



FORSCHUNGSBERICHT

AUSWIRKUNGEN VON KI, RECHENZENTREN UND HALBLEITERN AUF WASSERVERFÜGBARKEIT UND -QUALITÄT

JUNI 2025

FORSCHUNGSBERICHT

AUSWIRKUNGEN VON KI, RECHENZENTREN UND HALBLEITERN AUF WASSERVERFÜGBARKEIT UND -QUALITÄT

JUNI 2025

Impressum

Herausgeber:

Gesellschaft für Informatik e.V. (GI)
Weydingerstraße 14–16
10178 Berlin
Deutschland
Tel: +49 30 240 09-805
E-Mail: berlin@gi.de
Internet: <https://gi.de>

Autorinnen und Autoren:

Lena Hoffmann, GI
Teresa Zeck, GI
Nikolas Becker, GI

Stand:

Juni 2025

Gefördert durch:



Bundesministerium
für Forschung, Technologie
und Raumfahrt

Inhaltsverzeichnis

Executive Summary	8
1. Einleitung	11
2. Wissenschaftlicher Hintergrund	15
2.1. Methodische Ansätze zur Erfassung der digitalisierungsbedingten Umweltauswirkungen.....	15
2.2. Lebenszyklus von KI-Modellen und deren Umweltauswirkungen	16
2.3. Wassernutzung, Wasserverbrauch und Wasserqualität entlang des KI-Lebenszyklus	17
2.4. Wassernutzung in Rechenzentren.....	20
3. Methodologie	28
3.1. Systematische Literaturrecherche (SLR)	29
3.2. Analyse von Nachhaltigkeitsberichten	29
3.3. Experteninterviews und Workshop	30
3.4. Datenanalyse.....	30
4. Ergebnisse I: Wasserverbrauch durch KI.....	32
4.1. Quantifizierung des Wasserverbrauchs durch KI	32
4.2. Wasserverbrauch durch Rechenzentren	36
4.3. Wasserverbrauch durch Hardwareproduktion	40
4.4. Berichterstattung zu dem Wasserverbrauch aus der Praxis - Nachhaltigkeitsberichte	43
5. Ergebnisse II: Einflussfaktoren und Best Practices	44
5.1. Praktiken der Softwaregestaltung.....	45
5.2. Praktiken der Hardwaregestaltung.....	47
5.3. Rechenzentrumsdesign & Nutzung.....	48
5.4. Standortfaktoren.....	49
6. Ergebnis III: Hindernisse und Herausforderungen	51
6.1. Fehlende Transparenz und mangelnde Standardisierung	52
6.2. Geringe Sichtbarkeit und mangelnde Skalierung von Best Practices	53
6.3. Unzureichende Regularien und fehlende wirtschaftliche Anreize	54
6.4. Marktdynamiken und Machtkonzentration	54
6.5. Nachhaltigkeitsprobleme entlang der globalen Liefer- und Wertschöpfungsketten von KI und Rechenzentren	55
7. Handlungsempfehlungen für ressourceneffiziente KI und Rechenzentren	57
7.1. Einführung verbindlicher Transparenzvorgaben und Etablierung von Standards	58
7.2. Forschung und Innovationsförderung für ressourcen- und energieeffiziente KI-Technologien	59
7.3. Förderung von energie- und ressourceneffizienter Rechenzentrumsinfrastruktur und Betriebsweisen	60
7.4. Schaffung von Anreizen für energie- und ressourcenschonende Investitionen	61
7.5. Sichtbarmachung und Verbreitung von Best Practices	61
7.6. Förderung einer ganzheitlichen Betrachtung des direkten und indirekten Energie- und Ressourcenverbrauchs durch KI	62

7.7. Förderung ressourcenschonender Hardware und Stärkung der Kreislaufwirtschaft 63

8. Diskussion & Fazit 65

8.1. Kritische Reflexion der Ergebnisse & Limitationen 65

8.2. Fazit & Ausblick 66

Literaturverzeichnis 68

Danksagung 76

Abbildungsverzeichnis

Abbildung 1: Wasserstress und RZ in Europa.	13
Abbildung 2: Der Einfluss von KI auf den Wasserverbrauch entlang des Hardware-Lebenszyklus.....	18
Abbildung 3: Luftbasierte Kühlung - Computer Room Air-Conditioning.	21
Abbildung 4: Indirekte adiabate Kühlung für Serverräume: Sprühdüsen, Ventilatoren, Wärmetauscher und Serverraum.	23
Abbildung 5: Direkte Chipkühlung - Flüssigkeitsbasierte Kühlung.	25
Abbildung 6: Methodisches Vorgehen der Studie.	28
Abbildung 7: Direkter und indirekter Wasserverbrauch in Rechenzentren.	35
Abbildung 8: Einflussfaktoren und Best Practices.	45
Abbildung 9: Hindernisse und Herausforderungen für die ressourcenschonende KI und Rechenzentren.	51
Abbildung 10: Handlungsempfehlungen für ressourceneffiziente KI und Rechenzentren.....	58

Tabellenverzeichnis

Tabelle 1: Zentrale Studien zum Wasserverbrauch durch KI	34
Tabelle 2: Ausgewählte Studien zum Wasserverbrauch von Rechenzentren	37
Tabelle 3: Ausgewählte Studien zum Wasserverbrauch in der KI-Hardwareproduktion.	41

Abkürzungsverzeichnis

CO ₂	Kohlenstoffdioxid
CPU	Central Processing Unit (Prozessor)
CRAC	Computer Room Air Conditioning
CRAH	Computer Room Air Handler
DIN	Deutsches Institut für Normung
DLC	Direct Liquid Cooling (Direkte Flüssigkeitskühlung)
EU	Europäische Union
GB	Gigabyte
GPU	Graphics Processing Unit (Grafikprozessor)
GW	Gigawatt
IKT	Informations- und Kommunikationstechnologien
IoT	Internet of Things (Internet der Dinge)
IPCC	Intergovernmental Panel on Climate Change
IR	Information Retrieval
ISO	International Organization for Standardization
IT	Informationstechnologie
KI	Künstliche Intelligenz
kWh	Kilowattstunde
LCA	Life Cycle Assessment (Lebenzyklusanalyse)
l/kWh	Liter pro Kilowattstunde
m ³	Kubikmeter
ML	Machine Learning (Maschinelles Lernen)
Mt	Megatonne
OECD	Organisation for Economic Co-operation and Development (Organisation für wirtschaftliche Zusammenarbeit und Entwicklung)
PUE	Power Usage Effectiveness
RZ	Rechenzentrum
SLR	Systematische Literaturanalyse
SSD	Solid-State-Drive bzw. Solid-State-Disk
THG	Treibhausgase
TPU	Tensor Processing Unit
UPW	Ultrapure Water (Reinstwasser)
VN	Vereinte Nationen
WASH	Wasser, Sanitärversorgung und Hygiene
WRI	World Resources Institute
WUE	Water Usage Effectiveness

Executive Summary

Die rasante Entwicklung sowie die zunehmende Anwendung und Verbreitung von Künstlicher Intelligenz (KI) treiben den Bedarf an Rechenleistung und digitaler Infrastruktur – und damit auch den Ausbau hyperskalierter Rechenzentren – stark voran. Obwohl KI das Potenzial bietet, zur Lösung ökologischer Herausforderungen beizutragen, sind ihre eigenen Umweltauswirkungen – insbesondere im Hinblick auf den Wasserverbrauch – bislang nur unzureichend erforscht.

