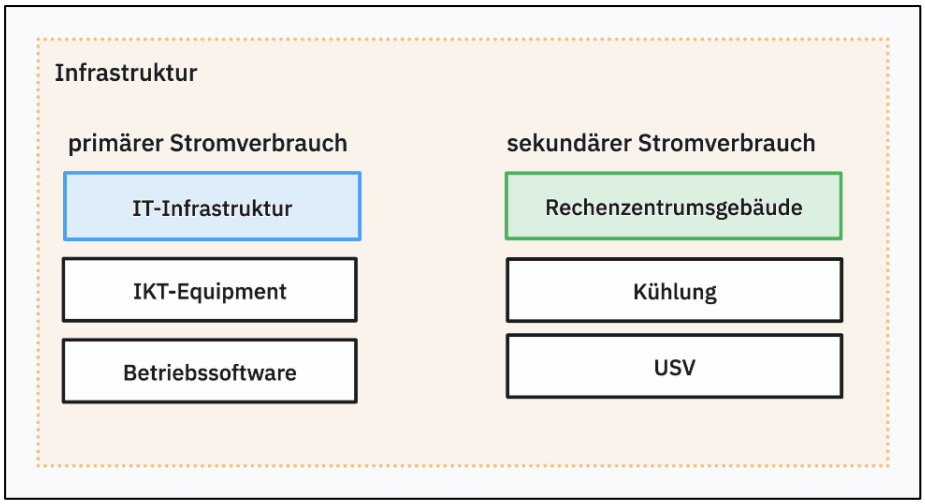


Abbildung 41: Übersicht über die Stromverbraucher-Infrastruktur in Rechenzentren - Darstellung der Aufteilung zwischen primärem Stromverbrauch (IT-Infrastruktur, IKT-Equipment, Betriebssoftware) und sekundärem Stromverbrauch (Rechenzentrumsgebäude, Kühlung, USV) in der digitalen Infrastruktur.

5.4 Technologische Entwicklung und Trends

Die technologischen Trends im Bereich der Infrastruktur eines Rechenzentrums lassen sich in zwei Teilbereiche aufteilen:

- IT-Infrastruktur (primärer Stromverbrauch)
    - IKT-Equipment (z. B. Server, Speicher- und Netzwerkgeräte)
    - Betriebssoftware (z. B. Virtualisierung, Orchestrierung) – wird zur Verbesserung des Nutzungsgrads eingesetzt.
- Gebäudeinfrastruktur (zusätzlicher Stromverbrauch von 20 bis 60 % pro kW IT-Infrastruktur, (Inc, 2022))
    - Kühlung (80 %) – um die Abwärme des IKT-Equipments abzuführen
    - USV (10 %) – um für das IKT-Equipment unterbrechungs- und störungsfrei Strom bereitzustellen.

5.4.1 Informationstechnik in Rechenzentren

In den letzten Jahren war die Technologieentwicklung der Informationstechnik in Rechenzentren vor allem durch die Virtualisierung, den Einzug von ARM-Prozessoren ins Rechenzentrum, Hochgeschwindigkeits-Ethernet/Fiber sowie die zunehmende Nutzung von High Performance GPUs

geprägt. Die typischen Rechenchips in Standardservern (x86) haben in den letzten Jahren hinsichtlich der Performance weiterhin stark zugenommen. Deutlich zeigen das Daten aus dem SPEC-power-Benchmark, welche insbesondere die Performance von CPU und Arbeitsspeicher messen:

Die Ergebnisse der Performance aktueller Systeme streuen sehr stark. Traut man jedoch den geometrischen Mittelwerten, hat sich die Performance zwischen 2014 und 2024 etwa um den Faktor 9 (1-CPU-Server) bis Faktor 10 (2-CPU-Server) verbessert. Ein sehr großer Teil dieser Leistungssteigerung basiert auf der Miniaturisierung von Bauteilen, was ermöglicht, dass pro Watt Leistung eine höhere Performance erzielt werden kann. Schaut man jedoch die Entwicklung der Stromaufnahme der zuvor betrachteten Server an, wird klar, dass auch die Stromaufnahme erheblich angestiegen ist:

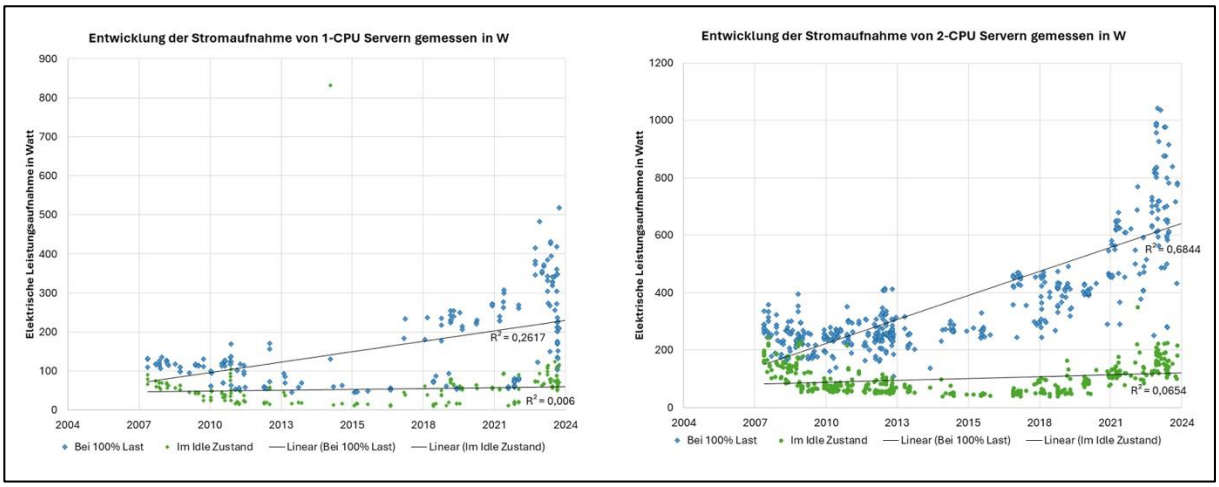


Abbildung 42: Entwicklung der Stromaufnahme von 1- und 2-Chip-basierten Servern in den Jahren 2004 bis 2024, eigene Darstellung, Daten von „SPEC.org“ (SPEC, 2024)

Im Bereich der 1-CPU-Server ist der Stromverbrauch im gleichen Zeitraum etwa um den Faktor 1,6 gestiegen, bei 2-CPU-Servern etwa um den Faktor 1,8. Beachtlich dabei ist, dass der Stromverbrauch im Idle-Zustand dagegen relativ gleichgeblieben ist, was vermutlich auch auf Effizienzvorschriften, wie die ECODESIGN-Verordnung in der EU (ENER-Lot. 9), zurückzuführen ist. Teilt man die Performance der individuellen Server durch ihre elektrische Leistungsaufnahme, erhält man den charakteristischen Wert für die Effizienz.

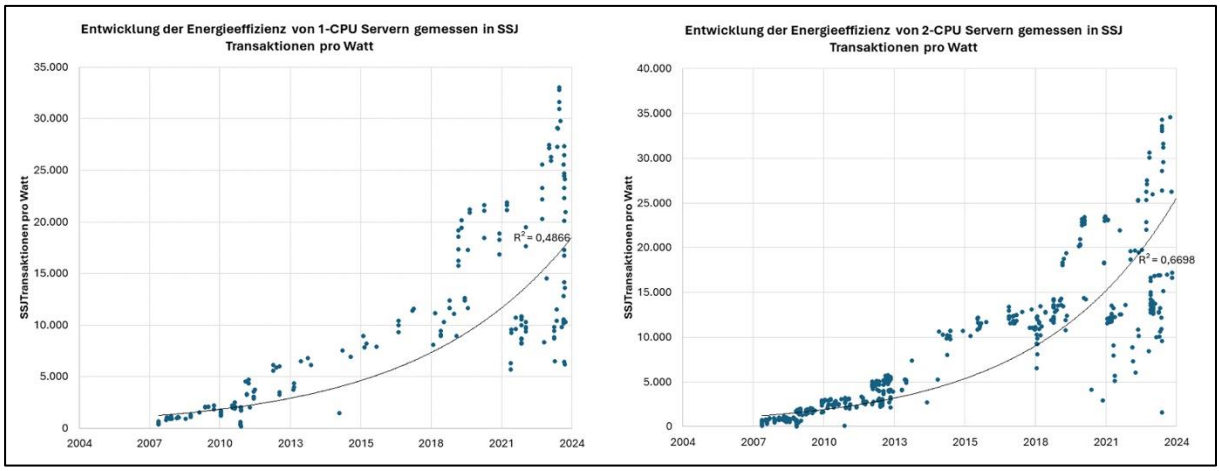


Abbildung 43: Entwicklung der Energieeffizienz (Transaktionen pro Watt) von 1- und 2-CPU-basierten Servern in den Jahren 2004 bis 2024, eigene Darstellung, Daten von „SPEC.org“ (SPEC, 2024)

Auch die Effizienz ist erwartungsgemäß stark angestiegen, bei 1-CPU-Servern etwa um den Faktor 5, bei 2-CPU-Servern etwa um den Faktor 6,3. In etwa entspricht das einer Steigerung der Rechenleistung pro Energie von durchschnittlich 20 - 22 % pro Jahr. Aufgrund der bereits heute sehr hohen Streuung der Effizienzwerte zwischen unterschiedlichen Servern/Chip-Herstellern sowie der Unsicherheiten hinsichtlich der Grenzen der Regeln von Moore (kontinuierliche Leistungssteigerung pro Chip) und Dennard (konstante Leistungsdichte) ergeben sich sehr hohe Unsicherheiten für die zukünftige Entwicklung in diesem Bereich. In einzelnen Anwendungssegmenten (z. B. Machine Learning) könnte eine Spezialisierung von Chips (vergleichbar mit den ASICS für Kryptomining) auch noch zu starken Effizienzsteigerungen führen. Die absolute (maximale) Leistungsaufnahme pro Chip bzw. pro Server zeigt derzeit jedoch einen klar ansteigenden Trend an, und es ist sehr wahrscheinlich, dass sich dieser auch noch in den kommenden Jahren fortsetzen wird.

5.4.2 Gebäudeinfrastruktur für Rechenzentren

Die größten Effizienzgewinne über die letzten knapp 20 Jahre, welche zu einer grundlegenden Absenkung des PUEs geführt haben, wurden im Bereich der Kühlung erschlossen (Uptime Institute, 2022). Dazu gehören kosteneffiziente Maßnahmen, wie die Kaltgangeinhausung oder auch eine Anhebung der Temperatur der Zuluft. Damit kann zugleich das Potenzial der freien Rückkühlung erhöht werden, womit der Einsatz einer energieintensiven Kältemaschine reduziert werden kann. Die erste Maßnahme gehört inzwischen zum Standard bei neueren Rechenzentren. Hinsichtlich der Zulufttemperatur zeigt sich ein anderes Bild. Auch wenn ASHRAE einen Temperaturbereich zwischen 18 und 27 °C empfiehlt (American Society of Heating, Refrigerating and Air-Conditioning Engineers, 2021), haben laut vorläufigen Zahlen aus dem RZReg 45 % der Rechenzentren in Deutschland eine Eintrittstemperatur von 22 bis 24 °C. Weniger als 24 °C haben 82 % der Rechenzentren. Pro 1 K Temperaturerhöhung können je nach klimatischen Bedingungen 2 bis 6 % Kühlenergie eingespart werden (Zhang, 2023). Hervorzuheben sind an dieser Stelle Google, Equinix (Wong, 2023) und Meta, die in einzelnen Rechenzentren Zulufttemperaturen von 27 bis 32 °C (Sustainability, 2024) praktisch erproben.

Um die jährlichen Betriebsstunden der freien Kühlung zu erhöhen, wird mittlerweile typischerweise ein Hybridkühler eingesetzt. Im Winter erfolgt ein trockener Betrieb, wohingegen an heißen Sommertagen

zusätzlich eine Wasserbenetzung erfolgt. Damit kann durch die Verdunstung von Wasser die Kühlwassertemperatur reduziert werden. Zwar wird auf diese Weise der spezifische Stromverbrauch der Kühlung reduziert, es erhöht sich aber gleichzeitig der Einsatz von Wasser. Ein Rechenzentrum mit 1 MW kann bei Einsatz der Verdunstungskühlung etwa 25,5 Mio. Liter Wasser pro Jahr verwenden (Heslin, 2016). Der WUE betrug bei den Hyperscalern 0,19 l/kWh (Amazon), 0,20 l/kWh (Meta) sowie 0,49 l/kWh (Microsoft) (Roundy, 2024). Auf globaler Ebene steigt der Wasserverbrauch aller Rechenzentren seit 2015 kontinuierlich an. Auch in Zukunft wird der Wasserbedarf vermutlich weiter ansteigen, vor allem im indirekten Bereich der Stromerzeugung, welche den Großteil des Gesamtwasserverbrauches ausmacht (Mytton, 2021).

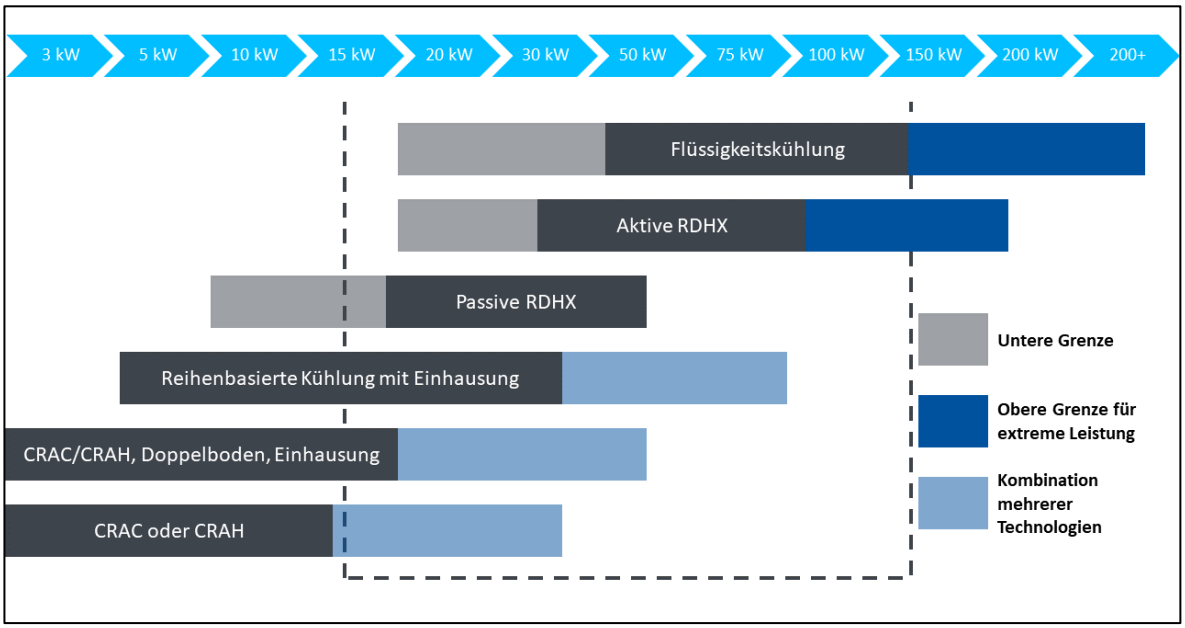


Abbildung 44: Kühltechnologie in Abhängigkeit der Leistungsdichte von Racks. Quelle: Eigene Darstellung basierend auf Daten von Vertiv

Innerhalb des Serverraums stellt die Entwicklung der Leistungsdichte der Racks die Raumkühlung vor Herausforderungen. Während mit der klassischen Luftkühlung (CRAH) Wärmelasten bis 12 kW/Rack (Dürr, 2018) abgeführt werden können und 2020 die durchschnittliche Rackleistungsdichte bei 8,2 kW/Rack lag (Dunnavant; Kätkä, 2024), bedingen vor allem die KI-, ML- und HPC-Anwendungen eine Erhöhung der Leistungsdichte. Für 2024 wird dadurch bereits eine durchschnittliche Rackleistungsdichte von 12 kW/Rack erwartet (McCord, 2024). Perspektivisch soll vor allem durch den KI-Boom die Rackleistungsdichte auf 40 bis 100 kW/Rack ansteigen (Goovaerts, 2024). Dies lässt sich mit der klassischen Luftkühlung nicht mehr kühlen, weswegen die Bedeutung der Hybridkühlung oder Flüssigkeitskühlung ansteigen wird. Dell'Oro geht davon aus, dass der Marktanteil der Flüssigkeitskühlung in diesem Zuge von 8 % im Jahr 2021 auf 33 % im Jahr 2028 ansteigt (Chernicoff & Vincent, 2024). Dabei lassen sich mit einer Rückwandkühlung je nach Konzept (passiv/aktiv) 15 bis 55 kW/Rack realisieren (Dürr, 2018).

Oberhalb dieser Grenze führt aus physikalischer Sicht kein Weg an der direkten oder totalen Flüssigkeitskühlung vorbei. Bis zu einem gewissen Grad können diese Limits aber auch durch eine Kombination von verschiedenen Kühltechnologien oder eine Änderung der Rackgröße verschoben werden. Ein Beispiel dafür ist die luftunterstützende Flüssigkeitskühlung von Meta, bei der im Server wasserdurchströmte Kühlplatten verbaut sind und über einen Wasser-Luft-Wärmeübertrager an der

Rückwand die Abwärme an die Raumluft abgegeben wird. Um die zu erwartenden Leistungsdichten der Racks mit der Flüssigkeitskühlung sicher kühlen zu können, gibt es inzwischen „Coolant Distribution Units", welche eine Kühlkapazität je nach Hersteller von bis zu 100 kW (Supermicro) oder 2 MW (CoolIT) (Super Micro, 2024; Uti, 2024) haben. In diesem Zuge stellt, unter Berücksichtigung der zukünftigen Leistungsentwicklungen der Server und Chips, auch die Wahl der Kühlwassertemperatur eine Herausforderung dar. Die OCP empfiehlt in diesem Zusammenhang eine Temperatur von 30 °C für das technische Kühlungssystem (Mills et al., 2024). Sie soll einen guten Kompromiss aus Kühlleistung, Nachhaltigkeit und Energieeffizienz ermöglichen. Damit ist immer noch eine Abwärmenutzung möglich.