Diese Studie analysiert den Wasserverbrauch entlang des gesamten Lebenszyklus von KI – von der Hardwareproduktion über das Training und den Betrieb bis hin zur Entsorgung von KI-Systemen. Ziel ist es, darzustellen, wie sich der wachsende Rechenleistungsbedarf auf den Wasserbedarf in Rechenzentren und globalen Lieferketten auswirkt, welche Praktiken bereits heute zur Reduktion des Ressourcenverbrauchs beitragen können und welche technologischen, regulatorischen und wirtschaftlichen Herausforderungen bestehen. Darauf aufbauend werden konkrete Handlungsempfehlungen für eine ökologisch nachhaltigere Gestaltung des KI- und Rechenzentrumssektors entwickelt.

KERNBEFUNDE

Die Studie zeigt, dass der steigende Einsatz von KI erhebliche Auswirkungen auf Rechenzentren und deren Ressourcenverbrauch hat. Dabei sind mehrere Faktoren entscheidend:

- **Wasser- und Energienutzung in KI-Trainings- und Inferenzprozessen:** Besonders große KI-Modelle benötigen für Training und Betrieb erhebliche Rechenleistung, die wiederum mit einem entsprechend hohen Energiebedarf verbunden sein kann. Inwiefern sich dies auf einen erhöhten Wasserverbrauch – sowohl für die Kühlung der Systeme als auch über die wasserintensive Stromerzeugung – auswirkt, hängt von einer Vielzahl an Faktoren wie dem Standort, der Energiequelle und der eingesetzten Kühltechnologie des Rechenzentrums ab. In Regionen mit angespannter Wassersituation kann dies zu Nutzungskonflikten mit anderen Sektoren führen.
- **Ressourceneffiziente KI-Entwicklung:** Es bestehen vielversprechende Ansätze zur Reduktion des Ressourcenverbrauchs über den gesamten Lebenszyklus von KI-Systemen hinweg. Dazu zählen energieeffiziente Algorithmen, adaptive Trainingsprozesse, spezialisierte Hardware mit geringerem Strombedarf sowie der Einsatz kleinerer, aufgabenspezifischer KI-Modelle. Diese Technologien weisen ein hohes Potenzial auf, werden bislang jedoch nur vereinzelt implementiert – zudem fehlt es häufig an geeigneten Bewertungsmaßstäben.
- **Wasserintensive Hardwareproduktion als relevanter Einflussfaktor:** Die Herstellung von KI-spezifischen Hardwarekomponenten – insbesondere Halbleitern – geht häufig mit einem hohen Wasserbedarf einher. Gleichzeitig ist die Datenlage zu diesen Verbräuchen begrenzt. Eine stärkere Berücksichtigung von Kreislaufwirtschaft, langlebigem Design und verbesserter Rückgewinnung von Ressourcen kann hier zur Entlastung beitragen.
- **Standortwahl als strategische Stellschraube:** Die Auswahl von Standorten für Rechenzentren ist entscheidend. Neben wirtschaftlichen Kriterien wie Strompreisen oder Netzinfrastruktur ist es wichtig, auch umweltbezogene Aspekte – etwa Wasserverfügbarkeit, Flächenbedarf oder potenzielle lokale Nutzungskonflikte – systematisch mit einzubeziehen.
- **Fehlende ganzheitliche Bewertungsmetriken:** Der weit verbreitete Power Usage Effectiveness (PUE)-Wert erlaubt Rückschlüsse auf die Energieeffizienz von Rechenzentren, bildet jedoch weder den Wasserverbrauch noch den Materialeinsatz über den gesamten Lebenszyklus adäquat ab. Der Water Usage Effectiveness (WUE) ist noch weniger verbreitet und bildet zudem nicht den gesamten Wasserverbrauch von Rechenzentren angemessen ab. Eine Erweiterung bestehender Kennzahlen und die Entwicklung neuer, lebenszyklusbasierter Bewertungsansätze erscheinen notwendig, um ökologische Auswirkungen besser erfassen und vergleichen zu können.

- **Geringe Sichtbarkeit technischer Lösungen und Best Practices:** Die Studie identifiziert eine Vielzahl vielversprechender Technologien und Ansätze zur Reduktion des Wasser- und Energieverbrauchs von KI und Rechenzentren – etwa im Bereich der Kühlung, der Softwareoptimierung oder der Abwärmenutzung. Diese werden bislang jedoch kaum systematisch erfasst, gefördert oder skaliert. Dadurch bleibt ihr Potenzial zur Ressourcenschonung vielfach ungenutzt.

HANDLUNGSEMPFEHLUNGEN

Auf Grundlage der Analyse formuliert die Studie sieben zentrale Handlungsfelder, um den Wasser- und Ressourcenverbrauch von KI-Systemen und Rechenzentren zu reduzieren:

- **Standards und Transparenz stärken:** Es braucht einheitliche Bewertungsstandards, die Umweltwirkungen – einschließlich Wasser- und Energieverbrauch – systematisch und vergleichbar erfassen. Diese sollten entlang des gesamten Lebenszyklus von KI-Infrastrukturen anwendbar sein. Ergänzend können transparente Berichtspflichten oder gezielte Anreize Unternehmen dabei unterstützen, ihren Ressourcenverbrauch zu messen und nachvollziehbar offenzulegen. So können Optimierungspotenziale identifiziert und genutzt werden.
- **Forschung zu ressourcenschonenden Technologien ausbauen:** Ressourcenschonende Technologien – etwa für energieeffiziente KI-Architekturen, neue Speicherlösungen, innovative Kühlmethoden oder adaptive Workload-Management-Systeme – bieten großes Potenzial, sind jedoch teilweise noch nicht ausreichend erprobt oder verbreitet. Die Weiterentwicklung und Verbreitung dieser Ansätze sollte gezielt gefördert werden, um ihre praktische Anwendbarkeit in Rechenzentren und KI-Systemen zu erhöhen.
- **Rechenzentren nachhaltig gestalten und betreiben:** Rechenzentren sollten von Beginn an ressourcenschonend geplant und betrieben werden. Dazu gehört eine Standortwahl, die unter anderem vorhandene Energie- und Wasserquellen sowie Potenziale zur Abwärmenutzung berücksichtigt. Zudem sollte jeweils die effizienteste Kühltechnologie in Abhängigkeit von den lokalen Bedingungen eingesetzt werden. Auch im Betrieb sind ergänzende Maßnahmen zur Ressourceneffizienz – wie intelligentes Lastmanagement, temperaturangepasste Steuerung oder der Einsatz alternativer Kühlmethoden – zu fördern.
- **Schaffung von Anreizen:** Um Investitionen in ressourcenschonende Rechenzentren und KI-Infrastrukturen attraktiver zu machen, sind steuerliche Vergünstigungen, gezielte Förderprogramme und vereinfachte Genehmigungsverfahren eine mögliche Option. Wirtschaftliche Anreize können ein Hebel zur Förderung nachhaltiger Rechenzentren sein. Gleichzeitig bedarf es einer fundierten Analyse, welche Anreizmechanismen tatsächlich zielgerichtet wirken und langfristig tragfähig sind.
- **Best Practices sichtbar machen und verbreiten:** Bereits existierende Lösungen aus Forschung, öffentlichem Sektor und Wirtschaft sollten systematisch erfasst, kommuniziert und zur Nachahmung bereitgestellt werden – etwa über Plattformen, Fallstudien oder Leitfäden. Dies erleichtert die Verbreitung technischer Innovationen und schafft Orientierung für Investitionen.
- **Ganzheitliche Umweltbewertung etablieren:** Es braucht neue Bewertungsrahmen, die nicht nur Energie- und Wasserverbräuche erfassen, sondern auch Rohstoffbedarf, Standortfolgen und soziale Auswirkungen entlang globaler Lieferketten einbeziehen. Nur ein systemischer Blick auf den gesamten Lebenszyklus von KI und Rechenzentren unter breiter Einbindung sozial-ökologischer Kriterien erlaubt tragfähige Entscheidungen und kann dazu beitragen, zukünftige Entwicklungen robuster und nachhaltiger zu gestalten.
- **Kreislaufwirtschaft für Hardware stärken:** Eine stärkere Einbindung von Hardware in kreislaufbasierte Konzepte ist eine Option, um den Ressourcenverbrauch zu senken. Im Fokus sollten Maßnahmen stehen, die eine möglichst lange und flexible Nutzung ermöglichen – etwa durch reparaturfreundliches Design, modulare Systeme oder die Förderung von Wiederverwendung und Refurbishing.