In Bezug auf die Stromversorgung verdrängen bei der Verwendung von elektrochemischen Speichersystemen die Lithium-Ionen-Akkus die bis dahin üblichen Blei-Säure-Batterien (Vertiv, 2018). Sie sollen 2025 bereits einen Marktanteil von 38,5 % erreichen. Die Preise für Lithium-Ionen-Akkus sind seit 2015 kontinuierlich am Sinken, vor allem bedingt durch die Fortschritte bei E-Fahrzeugen (Vertiv, 2018). Durch den anstehenden Technologiewechsel kann die gravimetrische Energiedichte der USV erhöht werden (Bansen et al., 2021). Zugleich kann die mögliche Lebensdauer höher sein und die Effizienz beim Laden und Entladen ist mindestens 10 % höher (Avelar, Zacho, 2018; Jung et al., 2015).

Hinsichtlich der Netzersatzanlagen gibt es zwei Entwicklungen. Auf der einen Seite gibt es bei den Dieselmotoren die Möglichkeit, den klassischen Dieselkraftstoff durch hydrobehandeltes Pflanzenöl (HVO) zu ersetzen. Amazon plant in Europa perspektivisch alle NEA auf HVO umzustellen (*Amazon, 2023b*). Auf der anderen Seite gibt es erste Pilotprojekte sowie Rechenzentrumsstandorte in den USA sowie vereinzelt in Europa, welche Brennstoffzellen als Alternative zu Dieselgeneratoren einsetzen. Dazu gehören Microsoft, Equinix und Rolls-Royce (Baringa, 2022; Robinson, 2024; Rolls-Royce, 2023).

Anderseits gibt es auch Projekte mit einer teilweisen Eigenerzeugung von Strom mittels KWK oder PV. Dazu zählen unter anderem das Kommunale Rechenzentrum Niederrhein (KWK) (Schödwell et al., 2013) oder JH Computers (KWK + PV) (JH-Computers, 2024). Für KWK muss nicht unbedingt Diesel eingesetzt werden, sondern es kann auch (Bio-)Gas genutzt werden. Gerade die PV-Stromerzeugung ist dann maximal, wenn der Stromverbrauch für die Kühlung erhöht ist. So kann insbesondere PV auch einen Beitrag zur Reduktion der Spitzenlast leisten (Hyvönen et al., 2024).

5.5 Standortkriterien für Deutschland

Grundsätzlich ist der deutsche Rechenzentrumsmarkt in Bezug auf ausländische Investitionen bereits gut aufgestellt. Mit einem starken, international renommierten Internetknoten in Frankfurt sowie einem hohen Digitalisierungspotenzial zieht Deutschland bereits heute Unternehmen aus der ganzen Welt an, die in Deutschland sowohl Rechenzentren als auch IT-Infrastruktur realisieren. Im Rahmen der Digitalisierung der deutschen Wirtschaft sehen viele internationale IT-Infrastrukturanbieter großes Potenzial, digitale Ressourcen in Deutschland zu produzieren und anzubieten. Zudem bietet Deutschland mit seinem starken Rechtsrahmen hohe Planungssicherheit. Die Potenziale auf dem deutschen Markt gleichen eventuelle Nachteile, wie z. B. einen höheren Strompreis oder einen längeren Planungs- und Bauprozess, aus. Um diesen Trend fortzusetzen, sollten konkrete Digitalisierungspläne verschiedener Industriezweige sowie der Regierung weiter ausgearbeitet und

präzisiert werden, so dass Marktteilnehmer und Infrastruktur-Investoren eine klare Nachfrage nach digitalen Ressourcen modellieren und daraus notwendige Gebäude und IT-Infrastruktur ableiten können.

Auf der Exportseite, also bezüglich digitaler Ressourcen, die aus deutschen Standorten von internationalen und nationalen Unternehmen ins europäische oder internationale Ausland exportiert werden, gibt es Nachholbedarf. Hier können die Niederlande als Vorbild gesehen werden, welche gleichermaßen mit hohen Energiepreisen und einem starken Internetknoten eine starke Exportfähigkeit von digitalen Ressourcen geschaffen haben. Dies wurde insbesondere durch starke Investitionen in die Ausbildung von IT-Infrastruktur-Know-how als auch durch einen klaren und sicheren Rechtsrahmen für IT-Infrastruktur realisiert. Zudem wurde massiv in Glasfaserausbau, insbesondere in Seekabel, investiert und international unter dem “Digital Gateway to Europe”[26] lanciert. Hinzu kommt ein Steuer- und Visaregime, welches die Ansiedlung von qualifizierten Fachkräften vereinfacht.

Mit der historisch starken Maschinenbau- und Energiesystem-Kompetenz sowie deutschen Marktführern im Bereich der Energieeffizienz-Technologien sollte sich Deutschland als Exportland für digitale Ressourcen über Nachhaltigkeit, Effizienz und Qualität differenzieren. Dabei kann die Nachhaltigkeitskomponente über gezielte Ansiedlung in Gebieten mit hoher erneuerbarer Energieproduktion sichergestellt werden, sowie durch nachhaltige Kühlsysteme inklusive Abwärmenutzung und eine hohe Lebensdauer von Rechenzentren in Deutschland.

Die Qualitätsdimension kann sowohl über einen starken Rechtsrahmen als auch eine hohe Veredelung der digitalen Ressourcen, zum Beispiel in branchenspezifischen IT-Diensten, geschehen, aber auch über die Zuverlässigkeit der zur Verfügung gestellten digitalen Infrastruktur. Hier sollte die weltweit anerkannte Marke „Made in Germany” auch auf digitale Ressourcen übertragen werden und damit IT-Anwendungen und Digitalunternehmen, die sich über Qualität, Datenschutz, Zuverlässigkeit, Nachhaltigkeit und Effizienz differenzieren möchten, angezogen werden.

5.6 Handlungsempfehlungen

Souverän

Deutschland sollte sich bewusst mit der Souveränitätsfrage auseinandersetzen und strategische Entscheidungen treffen, ab welcher Ebene der Wertschöpfung Souveränität erreicht werden soll. Auf Basis dieser strategischen Entscheidung sollten eine konkrete Strategie für die Entwicklung souveräner digitaler Infrastruktur abgeleitet und entsprechende internationale Partnerschaften angestrebt werden.

Im Rahmen der Globalisierung ist es nicht zu empfehlen, eine komplette Entkopplung anzustreben. Stattdessen sollten klare Abhängigkeiten definiert und Risiken in diesen minimiert werden. Gleichzeitig gilt es, die volkswirtschaftliche Wertschöpfung zu maximieren und insgesamt eine Steuerbarkeit der digitalen Infrastruktur in Deutschland sicherzustellen.

[26] https://www.digitalgateway.eu/

Auch sollte im Rahmen der Souveränitätsbetrachtung eine europäische Sicht entwickelt werden, welche gemeinschaftlich erarbeitet wird. Aus diesem gemeinschaftlichen Verständnis von Souveränität können konkrete Strategien für Deutschland und andere Mitgliedstaaten ausgearbeitet werden.

Europäisch

Mit einem offenen, transparenten Marktplatz für den Handel von digitalen Ressourcen können die europäische Digitalisierung und das Wachstum der europäischen Digitalwirtschaft teilweise souverän realisiert werden. Der Marktplatz unterliegt der europäischen Regulatorik und kann von Marktaufsichtsbehörden gesteuert und, falls nötig, korrigiert werden. Ein Marktplatz kann zudem den Wettbewerb und das Wachstum von europäischen IT-Infrastrukturanbietern fördern, die bereits heute sowohl über Qualität als auch über Kosten durchaus wettbewerbsfähig sind.

Dieser Marktplatz kann regionale Infrastruktur, insbesondere für die Digitalisierung der deutschen Industrie, nicht vollumfänglich ersetzen. Er schafft jedoch eine zentrale Anlaufstelle für den Großteil des digitalen Ressourcenbedarfs. Insbesondere für die digitale Wirtschaft kann ein solcher Marktplatz zu einer höheren Kosteneffizienz führen und damit zu einer höheren Wettbewerbsfähigkeit der digitalen Produkte und Dienstleistungen, die daraus veredelt werden. Zudem können in einen Marktplatz europäische Werte, wie Nachhaltigkeit, Effizienz und Sicherheit, über Regulatorik implementiert werden. Siehe dazu auch Kapitel 3.1.7.

Innovativ

Mit einer starken Expertise sowie einer vielfältigen Industrie in den Bereichen der elektrischen und mechanischen Infrastruktur kann Deutschland sich international als Innovationstreiber für Rechenzentrumsinfrastruktur – insbesondere bei der Gebäudetechnik, Automatisierung und Energiesystemintegration – platzieren. Hierfür sollte gezielt in Innovation und spezifische Bildung für Rechenzentrumsinfrastruktur investiert werden. Auch hohe regulatorische Anforderungen an Energieeffizienz und Energiesystem-Integration können für Innovationen förderlich sein, insbesondere wenn dafür neue Technologien geschaffen werden müssen.

Eine übergeordnete Rechenzentrumstrategie für Deutschland kann dabei richtungsweisend sein und über die Verknüpfung auch mit zukünftigen Regulierungsvorhaben Raum für mittelfristige Innovations- und Technologieentwicklung schaffen.

Auf Basis der technologischen Trends könnten Handlungsfelder unter anderem folgende sein:

- Stromversorgung
    - Zur Reduzierung der Umweltauswirkungen von Netzersatzanlagen ist bei ausreichender Verfügbarkeit ein Wechsel von Dieselkraftstoff zu HVO und langfristig (Neuanlagen) zu Wasserstoff/Wasserstoffderivaten sinnvoll und angebracht. Damit wird nicht nur ein Beitrag zur Dekarbonisierung, sondern auch zur Reduktion von lokalen Schadstoffemissionen geleistet.
    - Erhöhung des Wirkungsgrades von USV auf elektrochemischer Basis durch Wechsel von Blei-Säure auf Lithium-Ionen-Akkus. Daneben kann die gravimetrische Energiedichte und die Lebensdauer erhöht werden. Neue

Batterietechnologien können zudem erforscht und für Rechenzentren markttauglich gemacht werden.

- PV-Installation auf RZ-Gelände, um den Bezug von Netzstrom im Sommer zu reduzieren. Auf diese Weise kann der erhöhte Energiebedarf bei der Kühlung reduziert werden.
- Integrationskonzepte für weitere erneuerbare Energiequellen in und an Rechenzentren wie Windenergie weiter erforschen und fördern.

- Energiesystemintegration
  - Rechtliche Grundlagen für die Bereitstellung von Regelleistung am Strommarkt mittels der vorhandenen Speicher- (USV) und Erzeugungskapazitäten (NEA) eines Rechenzentrums schaffen.
  - Neue Technologien mit höherem Energieeffizienzgrad für die Auskopplung von Wärme erforschen und markttauglich machen.
- Kühlung
  - Weitgehende Adaption von Flüssigkühlung der Serverhardware im Markt realisieren. Dadurch lassen sich die vorhandenen Abwärmepotenziale des IKT-Equipments besser nutzbar machen. Die direkte Flüssigkeitskühlung ist aus technischer Perspektive bereits am weitesten entwickelt und ist die Standardkühltechnologie im HPC-Bereich.
  - Um den Energieaufwand für die Kühlung und damit den PUE noch weiter zu reduzieren, sollte dort, wo es technisch möglich ist, die Soll-Zulufttemperatur der Kühlluft noch weiter angehoben werden. Als Empfehlung gelten die Vorgaben von ASHRAE. Allein eine Temperaturerhöhung von 1 K führt zu Energieeinsparungen bei der Kühlung von 2 bis 6 % durch eine stärkere Ausnutzung der freien Kühlung. 82 % der Rechenzentren gemäß dem RZReg könnten ihre Zulufttemperatur um mehr als 3 K anheben. An dieser Stelle muss aber Tradeoff zwischen Energieaufwendungen bei der Kühlungsinfrastruktur und den Ventilatoren in den Servern berücksichtigt werden.
- Ressourceneffizienz
  - Es sollte eine strategische Festlegung bei der Verwendung von Wasser für Rückkühlung geben. Entweder kann Wasser intensiver (WUE steigt) für die Kühlung genutzt und dabei der Energiebedarf (PUE sinkt) für die Kühlung reduziert werden oder der Wassereinsatz ist stärkeren Restriktionen unterworfen, verbunden mit der Notwendigkeit eines höheren Energieeinsatzes.
  - Die Adaption von intelligenter Gebäudesteuerung in Rechenzentren sollte priorisiert werden. Damit können eine bessere Auslastung der Gebäudeinfrastruktur und eine dynamische Synchronisierung von IKT-Last, Kühlungsaufwand und Energieverfügbarkeit realisiert werden. Gleichzeitig besitzen sowohl die Bestands- als auch Neubaurechenzentren ein

Optimierungspotenzial im realen Betrieb, was so besser erschlossen werden kann.

Nachhaltig

Für die nachhaltige Entwicklung von digitaler Infrastruktur in Deutschland sollte sowohl der volkswirtschaftliche Aspekt als auch der Umweltaspekt mit einbezogen werden.

Volkswirtschaftlich nachhaltig ist digitale Infrastruktur primär in regionalen Ökosystemen, in denen ein Teil der Wertschöpfung systematisch angesiedelt wird. Siehe dazu auch Kapitel 3.1.7. Diese Strategie kann auch dem Souveränitätsgedanken dienlich sein.

Aus Sicht der Umweltauswirkung sollte digitale Infrastruktur in Regionen mit hoher Penetration erneuerbarer Energien gezielt angesiedelt werden. Im Idealfall sollte sie mit anderen Industrien oder Abnehmern gekoppelt werden, die einen hohen Bedarf an industrieller Abwärme haben. Gleichzeitig sollte der Bau von Rechenzentrumsgebäuden mit erneuerbaren Rohstoffen und einer Kreislaufwirtschaft für elektrisches, mechanisches und IKT-Equipment angestrebt werden.

6 Übergreifende Handlungsempfehlungen

6.1 Deutschland braucht eine Rechenzentrumsstrategie

Rechenzentren sind ein zentraler Bestandteil der digitalen Infrastruktur in Deutschland und stellen die notwendigen Ressourcen für den Einsatz digitaler Technologien bereit. Diese Technologien wiederum dienen als Basis für die Digitalisierung von Wirtschaft und Gesellschaft und das Wachstum der Digitalwirtschaft. Neue Technologien, wie das maschinelle Lernen und Künstliche Intelligenz, und eine boomende Digitalwirtschaft sorgen für immer weiter steigende Anforderungen an Rechen-, Speicher-, und Übertragungskapazität. Durch die steigende Nachfrage rücken Aspekte der Energieeffizienz und der Nachhaltigkeit der zugrundeliegenden digitalen Infrastrukturen zunehmend in den Mittelpunkt. Angesichts der schnellen technologischen Entwicklung und den damit verbundenen energetischen Herausforderungen (vor allem Flächen- und Stromverfügbarkeit), ist eine klare und strategische politische Steuerung sinnvoll, um Deutschland als wettbewerbsfähigen und nachhaltigen Standort für digitale Infrastrukturen zu stärken.

In der vorliegenden Studie werden die drängendsten Herausforderungen und Handlungsfelder im Zusammenhang mit der Entwicklung von Rechenzentren in Deutschland analysiert. Dabei wird deutlich, dass trotz bestehender Koordinierungsansätze, wie z. B. der Nationalen Digitalstrategie, eine gezielte Strategie für die Entwicklung von Rechenzentren fehlt. Ohne eine solche Strategie drohen nicht nur wirtschaftliche und technologische Nachteile, sondern es droht auch das Risiko, Potenziale im Bereich der Digitalisierung, der digitalen Souveränität und der digitalen Nachhaltigkeit ungenutzt zu lassen.

Eine Rechenzentrumsstrategie sollte dabei umfassend mit anderen strategischen Feldern der Digitalpolitik verknüpft werden, insbesondere mit Aspekten wie nationale Sicherheit, Datenschutz, Souveränität durch eigene Hard- und Softwarelösungen und robuste rechtliche Grundlagen. Dabei werden die in Deutschland bestehenden Rahmenbedingungen in Form der Datenschutz-Gesetzgebungen und ihrer Anwendung überwiegend nicht als restriktive Regulatorik bzw. Hemmnis, sondern als wichtiger Standortvorteil angesehen. Weitere sicherheitsrelevante Anforderungen ergeben sich etwa aus dem IT-Sicherheitsgesetz und der KRITIS-Verordnung. Diesen regulatorischen Rahmen gilt es im Zuge fortschreitender Digitalisierung konsequent weiterzuentwickeln, um das Vertrauen in den Rechenzentrumsstandort Deutschland und dessen digitale Souveränität zu stärken.

Um diesen Herausforderungen zu begegnen, werden im Folgenden zentrale Maßnahmen empfohlen, die als Grundlage für eine moderne und zukunftsfähige Strategie für Rechenzentren in Deutschland dienen können.

Abbildung 45: Handlungsempfehlungen für die strategische Entwicklung des RZ-Standorts Deutschland

Förderung der digitalen Souveränität, Wettbewerb und lokale Akteure auf nationaler und EU-Ebene: Politik und Wirtschaft sollten gemeinsam eine Roadmap für digitale Infrastruktur entwerfen, die auch die europäische Perspektive berücksichtigt – darunter gemeinsame Standards, grenzübergreifenden Ressourcenaustausch und Handel. Es sollte strategisch entschieden werden, welche Menge und Art von digitalen Ressourcen durch Importe gedeckt werden sollen, was wiederum die Wettbewerbspolitik gegenüber globalen Akteuren beeinflussen wird. Um die Bereitstellung von Rechenzentren und digitalen Ressourcen strategisch zu steuern und zu lokalisieren und die grundsätzlich bestehende Importabhängigkeit zu reduzieren, sollten geschlossene Marktplätze für digitale Ressourcen durch einen europäischen Marktplatz oder eine Verordnung reguliert werden. Dies würde gleiche Wettbewerbsbedingungen auf dem Markt für IT-Infrastrukturen gewährleisten und die geschlossenen Märkte für alle Marktteilnehmer öffnen. Eine Option ist die Einrichtung eines europäischen Marktplatzes, um den Handel mit digitalen Ressourcen innerhalb Europas zu fördern. Der Aufbau einer eigenen nationalen digitalen Infrastruktur könnte auch durch strategische Partnerschaften und entsprechende Importe, insbesondere aus anderen europäischen Ländern, erreicht werden. Die Zusammenarbeit sollte die folgenden Ansätze in Betracht ziehen:

1. Einführung eines offenen europäischen Marktplatzes für digitale Ressourcen, um Wettbewerb zu fördern und Abhängigkeiten von globalen Anbietern zu reduzieren
2. Investitionen in Innovationen, insbesondere in Virtualisierungs-Technologien und Technologien zur Erhöhung des Nutzungsgrades in lokalen IT-Infrastrukturen

3. Aufbau nationaler Rechenzentrums-Infrastrukturen, um der zukünftigen Abhängigkeit von Importen aus anderen Märkten (USA, China) entgegenzuwirken
4. Zugang zu Kapital für den Ausbau regionaler digitaler Infrastruktur für kleine und mittelständische IT-Infrastruktur-Betriebe erleichtern.