FAZIT

Die zunehmende Nutzung von KI hat spürbare Auswirkungen auf digitale Infrastrukturen und ökologische Ressourcen. Insbesondere der wachsende Rechenbedarf trägt zu steigenden Energie- und Wasserverbräuchen bei, während die Produktion spezialisierter Hardware neue Anforderungen an globale Rohstoff- und Versorgungsketten stellt. Die vorliegende Analyse verdeutlicht, dass ein nachhaltiger Umgang mit diesen Herausforderungen ein Zusammenspiel aus technologischer Innovation, politischer Steuerung und wirtschaftlicher Anreizsetzung erfordert.

Ein rein sektoraler Ansatz – etwa durch die Optimierung einzelner Rechenzentren – greift dabei zu kurz. Vielmehr braucht es eine systemische Perspektive, die KI-Modelle, Hardwareproduktion, Standortwahl und Infrastrukturplanung ganzheitlich in den Blick nimmt. Bestehende Technologien und Ansätze zur Ressourcenschonung zeigen bereits vielversprechende Potenziale, ihre Umsetzung bleibt bislang jedoch fragmentiert. Um Fortschritte zu erzielen, ist eine stärkere Verbreitung bewährter Lösungen ebenso notwendig wie die Weiterentwicklung geeigneter Bewertungsmethoden und eine intensive interdisziplinäre Zusammenarbeit.

1. Einleitung

Eine der prägendsten technologischen Entwicklungen der letzten Jahre ist die rasante Zunahme der Entwicklung und Nutzung von Künstlicher Intelligenz (KI). KI-Systeme beeinflussen bereits heute zahlreiche gesellschaftliche, wirtschaftliche und technologische Prozesse – sie bestimmen verstärkt, welche Informationen wir konsumieren, wie Unternehmen Entscheidungen treffen und wie automatisierte Abläufe gestaltet werden.

Ein wesentlicher Treiber dieser Entwicklung war lange Zeit die exponentielle Zunahme der Transistordichte, ein Phänomen, das unter dem Namen Moore's Law bekannt ist. Dieses besagt, dass sich die Anzahl der Transistoren auf einem Mikrochip etwa alle 18 bis 24 Monate verdoppelt – was historisch mit einer Steigerung der Rechenleistung einherging. Inzwischen stößt dieses Wachstum jedoch zunehmend an physikalische und technologische Grenzen, da die Miniaturisierung von Transistoren technisch immer anspruchsvoller wird. Dennoch wird der Fortschritt in der KI-Entwicklung durch algorithmische Effizienzsteigerungen und spezialisierte Hardware weiterhin vorangetrieben (Hernandez & Brown, 2020; Leiserson et al., 2020).

Der KI-Trend trägt zu einer rapiden Nachfrage nach Rechenzentren (RZ) bei, insbesondere an Hyperscale-RZ. Expert*innen gehen davon aus, dass der zunehmende Einsatz von KI diesen Trend weiter verstärken wird (Bitkom, 2024b). Eine McKinsey-Analyse zeigt, dass die globale RZ-Nachfrage zwischen 2023 und 2030 jährlich um 19 bis 22 % wachsen könnte und dabei 171 bis 219 Gigawatt (GW) erreichen würde – im Vergleich zu derzeit 60 GW Nachfrage. Besonders stark wird die Nachfrage nach KI-fähigen RZ steigen: Diese könnte im mittleren Szenario um durchschnittlich 33 % pro Jahr wachsen, sodass bis 2030 etwa 70 % der gesamten RZ-Kapazitäten für fortgeschrittene KI-Workloads benötigt werden. Generative KI, derzeit der am schnellsten wachsende KI-Anwendungsbereich, würde dabei rund 40 % des Gesamtbedarfs ausmachen. Den größten Teil der steigenden Nachfrage treiben Cloud-Service-Provider wie Amazon Web Services, Google Cloud, Microsoft Azure und Baidu voran, die enorme Kapazitäten benötigen, um große KI-Modelle wie Googles Gemini oder OpenAIs ChatGPT zu betreiben (McKinsey, 2024).

ENERGIE- UND RESSOURCENVERBRÄUCHE VON KI UND RZ

In den letzten Jahren rückten die Umweltwirkungen von KI und RZ verstärkt in den Fokus politischer und wissenschaftlicher Diskurse. Im Zentrum stehen das Potenzial von KI für ökologische Nachhaltigkeit und ihre eigenen negativen Umweltauswirkungen. Einerseits können KI-Technologien eine effizientere Ressourcennutzung ermöglichen und Entscheidungsprozesse optimieren, um globale Herausforderungen wie den Klimawandel besser zu bewältigen (Bitkom, 2024a; Rolnick et al., 2023). Andererseits zeigen immer mehr wissenschaftliche Studien, dass die Entwicklung und der Betrieb von KI, insbesondere von generativen KI-Systemen, mit einem erheblichen Energie- und Ressourcenverbrauch verbunden sind (Berthelot et al., 2024; Luccioni et al., 2022; Strubell et al., 2019). Lange Zeit konzentrierte sich die wissenschaftliche Debatte vor allem auf den Energieverbrauch und die damit verbundenen Kohlenstoffdioxid (CO₂)-Emissionen von KI-Modellen (Luccioni et al., 2022; Strubell et al., 2019). Erst in jüngerer Zeit rückt auch der Wasserverbrauch von KI als bedeutender Umweltfaktor verstärkt in den Fokus (Li et al., 2023; Tagesschau, 2023; Verma & Tan, 2024).

Der World Water Development Report 2024 der Vereinten Nationen unterstreicht die ambivalente Rolle der KI: Während KI-Technologien zur Optimierung der Wasserversorgung beitragen können, sind ihre eigenen Wasserverbräuche bislang unzureichend erforscht (United Nations, 2024). Eine erste Schätzung des Wasserbedarfs von KI stammt von Li et al. (2023) und geht davon aus, dass die weltweite Wassernutzung durch KI-Anwendungen bis 2027 auf 4,2 bis 6,6 Milliarden Kubikmeter (m³) ansteigen könnte – das entspricht mehr als dem Vier- bis Sechsfachen des jährlichen Wasserverbrauchs von Dänemark. Am Beispiel des Sprachmodells GPT-3 mit 175 Milliarden Parametern zeigt die Studie, dass dessen Training in Microsofts hochmodernen RZ in den USA einen Wasserverbrauch von insgesamt rund 5,4 Millionen Litern verursachen kann. Diese Schätzungen sind jedoch mit erheblichen Unsicherheiten verbunden: Zum einen variiert der Wasserverbrauch von

Rechenzentren räumlich und zeitlich stark je nach verwendeten Kühlungstechnologien und klimatischen Bedingungen, zum anderen berücksichtigen die Projektionen von Li et al. aufgrund mangelnder Datenverfügbarkeit den Wasserverbrauch der Lieferkette, etwa für die Chip-Herstellung, nicht.

GLOBALER WASSERSTRESS UND KI

Diese ersten globalen Schätzungen und die weiterhin bestehenden Unsicherheiten über die tatsächlichen Verbräuche sind besonders kritisch vor dem Hintergrund des wachsenden globalen Wasserstresses, der das Verhältnis zwischen Wassernachfrage und erneuerbarer Wasserversorgung beschreibt. In den letzten 40 Jahren ist die weltweite Wassernutzung jährlich um etwa 1 % gestiegen, was die Belastung bestehender Wasserressourcen weiter verstärkt. Bereits heute sind rund 50 % der Weltbevölkerung von Wasserstress betroffen (United Nations, 2023). Laut aktuellen Daten des Aqueduct Water Risk Atlas vom World Resources Institute (WRI) leiden 25 Länder – in denen etwa ein Viertel der globalen Bevölkerung lebt – jedes Jahr unter extremem Wasserstress und verbrauchen nahezu ihre gesamte verfügbare Wassermenge. Mindestens vier Milliarden Menschen erleben mindestens einen Monat pro Jahr eine hohe Wasserknappheit (WRI, 2023). Prognosen zufolge könnte sich die Situation bis 2050 erheblich verschärfen, insbesondere in urbanen Regionen (He et al., 2021; Munia et al., 2020; United Nations, 2023). Der Klimawandel verschärft die Problematik zusätzlich, da veränderte Niederschlagsmuster und häufigere Extremwetterereignisse die Wasserverfügbarkeit weiter beeinträchtigen (IPCC, 2023). Selbst in einem optimistischen Szenario, in dem der globale Temperaturanstieg bis 2100 auf 1,3 bis 2,4 Grad Celsius begrenzt wird, könnten bis 2050 dennoch eine Milliarde Menschen zusätzlich unter extremem Wasserstress leben (WRI, 2023).

Auch Europa bleibt von dieser Entwicklung nicht verschont. Prognosen des Water Risk Atlas vom WRI zufolge werden bis 2030 insbesondere Spanien, Italien, Belgien und Griechenland stark von Wasserstress betroffen sein. Auch Teile Deutschlands, insbesondere Brandenburg sowie Regionen in Mitteldeutschland, darunter Teile des Rhein-Main-Gebiets um Frankfurt, werden voraussichtlich unter mittlerem bis hohem Wasserstress leiden (WRI, n.y.). Trotz der alarmierenden Entwicklung wird keines der Unterziele des Sustainable Development Goals „Clean Water and Sanitation for All“ erreicht, das von den Vereinten Nationen im Jahr 2015 beschlossen wurde (United Nations, 2023).

Vor diesem Hintergrund verschärft der stark gestiegene Neubau von RZ die Problematik zusätzlich. Die Anzahl von RZ hat in den letzten Jahren europaweit deutlich zugenommen (Statista, 2024), wobei viele dieser Anlagen in Regionen errichtet wurden, die bereits unter Wasserknappheit leiden oder davon bedroht sind (siehe Abbildung 1). Dies kann zu Nutzungskonflikten zwischen dem steigenden Wasserbedarf der digitalen Infrastruktur und anderen gesellschaftlichen sowie ökologischen Anforderungen führen.

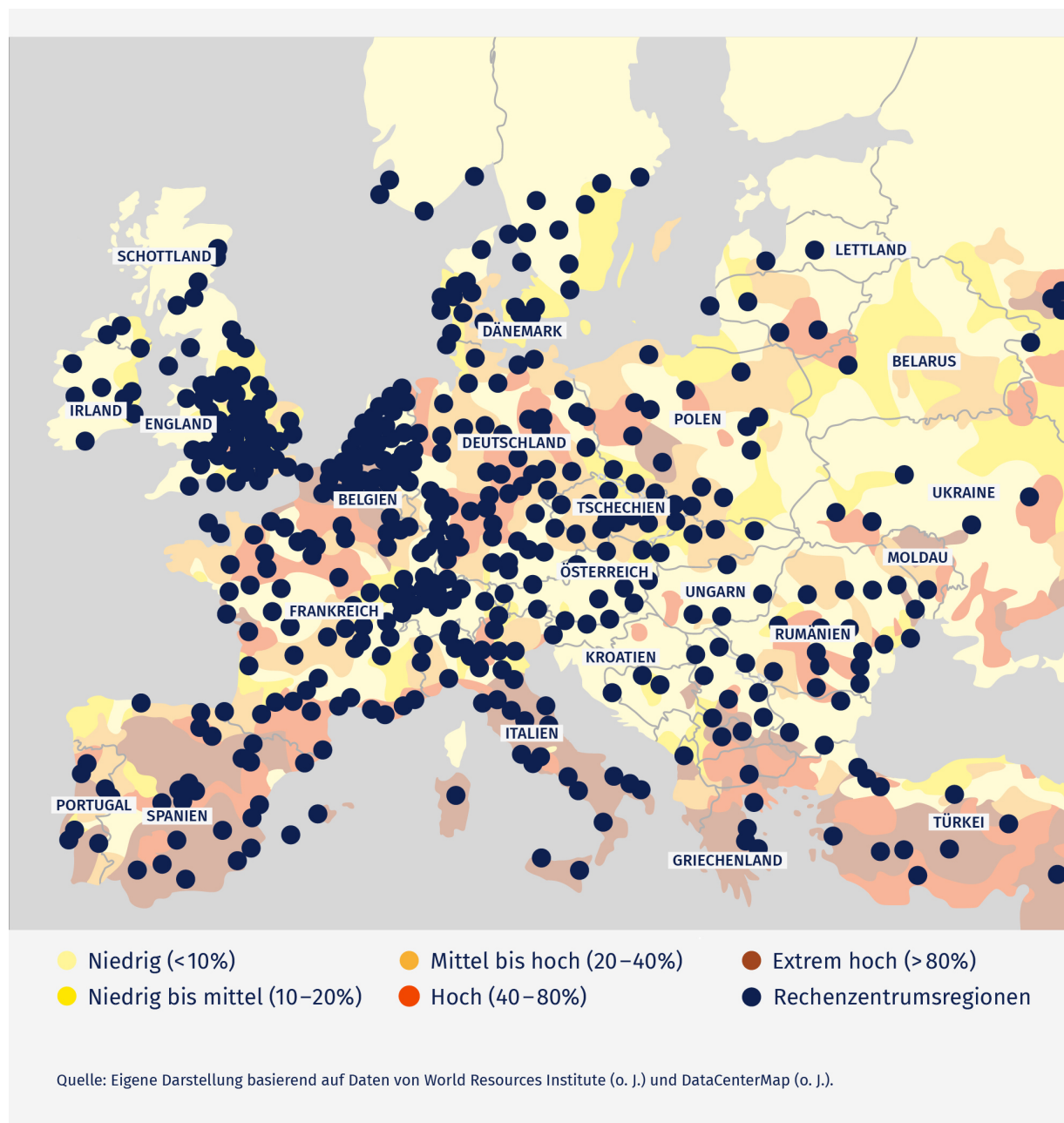


Abbildung 1: Wasserstress und RZ in Europa.