Empfehlung 1: Intensivierung eines politischen Dialogprozesses, um die Entwicklung eines souveränen Rahmens für digitale Infrastruktur in der EU in Zusammenarbeit mit europäischen Partnern zu sichern.

Stimulierung der Nachfrageseite: Ein zentrales Hindernis für die Wettbewerbsfähigkeit lokaler Rechenzentrumsanbieter ist der Mangel an Skaleneffekten, der zu einer teilweise geringen Auslastung der bestehenden IT-Infrastruktur führt. Um dieser entgegenzuwirken, sollte der Nutzungsgrad der digitalen Infrastruktur erhöht werden. Aktuell werden aufgrund der wirtschaftlichen Lage gerade bei kleineren, unternehmenseigenen Rechenzentren häufig Neuinvestitionen zurückgestellt. Eine klare und umfassende Rechenzentrumsstrategie, die sowohl von der Bundesregierung, als auch von der Verwaltung und den verschiedenen Industriezweigen getragen wird, kann entscheidend zur Stärkung und zum Wachstum der digitalen Infrastruktur in Deutschland beitragen. Insbesondere wirken sich Investitionen in öffentliche und private Digitalisierungsprojekte positiv auf die Nachfrage nach digitalen Ressourcen aus und fördern dadurch auch die Marktentwicklung der Rechenzentren. Gleichzeitig können gezielte Anpassungen in den Einkaufskriterien der öffentlichen Hand dazu beitragen, regionale digitale Ökosysteme zu stärken und die volkswirtschaftliche Wertschöpfung, insbesondere auf kommunaler Ebene, nachhaltig zu verbessern. Darüber hinaus spielt die Entwicklung einheitlicher Standards eine Schlüsselrolle. Dies könnte durch die Einführung einheitlicher Einkaufskriterien und maßgeschneiderter Förderprogramme erreicht werden, die lokale und nachhaltigere Anbieter stärken und gleichzeitig die lokale sowie regionale Wertschöpfung unterstützen. Solche Standards vereinfachen nicht nur den Handel und den Einkauf digitaler Ressourcen für Wirtschaft und Regierung, sondern definieren und implementieren explizit Anforderungen an Qualität, Nachhaltigkeit und Effizienz im Rechenzentrumsbetrieb. Dies wiederum fördert die Diversifizierung des Angebots auf dem Markt für digitale Ressourcen und schafft Vorteile für Marktteilnehmer. Ein abgestimmtes und integriertes Maßnahmenpaket in diesen Bereichen kann die Wettbewerbsposition der lokalen Rechenzentrumsanbieter stärken und die digitale Transformation in Deutschland beschleunigen.

Empfehlung 2: Initiierung eines übergeordneten Dialogs von Bundesregierung, Landesregierungen, Behörden, Organisationen, Unternehmen und Kommunen, um gezielte Maßnahmen zu entwickeln, die Skaleneffekte und Effizienzpotenziale bei der Nutzung der Rechenzentren-Kapazitäten fördern.

Regionale Netzwerke und Ökosysteme fördern: Die Ansiedlung und die Bedeutung von Rechenzentren hängt maßgeblich von Faktoren wie Bevölkerungsdichte, Wirtschaftsleistung und Digitalisierungsgrad der Regionen ab. Wirtschaftlich nachhaltige Rechenzentren entfalten ihr Potenzial in Ökosystemen, die gezielt auf lokale Wertschöpfung ausgerichtet sind. Um solche Effekte zu fördern, sollten Beschaffungsstrategien auf kommunaler, Landes- und Bundesebene Regionalität stärker berücksichtigen. Entwicklungskonzepte für Groß- und Industriestädte, mittelgroße Gemeinden und ländliche Räume sollten in diesem Zusammenhang einbezogen werden. Die Modernisierung und der Ausbau der Netzinfrastruktur – insbesondere für dezentrale Einspeisung und effizienten Stromtransport – sind essenzielle Faktoren. Dies erfordert die gezielte Ausweisung von Standorten in der Nähe von erneuerbaren Energiequellen und Netzknotenpunkten für maximale Synergieeffekte. Gleichzeitig sollten Rechenzentren in Regionen angesiedelt werden, die eine Aufnahmefähigkeit für

Abwärme aus Rechenzentren haben. Diese beiden Aspekte ganzheitlich zu denken und beispielsweise mit dem Netzentwicklungsplan und der KWP zu verbinden, setzt eine abgestimmte Planung voraus. Der Bau solcher Rechenzentrumsgebäude sollte erneuerbare Energien und Prinzipien der Kreislaufwirtschaft einbeziehen, um ökologische Anforderungen zu erfüllen und die Ressourceneffizienz zu optimieren. Ergänzend wäre die Förderung der Eigenversorgung von Rechenzentren durch die Integration lokaler Photovoltaik- oder Windkraftanlagen auf Rechenzentrumsgeländen ein weiterer Baustein, um die Standorte energetisch zukunftsfähig zu gestalten. Durch diese Maßnahmen können Effizienz, Nachhaltigkeit und regionale Wertschöpfung in Einklang gebracht werden, während die Belastung des Stromnetzes gezielt gemanagt und reduziert wird.

Empfehlung 3: Entwicklung von raumordnenden Planungen und Übersichten in Zusammenarbeit mit Bundesländern, Regionen und Kommunen, die die regionale Ansiedlung von Rechenzentren in besonders geeigneten Regionen fördern und gleichzeitig die lokale Wirtschaft sowie die nachhaltige Entwicklung stärken.

Verbesserung der Infrastrukturentwicklung: Ein integrierter Ansatz, der Glasfaser, Rechenzentren, Strom- und Wärmenetze verbindet, ist notwendig, um eine nachhaltige digitale Infrastruktur flächendeckend voranzutreiben. Trotz bestehender Vorgaben im Bereich erneuerbarer Energien bleibt der Mangel an Highspeed-Internet und flächendeckender Glasfaser-Konnektivität in vielen Regionen ein Hindernis. Teilweise können Synergieeffekte bei der Planung und Genehmigung, aber auch bei den Baumaßnahmen gehoben werden, wenn mehrere Infrastrukturen gemeinsam gedacht werden. Ergänzend dazu könnten bundesweit einheitliche Baunormen für Rechenzentren die Planungssicherheit verbessern und Genehmigungsverfahren beschleunigen. Die Anerkennung von Rechenzentren als kritische Infrastruktur würde außerdem die Priorisierung solcher Projekte gewährleisten und die langfristige Wettbewerbsfähigkeit Deutschlands als Standort sichern.

Empfehlung 4: Entwicklung eines integrierten Ansatzes der Infrastrukturplanung in Zusammenarbeit mit Bundesländern, Regionen und Kommunen, der Glasfasernetze, Breitbandinfrastruktur, Strom- und Wärmenetze als verknüpfte Schlüsselkomponenten für nachhaltige Infrastrukturen verbindet.

Empfehlung 5: Beschleunigung und Vereinheitlichung von Genehmigungsverfahren und Priorisierung von Rechenzentren als kritische Infrastruktur.

Integration in Energiesysteme und Förderung der Energie- und Ressourceneffizienz: Der Anschluss großer Stromverbraucher wie Rechenzentren an bestehende Stromnetze erfordert erhebliche Abstimmungsprozesse, da diese Anlagen erheblichen Einfluss auf die Netzbelastung haben. Stromanschlussanträge sollten daher flexibler gestaltet und an neue Anforderungen wie Demand-Response-Mechanismen angepasst werden. Gleichzeitig muss auch weiter gewährleistet werden, dass Rechenzentren im Rahmen der Netzanschlussbegehren klare Zusagen durch die Netzbetreiber erhalten und große Energieverbraucher wie neue Rechenzentren bei der Kapazitätszuteilung nicht benachteiligt werden. Entsprechende Reformen könnten die Netzintegration deutlich verbessern. Die Entwicklung einheitlicher Standards für Strom- und Wärmenetzanschlüsse sowie die Priorisierung des Wärmenetzausbaus an wichtigen Rechenzentrumsstandorten kann dabei maßgeblich zur Energieeffizienzsteigerung und effizienten Ressourcennutzung beitragen. Dies sollte durch eine Optimierung der Planungs- und Anschlussprozesse im Bereich der Stromnetze und eine Konkretisierung der Wärmenetz(anschluss)planung ergänzt werden, um zügige Implementierungen

sicherzustellen. Prioritäten sollten auf der Förderung von Forschung und Entwicklung für energieeffiziente Technologien sowie den Ausbau von modernen Speicherkapazitäten zur Entlastung der Stromnetze und zur Gestaltung einer nachhaltigen USV liegen. Besonders dringend ist aber das Beschleunigen effizienter Genehmigungsverfahren zum Anschluss der bestehenden Projektierungen.

Empfehlung 6: Verbesserung der Ansiedlungsbedingungen für Rechenzentren, z. B. durch vereinfachte Planungs- und Genehmigungsverfahren für Errichtung und Anschluss, beschleunigten Ausbau erneuerbarer Energiekapazitäten und Unterstützung nachhaltiger USV.

Einrichtung von Anreizmodellen für die Abwärmenutzung sowie energieeffiziente Kühlsysteme: Ein entscheidendes, bislang meist ungenutztes Energieeffizienzpotenzial liegt in der Abwärme von Rechenzentren: Trotz allgemeiner Vorgaben gemäß EnEfG sowie vielfältiger technischer Möglichkeiten wird die tatsächliche Nutzbarkeit durch strukturelle, technische und wirtschaftliche Hürden begrenzt. Um das Potenzial möglichst weitgehend zu erschließen, sollten Intermediäre für Abwärmenutzung aufgebaut und unterstützt werden, aber auch spezifische finanzielle Anreize für wenig lukrative Investitionen in außerbetriebliche Abwärmenutzung geschaffen werden. Denn heute sind andere Wärmeversorgungsoptionen für EVU meist attraktiver. Darüber hinaus würden angepasste Amortisationszeiträume für solche Investitionen konkrete wirtschaftliche Anreize setzen und die Umsetzung signifikant beschleunigen. Die Förderung von Pilotprojekten und Forschungsinitiativen im Bereich innovativer Energie- und Infrastrukturkonzepte kann ebenfalls einen Beitrag leisten, um Good Practice zu etablieren und zu kommunizieren. Ein weiterer wichtiger Punkt ist die Differenzierung der Abwärmenutzung von Rechenzentren. Ältere Rechenzentren in urbanen Metropolregionen sind oft durch ineffizientere Kühlsysteme geprägt, die nur mit hohem Kapitalaufwand zu ersetzen sind. Hier sollten Anreize zur Abwärmenutzung geschaffen werden, welche die Energierückgewinnung vereinfachen und somit die Ineffizienz der Kühlsysteme, eventuell auch ohne diese zu ersetzen, energetisch nutzbar machen. Kommunen sollten zusätzlich unterstützt werden, wenn sie Abwärme aus Rechenzentren integrieren, z. B: bei der konkreten Planung. Die Nutzung von industrieller Abwärme (z. B. aus Rechenzentren), die ansonsten in die Umwelt abgeführt wird, sollte, wenn aus erneuerbaren Energien stammend, immer als klimaneutral gelten, während für fossile Kraft-Wärme-Kopplung ehrliche Emissionsfaktoren hinterlegt werden sollten. Hierzu sollte auch das etablierte Bewertungsverfahren der Fernwärme (insb. die Berechnung anhand eines Verdrängungsstrommixes) kritisch betrachtet werden, welches der fossilen KWK bis heute ermöglicht, die hauptsächlich fossil erzeugte Wärme als weitgehend CO_2-neutral zu deklarieren.

Empfehlung 7: Aufbau, Bereitstellung und Pflege eines Zentralen Wärmenetzregisters gemeinsam mit Bundesländern und Kommunen, um die Abwärme von Rechenzentren und weiteren Industrien und die lokalen Wärmenetze und Wärmenetzplanungen miteinander zu verbinden und Konzeption und Umsetzung weiterer Maßnahmen zur Generierung von Good Practice-Beispielprojekten der Abwärmenutzung an Rechenzentren.

Empfehlung 8: Entwicklung bzw. Ausweitung von Unterstützungsangeboten für Kommunen, EVU und Rechenzentren, um Pilotprojekte und Best Practice zu generieren und auch weniger wirtschaftliche Abwärmeprojekte zu realisieren. Besserstellung der Abwärme aus mit erneuerbarem Strom versorgten Rechenzentren gegenüber klassischen Wärmeerzeugern in Nah- und Fernwärmenetzen.

Förderung der Flexibilisierung von Stromverbrauch und -einspeisung durch Nachfrage-Anreizsysteme: Zugleich sollte die Flexibilisierung und Integration von Rechenzentren in den Strommarkt vorangetrieben werden, unterstützt durch gezielte Anreize für eine geografisch sinnvolle Ansiedlung. Dies umfasst z. B. die Steigerung der Nachfrageflexibilität durch deutlich variierende Strompreise in Abhängigkeit vom lokalen EE-Stromangebot. Ein weiterer Ansatzpunkt kann dabei sein, die in Rechenzentren verfügbaren Stromerzeugungs- und Stromspeicherkapazitäten netzdienlich bereitzustellen. Die Herausforderungen, die durch neue Stromgroßverbraucher entstehen – Rechenzentren gehören hier zu den prominentesten Akteuren – sind bereits im Netzentwicklungsplan adressiert, erfordern allerdings eine genauere Abstimmung. Geplante Projekte von Stromgroßverbrauchern sollten stärker in die strategische Netzplanung integriert werden, indem Bottom-up-Erkenntnisse zu geplanten Anschlussprojekten mit den Top-down-Prognosen der Stromverbrauchsmodellierung abgeglichen werden. Insbesondere bei besonders großen Rechenzentren, wie Hyperscale-Rechenzentren mit einer Leistung von mehr als 30 MW, sollten Netzanschlussanträge angepasst werden, um sowohl die Netzverträglichkeit als auch die Flexibilisierungspotenziale dieser Anlagen zu bewerten. Dabei sollten Mindestanforderungen an ihre Fähigkeit zur Demand-Response (Nachfrageflexibilität) berücksichtigt werden.

Empfehlung 9: Forcierung der Einführung von Flexibilitätsmechanismen für Rechenzentren und weitere Industrien, nicht nur mittels variabler Stromtarife, sondern auch durch die Motivation der Rechenzentren und der Stromnetzbetreiber, die USV stärker zur Bereitstellung von Flexibilität einzubeziehen.

Nachhaltigkeitsstandards für digitale Infrastruktur weiter fördern, fordern und entwickeln: Die Schaffung eines Wettbewerbsvorteils für nachhaltige Anbieter von Rechenzentrumsleistungen und anderer digitaler Infrastruktur ist möglich. Der Wettbewerb um die Produktion von nachhaltigen und kosteneffizienten digitalen Ressourcen sollte durch verschiedene Maßnahmen unterstützt werden. Der beschleunigte Ausbau erneuerbarer Energien zur Deckung des steigenden Strombedarfs von Rechenzentren sollte durch längerfristig attraktive politische Rahmenbedingungen für erneuerbare Energien unterstützt werden. Mit seiner starken und vielfältigen Expertise und der guten industriellen Basis in den Bereichen der elektrischen und mechanischen Infrastruktur kann Deutschland sich international als Innovationstreiber für Rechenzentrumsinfrastruktur – insbesondere bei der Gebäudetechnik, der Automatisierung und der Energiesystem-Integration – platzieren. Hierfür sollten gezielt Forschung, Innovationen und spezifische Aus- und Weiterbildung für Rechenzentrums-Infrastruktur unterstützt werden. Auch hohe regulatorische Anforderungen an die Energieeffizienz und die Energiesystem-Integration von Rechenzentren können für Innovationen in der Branche förderlich sein, insbesondere wenn dafür neue Technologien entwickelt und eingesetzt werden müssen. Es sollte dabei jedoch darauf geachtet werden, nur das praktikabel Machbare zu fördern, um Märkte nicht zu überfordern bzw. zu stark zu regulieren. Trotzdem ist es ratsam, ehrgeizige Ziele für die Energieeffizienz von Rechenzentren zu setzen. Allerdings fehlt es aktuell an präzisen Indikatoren und Messgrößen, die über standardisierte Werte wie PUE hinausgehen und die eigentliche Effizienz des Rechenzentrumsbetriebs der IT-Infrastruktur beschreiben (wie z. B. SIC, ITEU, ITEE etc.). Hier sollten auch die Energieeffizienz, die Auslastung und der Nutzungsgrad der IT-Infrastruktur mit in die Betrachtung der Energieeffizienz von Rechenzentren einbezogen werden. Um die globale Wettbewerbsfähigkeit Deutschlands im Bereich nachhaltiger digitaler Infrastruktur zu stärken, könnte der Übergangsprozess hin zu höheren Effizienzstandards durch mehr Flexibilität, insbesondere beim Umgang mit den PUE-Grenzwerten, gestaltet werden. Denkbar wären flexiblere

Anforderungssysteme, welche auch Maßnahmen wie IT-Energieeffizienz-Verbesserungen, Abwärmenutzung, über die Anforderungen hinausgehende Nutzung von Erneuerbaren Energien oder netzdienliches Verhalten berücksichtigen. Hierbei bieten internationale Ansätze wertvolle Orientierung: So zeigt Irland beispielsweise, wie die Integration von Rechenzentren in das Stromsystem durch verpflichtende Flexibilitätsmaßnahmen die Resilienz des Systems deutlich erhöhen kann. Deutschland könnte durch Prüfung und ggf. Übernahme der stromnetzdienlichen Anforderungen für Rechenzentren aus Ländern wie Irland und Dänemark profitieren. Schließlich sind diese Länder in Europa Vorreiter in der erfolgreichen Integration von Rechenzentren in Energie- und Nachhaltigkeitssysteme. Auch EU-weite Ökodesign-Standards sollten weiter und stärker genutzt werden, um die Energieeffizienz in Rechenzentren zu befördern.

Empfehlung 10: Einforderung und Förderung der Nachhaltigkeit in Rechenzentren, durch staatliche Maßnahmen zur Forcierung von Energieeffizienz, Erneuerbaren Energien, Flexibilität, Abwärmenutzung, Ressourceneffizienz, Aus- und Weiterbildung, Innovationen etc. auf nationaler und europäischer Ebene inkl. der Übernahme von Best Practices aus anderen Ländern.

Empfehlung 11: Forcierung der Energieeffizienz in Rechenzentren durch Erhöhung der Datenverfügbarkeit und Transparenz zu spezifischem Energieverbrauch und Energieeffizienz, Leistung und Auslastung von IT-Systemen unter Berücksichtigung weiterer Indikatoren von Energieeffizienz-Metriken.

Integrale Planung von Flächen: Der Bau eines Rechenzentrums setzt heute einen sehr umfangreichen Planungs- und Genehmigungsprozess voraus. Um sowohl Bürokratie zu reduzieren als auch einen Anreiz für eine flexible Ansiedlung zu bieten, könnten Kommunen, unterstützt durch die Länder/den Bund lokal verschiedene Flächenprüfungen vorwegnehmen, eng verzahnt mit der KWP. Ein Standort, der sowohl über Erneuerbare Stromerzeugungs-, als auch über Stromnetzkapazitäten verfügt als auch für ein Potenzial für Abwärmenutzung in Wärmenetzen, könnte vorab auf weitere Kriterien (Umweltverträglichkeit, TA-Lärm etc.) geprüft werden, um als Vorranggebiet ausgewiesen zu werden. Bei konkretem Bedarf könnte dann zügiger und mit einem vereinfachten Genehmigungsverfahren ein Rechenzentrumsbau erfolgen. Mit klaren und bundesweit stärker vereinheitlichten Standards und digitalisierten Prozessen könnte dabei die Planungssicherheit für Investoren und Betreiber deutlich erhöht werden. Eine zentrale Koordinierungsstelle könnte bei Prozessoptimierung und Flächenzuweisung unterstützen. Die Ansiedlung neuer Großrechenzentren in ländlichen Gebieten könnte gezielt mit umfassenden Abwärmekonzepten, erneuerbarer Energieerzeugung und Flexibilitätsbereitstellung gekoppelt werden.

Empfehlung 12: Errichtung einer Zentralen Koordinierungsstelle für Flächenausweisung, um die Identifizierung und Ausweisung besonders geeigneter Flächen für Rechenzentren zu erleichtern und weitere, dafür erforderliche kommunale Entwicklungs- und Wärmepläne anzureizen.

Regulatorische Innovationen: Die Regulierung und Steuerung der Entwicklung von Rechenzentren sind globale Herausforderungen, an denen viele Länder arbeiten. Deutschland hat hier bereits mit dem EnEfG eine Vorreiterrolle eingenommen – es sollte weiter an innovativen Ansätzen für die Regulierung gearbeitet werden, um nachhaltiges Wachstum von digitaler Infrastruktur zu realisieren. Dialogplattformen sind der Schlüssel zur Weiterentwicklung. Die Zusammenarbeit zwischen Rechenzentrumsbetreibern, Kommunen, EVU und politischen Akteuren ist derzeit noch unzureichend strukturiert. Regelmäßige Austauschforen könnten zentrale Herausforderungen adressieren und

gemeinsame Strategien entwickeln. Diese Plattformen sollten als Schnittstellen dienen, um Wissenstransfer, Planungssicherheit und politische Unterstützung zu fördern. Auch mit Blick auf die Weiterentwicklung regulatorischer Rahmenbedingungen können breitere Konsultationsprozesse und Dialogformate die nötigen Strukturen bieten, um die Bedürfnisse und Expertise der Rechenzentrumsbranche bei der nachhaltigen Gestaltung gesetzlicher Vorgaben zu nutzen. Hierdurch würden Marktakteure frühzeitig in Policy-Prozesse einbezogen, deren Planungssicherheit gestärkt und regulatorische Unsicherheiten abgebaut. Dabei sollten insbesondere auch technische Rahmenbedingungen und Restriktionen stärker im Rahmen politischer Prozesse berücksichtigt werden. Zentrale Austauschplattformen, Arbeitsgruppen für spezifische Herausforderungen und Kooperationsprojekte sind demnach wichtige Bausteine für eine erfolgreiche Stärkung des Rechenzentrumsstandorts Deutschland.

Empfehlung 13: Bereitstellung einer Dialogplattform zwischen Unternehmen und Verbänden, Kommunen, Strom- und Wärmenetzbetreibern sowie Bundes- und Landesregierungen, um einen regelmäßigen Austausch über bestehende und denkbare hemmende und fördernde Faktoren der Rechenzentrumsentwicklung in Deutschland zu sprechen und gemeinsam an der Verbesserung der Rahmenbedingungen zu arbeiten.

Marktverständnis entwickeln und Wachstum fördern: Rechenzentren sind eine Wachstumsbranche mit hoher Bedeutung für die Wirtschaftsentwicklung in Deutschland. Mit dem EnEfG wurde im Jahr 2023 eine Meldepflicht für Betreiber von Rechenzentren mit einer nichtredundanten Anschlussleistung von 300 kW eingeführt. Auf Basis der Ergebnisse der ersten Jahresmeldungen für das Jahr 2023 zeigt sich, dass ein erheblicher Anteil der Betreiber von Rechenzentren ihrer Meldepflicht noch nicht oder nicht vollständig nachgekommen sind, so dass bei der Analyse dieser Daten aufgrund der Unvollständigkeit weiter hohe Unsicherheiten bestehen. Im Rahmen der zum 1. März 2025 fälligen verpflichtenden Übermittlung der Jahresdaten 2024 bzw. für Rechenzentren zwischen 300 - 500 kW bis zum 1. Juli 2025, sollte zusammen mit den Verbänden der Rechenzentrumsbrache erneut auf die Verpflichtung zur Meldung der Daten hingewiesen werden. Hierbei sollten auch Industrieunternehmen, die firmeneigene Rechenzentren betreiben, erneut über die gesetzliche Verpflichtung aufgeklärt werden. Die ursprünglich im Entwurf des EnEfG enthaltene Verpflichtung für IT-Betreiber, ebenfalls Daten an das RZReg zu melden, könnte wichtige zusätzliche Informationen für das Verständnis der Digitalbranche liefern. Wenn eine Verpflichtung für eine Datenerhebung aktuell politisch nicht umsetzbar ist, bleibt als weitere Möglichkeit eine freiwillige Erhebung sowie eine aktive Erhebung mittels Marktbefragung. Denn diese Informationen und Kennzahlen liefern die Grundlagen für eine zielgerichtete Weiterentwicklung des Digitalstandortes Deutschland. Dieses dynamische Wachstum gilt es zu begleiten und zu unterstützen, damit Deutschland seine Vorreiterrolle in Europa weiter ausbauen kann.

Empfehlung 14: Verstärkte Kommunikation zu den Datenlieferpflichten der RZ-Betreiber gemäß EnEfG und Erhebung und Bereitstellung aktueller, zuverlässiger und statistisch validierter Daten zu betriebenen Rechenzentren und IT-Systemen, um eine zielgerichtete Begleitung und Unterstützung der Entwicklung der Rechenzentrumsbranche sicherzustellen.

Ausbau der Ausbildungs- und Weiterbildungsangebote: Der Betrieb und Ausbau von Rechenzentren erfordert hochqualifiziertes Personal, sowohl im technischen als auch im unterstützenden Bereich. Da der RZ-Markt einer der am stärksten wachsenden Märkte ist, bietet er bei ausreichend verfügbaren Fachkräften große Chancen für Beschäftigungs- und

Wertschöpfungseffekte in Deutschland. Deshalb sollte in diesem Themenfeld verstärkt in die Ausbildung von Fachkräften für Rechenzentrumstechnik und IT-Infrastrukturbetrieb sowie in die Weiterbildung von Handwerkern in Bereichen wie Elektrotechnik und Klimatechnik investiert werden. Ergänzend könnten gezielte Zuwanderungsinitiativen helfen, internationale Experten für den deutschen Markt zu gewinnen. Begleitend sollte auch in den Kompetenzauf- und -ausbau hinsichtlich digitaler Infrastruktur in Regierungs- und Verwaltungseinheiten, sowohl auf Bundes-, Landes- als auch kommunaler Ebene, investiert werden. Die Aus- und Weiterbildung von Fachkräften sollte durch die Verankerung entsprechender Inhalte in bestehende Aus- und Fortbildungsstrukturen, als auch durch eine dahingehende Schwerpunktsetzung bei der gezielten Anwerbung von Fachkräften aus dem Ausland, gestärkt werden. Auch internationale Austauschprogramme und die Entwicklung spezialisierter Bildungsprogramme können insbesondere der Entwicklung von IT-Infrastruktur-Kompetenz zuträglich sein. Dabei bedarf es auch disziplinübergreifender Studiengänge, um die gesamte Themenbreite im Bereich der digitalen Infrastrukturen zukunftsorientiert abzudecken.

Empfehlung 15: Stärkung der Aus- und Weiterbildung sowie Fachkräftesicherung im IKT-Themenfeld gemeinsam mit Bundesländern, Wissenschaft und Forschung, um nachhaltige IKT und Digitalisierung als fundamentale Bausteine der Zukunftsfähigkeit Deutschlands zu berücksichtigen.

Diese Maßnahmen und Empfehlungen bilden eine zentrale Grundlage für eine umfassende, koordinierte und zukunftssichere Weiterentwicklung des Rechenzentrumssektors in Deutschland. Sie berücksichtigen gleichermaßen die ökonomischen, ökologischen und technologischen Anforderungen, denen sich der Standort Deutschland stellen muss. Durch ihre Umsetzung könnte nicht nur die Wettbewerbsfähigkeit Deutschlands im Bereich digitaler Infrastrukturen entscheidend gestärkt, sondern auch ein Beitrag zur europäischen digitalen Souveränität und zur globalen Nachhaltigkeit geleistet werden.

Abbildungsverzeichnis

Abbildung 1: Entwicklung der Kapazitäten von Rechenzentren und kleineren IT-Installationen in Deutschland in den Jahren 2010 bis 2024 und Prognose bis 2030. Quelle: Hintemann et al. (2024a) 5

Abbildung 2: Aktuelle und in konkreter Planung befindliche IT-Anschlussleistung von Rechenzentren in verschiedenen deutschen Regionen. Quelle: Hintemann et al. (2024a) 6

Abbildung 3: Abschätzung der Entwicklung der RZ-Kapazitäten in den USA, in China, Europa und Deutschland 2024 und 2030. Quelle: Hintemann et al. (2024a) 7

Abbildung 4: Entwicklung des Strombedarfs der Rechenzentren und kleineren IT-Installationen in Deutschland in den Jahren 2010 bis 2024 und Langfristprognose bis 2045. Quelle: Berechnungen Borderstep 9

Abbildung 5: Vereinfachte Darstellung der Wertschöpfung durch digitale Infrastruktur. Quelle: Eigendarstellung SDIA 18

Abbildung 6: Abschätzung der Entwicklung der weltweiten Kapazitäten der Rechenzentren und kleineren IT-Installationen in den Jahren 2022 bis 2030. Quelle: Berechnungen Borderstep 24

Abbildung 7: Geschätzte Stromverbräuche der Rechenzentren in den USA, der EU und China in den Jahren 2022 und 2026. Quelle: IEA (2024) 27

Abbildung 8: Entwicklung der Kapazitäten von Rechenzentren und kleineren IT-Installationen in Deutschland in den Jahren 2010 bis 2024 und Prognose bis 2030. Quelle: Hintemann et al. (2024a) 29

Abbildung 9:Entwicklung der Kapazitäten von Rechenzentren und kleineren IT-Installationen in Deutschland in den Jahren 2010 bis 2024 - Aufteilung in RZ-Größenklassen. Quelle: Hintemann et al. (2024a) 30

Abbildung 10: Aktuelle und in konkreter Planung befindliche IT-Anschlussleistung von Rechenzentren in verschiedenen deutschen Regionen. Quelle: Hintemann et al. (2024a) 31

Abbildung 11: Entwicklung der Kapazitäten von Rechenzentren in Deutschland mit Anteilen von Cloud, Colocation, On Premise und Edge Rechenzentren (IT-Anschlussleistung) in den Jahren 2010 bis 2024. Quelle: Hintemann et al. (2024a) 33

Abbildung 12: Entwicklung der Kapazitäten der Rechenzentren in den Jahren 2016 bis 2024 in ausgewählten europäischen Ländern. Quelle: Berechnungen Borderstep 34

Abbildung 13: Abschätzung der Entwicklung der RZ-Kapazitäten in den USA, in China, Europa und Deutschland 2024 und 2030. Quelle: Hintemann et al. (2024a) 35

Abbildung 14: Entwicklung des Stromverbrauchs der Rechenzentren und kleineren IT-Installationen in Deutschland in den Jahren 2010 bis 2024. Quelle: Hintemann et al. (2024a) 39

Abbildung 15: Entwicklung des Stromverbrauchs der Rechenzentren und kleineren IT-Installationen in Deutschland in den Jahren 2010 bis 2024 und Prognose bis 2030 inkl. Abschätzung der Auswirkungen des EnEfG..... 41

Abbildung 16: Entwicklung des Strombedarfs der Rechenzentren in Deutschland in den Jahren 2010 bis 2024 und Langfristprognose bis 2045. Quelle: Berechnungen Borderstep..... 42

Abbildung 17: Entwicklung des PUE-Werts in Deutschland in den Jahren 2010 bis 2024 und Prognose bis 2030. Quelle: Berechnungen Borderstep..... 44

Abbildung 18: Aufteilung der Treibhausgasemissionen (in $CO_{2äq}$) eines Rechenzentrums auf verschiedene Emissionsquellen (Beispiel-RZ mit 2,5 MW IT-Anschlussleistung). Quelle: Hintemann (2024a)..... 46

Abbildung 19: Entwicklung der durch den Stromverbrauch der Rechenzentren und kleineren IT-Installationen bedingten Treibhausgasemissionen in Deutschland in den Jahren 2010 bis 2024 und Prognose bis 2030 (* 2024 Schätzwert für den nationalen Strommix). Quelle: Hintemann et al. (2024a). 47

Abbildung 20: Verläufe der 110kV Hochspannungsleitungen in DE. Quelle: Bundesnetzagentur (2023) 50

Abbildung 21: Deutsches Glasfasernetz. Quelle: NGN Fiber Network GmbH & Co. KG (2024)... 50

Abbildung 22: Einbindungskosten von Abwärme aus Rechenzentren in € pro MWh in Abhängigkeit der Entfernung (in Meter), Quelle: Berechnung Fraunhofer ISI 65

Abbildung 23: Vereinfachte Darstellung des Digitalsektors. Quelle: Eigendarstellung SDIA..... 70

Abbildung 24: Entwicklung des Umsatzes in der IKT-Branche in Deutschland in den Jahren 2009 bis 2022 (in Milliarden Euro). Quelle: Berechnungen des ZEW, 2023. Werte am aktuellen Rand approximiert. 71

Abbildung 25: Übersicht der Kennzahlen zur volkswirtschaftlichen Wertschöpfung in Mitarbeitenden in Rechenzentren. Quelle: Eigendarstellung SDIA 78

Abbildung 26: Wertschöpfungskette und Spillover-Effekte der Digitalwirtschaft: Von der Infrastruktur über die Veredelung durch digitale Ressourcen bis hin zu den resultierenden Produkten und deren Auswirkungen auf die Gesamtwirtschaft. Quelle: Eigendarstellung SDIA... 84

Abbildung 27: Prozentualer Anteil der Kosten für digitale Ressourcen im Verhältnis zum Umsatz großer Digitalunternehmen. Quelle: Wang, Casado (2021)..... 85

Abbildung 28: Screenshot Amazon mit Auswahloption für die Miete einer EC2-Instanz (Rechenleistung mit optionalem Datenspeicher) aus dem Kostenkalkulator von Amazon. Dargestellt sind die Zahlungsmodelle (On-Demand, Spot, Reserved). Quelle: Amazon (2024) ... 86

Abbildung 29: Quartalsentwicklung in Prozent der führenden Cloud-Anbieter in den Jahren 2018 bis 2022. Quelle: Ofcom-Analyse 87

Abbildung 30: Modelle für digitale Ökosysteme: Vergleichende Darstellung der Entwicklungskonzepte für Groß- und Industriestädte, regionale Zentren und ländlichen Raum mit unterschiedlicher Ausprägung der IT-Infrastruktur und Vernetzung. Quelle: Eigendarstellung SDIA 89

Abbildung 31: Server-Stromverbrauch in Watt in Abhängigkeit vom Nutzungsgrad in Prozent. Die Grafik verdeutlicht den Zusammenhang zwischen Serverauslastung und Energieverbrauch eines HPE Synergy 480 Gen10 Plus Servers, basierend auf Daten der SPEC Organisation. Quelle: Green Coding Solution, Cloud Energy Modell (2022) 92

Abbildung 32: Weltweite durchschnittliche jährliche PUE-Werte in den Jahren 2007 bis 2024 (Uptime Institute, 2024) 108

Abbildung 33: Weltweite gewichtete durchschnittliche PUE-Werte nach IT-Kapazität in kW/ MW (Uptime Institute, 2024) 109

Abbildung 34: Verlauf des Digitalisierungsgrades der verschiedenen Wirtschaftsabschnitte in Deutschland von 2024 bis 2045. Quelle: Berechnungen der SDIA 121