Die wachsenden Wasserverbräuche der IT-Branche führen zunehmend zu Konflikten: Weltweit entstanden Proteste gegen neue Rechenzentren, etwa 2019 in South Carolina und 2023 in Uruguay, wo Google-Projekte aufgrund lokaler Bedenken über Wassernutzung angepasst wurden (Copley, 2022; Livingstone, 2023). Parallel dazu dokumentieren die Nachhaltigkeitsberichte großer IT-Unternehmen deutliche Anstiege: Microsoft

verzeichnete zwischen 2021 und 2022 einen Anstieg des Wasserverbrauchs um 34 % (Microsoft, 2022), Google um 20 % (Google, 2023).

Laut Forscher*innen der University of California, Riverside, steht dieser Anstieg in direktem Zusammenhang mit der wachsenden Rechenkapazität von Google, die maßgeblich durch den verstärkten Einsatz von KI-Technologien vorangetrieben wurde. Der steigende Wasserverbrauch deutet darauf hin, dass der zunehmende Einsatz energieintensiver KI-Modelle in RZ direkte Auswirkungen auf den Wasserbedarf hat (Langley, 2023). Diese Entwicklung verdeutlicht die Dringlichkeit, die Auswirkungen von KI auf den Wasserverbrauch systematisch zu untersuchen.

ZIEL UND AUFBAU DER STUDIE

Vor diesem Hintergrund untersucht die vorliegende Studie die Auswirkungen der KI-Entwicklung auf den Wasserverbrauch und analysiert technologische, ökonomische und regulatorische Hindernisse, die einer nachhaltigen Wassernutzung entgegenstehen. Dabei wird insbesondere analysiert, welche Faktoren den Wasserverbrauch beeinflussen und welche Maßnahmen dazu beitragen können, ihn zu reduzieren. Die zentralen Forschungsfragen lauten:

1. Wie beeinflusst die zunehmende Nutzung von KI den globalen Wasserverbrauch und welche Faktoren bestimmen den Wasserverbrauch von KI in verschiedenen Phasen des Lebenszyklus?
2. Welche Möglichkeiten gibt es, den direkten Wasserfußabdruck von KI und insbesondere von RZ zu reduzieren?
3. Welche technologischen, ökonomischen und regulatorischen Hindernisse stehen der Umsetzung wassersparender Lösungen und Innovationen entgegen?

Diese Fragen bilden den Rahmen der Studie und dienen als Grundlage für die Analyse der ökologischen Auswirkungen von KI sowie für die Ableitung konkreter Handlungsempfehlungen zur Reduktion des Wasserverbrauchs.

Der Bericht gliedert sich in die folgenden Abschnitte: Kapitel 2 liefert die wissenschaftliche Grundlage. In Kapitel 3 wird die methodische Vorgehensweise beschrieben. Die Ergebnisse der Untersuchung sind zur besseren Übersichtlichkeit in drei Kapitel unterteilt. Kapitel 4 befasst sich hauptsächlich mit den Ergebnissen der Literaturrecherche und beantwortet die erste Forschungsfrage, die sich mit den First-Order-Effekten von KI auf den Wasserverbrauch auseinandersetzt. Kapitel 5 beschreibt die in der Studie identifizierten Einflussfaktoren auf den Wasserverbrauch von KI sowie dazugehörige Best Practices, während Kapitel 6 die identifizierten Herausforderungen behandelt. Aufbauend auf diesen Ergebnissen gibt Kapitel 7 Handlungsempfehlungen für Wirtschaft, Politik und Forschung. Abschließend fasst Kapitel 8 im Fazit die zentralen Erkenntnisse zusammen, reflektiert die Limitationen der Studie und gibt einen Ausblick auf zukünftige Forschungsbedarfe.

2. Wissenschaftlicher Hintergrund

2.1. METHODISCHE ANSÄTZE ZUR ERFASSUNG DER DIGITALISIERUNGSBEDINGTEN UMWELTAUSWIRKUNGEN

Die Wahrnehmung digitaler Technologien als immaterielle Güter vernachlässigte lange die physische Infrastruktur, die für ihren Betrieb erforderlich ist. Tatsächlich basiert jede Form der Digitalisierung auf einer materiellen Grundlage: Hardwarekomponenten wie Prozessoren, Speichergeräte und Netzwerkinfrastrukturen sind mit erheblichen Ressourcenverbräuchen verbunden – wobei unterschiedliche Software-Anwendungen diese Ressourcen in sehr unterschiedlichem Maße beanspruchen (Gröger et al., 2018). Diese software-bedingten Anforderungen bestimmen maßgeblich den Energie-, Rohstoff- und Wasserbedarf der digitalen Infrastruktur – sei es für die Herstellung von Halbleitern, den Betrieb von RZ oder die Kühlung der darin betriebenen Server. In den letzten Jahren hat die wissenschaftliche Auseinandersetzung mit diesen software-induzierten Umweltauswirkungen zugenommen. Dabei kommen unterschiedliche methodische Ansätze zum Einsatz, die auf bestehenden Konzepten zur Erfassung der Umweltwirkungen von Informations- und Kommunikationstechnologien (IKT) basieren.

DIREKTE UND INDIREKTE EFFEKTE

Die Umweltauswirkungen digitaler Technologien lassen sich in direkte und indirekte Effekte und Umweltauswirkungen unterteilen.

- **Direkte Effekte** beziehen sich auf den unmittelbaren Ressourcenverbrauch und die Emissionen, die bei Herstellung, Nutzung und Entsorgung von IKT entstehen – etwa der Energiebedarf von RZ oder die CO₂-Emissionen bei der Halbleiterproduktion.
- **Indirekte Effekte** ergeben sich aus veränderten Konsum- und Produktionsmustern, die durch digitale Technologien angestoßen werden (Bieser & Hilty, 2018).

FIRST-, SECOND- UND THIRD-ORDER-EFFEKTE

Ein weiteres etabliertes methodisches Konzept ist die Unterscheidung in First-, Second- und Third-Order-Effekte von IKT (Hilty & Aebischer, 2015), die mittlerweile auch auf Software- und KI-Systeme übertragen wurde (Ligozat et al., 2022).

- **First-Order-Effekte:** Direkte Umweltauswirkungen, die durch die Herstellung, Nutzung und Entsorgung von Hardware und Software entstehen. Dazu zählt z. B. der Wasserverbrauch während der Produktion von Halbleitern.
- **Second-Order-Effekte:** Indirekte Umweltfolgen, die sich aus der Nutzung digitaler Technologien ergeben, etwa Effizienzsteigerungen oder verändertes Nutzerverhalten. Ein Beispiel hierfür ist der Einsatz von KI zur Optimierung des Energieverbrauchs, etwa durch intelligente Steuerung von Heizungs-, Beleuchtungs- oder Produktionssystemen.
- **Third-Order-Effekte:** Langfristige, systemische Veränderungen durch Digitalisierung – z. B. Rebound-Effekte, bei denen Effizienzgewinne durch steigende Nachfrage überkompensiert werden. So kann etwa die zunehmende Effizienz von KI-Hardware leistungsfähigere KI-Modelle ermöglichen, was wiederum die Nachfrage nach rechenintensiven Anwendungen erhöht.