Abbildung 35: Berechneter Bedarf an digitalen Ressourcen (Standard-Server-Äquivalente) auf Basis der Schätzung des Digitalisierungsgrades der Wirtschaftsabschnitte. Quelle: Berechnungen SDIA 122

Abbildung 36: Prognose der Entwicklung des Bedarfs an digitalen Ressourcen (Standard-Server-Äquivalente) in den Jahren 2024 bis 2045 auf Basis der Schätzung des Digitalisierungsgrades der Wirtschaftsabschnitte. Quelle: Berechnungen SDIA 123

Abbildung 37: Prognose der Entwicklung der Unterschiede des Bedarfs an Standard-Server-Äquivalenten durch verschiedene Effizienzmaßnahmen in den Jahren 2024 bis 2045. Quelle: Berechnungen SDIA 124

Abbildung 38: Prognose der Entwicklung der Nachfrage an Standard-Server-Äquivalenten in den Jahren 2024 bis 2045, die aus dem vorhergesagten Wachstum der Digitalwirtschaft bis 2045 entstehen könnten. Quelle: Berechnungen SDIA 125

Abbildung 39: Wertschöpfungsstufen der digitalen Transformation: Von grundlegenden Inputs über physische Infrastruktur zu IT-Diensten bis hin zur Digitalwirtschaft. Die Darstellung zeigt den mehrstufigen Veredelungsprozess digitaler Ressourcen und deren Transformation in Digitalisierungsdienstleistungen und digitale Wirtschaftsaktivitäten. 128

Abbildung 40: Entwicklung des Außenhandels mit IKT-Gütern in Deutschland mit ausgewählten Ländern in den Jahren 2020 bis 2023. Quelle: Statistisches Bundesamt 130

Abbildung 41: Übersicht über die Stromverbraucher-Infrastruktur in Rechenzentren - Darstellung der Aufteilung zwischen primärem Stromverbrauch (IT-Infrastruktur, IKT-Equipment,

Betriebssoftware) und sekundärem Stromverbrauch (Rechenzentrumsgebäude, Kühlung, USV) in der digitalen Infrastruktur. 133

Abbildung 42: Entwicklung der Stromaufnahme von 1- und 2-Chip-basierten Servern in den Jahren 2004 bis 2024, eigene Darstellung, Daten von „SPEC.org “ (SPEC, 2024) 134

Abbildung 43: Entwicklung der Energieeffizienz (Transaktionen pro Watt) von 1- und 2-CPU-basierten Servern in den Jahren 2004 bis 2024, eigene Darstellung, Daten von „SPEC.org “ (SPEC, 2024) 135

Abbildung 44: Kühltechnologie in Abhängigkeit der Leistungsdichte von Racks. Quelle: Eigene Darstellung basierend auf Daten von Vertiv 136

Abbildung 45: Handlungsempfehlungen für die strategische Entwicklung des RZ-Standorts Deutschland 143

Literatur

Althoff, E.; Hobohm, J.; Krampe, L.; Wendring, P.; Wünsch, A. (2024). Prognos Energieatlas 2024: Grüner Strom Atlas. in Kooperation mit der Süddeutschen Zeitung. Online verfügbar unter https://www.prognos.com/sites/default/files/2024-01/Prognos%20Energieatlas_1-2024_0.pdf, zuletzt geprüft am 14.08.2024.

Amazon (Hg.) (2023a, Mai 18). AWS to invest INR 1,05,600 crores (US $12.7 billion) into cloud infrastructure in India. Abgerufen 16. November 2024, von IN Press Center website: https://press.aboutamazon.com/in/2023/5/aws-to-invest-inr-1-05-600-crores-us-12-7-billion-into-cloud-infrastructure-in-india.

Amazon (Hg.) (2023b). Harnessing the power of plants to decarbonise our data centres. Online verfügbar unter https://www.aboutamazon.eu/news/sustainability/harnessing-the-power-of-plants-to-decarbonise-our-data-centres, zuletzt aktualisiert am 01.11.2023, zuletzt geprüft am 06.12.2024.

Amazon (Hg.) (2024, Mai 15). AWS plant langfristige Investitionen in Höhe von 7,8 Milliarden Euro in die AWS European Sovereign Cloud in Brandenburg. Abgerufen 8. Juli 2024, von DE Press Center website: https://press.aboutamazon.com/de/arbeitsplaetze-und-investitionen/2024/5/aws-plant-langfristige-investitionen-in-hoehe-von-7-8-milliarden-euro-in-die-aws-european-sovereign-cloud-in-brandenburg

American Society of Heating, Refrigerating and Air-Conditioning Engineers (Hg.) (2016). ASHRAE TC9.9 Data Center Power Equipment Thermal Guidelines and Best Practices. Online verfügbar unter https://www.ashrae.org/file%20library/technical%20resources/bookstore/ashrae_tc0909_power_white_paper_22_june_2016_revised.pdf, zuletzt geprüft am 10.10.2024.

American Society of Heating, Refrigerating and Air-Conditioning Engineers (Hg.) (2021). 2021 Equipment Thermal Guidelines for Data Processing Environments. ASHRAE TC 9.9 Reference Card. Online verfügbar unter https://www.ashrae.org/file%20library/technical%20resources/bookstore/supplemental%20files/thermgdlns-5th-r-e-refcard.pdf, zuletzt geprüft am 06.08.2024.

American Society of Heating, Refrigerating and Air-Conditioning Engineers (Hg.) (2011). ASHRAE TC 9.9 2011 Thermal Guidelines for Data Processing Environments – Expanded Data Center Classes and Usage Guidance. Appendix C. IT Equipment Reliability Data. Online verfügbar unter https://airatwork.com/wp-content/uploads/ASHRAETC99.pdf, zuletzt geprüft am 10.10.2024.

Anderson, D.; Cader, T.; Darby, T.; Gruendler, N.; Hariharan, R.; Holler, A. et al. (2008). A Framework for Data Center Energy Productivity. White Paper #13. Hg. v. The Green Grid. Online verfügbar unter https://www.greenbiz.com/sites/default/files/document/GreenGrid-Framework-Data-Center-Energy-Productivity.pdf, zuletzt geprüft am 31.07.2024.

Arizton (Hg.) (2024). United Kingdom Data Center Market Size & Share Analysis, Growth. Abgerufen 11. Juli 2024, von Arizton Advisory & Intelligence website: https://www.arizton.com/market-reports/uk-data-center-market-investment-analysis

Avelar, V.; Martin Z. (2018). Battery Technology for Data Centers: VRLA vs. Li-ion. International Journal of Science and Innovative Technology 1, no. 1 (2018). Online verfügbar unter https://ph01.tci-thaijo.org/index.php/IJSIT/article/view/146544/108039.

Babar, K. (2024, Mai 8). India’s data center industry set to double capacity by 2026, attract investments. The Economic Times. Abgerufen von https://economictimes.indiatimes.com/industry/services/property-/-cstruction/indias-data-center-industry-set-to-double-capacity-by-2026-attract-investments/articleshow/109952430.cms?from=mdr

Ballandt, J. (2017). Datacenterkühlung mit hybriden Rückkühlern. Hg. v. JAEGGI Hybridtechnologie AG. Online verfügbar unter https://www.jaeggi-hybrid.ch/fileadmin/sites/jaeggi/ch/Fachartikel/Datencenterk%C3%BChlung.pdf, zuletzt geprüft am 06.08.2024.

Bansen, R.; Clasen, K. (2021). Blei- und Lithium-Ionen-Batterien in Rechenzentren. Ein Vergleich. Hg. v. Bitkom e.V. Online verfügbar unter https://www.bitkom.org/sites/default/files/2021-01/20210119_bitkom_leitfaden_blei-_li-ionen-batterien_in_rechenzentren.pdf, zuletzt geprüft am 08.07.2024.

Baringa (Hg.) (2022). Green Data: A vision for sustainable data centres in Ireland. Online verfügbar unter https://bitpower.ie/images/Reports/Baringa_Green_Data_-_Full_Report.pdf

Barroso, L. A.; Hölzle, U. (2007). The Case for Energy-Proportional Computing. In: Computer 40 (12), S. 33–37. DOI: 10.1109/MC.2007.443.

Bartos, K. E.; Dede, C.; Brehm, C.; Kaiser, P. (2024). Szenariorahmen zum Netzentwicklungsplan Strom 2037/2045, Version 2025. Entwurf der Übertragungsnetzbetreiber. Online verfügbar unter https://www.netzentwicklungsplan.de/sites/default/files/2024-07/Szenariorahmenentwurf_NEP2037_2025_1.pdf, zuletzt geprüft am 01.10.2024.

Beucker, S. ; Doderer, H. ; Funke, A. ; Kondziella, H. ; Koch, C. ; Hartung, J.; Maeding, S.; Medert, H.; Meyer-Braune, G.; Rath, M.; Rogler N. (2021). Flexibilität, Markt und Regulierung. [Synthesebericht]. WindNODE-Konsortium. Online verfügbar unter https://www.borderstep.de/wp-content/uploads/2021/01/FMR_ES.pdf

Bitkom e.V. (2013). Betriebssicheres Rechenzentrum: Leitfaden. Bitkom e.V. Online verfügbar unter https://www.bitkom.org/Bitkom/Publikationen/Betriebssicheres-Rechenzentrum.html

Bitkom e.V. (2015). Energieeffizienz in Rechenzentren: Leitfaden. Bitkom e.V. Online verfügbar unter https://www.bitkom.org/Bitkom/Publikationen/Leitfaden-Energieeffizienz-in-Rechenzentren.html

Bizo, D. (2023). Global PUEs — are they going anywhere? Hg. v. Uptime Institute. Online verfügbar unter https://journal.uptimeinstitute.com/global-pues-are-they-going-anywhere/, zuletzt aktualisiert am 04.12.2023, zuletzt geprüft am 12.11.2024.

Blauer Engel für Rechenzentren: Power Usage Effectiveness (PUE). Online verfügbar unter https://be-rechenzentren.de/vergabekriterien/power-usage-effectiveness-pue/

BloombergNEF, Statkraft & Eaton (2021). Data Centers and decarbonization: Unlocking Flexibility in Europe’s Data Centers. Online verfügbar unter https://www.eaton.com/gb/en-gb/company/news-insights/energy-transition/bnef-data-centres-and-decarbonisation-study.html?source=post:1427248746593282584

Boulos, S.; Nuttall, C.; Harrison, B.; Moura, P.; Jehle, C. (2014). ErP Lot 27 – Uninterruptible Power Supplies Preparatory Study - Final Report for European Commission, DG Energy. Ricardo-AEA/R/ED56828, Issue Number 1. Online verfügbar unter https://netzwerke.bam.de/Netzwerke/Content/DE/Standardartikel/Evpg/Evpg-Produktgruppen/Elektronik/ener-27-usv.html

Brinkhaus, M.; Schwartz, J. (2024). CFD Policy Briefing—Analyse der Stellschrauben im CfD Design. Energy Brainpool. Online verfügbar unter https://green-planet-energy.de/fileadmin/docs/publikationen/Studien/policy-paper-cfd-policy-energybrainprool-green-planet-energy.pdf

Brochard, L.; Kamath, V.; Corbalán, J.; Holland, S.; Mittelbach, W.; Ott, M. (Hg.) (2019): Energy-Efficient Computing and Data Centers. Newark: John Wiley & Sons Incorporated.

Brocklehurst, F. (2022). International review of energy efficiency in Data Centres for IEA EBC Building Energy Codes Working Group. Hg. v. IEA EBC. Online verfügbar unter https://www.iea-ebc.org/Data/publications/EBC_WG_BECs_Data_Centers_March_2022.pdf, zuletzt geprüft am 10.10.2024.

Brocklehurst, F. (2024, Februar). Policy development on energy efficiency of data centres. Hg. v. The Technology Collaboration Programme on Energy Efficient End-Use Equipment (4E TCP). Online verfügbar unter https://www.iea-4e.org/wp-content/uploads/2024/02/Policy-development-on-energy-efficiency-of-data-centres-draft-final-report-v1.05.pdf

Bundesamt für Wirtschaft und Ausfuhrkontrolle, Bundesstelle für Energieeffizienz (Hg.) (2024). Leitfaden zu den Datenpunkten im Rechenzentrumsregister (RZReg): Informationen für Betreiber von Rechenzentren gemäß der §§ 13, 14 Energieeffizienzgesetz. Eschborn. Online verfügbar unter https://www.bmwk.de/RZReg/Downloads/leitfaden-datenpunkte-im-rechenzentrumsregister.pdf?__blob=publicationFile&v=4, zuletzt geprüft am 06.08.2024.

Bundesnetzagentur für Elektrizität, Gas, Telekommunikation, Post und Eisenbahnen (Hg.) (2022). Bedarfsermittlung 2021-2035. Fact Sheet zur Bestätigung des Netzentwicklungsplans Strom. Bonn. Online verfügbar unter https://www.netzausbau.de/SharedDocs/Downloads/DE/Bedarfsermittlung/2035/NEP/Fact_Sheet.pdf?__blob=publicationFile, zuletzt geprüft am 01.10.2024.

Bundesnetzagentur für Elektrizität, Gas, Telekommunikation, Post und Eisenbahnen (Hg.) (2023): Strom Bericht. Zustand und Ausbau der Verteilernetze 2022. Bonn. Online verfügbar unter https://www.bundesnetzagentur.de/SharedDocs/Downloads/DE/Sachgebiete/Energie/Unternehmen_Institutionen/NetzentwicklungUndSmartGrid/ZustandAusbauVerteilernetze2022.pdf?__blob=publicationFile&v=1, zuletzt geprüft am 01.10.2024.

Bundestag: Gesetz zur Steigerung der Energieeffizienz in Deutschland. Energieeffizienzgesetz - EnEfG. Online verfügbar unterhttps://www.gesetze-im-internet.de/enefg/EnEfG.pdf, zuletzt geprüft am 09.01.2023.

CBRE (2023). Europe Data Centres: Frankfurt, London, Amsterdam, Paris and Dublin. Abgerufen von https://www.cbre.com/insights/figures/europe-data-centres-figures-q4-2022

CBRE (2024a). Europe Data Centres Q1 2024. Abgerufen von https://mktgdocs.cbre.com/2299/b287f590-4c28-4208-b1d4-0240db8a47e0-2046017529/v032024/europea-data-centres-figures-q1-2024.pdf

CBRE (2024b, Juni 24). Global Data Center Trends 2024. Abgerufen 18. November 2024, von https://www.cbre.com/insights/reports/global-data-center-trends-2024

CBRE. (2024c). Market Outlook Deutschland 2024. Abgerufen 23. Mai 2024, von https://mediaassets.cbre.com/-/media/cbre/countrygermany/insights/2024_cbre_real_estate_market_outlook_de.pdf?rev=b45a1d9a0750458cbd5e53e0d0245f2f

Chernicoff, D.; Vincent, M. (2024, Juli 25). Supermicro Rises to Meet Massive Growth in Data Center Liquid Cooling Demand. Data Center Frontier. https://www.datacenterfrontier.com/machine-learning/article/55128681/supermicro-rises-to-meet-massive-growth-in-data-center-liquid-cooling-demand?o_eid=7516C0132945C1R&oly_enc_id=7516C0132945C1R&rdx.ident[pull]=omeda|+7516C0132945C1R&utm_campaign=CPS240802017&utm_medium=email&utm_source=DCF+Newsletter

Copenhagen Economics (Hg.) (2020). Inside Finland. Google’s European hyperscale data centres and infrastructure ecosystem. Online verfügbar unter https://copenhageneconomics.com/wp-content/uploads/2021/12/copenhagen-economics-hyperscale-data-centres-and-related-infrastructures-the-case-of-finland_november2020.pdf

Cushman & Wakefield (2024). Global Data Center Market Comparison 2024. Abgerufen 23. Mai 2024, von https://cushwake.cld.bz/2024-Global-Data-Center-Market-Comparison

Data Bridge (2022, Dezember). Marktgröße Und Wachstumspotenzial Für Rechenzentren In Afrika Bis 2030. Abgerufen 16. November 2024, von https://www.databridgemarketresearch.com/de/reports/africa-data-centre-facility-market

Datacentermap.com (2024). Datenbank für Colocation-Rechenzentren. Abgerufen 10. Juli 2014, von https://www.datacentermap.com/

Davies, G. F.; Maidment, G. G.; Tozer, R. M. (2016). Using data centres for combined heating and cooling: An investigation for London. In: Applied Thermal Engineering 94, S. 296–304. DOI: 10.1016/j.applthermaleng.2015.09.111.

Davis, J. (2024). Large data centers are mostly more efficient, analysis confirms. Hg. v. Uptime Institute. Online verfügbar unter https://journal.uptimeinstitute.com/large-data-centers-are-mostly-more-efficient-analysis-confirms/, zuletzt aktualisiert am 07.02.2024, zuletzt geprüft am 07.07.2024.

dena (2022). „dena-Netzstudie III – Stakeholderdialog zur Weiterentwicklung der Planungsverfahren für Energieinfrastrukturen auf dem Weg zum klimaneutralen Energiesystem“. Online verfügbar unter https://www.dena.de/fileadmin/dena/Publikationen/PDFs/2022/Abschlussbericht_dena-Netzstudie_III.pdf

dena (2023). Regulatorische Modelle für eine klimaneutrale Fernwärme in Deutschland. Online verfügbar unter https://www.dena.de/fileadmin/dena/Publikationen/PDFs/2023/Regulatorische_Modelle_fuer_eine_klimaneutrale_Fernwaerme_in_Deutschland.pdf

DER STANDARD (2022, August 25). Irland hat nicht mehr genug Strom für neue Rechenzentren von Amazon und Microsoft. Abgerufen 17. Januar 2024, von DER STANDARD website: https://www.derstandard.de/story/2000138547311/irland-hat-nicht-mehr-genug-strom-fuer-neue-rechenzentren-von

Deutscher Wetterdienst (Hg.) (2024). Klimakarten Deutschland. Online verfügbar https://www.dwd.de/DE/leistungen/klimakartendeutschland/klimakartendeutschland.html, zuletzt geprüft am 20.11.2024.

Digital Chiefs (2023, September 27). Rechenzentrum: Was macht eine Netzersatzanlage und welche Rolle spielen Brennstoffzellen dabei? Online verfügbar unter https://www.digital-chiefs.de/rechenzentrum-was-macht-eine-netzersatzanlage-und-welche-rolle-spielen-brennstoffzellen-dabei/

Digital Realty (Hg.) (2019). Environmental, Social and Governance Report. 2019. Online verfügbar unter https://go2.digitalrealty.com/rs/087-YZJ-646/images/Report_Digital_Realty_2019_ESG_Report.pdf, zuletzt geprüft am 22.11.2024.

DIN (Hg.) (2019). DIN EN 50600-4-2 (VDE 0801-600-4-2): 2019 -08. Online verfügbar unter https://www.normenbibliothek.de/sso/headless/authenticate/login?token=normenbibliothek&requestid=&nonce=f1da629f-7f42-463f-8354-1a93f6c976ee&timestamp=2024-11-21T12%3A39%3A46.635%2B01%3A00&return_url=https://www.normenbibliothek.de/vde-xaveropp/normenbibliothek/start.xav%3F, zuletzt geprüft am 20.11.2024.

Dunnavant, K.; Kätkä, A. Data Center Innovation – wie durch Messgenauigkeit Energieeffizienz ermöglicht wird. Hg. v. Vaisala. Online verfügbar unter https://www.vaisala.com/de/expert-article/data-center-innovation-how-measurement-accuracy-enables-energy-efficiency, zuletzt geprüft am 20.11.2024.

Dürr, B. (2018). IT-Räume und Rechenzentren planen und betreiben. Handbuch der baulichen Maßnahmen und Technischen Gebäudeausrüstung. 2nd. Erkrath: Verlag Bau+Technik.

Ebrahimi, K.; Jones, G. F.; Fleischer, A. S. (2014). A review of data center cooling technology, operating conditions and the corresponding low-grade waste heat recovery opportunities. In: Renewable and Sustainable Energy Reviews 31, S. 622–638. DOI: 10.1016/j.rser.2013.12.007.

EirGrid (Hg.) (2020, Juli 17). Data Centre Connection Offer Process and Policy. Version 2.0. Online verfügbar unter https://cms.eirgrid.ie/sites/default/files/publications/Data-Centre-Connection-Offer-Process-and-Policy_v2_July-2020.pdf

El-Sayed, N.; Stefanovici, I.; Amvrosiadis, G.; Hwang, A. A.; Schroeder, B. (2012). Temperature Management in Data Centers: Why Some (Might) Like It Hot. Online verfügbar unterhttps://www.cs.toronto.edu/~bianca/papers/temperature_cam.pdf, zuletzt geprüft am 10.10.2024.

Emerson Network Power (Hg.) (2011). Energy Logic: Reducing Data Center Energy Consumption by Creating Savings that Cascade Across Systems. Online verfügbar unter https://media.ptsdcs.com/product/emerson-network-power-white-paper-energy-logic-reducing-data-center-energy-consumption-by-creating-savings-that-cascade-across-systems/, zuletzt geprüft am 11.03.2024.

Etheredge, J. (2024). Fun Friday: How Much Solar Power Would It Actually Take to Power a Hyperscale Data Center. Hg. v. Simple Thread. Online verfügbar unter https://www.simplethread.com/fun-friday-how-much-solar-power-would-it-actually-take-to-power-a-hyperscale-data-center/, zuletzt aktualisiert am 09.02.2024, zuletzt geprüft am 20.11.2024.

ETCIO.com (2024, Mai 25). India’s data center boom to drive USD 5.7 billion investments by 2026—ET CIO. Abgerufen 16. November 2024, von ETCIO.com website: https://cio.economictimes.indiatimes.com/news/data-center/indias-data-center-boom-to-drive-usd-5-7-billion-investments-by-2026/110404441

ETSI (Hg.) (2019). Environmental Engineering (EE); Energy Efficiency measurement methodology and metrics for servers. ETSI EN 303 470.

Europäische Kommission (2019). Verordnung (EU) 2019/424 der Kommission vom 15. März 2019 zur Festlegung von Ökodesign-Anforderungen an Server und Datenspeicherprodukte gemäß der Richtlinie 2009/125/EG des Europäischen Parlaments und des Rates und zur Änderung der Verordnung (EU) Nr. 617/2013 der Kommission. Online verfügbar unter https://netzwerke.bam.de/Netzwerke/Content/DE/Standardartikel/Evpg/Evpg-Produktgruppen/Elektronik/grow-09-enterprise-server.html

Europäische Kommission (2024). COMMISSION DELEGATED REGULATION (EU) 2024/1364 of 14 March 2024 on the first phase of the establishment of a common Union rating scheme for data centres. 2024/1364. Online verfügbar unter https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L_202401364, zuletzt geprüft am 31.07.2024.

Eurostat (2024). Cloud computing—Statistics on the use by enterprises. Abgerufen 11. Oktober 2024, von https://ec.europa.eu/eurostat/statistics-explained/index.php/Cloud_computing_-_statistics_on_the_use_by_enterprises#Types_of_cloud_computing:_public_and_private_cloud.

Fichter, K.; Hintemann, R. (2014). Beyond Energy. The Quantities of Materials Present in the Equipment of Data Centers. Journal of Industrial Ecology, 18(6), 846–858.

Fortune Business Insights (2024, Oktober 28). Größe, Anteil und Trends des Rechenzentrumsmarktes: Wachstumsbericht [2032]. Abgerufen 16. November 2024, von https://www.fortunebusinessinsights.com/de/markt-f-r-rechenzentren-109851.

Galabov, V.; Sukumaran, M. (2023). Cloud and Data Center Market Snapshot - November 2023. Hg. v. OMDIA. Online verfügbar unter https://omdia.tech.informa.com/om033574/cloud-and-data-center-market-snapshot---november-2023, zuletzt geprüft am 31.07.2024.

Geng, H. (Hg.) (2021). Data Center Handbook. Plan, Design, Build, and Operations of a Smart Data Center. 2nd ed. Newark: John Wiley & Sons Incorporated.

German Datacenter Association (2024). Data Center Impact Report Deutschland 2024. Frankfurt am Main. Abgerufen von https://www.germandatacenters.com/dcird-24/.

Gholami, G. (2022). KWKK zur Energieversorgung von Rechenzentren. Berlin: Erich Schmidt Verlag GmbH & Co. KG (6).

Gimarc, R. (2015). Quantifying Trends in Server Power Usage. Hg. v. CA Technologies. Online verfügbar unter https://www.cmg.org/wp-content/uploads/2015/10/STLCMG-Quantifying_Trends_Power_Usage.pdf, zuletzt geprüft am 23.10.2024.

Global Market Insights (Hg.) (2024). Data Center Cooling Market - By Component (Solution, Service), By Cooling Technique (Rack/Row-based, Room-based), By End Use (BFSI, Colocation, Energy, Government, Healthcare, Manufacturing, IT & Telecom), By Data Center Size, Forecast 2024 – 2032. Online verfügbar unter https://www.gminsights.com/industry-analysis/data-center-cooling-market, zuletzt geprüft am 31.07.2024.

Goecke, H.; Kempermann, H.; Kestermann, C; Ewald, J.; van Baal, S. (2024). Studie: Spillover-Effekte von Rechenzentren – Rückgrat der KI-Revolution in Deutschland. Abgerufen 31. Oktober 2024, von Eco website: https://www.eco.de/studie_spillover-effekte-von-rechenzentren/.

Goldman Sachs (Hg.) (2024, Mai 14). AI is poised to drive 160% increase in data center power demand. Online verfügbar unter https://www.goldmansachs.com/insights/articles/AI-poised-to-drive-160-increase-in-power-demand, zuletzt aktualisiert am 14.05.2024.

Granskog, A.; Diaz, D. H.; Noffsinger, J.; Milanesi, L. M.; Sachdeva, P.; Bhan, A.; von Schantz S. (2024). The role of power in unlocking the European AI revolution. Abgerufen von https://www.mckinsey.com/industries/electric-power-and-natural-gas/our-insights/the-role-of-power-in-unlocking-the-european-ai-revolution.

Goovaerts, D. (2024). Cloud providers want to crank up rack power 10X for AI. Online verfügbar unter https://www.fierce-network.com/cloud/cloud-providers-want-crank-rack-power-10x-ai, zuletzt aktualisiert am 08.10.2024.

Green, A.; Tai, H.; Noffsinger, J.; Sachdeva, P.; Bhan, A.; Sharma, R. (2024). How data centers and the energy sector can sate AI’s hunger for power. Hg. v. McKinsey & Company. Online verfügbar unter https://www.mckinsey.com/industries/private-capital/our-insights/how-data-centers-and-the-energy-sector-can-sate-ais-hunger-for-power#/, zuletzt geprüft am 01.10.2024.

Gröger, J.; Behrens, F.; Liu, R.; Ackerman, L.; Chwoyka, F.; Schlösser, A. (2023). Umweltzeichen Blauer Engel für Rechenzentren. Hintergrundbericht zur Erarbeitung der Vergabekriterien DE-UZ 228. Hg. v. Umweltbundesamt. Online verfügbar unter https://www.umweltbundesamt.de/publikationen/umweltzeichen-blauer-engel-fuer-rechenzentren, zuletzt geprüft am 14.07.2024.

Hamann, H.; Marianno, F.; Klein, L. (2021). Energy Efficency Assessment of Data Centers Using Measurement and Managment Technology. In: Hwaiyu Geng (Hg.): Data Center Handbook. Plan, Design, Build, and Operations of a Smart Data Center. 2nd ed. Newark: John Wiley & Sons Incorporated, S. 657–668.

Haring, M. (2023, Juli 14). In Indien entstehen reihenweise neue Rechenzentren | Branchen | Indien | IKT. Abgerufen 16. November 2024, von https://www.gtai.de/de/trade/indien/branchen/in-indien-entstehen-reihenweise-neue-rechenzentren-1016370

Harryvan, D. J. (2021). The Idle Coefficients. KPIs to assess energy wasted in servers and data centres. Online verfügbar unter https://www.iea-4e.org/wp-content/uploads/2021/10/Server-Idle-Coefficients-FINAL-1.pdf, zuletzt geprüft am 14.07.2024.

Harryvan, D.; Verzijl, M.; Amzarakov, M. (2021). Analysis LEAP Track 1 ‘Powermanagement'. Hg. v. Netherlands Enterprise Agency. Utrecht.

He, Z.; Ding, T.; Liu, Y.; Li, Z. (2018). Analysis of a district heating system using waste heat in a distributed cooling data center. Applied Thermal Engineering, 141, 1131-1140.

Hermann, C. (2024, April 12). KI belastet Stromversorgung: Amazon zahlt 650 Millionen Dollar an ein AKW - n-tv.de. Abgerufen 31. Oktober 2024, von https://www.n-tv.de/wirtschaft/Amazon-zahlt-650-Millionen-Dollar-an-ein-AKW-article24853002.html.

Heslin, K. (2016, 17. Juni). Ignore Data Center Water Consumption at Your Own Peril. Uptime Institute Blog. https://journal.uptimeinstitute.com/dont-ignore-water-consumption/.

Hintemann, R.; Clausen, J. (2014). Rechenzentren in Deutschland: Eine Studie zur Darstellung der wirtschaftlichen Bedeutung und der Wettbewerbssituation. Im Auftrag des BITKOM - Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e.V. Borderstep Institut für Innovation und Nachhaltigkeit gemeinnützige GmbH.

Hintemann, R.; Hinterholzer, S. (2020). Rechenzentren in Europa – Chancen für eine nachhaltige Digitalisierung—Teil 1. Berlin: Allianz zu Stärkung digitaler Infrastrukturen in Deutschland. Abgerufen von Allianz zu Stärkung digitaler Infrastrukturen in Deutschland website: https://digitale-infrastrukturen.net/studie-nachhaltige-digitalisierung-in-europa/.

Hintemann, R.; Graß, M.; Hinterholzer, S.; Grothey, T. (2022). Rechenzentren in Deutschland: Aktuelle Marktentwicklungen, Stand 2022. Berlin: Bitkom. Abgerufen von Bitkom website: https://www.bitkom.org/Bitkom/Publikationen/Rechenzentren-in-Deutschland-2022.

Hintemann, R.; Hinterholzer, S. (2023a). Steigender Energie- und Ressourcenbedarf der Rechenzentrumsbranche. Rechenzentren rücken in den Fokus von Politik und Regulierung. Hg. v.

Borderstep Institut für Innovation und Nachhaltigkeit gemeinnützige GmbH. Online verfügbar unter https://www.borderstep.de/wp-content/uploads/2023/08/Borderstep_Rechenzentren2022_final.pdf, zuletzt geprüft am 31.07.2024.

Hintemann, R.; Hinterholzer, S.; Seibel, H. (2023b). Rechenzentren in Deutschland: Aktuelle Marktentwicklungen (Update 2023). Berlin: Bitkom e.V. Abgerufen von Bitkom e.V. Online verfügbar unter https://www.bitkom.org/sites/main/files/2023-05/BitkomStudieRechenzentreninDeutschland2023.pdf, zuletzt geprüft am 23.07.2024.

Hintemann, R.; Hinterholzer, S. (2023c). Rechenzentren 2022. Steigender Energie- und Ressourcenbedarf der Rechenzentrumsbranche. Berlin: Borderstep Institut. Abgerufen von Borderstep Institut website: https://www.borderstep.de/wp-content/uploads/2023/08/Borderstep_Rechenzentren2022_final.pdf.

Hintemann, R.; Hinterholzer, S.; Progni, K. (2024a). Rechenzentren in Deutschland – Aktuelle Marktentwicklungen 2024. Berlin: Bitkom e.V. Abgerufen von Bitkom e.V. website: https://www.bitkom.org/Bitkom/Publikationen/Studie-Rechenzentren-in-Deutschland

Hintemann, R.; Hinterholzer, S.; Konrat, F. (2024b, Juni 18). Server stock data – A basis for determining the energy and resource requirements of data centres. Electronic Goes Green 2024+, Berlin.

Hopf, S. (2018). Planung der elektrischen Energieverteilung: Technische Grundlagen. Hg. v. Siemens AG. Siemens AG Energy Management.

Huawei (Hg.) (2021). Analysis of Lithium Ion Battery in Data Centres. Online verfügbar unter https://e.huawei.com/en/material/networkenergy/dc-energy/86d73ab583eb405cbb24b08b9a4a01c0, zuletzt geprüft am 20.11.2024.

Huawei (Hg.) (2024). Lithium-Ionen-Technologie: Die Zukunft der Rechenzentren. 2024 Huawei Digital Power Technologies Co., Ltd.

Hyvönen, J.; Mori, T.; Saunavaara, H., P.; Pärssinen, M.; Syri, S. (2024). Potentail of solar photovoltaics and waste heat utilization in cold climate data centers, Case stuy: Finland and northern Japan. Renewsable and Sustainable Energy Reviews, Volume 201, 114619, ISSN 1364-0321.

IBISWorld (2024). IT-Beratung - Market Research Report (2014-2029). Online verfügbar unter https://www.ibisworld.com/germany/industry/it-beratung/355/

IDC (2024, März 27). Servers Market Insights. Abgerufen 15. April 2024, von IDC: The premier global market intelligence company website: https://www.idc.com/promo/servers

IEA (Hg.) (2024a). Data Centres and Data Transmission Networks. Online verfügbar unter https://www.iea.org/energy-system/buildings/data-centres-and-data-transmission-networks, zuletzt geprüft am 20.11.2024.

IEA (Hg.) (2024b). Electricity 2024 – Analysis and forecast to 2026. Abgerufen von https://iea.blob.core.windows.net/assets/6b2fd954-2017-408e-bf08-952fdd62118a/Electricity2024-Analysisandforecastto2026.pdf

IEA (Hg.) (2024c). Electricity Mid-Year Update. July 2024. Online verfügbar unter https://www.iea.org/reports/electricity-mid-year-update-july-2024, zuletzt geprüft am 22.07.2024.

ifo Institut (2024, September 26). Gemeinschaftsdiagnose Herbst 2024: Deutsche Wirtschaft im Umbruch – Konjunktur und Wachstum schwach. Online verfügbar unter https://www.ifo.de/fakten/2024-09-26/gemeinschaftsdiagnose-herbst-2024-deutsche-wirtschaft-im-umbruch-konjunktur-und

Inc, P. S. P. D. C. S. (2022, 22. Juni). Emerson Network Power White Paper: Energy Logic – Reducing Data Center Energy Consumption by Creating Savings that Cascade Across Systems - e-Store & Library. e-Store & Library. https://media.ptsdcs.com/product/emerson-network-power-white-paper-energy-logic-reducing-data-center-energy-consumption-by-creating-savings-that-cascade-across-systems/

Institute of Electrical and Electronics Engineers (2015). 2015 IEEE 15th International Conference on Environment and Electrical Engineering (EEEIC 2015). Rome, Italy, 10 - 13 June 2015. 2015 IEEE 15th International Conference on Environment and Electrical Engineering (EEEIC). Rome, Italy, 6/10/2015 - 6/13/2015. Piscataway, NJ: IEEE.

Investmentweek (2024, Oktober 2). Nvidia: Auf Wachstumskurs dank Künstlicher Intelligenz. Abgerufen 16. November 2024, von InvestmentWeek website: https://www.investmentweek.com/nvidia-auf-wachstumskurs-dank-kunstlicher-intelligenz/

IW Consult (2024). Spillover-Effekte von Rechenzentren: Rückgrat der KI-Revolution in Deutschland, [online] https://www.eco.de/studie_spillover-effekte-von-rechenzentren/, abgerufen am 05.11.2024.

JAEGGI Hybridtechnologie AG (Hg.) (2022). IT Cooling Solutions. Nachhaltige Rückkühlung von Datacentern. Online verfügbar unter https://www.jaeggi-hybrid.eu/fileadmin/literature/europe/JAEGGI/Application_IT/JAEGGI_Application_IT_DE.pdf, zuletzt geprüft am 21.08.2024.

Jerez Monsalves, J. J.; Bergaentzlé, C.; Backer, M. (2022). Regulatory Frameworks and Business Models for Data Centres Integrated to the Energy System: A comprehensive review of the rules and incentives affecting the provision of flexibility and waste-heat recovery by data centres in Denmark. Report. Hg. v. Technical University of Denmark. 10.11581/dtu.00000242. Online verfügbar unter https://orbit.dtu.dk/en/publications/regulatory-frameworks-and-business-models-for-data-centres-integr

JH-Computers (Hg.) (2024). Nachhaltigkeit. Online verfügbar unter https://jh-computers.de/nachhaltigkeit/, zuletzt geprüft am 20.11.2024.