KATEGORIEN AN UMWELTAUSWIRKUNGEN DURCH KI

Ein weiterer methodischer Rahmen zur Einordnung der Umweltauswirkungen von KI wurde von Kaack et al. (2022) entwickelt. Dieser unterscheidet zwischen drei Wirkungskategorien, die sich entlang des Lebenszyklus und der Nutzung von KI-Systemen verorten lassen:

- **Rechenbezogene Effekte (*Compute-related impacts*):** Diese umfassen die Umwelteffekte, die durch Rechenprozesse entlang des KI-Lebenszyklus verursacht werden – sowohl durch den Stromverbrauch bei maschinellen Lernprozessen als auch durch die *embodied*-Emissionen die im Lebenszyklus der eingesetzten Hardware entstehen.
- **Anwendungsbezogene Effekte (*Immediate application impacts*):** Dies sind Effekte, die unmittelbar mit der kurzfristigen Anwendung von KI verbunden sind, beispielsweise Effizienzgewinne in der Energienutzung, Logistik oder industriellen Produktion.
- **Systemische Effekte (*System-level impacts*):** Dazu zählen langfristige Veränderungen auf gesamtgesellschaftlicher Ebene – etwa durch automatisierte Entscheidungsprozesse oder personalisierte Werbung, die Konsummuster beeinflussen. Auch Rebound-Effekte fallen in diese Kategorie (Kaack et al., 2022).

2.2. LEBENSZYKLUS VON KI-MODELLEN UND DEREN UMWELTAUSWIRKUNGEN

Die Umweltwirkungen eines Produkts entstehen entlang seines gesamten Lebenszyklus – von der Rohstoffgewinnung über Produktion, Transport und Nutzung bis zur Entsorgung. Diese Lebenszyklusperspektive gilt auch für Softwareprodukte, die in ihren Lebenszyklus unterschiedliche Ansprüche an Hardware stellen. Welche wiederum produziert, mit Energie versorgt und schließlich entsorgt werden muss. Somit trägt jedes Softwareprodukt zu den Umweltfolgen der dafür benötigten Hardware bei (Hilty & Aebischer, 2015). Ein bewährtes Instrument zur Erfassung dieser Belastungen ist die Lebenszyklusanalyse (*Life-Cycle Assessment*, LCA), mit der die Umweltauswirkungen eines Produkts oder einer Dienstleistung über ihren gesamten Lebenszyklus hinweg systematisch erfasst und bewertet werden. Die Umweltauswirkungen von KI lassen sich ebenfalls entlang des gesamten KI-Lebenszyklus analysieren. Je nach Quelle werden die Lebensphasen unterschiedlich eingeteilt, jedoch umfassen gängige Modelle folgende Phasen (siehe auch Abbildung 2): (1) Planung und Datenerfassung, (2) Modellentwicklung und Training (3) Bereitstellung und Inferenz, (4) Monitoring und Wartung, sowie (5) End-of-Life.

Die bisherigen wissenschaftlichen Untersuchungen zu den Umweltauswirkungen von KI konzentrieren sich vorrangig auf den Energieverbrauch und die damit verbundenen Treibhausgasemissionen (Luccioni et al., 2022; Strubell et al., 2019). Besonders die Trainingsphase großer KI-Modelle, etwa neuronaler Netzwerke für Sprachverarbeitung und Bildanalyse, erfordert erhebliche Rechenkapazitäten und ist mit einem hohen Energieaufwand verbunden. Strubell et al. (2019) schätzen, dass das Training und Feintuning eines großen Transformer-Modells unter Einsatz von Neural Architecture Search (NAS) Emissionen von 284.019 kg CO₂ verursachte – ein Wert, der den lebenslangen Emissionen von fünf US-amerikanischen Autos entspricht. Die CO₂-Emissionen variieren dabei erheblich zwischen verschiedenen KI-Modellen (Luccioni & Hernandez-Garcia, 2023). Eine weitere Schätzung der CO₂-Emissionen von KI-Modellen lieferten Luccioni et al. (2022), die schätzen, dass das Training des BLOOM-Modells mit 176 Milliarden Parametern etwa 24,7 Tonnen CO₂-Äquivalent verursacht hat, wenn nur der dynamische Energieverbrauch berücksichtigt wird – und 50,5 Tonnen, wenn alle Prozesse einbezogen werden, von der Herstellung der Hardware bis hin zum energiebedingten Verbrauch während des Betriebs.

Neben den während der Trainingsphase verursachten Energie- und Ressourcenverbräuchen zeigen weitere Studien, dass insbesondere generative KI, wie große Sprachmodelle (z.B. ChatGPT), auch in der Inferenzphase

einen erheblichen Anteil am Gesamtenergieverbrauch ausmacht. Während eine einzelne Inferenz deutlich weniger Rechenleistung erfordert als das Training desselben Modells, kann der kumulierte Energieverbrauch aufgrund der hohen Nutzungshäufigkeit beträchtlich sein. Der Energieaufwand pro Inferenz variiert je nach Modellgröße und Architektur (Luccioni et al., 2024). Insgesamt sind die vorhandenen Messungen jedoch nur eingeschränkt vergleichbar, da sie auf unterschiedlichen Systemgrenzen, Maßeinheiten und Methoden basieren.

2.3. WASSERNUTZUNG, WASSERVERBRAUCH UND WASSERQUALITÄT ENTLANG DES KI-LEBENSZYKLUS

Im Gegensatz zum Energieverbrauch gewann der Wasserbedarf durch KI erst kürzlich Aufmerksamkeit in Wissenschaft und Medien. Aus einer Lebenszyklusperspektive zeigt sich, dass KI-Systeme in jeder Phase unterschiedliche Hardware-Anforderungen stellen, die jeweils mit spezifischem Wasserverbrauch einhergehen – von der Chipproduktion über die Montage bis hin zum laufenden Serverbetrieb.