JLL (2024). Marktüberblick Deutschland 2024: Rechenzentren in Deutschland. Abgerufen von https://www.jll.de/de/trends-and-insights/research/rechenzentren-in-deutschland.

Judge, P. (2023, Juli 21). Microsoft granted permission to run its Dublin data center on gas. Abgerufen 17. Januar 2024, von https://www.datacenterdynamics.com/en/news/microsoft-granted-permission-to-run-its-dublin-data-center-on-gas/

Jung, S.; Ricci, B.; Chung, G. (2015). Lithium Ion Battery System in data centers. In: 2015 IEEE 15th International Conference on Environment and Electrical Engineering (EEEIC 2015). Rome, Italy, 10 - 13 June 2015. 2015 IEEE 15th International Conference on Environment and Electrical Engineering (EEEIC). Rome, Italy, 6/10/2015 - 6/13/2015. Institute of Electrical and Electronics Engineers. Piscataway, NJ: IEEE, S. 968–973.

Kamiya, G.; Bertoldi, P. (2024). Energy Consumption in Data Centres and Broadband Communication Networks in the EU. https://interactdc.com/static/images/documents/JRC135926_01.pdf

Kavanagh, C.; Turner, M. (2022). Green Data: A vision for sustainable data centres in Ireland. Online verfügbar unter https://bitpower.ie/images/Reports/Baringa_Green_Data_-_Full_Report.pdf, zuletzt geprüft am 01.10.2024.

Khan, F. (2024). Omdia research predicts data center cooling market to reach $16.87 billion in 2028. Hg. v. PR Newswire. Online verfügbar unter https://www.prnewswire.com/news-releases/omdia-research-predicts-data-center-cooling-market-to-reach-16-87-billion-in-2028--302175463.html, zuletzt aktualisiert am 18.06.2024, zuletzt geprüft am 31.07.2024.

Koningstein, R. (2021, Mai 18). We now do more computing where there’s cleaner energy. Online verfügbar unter https://blog.google/outreach-initiatives/sustainability/carbon-aware-computing-location/

Konstantin, P.; Konstantin, M. (Hg.) (2024): Praxisbuch der Fernwärme- und Fernkälteversorgung. Berlin, Heidelberg: Springer Berlin Heidelberg.

KPMG (2023). Cloud-Monitor 2023. Abgerufen von https://hub.kpmg.de/de/cloud-monitor-2023?utm_campaign=TECH%20-%20STUDIE%20-%20Cloud%20Monitor%202023&utm_source=AEM_DE.

Krempl, S. (2022, Januar 3). Blackout-Risiko: Rechenzentren-Boom gefährdet Irlands Stromnetz. Abgerufen 18. November 2024, von Heise online website: https://www.heise.de/news/Blackout-Risiko-Rechenzentren-Boom-gefaehrdet-Irlands-Stromnetz-6316877.html.

Kutzer, K. (2024). Server heat generation and cooling. BTGA Taskgruppe Abwärme und Rechenzentren, 11.07.2024, zuletzt geprüft am 11.07.2024.

Lampe, F. (2010). Green-IT, Virtualisierung und Thin Clients. Mit neuen IT-Technologien Energieeffizienz erreichen, die Umwelt schonen und Kosten sparen. 1. Aufl. (1).

Land NRW (2024, Februar 15). Microsoft investiert Milliarden im Rheinischen Revier | Land.NRW. Abgerufen 6. März 2024, von https://www.land.nrw/startseite/microsoft-investiert-milliarden-im-rheinischen-revier.

Latina Press, A. (2024, August 31). Rückführung von Rechenzentren: Lateinamerika holt seine Daten zurück » latinapress Nachrichten. Abgerufen 16. November 2024, von https://latina-press.com/news/328869-rueckfuehrung-von-rechenzentren-lateinamerika-holt-seine-daten-zurueck/

Lawrence, A. (2020). Which regions have the most energy efficient data centers? Hg. v. Data Centre Dynamics. Online verfügbar unter unterhttps://www.datacenterdynamics.com/en/opinions/which-

regions-have-most-energy-efficient-data-centers/, zuletzt aktualisiert am 04.08.2020, zuletzt geprüft am 07.07.2024.

LBNL (2024). 2024 Report to Congress on U.S. Data Center Energy Use—Preliminary Report DraftAugust 2024. LBNL Center of Expertise for Energy Efficiency in Data Centers.

Lei, N.; Masanet, E. (2022). Climate- and technology-specific PUE and WUE estimations for U.S. data centers using a hybrid statistical and thermodynamics-based approach. In: Resources, Conservation and Recycling 182, S. 106323. DOI: 10.1016/j.resconrec.2022.106323.

Li, G.; Sun, Z.; Wang, Q.; Wang, S.; Huang, K.; Zhao, N. et al. (2023). China's green data center development: Policies and carbon reduction technology path. In: Environmental research 231 (Pt 3), S. 116248. DOI: 10.1016/j.envres.2023.116248.

Li, H.; Hou, J.; Hong, T.; Ding, Y.; Nord, N. (2021). Energy, economic, and environmental analysis of integration of thermal energy storage into district heating systems using waste heat from data centres. Energy, 219, 119582.

Liu, Y. (2023, Mai 2). China sets requirements for government procurement of data center equipment, services. Online verfügbar unter **Fehler! Linkreferenz ungültig.**https://enviliance.com/regions/east-asia/cn/report_10060

Mahnke, E.; Mühlenhoff, J. (2012). Strom speichern: Hintergrundinformation der Agentur für Erneuerbare Energien (Renews Spezial, 57). Online verfügbar unter https://www.unendlich-viel-energie.de/media/file/160.57_Renews_Spezial_Strom_speichern_mar13_online.pdf, zuletzt geprüft am 31.07.2024.

Malone, C.; Artman, P.; Wildstone, K. (2023). Coolant Temperatures for Next Generation IT and Durable Data Center Designs. Hg. v. Open Compute Project. Online verfügbar unter https://drive.google.com/file/d/1CxcjdtdIznn_0AC2cJeVbhTLPar3E9P8/view?pli=1, zuletzt geprüft am 20.11.2024.

manage IT (2024, Mai 31). Gartner-Prognose: Weltweiter Umsatz mit KI-Chips wächst 2024 um 33 Prozent. Abgerufen 16. November 2024, von https://ap-verlag.de/gartner-prognose-weltweiter-umsatz-mit-ki-chips-waechst-2024-um-33-prozent/88596/.

Martins, F.; Kobylinska, A. (2023). Gleich(strom) und Gleich(strom) gesellt sich gern. Spannung aufbauen. Online verfügbar https://www.datacenter-insider.de/gleichstrom-und-gleichstrom-gesellt-sich-gern-a-69dce513334ef79c62551d3f58aa7ef2/, zuletzt aktualisiert am 28.12.2023, zuletzt geprüft am 23.10.2024.

Masanet, E. R.; Brown, R. E.; Shehabi, A.; Koomey, J. G.; Nordman, B. (2011). Estimating the Energy Use and Efficiency Potential of U.S. Data Centers. Online verfügbar unter https://eta-publications.lbl.gov/sites/default/files/lbnl_version_procieee_embargoed-1-1.pdf, zuletzt geprüft am 11.10.2024.

McCord, M. (2024). Data Center Cooling: The Unexpected Challenge to AI. Hg. v. Spectra. Online verfügbar unter https://spectra.mhi.com/data-center-cooling-the-unexpected-challenge-to-ai, zuletzt aktualisiert am 09.07.2024, zuletzt geprüft am 20.11.2024.

McKinsey (2024, Oktober 29). AI power: Expanding data center capacity to meet growing demand | McKinsey. Abgerufen 16. November 2024, von https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/ai-power-expanding-data-center-capacity-to-meet-growing-demand.

Menn, A. (2024, Mai 8). Wirtschaft von oben: In diesen Hallen trainiert Microsoft die Künstliche Intelligenz ChatGPT. Abgerufen 8. Juli 2024, von https://www.wiwo.de/technologie/wirtschaft-von-oben/wirtschaft-von-oben-254-kuenstliche-intelligenz-in-diesen-hallen-trainiert-microsoft-die-kuenstliche-intelligenz-chatgpt-/29714212.html.

Ministry of Economy, Trade and Industry (METI) (2019, März, 29). New Energy Conservation Standards for Electronic Computers Formulated. Online verfügbar unter https://www.meti.go.jp/english/press/2019/0329_009.html

Mills, S.; Fernandes, J.; Hang, Y.; Artman, P.; Kim, J.; Steinke, M.; Moscheni, A.; Malone, C.; Carte, C. (2024, Juni 4). 30 °C Coolant – A Durable Roadmap for the Future. Online verfügbar unter https://ocp-all.groups.io/g/ocp-acf/attachment/379/0/30C%20Durable%20Coolant%20Whitepaper%20Rev0.2.pdf, zuletzt geprüft am 15.11.2024.

Ministerium für Umwelt, Klima und Energiewirtschaft Baden-Württemberg: Schwungradspeicher zur unterbrechungsfreien Stromversorgung. Online verfügbar unter https://um.baden-wuerttemberg.de/fileadmin/redaktion/m-um/intern/Dateien/Dokumente/1_Ministerium/Aufgaben_und_Organisation/Green-IT/Green-IT-Schwungrad.pdf

Ministerium für Wirtschaft, Industrie, Klimaschutz und Energie des Landes Nordrhein-Westfalen (2024). Garantiezahlungen und Gewerbesteuer bei Ansiedlungsvorhaben von Hyperscale-Rechenzentren. Gutachterliche Stellungnahme. Online verfügbar unter https://broschuerenservice.justiz.nrw/mwike/shop/garantiezahlungen-und-gewerbesteuer-bei-ansiedlungsvorhaben-von-hyperscale-rechenzentren%7C2190/8

Mollenkopf, W. (2020). Effizienzsteigerung im Rechenzentrum durch optimierte Technologien. Hg. v. Nachhaltige Rechenzentren Baden-Württemberg. HFT Stuttgart.

Mordor Intelligence (2024a). Analyse der Marktgröße und Marktanteile von Rechenzentren in Afrika – Branchenforschungsbericht – Wachstumstrends. Abgerufen 16. November 2024, von https://www.mordorintelligence.com/de/industry-reports/africa-data-center-market.

Mordor Intelligence (2024b). Information Technology Research Reports & Market Industry Analysis. Abgerufen 8. Juli 2024, von https://www.mordorintelligence.com/market-analysis/information-technology.

Mordor Intelligence (2024c). Marktgröße und Marktanteil im Bau von Rechenzentren in Südamerika – Branchenforschungsbericht – Wachstumstrends. Abgerufen 16. November 2024, von

https://www.mordorintelligence.com/de/industry-reports/south-america-data-center-construction-market

Murley, L.; Ferris, J. (2024, März). 2023 Market Monitor for Demand Side Flexibility. Hg. v. LCP Delta und smartEn. Online verfügbar unter https://www.lcp.com/en/insights/publications/5th-edition-of-the-market-monitor-for-demand-side-flexibility

Mytton, D. (2021). Data centre water consumption. Npj Clean Water, (1). Online verfügbar unter https://doi.org/10.1038/s41545-021-00101-w

Newmark, R.; Winkler, W. (2017). ICT capacity and utilization metrics. Hg. v. The Green Grid. Online verfügbar unter https://www.thegreengrid.org/en/resources/library-and-tools/436-WP, zuletzt geprüft am 20.11.2024.

NGN FIBER NETWORK GmbH & Co. KG (Hg.) (2024). Netzwerkkarte Deutschland. Online verfügbar unter https://www.ngn-fibernetwork.de/en/about-ngn, zuletzt geprüft am 06.12.2024.

Nief, A. (2017). Wirtschaftlichkeitsanalyse von Klimatisierungskonzepten und energieeffizienter Maßnahmen für Rechenzentren mittels Gebäude-und Anlagensimulation. Diplomarbeit. Institut für Energiesysteme und Elektrische Antriebe, Technische Universität Wien. Online verfügbar unter https://repositum.tuwien.at/handle/20.500.12708/2166.

OECD (2024). OECD-Bericht zu Künstlicher Intelligenz in Deutschland: OECD.

Ofcom (2023). Statement: Cloud services market study (final report). Online verfügbar unter https://www.ofcom.org.uk/internet-based-services/cloud-services/cloud-services-market-study?showall=1?showall%3D1?showall%3D1

online, heise (2024, September 3). KI-Boom: Enormes Wachstum bei US-Rechenzentren. Abgerufen 13. September 2024, von iX Magazin website: https://www.heise.de/news/KI-Boom-Enormes-Wachstum-bei-US-Rechenzentren-9856469.html.

Orozaliev, J. (2023, Januar). Machbarkeitsuntersuchung Abwärmenutzung aus Rechenzentren in Eschborn und Frankfurt Sossenheim. Gehalten auf der Impact Dialog zur Abwärmenutzung aus Rechenzentren, Frankfurt am Main. Frankfurt am Main.

Ostler, U. (2023). Wasserstoff-Brennstoffzellen als Notstromaggregat im Datacenter. Rolls Royce, Plug Power und Microsoft melden Erfolge. Hg. v. Datacenter Insider. Online verfügbar unter https://www.datacenter-insider.de/wasserstoff-brennstoffzellen-als-notstromaggregat-im-datacenter-a-9f697d860f077a76b36d920063cd6f66/, zuletzt aktualisiert am 30.01.2023, zuletzt geprüft am 20.11.2024.

Ott, B.; Wenzel, P. M.; Radgen, P. (2024). Analysis of Cooling Technologies in the Data Center Sector on the Basis of Patent Applications. Energies, 17(15), 3615. Online verfügbar unter https://doi.org/10.3390/en17153615.

Patterson, M.; Tschudi, B.; Vangeet, O.; Cooley, J.; Azevedo, D. (2010). ERE: A Metric for Measuring the Benefit of Reuse Energy from a Data Center. WHITE PAPER #29. Hg. v. The Green Grid. Online

verfügbar unter https://www.thegreengrid.org/en/resources/library-and-tools/242-ERE%3A-A-Metric-for-Measuring-the-Benefit-of-Reuse-Energy-From-a-Data-Center, zuletzt geprüft am 31.07.2024

Peasley, J. R. (2024). Ranked: Top 50 Data Center Markets by Power Consumption. Online verfügbar unter https://www.visualcapitalist.com/cp/top-data-center-markets/, zuletzt aktualisiert am 10.01.2024, zuletzt geprüft am 07.07.2024.

Pelda, J.; Holler, S.; Persson, U. (2021). District heating atlas - Analysis of the German district heating sector. In: Energy 233, S. 121018. DOI: 10.1016/j.energy.2021.121018.

Pinheiro, E.; Weber, W.-D.; Barroso, L. A. (2007). Failure Trends in a Large Disk Drive Population. Online verfügbar unter https://static.googleusercontent.com/media/research.google.com/de//archive/disk_failures.pdf, zuletzt geprüft am 10.10.2024.

Precedence Research (2024, Oktober). Data Center Market Size To Hit Around USD 364.62 Bn by 2034. Abgerufen 15. November 2024, von https://www.precedenceresearch.com/data-center-market.

Radgen, P.; Turek, D. (2021). Drivers for future data centre demand – historical trends and prospects for 2030. Hg. v. University of Stuttgart, Institute for Energy Economics and the Rational Use of Energy. Stuttgart.

Radgen, P.; Turek, D.; Wesner, S.; Held, F.; Mollenkopf, W.; Narayanan, M. et al. (2020). Nachhaltige Rechenzentren (EcoRZ). Forschungsbericht BWPLUS.

Reddy, V. D.; Setz, B.; Rao, G. S. V.R.K.; Gangadharan, G. R.; Aiello, M. (2017). Metrics for Sustainable Data Centers. In: IEEE Trans. Sustain. Comput. 2 (3), S. 290–303. DOI: 10.1109/TSUSC.2017.2701883.

REOS (2019). Ruimtelijke Strategie Datacenters. Routekaart 2030 voor de groei van datacenters in Nederland. Online verfügbar unter https://www.rijksoverheid.nl/documenten/rapporten/2019/03/15/ruimtelijke-strategie-datacenters

Robinson, D. (2024). Equinix pilots use of fuel cell in 'shipping container' outside datacenter. Hg. v. The Register. Online verfügbar unter https://www.theregister.com/2024/08/16/equinix_shows_off_demo_fuel/, zuletzt aktualisiert am 16.08.2024, zuletzt geprüft am 20.11.2024.

Rohde, C.; Arnold-Keifer, S.; Hirzel, S.; Schlomann, B.; Brugger, H.; Reinfandt, N. (2023). Erhebung des Endenergieverbrauchs im Sektor Gewerbe, Handel, Dienstleistungen (GHD) für das Jahr 2019. Endbericht mit Sonderauswertung Digitalisierung. Hg. v. Fraunhofer ISI. Karlsruhe: Online verfügbar unter https://ag-energiebilanzen.de/wp-content/uploads/2023/06/Endbericht-Energieverbrauch-GHD-Befragung-2019.pdf, zuletzt geprüft am 01.10.2024.

Rolls-Royce (Hg.) (2023). Rolls-Royce testet mtu-Brennstoffzellensystem erfolgreich für Blackoutfall und Spitzenstrom. Online verfügbar unter https://www.mtu-solutions.com/eu/de/pressreleases/2023/rolls-royce-successfully-tests-mtu-fuel-cell-system-for-blackout.html, zuletzt aktualisiert am 17.04.2023, zuletzt geprüft am 20.11.2024.

Roundy, J. (2024). How to manage data center water usage sustainably. Hg. v. TechTarget. Online verfügbar unter https://www.techtarget.com/searchdatacenter/tip/How-to-manage-data-center-water-usage-sustainably, zuletzt aktualisiert am 17.01.2024, zuletzt geprüft am 26.11.2024.

Rüdiger, A. (2021). Mehr Datacenter-Effizienz durch Gleichstrom. Ein neuer, altbekannter Weg der Stromversorgung. Hg. v. Datacenter Insider. Online verfügbar unter https://www.datacenter-insider.de/mehr-datacenter-effizienz-durch-gleichstrom-a-1075018/, zuletzt aktualisiert am 10.12.2021, zuletzt geprüft am 23.10.2024.