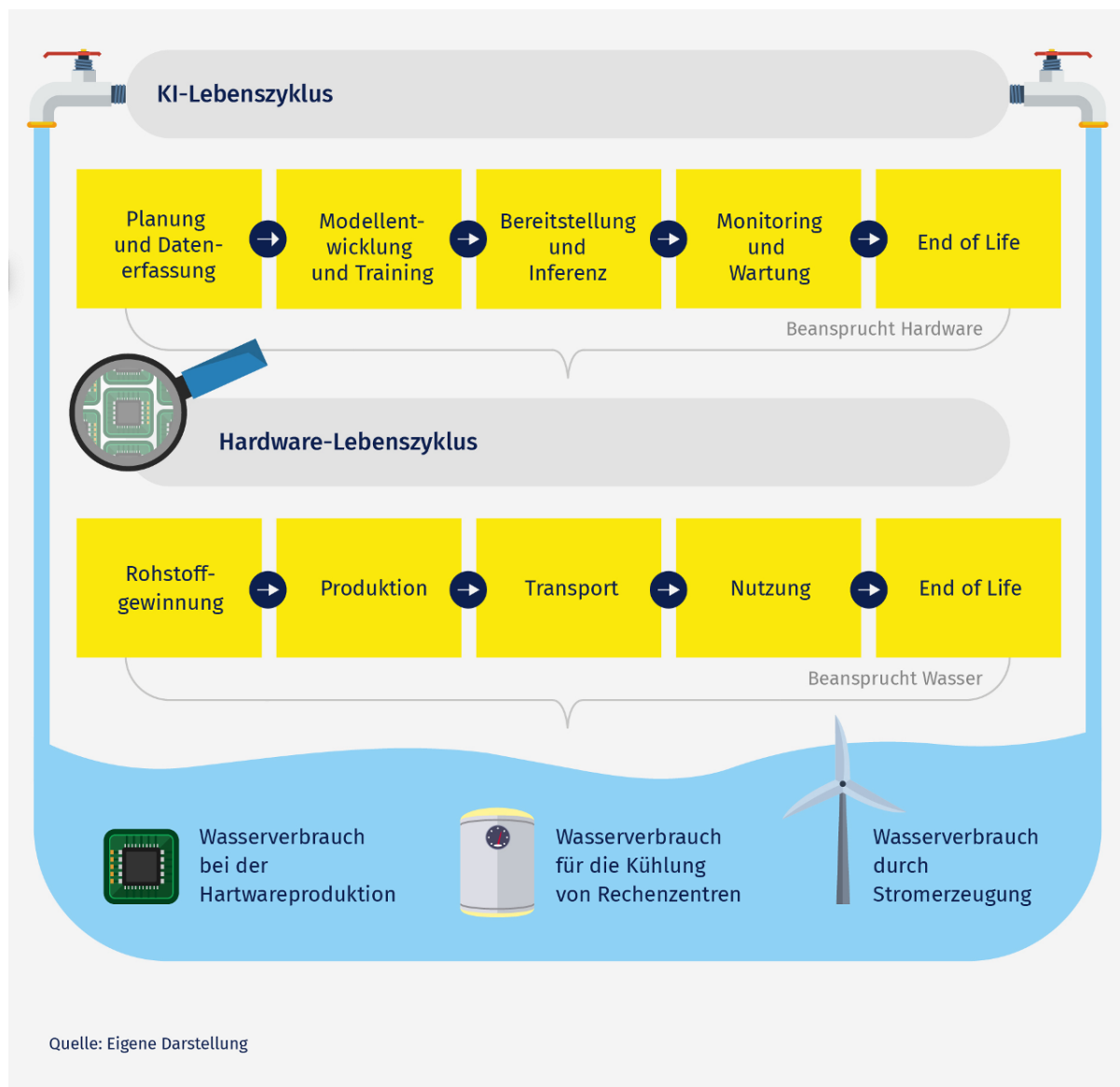


Abbildung 2: Der Einfluss von KI auf den Wasserverbrauch entlang des Hardware-Lebenszyklus.

Wie Abbildung 2 verdeutlicht, erstreckt sich der Wasserverbrauch über den gesamten KI-Lebenszyklus. Bisherige Untersuchungen fokussieren sich dabei auf drei Hauptbereiche: den direkten Wasserbedarf (*On-site Water*) durch die Kühlung von Rechenzentren, den indirekten Wasserbedarf (*Off-site Water*) durch die Energieerzeugung für KI-Systeme sowie den Wasserverbrauch während der Hardwareproduktion, insbesondere bei der Herstellung von Computerchips.

WASSERNUTZUNG UND WASSERVERBRAUCH UND WASSERQUALITÄT

In diesem Zusammenhang ist zwischen den Begriffen Wassernutzung und Wasserverbrauch zu unterscheiden: Wassernutzung (*water use*) bezeichnet die gesamte Menge an Wasser, die einem natürlichen oder künstlichen

Reservoir entnommen (*water withdrawal*) wird – unabhängig davon, ob sie anschließend zurückgeführt wird. Wasserverbrauch (*water consumption*) hingegen beschreibt jenen Anteil, der – etwa durch Verdunstung – nicht mehr weiter zur Verfügung steht, also die genutzte Wassermenge abzüglich des abgeleiteten Wassers. Der Wasserverbrauch in Verbindung mit der Qualität des abgeleiteten Wassers ist ein wichtiger Indikator für den Wasserstress in einer Region (Reig, 2013). Aufgrund der häufig unzureichenden Datenlage zum abgeleiteten Wasser wird jedoch oftmals auf die Wassernutzung zurückgegriffen.

WASSERQUALITÄT UND RECHENZENTREN

Neben der genutzten Wassermenge spielt auch die Qualität des eingesetzten und abgeleiteten Wassers eine entscheidende Rolle. Für den zuverlässigen und energieeffizienten Betrieb von RZ muss Kühlwasser in der Regel bestimmte physikalisch-chemische Anforderungen erfüllen, um Korrosion, Ablagerungen oder mikrobielles Wachstum zu verhindern. Diese Anforderungen hängen vom Design des Rechenzentrums, der eingesetzten Kühltechnologie, dem Standort sowie der jeweiligen Betriebsweise ab (Zucon et al., 2023). Je schlechter die Wasserqualität, desto häufiger sind sogenannte Blowdown-Prozesse erforderlich – also das Spülen von Leitungen, um Ablagerungen zu vermeiden.

Wegen dieser Anforderungen greifen viele RZ auf Trinkwasser zurück, da alternative Wasserquellen wie Regen-, Grau- oder Oberflächenwasser häufig als ungeeignet gelten (Mytton, 2021). Zwar gibt es Fortschritte beim Einsatz von recyceltem oder nicht-trinkbarem Wasser, doch deren Anteil ist bislang gering. Ein Beispiel für die Nutzung alternativer Wasserquellen ist das Google-Rechenzentrum im finnischen Hamina: Dort wird seit der Eröffnung Meerwasser zur Kühlung eingesetzt. Mithilfe bestehender Rohrleitungen aus der Zeit als Papierfabrik wird kaltes Meerwasser in Wärmetauscher des Rechenzentrums gepumpt. Das Meerwasser bleibt dabei vom Süßwasser getrennt, das im Wärmetauscher zirkuliert. Nach der Wärmeaufnahme wird das erhitzte Wasser mit kälterem Meerwasser vermischt und anschließend ins Meer zurückgeführt (Mytton, 2021).

Insgesamt fehlen detaillierte Daten zu Herkunft, Nutzung und Qualität des verwendeten Wassers – sowohl hinsichtlich der Wasserentnahme als auch der Abgabe von Kühl- oder Reinigungswasser. Es besteht ein erheblicher Forschungsbedarf, insbesondere um Unterschiede zwischen einzelnen Rechenzentrumsdesigns und Standorten berücksichtigen zu können (Mytton, 2021).

WASSERQUALITÄT IN DER HALBLEITERPRODUKTION

Auch in der Halbleiterindustrie ist die Qualität des verwendeten Wassers von zentraler Bedeutung, da selbst geringste Verunreinigungen die Funktion empfindlicher Bauteile beeinträchtigen können. Besonders bei der Reinigung von Wafern und der Herstellung von Ultrapure Water (UPW) – extrem reinem Wasser ohne organische oder anorganische Rückstände – gelten höchste Qualitätsstandards (Wang et al., 2023).