Shankar, D. (2022). Under the lens: India's data center explosion. Abgerufen von https://api.anarock.com/uploads/research/Binswanger%20ANAROCK_DC%20Report_Aug%202022_Online.pdf.

Sankar, S.; Shaw, M.; Vaid, K.; Gurumurthi, S. (2013). Datacenter Scale Evaluation of the Impact of Temperature on Hard Disk Drive Failures. In: ACM Trans. Storage 9 (2), S. 1–24. DOI: 10.1145/2491472.2491475.

Schödwell, B.; Drenkelfort, G.; Pröhl, T.; Erek, K.; Zarnekow, R. (2013). Smart Data Centers — intelligente Energieversorgung für Rechenzentren. In: HMD Praxis der Wirtschaftsinformatik 50 (3), S. 40–51. DOI: 10.1007/BF03340814.

Schödwell, B.; Zarnekow, R.; Liu, R.; Gröger, J.; Wilkens, M. (2018). Kennzahlen und Indikatoren für die Beurteilung der Ressourceneffizienz von Rechenzentren und Prüfung der praktischen Anwendbarkeit. Abschlussbericht. Hg. v. Umweltbundesamt. Dessau-Roßlau. Online verfügbar unter https://www.umweltbundesamt.de/sites/default/files/medien/1410/publikationen/2018-02-23_texte_19-2018_ressourceneffizienz-rechenzentren.pdf, zuletzt geprüft am 01.10.2024.

Schramm, D. (2008). Data Center of the Future. Von der CO_2-Neutralität zur Energieeffizienz. Online verfügbar unter https://www.eco.de/wp-content/uploads/2014/04/080911_keynote1_schramm_dell.pdf, zuletzt geprüft am 23.10.2024.

SDIA (Hg.) (2024). Data Center Metrics. Online verfügbar unter https://knowledge.sdialliance.org/data-center-metrics, zuletzt geprüft am 07.07.2024.

Seiter, F.; Zimmermann, H.; Babilon, L.; Richard, P.; Müller, K.; Samek, W. et al. (2024). Energieeffiziente Künstliche Intelligenz für eine klimafreundliche Zukunft. Neue Erkenntnisse über Energieeinsparpotenziale bei KI-Anwendungen. Hg. v. Deutsche EnergieAgentur GmbH. Berlin. Online verfügbar unter https://future-energylab.de/app/uploads/2024/04/Studie_Energieeffiziente_kuenstliche_Intelligenz_fuer_eine_klimafreu ndliche_Zukunft.pdf, zuletzt geprüft am 02.05.2024.

Shao, X.; Zhang, Z.; Song, P.; Feng, Y.; Wang, X. (2022). A review of energy efficiency evaluation metrics for data centers. In: Energy and Buildings 271, S. 112308. DOI: 10.1016/j.enbuild.2022.112308.

Shehabi, A.; Smith, S.J.; Horner, N.; Azevedo, I.; Brown, R.; Koomey, J.; Masanet, E.; Sartor D. (2016). United States Data Center Energy Usage Report. Hg. v. Lawrence Berkeley National Laboratory. Berkeley, California (LBNL-1005775). Online verfügbar unter https://www.iea-4e.org/wp-

content/uploads/publications/2016/06/05j_-_LBNL_-_US_Data_Centres_Energy_USe.pdf, zuletzt geprüft am 05.11.2024.

Sopra Steria (2024, September 17). Studie: Globaler KI-Markt wächst dreimal so schnell wie der Tech-Markt. Abgerufen 16. November 2024, von Sopra Steria DE website: https://www.soprasteria.de/newsroom/publikationen/managementkompass/details/studie-globaler-ki-markt-wachst-dreimal-so-schnell-wie-der-tech-markt.

SPEC (2024). SPECpower_ssj2008 results. Online verfügbar unter https://www.spec.org/power_ssj2008/results/

Spherical Insights (2023). Global Data Center Market Forecast To 2032. Abgerufen 15. November 2024, von Spherical Insights website: https://www.sphericalinsights.com/reports/data-center-market.

Spiegel (2024, Oktober 17). Amazon steckt halbe Milliarde Dollar in Entwicklung von Atomenergie wegen Strombedarf von KI. Der Spiegel. Abgerufen von https://www.spiegel.de/netzwelt/amazon-steckt-halbe-milliarde-dollar-in-entwicklung-von-atomenergie-wegen-strombedarf-von-ki-a-223ae965-f38b-4e9c-8376-975d6a93286c.

Statista (2024a). Servers—Europe. Abgerufen von https://de.statista.com/prognosen/970140/prognose-zum-umsatz-mit-serversystemen-in-europa.

Statista (2024b, Juli). Statista Market Insights Data Center. Abgerufen 7. Juli 2024, von Statista website: https://www.statista.com/outlook/tmo/data-center/servers/germany.

Statista (2024c, Juli 16). Rechenzentren in Europa 2024. Abgerufen 18. November 2024, von Statista website: https://de.statista.com/statistik/daten/studie/1405175/umfrage/rechenzentren-in-europa-nach-laendern/.

Statistisches Bundesamt (2024a, Februar 23). Entwicklung des Außenhandels mit IKT-Gütern in Deutschland nach wichtigsten Importländern. https://www.destatis.de/DE/Themen/Branchen-Unternehmen/Unternehmen/IKT-in-Unternehmen-IKT-Branche/Tabellen/iktb-01-importlaender.html

Statistisches Bundesamt (2024b, Juli 24). Anteile kleiner und mittlerer Unternehmen an ausgewählten Merkmalen 2022 nach Größenklassen. https://www.destatis.de/DE/Themen/Branchen-Unternehmen/Unternehmen/Kleine-Unternehmen-Mittlere-Unternehmen/Tabellen/wirtschaftsabschnitte-insgesamt.html

Statistisches Bundesamt (2024c, August 5). Unternehmen und abhängig Beschäftigte nach Wirtschaftsabschnitten. https://www.destatis.de/DE/Themen/Branchen-Unternehmen/Unternehmen/Unternehmensregister/Tabellen/stat-unternehmen-beschaeftigten-groessenklassen-wz08.html

Statistisches Bundesamt (2024d, Dezember 2). Rechtliche Einheiten, Beschäftigte und Umsatz nach Wirtschaftsabschnitten. https://www.destatis.de/DE/Themen/Branchen-Unternehmen/Unternehmen/Unternehmensregister/Tabellen/unternehmen-beschaeftigte-umsatz-wz08.html

Stobbe, L. (2024, Mai 23). Facts & Trends on Server/CPU/GPU Power Consumption, Performance and Energy Efficiency + Data Center Cooling.

Stobbe, L.; Nissen, N.; Proske, M.; Schulz, A.; Aslan, T. (2024, Juni 18). Methodical challenges of prognostic lifecycle assessment—An exemplary study of ICT's environmental impact in Germany 2030. Proceedings Electronics Goes Green 2024+. Electronic Goes Green 2024+, Berlin.

Stobbe, L.; Proske, M.; Zedel, H.; Hintemann, R.; Clausen, J.; Beucker, S. (2015). Entwicklung des IKT-bedingten Strombedarfs in Deutschland. Abschlussbericht. Studie im Auftrag des Bundesministeriums für Wirtschaft und Energie. Online verfügbar unter https://www.bmwk.de/Redaktion/DE/Downloads/E/entwicklung-des-ikt-bedingten-strombedarfs-in-deutschland-abschlussbericht.pdf?__blob=publicationFile&v=1, zuletzt geprüft am 05.11.2024.

Strutt, Steve; Kelley, Chris; Singh, Harkeeret; Smith, Vic (2012): Data Center Efficiency and IT Equipment Reliability at wider Operating Temperature and Humidity Ranges. Hg. v. The Green Grid. Online verfügbar unter https://www.energy.gov/sites/prod/files/2013/12/f5/data_center_efficiency_and_reliabilit_at_wide, zuletzt geprüft am 10.10.2024.

Sustainability, M. (2024, 28. August). Data centers – meta sustainability. Meta Sustainability. https://sustainability.atmeta.com/data-centers/#:~:text=In%202022%2C%20we%20tested%20increasing,F%20at%20select%20pilot%20campuses.

Super Micro (2024). https://www.supermicro.com/solutions/Solution-Brief_Supermicro_Liquid_Cooling_Solution_Guide.pdf

Suthar, S. (2024). Strategies For Efficient Server Utilization & Energy Savings. Online verfügbar unter https://middleware.io/blog/efficient-server-utilization/, zuletzt aktualisiert am 28.08.2024, zuletzt geprüft am 20.11.2024.

Swinhoe, D. (2024, Juni 10). Amazon, Google, Microsoft want to build own power lines in Ireland—Report. https://www.datacenterdynamics.com/en/news/amazon-google-microsoft-want-to-build-own-power-lines-in-ireland-report/

Synergy Research (2024a, August 21). Cloud is a Global Market—Apart from China | Synergy Research Group. Abgerufen 18. November 2024, von https://www.srgresearch.com/articles/cloud-is-a-global-market-apart-from-china.

Synergy Research (2024b, September 24). A Robust and Flourishing Cloud Ecosystem Hits $427 Billion in the First Half | Synergy Research Group. Abgerufen 18. November 2024, von https://www.srgresearch.com/articles/first-half-numbers-show-a-427-billion-cloud-ecosystem-in-rude-good-health.

Tagesschau (2024a, Juni 19). Amazon investiert weitere 10 Milliarden Euro in Deutschland. Abgerufen 8. Juli 2024, von Tagesschau.de website: https://www.tagesschau.de/wirtschaft/digitales/amazon-10-milliarden-deutschland-investition-100.html.

Tagesschau (2024b, September 20). Störfall-AKW in den USA soll für Microsoft wieder in Betrieb gehen. Abgerufen 31. Oktober 2024, von Tagesschau.de website: https://www.tagesschau.de/ausland/amerika/usa-atomkraftwerk-microsoft-100.html.

Taylor, P. (2024). Leading countries by number of data centers as of March 2024 (Statista). Abgerufen von https://www.statista.com/statistics/1228433/data-centers-worldwide-by-country/.

The Green Grid (2007). The Green Grid Data Center Power Efficiency Metrics: PUE & DCiE. https://de.scribd.com/document/725765924/TGG-Data-Center-Power-Efficiency-Metrics-PUE-and-DCiE.

The Green Grid (2024). The Green Grid®'s TGG Server Energy Efficiency Database v02_00 of SPEC SERT® benchmark results (Version 02.00) [Dataset]. https://www.thegreengrid.org

Triebs, M. S.; Papadis, E.; Cramer, H.; Tsatsaronis, G. (2021). Landscape of district heating systems in Germany – Status quo and categorization. In: Energy Conversion and Management: X 9, S. 100068. DOI: 10.1016/j.ecmx.2020.100068.

Turek, D. (2022). Die Standortwahl von Rechenzentren unter Berücksichtigung der Total Cost of Ownership : ein Framework zur multikriteriellen Entscheidungsanalyse. Dissertation, Universität Stuttgart, http://dx.doi.org/10.18419/opus-12825

Umweltbundesamt; Borderstep Institut für Innovation und Nachhaltigkeit; Europäische Kommission (2020). Energy-efficient cloud computing technologies and policies for an eco-friendly cloud market. Final study report. 1st edition, revised. Luxembourg: Publications Office of the European Union. Online verfügbar unter https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=71330, zuletzt geprüft am 29.10.2024.

Uptime Institute (Hg.) (2020). Uptime Institute global data center survey 2020. Online verfügbar unter https://uptimeinstitute.com/uptime_assets/8160f3a6d5d83de52d132535575efcf1a5614c6fb69515b5ef91c8c8c1d611d1-2020-data-center-industry-survey.pdf, zuletzt geprüft am 07.07.2024.

Uptime Institute (Hg.) (2021). Uptime Institute Global Data Center Survey 2021. Growth stretches an evolving sector. Online verfügbar unter https://uptimeinstitute.com/uptime_assets/4d10650a2a92c06a10e2c70e320498710fed2ef3b402aa82fe7946fae3887055-2021-data-center-industry-survey.pdf, zuletzt geprüft am 07.07.2024.

Uptime Institute (Hg.) (2022, 14. September). Global Data Center Survey 2022. Resiliency remains critical in a volatile world. Online verfügbar unter https://uptimeinstitute.com/uptime_assets/6768eca6a75d792c8eeede827d76de0d0380dee6b5ced20fde45787dd3688bfe-2022-data-center-industry-survey-en.pdf

Uptime Institute (Hg.) (2023). Uptime Institute Global Data Center Survey 2023. Executive summary. Online verfügbar unter https://uptimeinstitute.com/uptime_assets/74fd7ed906aad2b6df2a96dfeb803dde83d52ee3dffdd8ae41a50fab4e23182f-uptime-institute_global-data-center-survey-2023_executive-summary.pdf, zuletzt geprüft am 07.07.2024.

Uptime Institute (Hg.) (2024). Uptime Institute Global Data Center Survey 2024. Executive summary. Online verfügbar unter https://uptimeinstitute.com/uptime_assets/7425ec68d479c5d78a743df94a79b114ed9f9c73f13b6460949d2b8e73373209-GA-2024-07-uptime-institute-global-data-center-survey-results-2024.pdf, zuletzt geprüft am 12.11.2024.

Urgaonkar, R.; Urgaonkar, B.; J. Neely, M.; Sivasubramaniam, A. (2011). Optimal power cost management using stored energy in data centers. DOI: 10.1145/1993744.1993766.

Uti, T. (2024, 9. Juli). CoolIT Systems launches three High-Density Coolant Distribution Units. CoolIT Systems. https://www.coolitsystems.com/coolit-systems-launches-three-high-density-coolant-distribution-units/.

van Geet, O.; Sickinger, D. (2024). Best Practices Guide for Energy-Efficient Data Center Design. Online verfügbar unter https://www.energy.gov/sites/default/files/2024-07/best-practice-guide-data-center-design.pdf.

Vertiv (Hg.). Understanding Coolant Distribution Units (CDUs) for Liquid Cooling. Online verfügbar unter https://www.vertiv.com/en-asia/about/news-and-insights/articles/educational-articles/understanding-coolant-distribution-units-cdus-for-liquid-cooling/, zuletzt geprüft am20.11.2024.

Vertiv (Hg.) (2018). The Emergence Of Lithium-Ion Batteries Within The Data Center. A Vertiv Application Report. Online verfügbar unter https://www.vertiv.com/4aa16f/globalassets/products/critical-power/uninterruptible-power-supplies-ups/vertiv-lithium-ion-battery-wp-en-na-sl-24692_138606_00.pdf, zuletzt geprüft am 20.11.2024.

Voigt, E., dpa, & AFP. (2024, Februar 22). Künstliche Intelligenz: KI-Entwicklung sorgt für Umsatzrekord beim Chipkonzern Nvidia. Die Zeit. Abgerufen von https://www.zeit.de/wirtschaft/2024-02/nvidia-kuenstliche-intelligenz-wachstum.

Wang S.; Casado M. (2021, Mai 27). The Cost of Cloud, a Trillion Dollar Paradox. Andreessen Horowitz. Online verfügbar unter https://a16z.com/the-cost-of-cloud-a-trillion-dollar-paradox/.

Weber, J. (2009). Hybridkühlung auf dem Vormarsch. Hg. v. JAEGGI Hybridtechnologie AG. Online verfügbar unter https://www.jaeggi-hybrid.ch/fileadmin/sites/jaeggi/ch/Fachartikel/Hybridk%C3%BChlung_auf_dem_Vormarsch.pdf, zuletzt geprüft am 06.08.2024.

Wenzel, P. M.; Mühlen, M.; Radgen, P. (2023). Free Cooling for Saving Energy: Technical Market Analysis of Dry, Wet, and Hybrid Cooling Based on Manufacturer Data. In: Energies 16 (9), S. 3661. DOI: 10.3390/en16093661.

Wirtschaftswoche (2023, Oktober 6). Informationstechnologie: Warum Google eine Milliarde in die deutsche Cloud-Infrastruktur steckt. Abgerufen 6. Januar 2024, von https://www.wiwo.de/unternehmen/it/rechenzentrum-warum-google-eine-milliarde-in-die-deutsche-cloud-infrastruktur-steckt-/29430892.html

Wong, W. (2023, 3. Januar). Hot in Here: Is Raising Temperatures in Data Centers Good for Hardware? https://www.datacenterknowledge.com/data-center-hardware/hot-in-here-is-raising-temperatures-in-data-centers-good-for-hardware-.

Xie, X.; Han, Y.; Tan, H. (2024). Greening China’s digital economy: exploring the contribution of the East–West Computing Resources Transmission Project to CO_2 reduction. Humanit Soc Sci Commun 11, 466. https://doi.org/10.1057/s41599-024-02963-0

Zaloumis, C. (2022). Are Your Data Centers Keeping You From Sustainability? Hg. v. IBM. Online verfügbar unter https://www.ibm.com/think/insights/are-your-data-centers-keeping-you-from-sustainability, zuletzt aktualisiert am 07.06.2022, zuletzt geprüft am 20.11.2024.

Zhang, C.; Slijkerman, J. F.; Stadig, D. (2023). Growth in water consumption of data centres needs more attention. Online verfügbar unter https://think.ing.com/articles/data-centres-growth-in-water-consumption-needs-more-attention/, zuletzt aktualisiert am 24.10.2023, zuletzt geprüft am 26.11.2024.

Zhang, M. (2024). Data Center Water Usage: A Comprehensive Guide. Hg. v. Dgtl Infra. Online verfügbar unter https://dgtlinfra.com/data-center-water-usage/, zuletzt aktualisiert am 17.01.2024, zuletzt geprüft am 26.11.2024.

Zhang, Y.; Li, H.; Wang, S. (2023). The global energy impact of raising the space temperature for high-temperature data centers. Cell Reports Physical Science, 4(10), 101624. https://doi.org/10.1016/j.xcrp.2023.101624.

ZincFive (Hg.) (2024). 2024 Data Center Energy Storage Industry Insights Report. Online verfügbar unter https://creative.endeavorb2b.com/ClientMarketing/processing/Data-Center-Energy-Storage-Industry-Insights-Report-2024.pdf, zuletzt geprüft am 09.10.2024.