Ein erheblicher Teil des verwendeten Wassers wird nach seiner Nutzung wieder abgeführt, ist dann jedoch häufig chemisch oder thermisch belastet und muss entsprechend aufbereitet werden. Moderne Halbleiterfabriken investieren daher in fortschrittliche Recycling- und Wasseraufbereitungssysteme, um eine Mehrfachnutzung zu ermöglichen. Dennoch sind auch hier öffentlich zugängliche, verlässliche Daten zur Qualität und Behandlung des verwendeten Wassers bislang begrenzt.

2.4. WASSERNUTZUNG IN RECHENZENTREN

RZ sind essenzielle Infrastrukturen für die Entwicklung und den Betrieb von KI, sie ermöglichen die Verarbeitung und Speicherung großer Datenmengen. Die leistungsstarken Server erzeugen erhebliche Mengen an Abwärme, die ohne effiziente Kühlung zu Überhitzung, Systemausfällen, Leistungsverlusten und einer verkürzten Hardware-Lebensdauer führen können. Untersuchungen zeigen, dass bereits eine Temperaturerhöhung um 10°C die Lebensdauer von IT-Hardware um bis zu 50% reduzieren kann (J. Zhang et al., 2019). Mit der zunehmenden Entwicklung und Nutzung von KI-Technologien müssen sich RZ an neue Anforderungen anpassen. KI-Modelle erfordern eine hohe Rechenleistung, die meist durch spezialisierte Hardware wie Grafikprozessoren (GPUs) und Tensor Processing Units (TPUs) bereitgestellt wird. Diese Komponenten erzeugen deutlich mehr Wärme als herkömmliche Central Processing Unit (CPU)-Workloads, weshalb effiziente Kühlsysteme unverzichtbar sind (Google, 2023; Zheng et al., 2022). Bereits 2016 prognostizierte Shehabi et al., dass RZ in den USA im Jahr 2020 etwa 660 Milliarden Liter Wasser verbrauchen würden – davon entfallen durchschnittlich 20% auf die direkte Kühlung vor Ort und 80% auf die Stromerzeugung. Der direkte Wasserverbrauch ist größtenteils auf die Kühlung zurückzuführen, die je nach Studie und untersuchtem RZ auch etwa 30-40% des Gesamtenergieverbrauchs ausmacht (Shehabi et al., 2016; Q. Zhang et al., 2021).

Kühlsysteme von RZ lassen sich nach verschiedenen Kriterien klassifizieren. Eine verbreitete Unterteilung erfolgt nach dem primären Kühlmedium, das die Wärme direkt von den IT-Komponenten abführt. Grundsätzlich unterscheidet man zwischen Luftkühlung und Flüssigkeitskühlung (Chang et al., 2024; Ott et al., 2024; UBA, 2024; Y. Zhang et al., 2022). Beide Varianten haben eine Reihe von unterschiedlichen Unterkategorien. Der Luftkühlung wird auch oftmals die Verdunstungskühlung wie der Varianten wie der adiabaten Kühlung zugeordnet (UBA, 2024). Zudem gibt es noch RZ die auf Freier Kühlung entweder basieren und in der Praxis kommen häufig hybride Systeme zum Einsatz, die verschiedene Kühlprinzipien kombinieren – beispielsweise luftgekühlte Server mit wassergekühlten Rückkühlern (Huang et al., 2022). Wichtig ist dabei zu beachten, dass auch primär luftbasierte Systeme sekundär Wasser zur Wärmeabfuhr nutzen können, etwa in Kühltürmen.

Die Kühlsysteme von RZ bestehen nach Zhang et al. (2021) aus zwei Hauptkomponenten: Das mechanische Kältesystem (Mechanical Refrigeration Sub-System, MRSS) umfasst Geräte wie Kaltwassersätze (Chiller), Pumpen und Kühltürme und stellt die Kältebereitstellung sicher. Chiller lassen sich in luftgekühlte Kondensatoren (air-cooled chiller) und wassergestützte Kondensatoren mit Kühltürmen (water-cooled chiller) unterteilen. Luftgekühlte Chiller nutzen Außenluft zur Wärmeabfuhr und sind energieintensiver (nahezu kein Wasserverbrauch), während wassergekühlte Chiller Kühlwasser aus Kühltürmen zur Wärmeabfuhr verwenden und dadurch einen höheren Wasserverbrauch aufweisen (Kiriimi et al., 2021). Das terminale Kühlsystem (Terminal Cooling Sub-System, TCSS) überträgt die Wärme aus dem Innenraum des Rechenzentrums über das MRSS an die Außenumgebung. Diese Zweiteilung ermöglicht eine effiziente Wärmeabfuhr bei gleichzeitig stabilen Betriebsbedingungen für die IT-Systeme.

Die folgenden Abschnitte geben wir einen kurzen Überblick über drei verbreitete Kühltechnologien von RZ.

LUFTBASIERTE KÜHLUNG

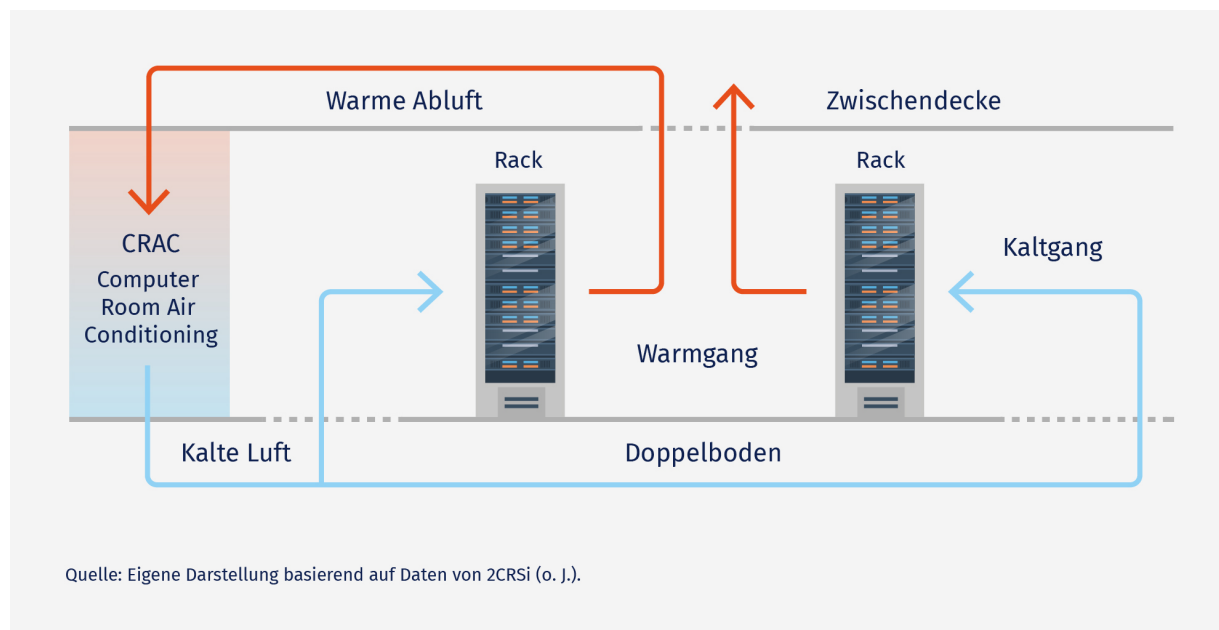


Abbildung 3: Luftbasierte Kühlung - Computer Room Air-Conditioning.

Luftkühlung ist die am weitesten verbreitete Kühltechnologie in Rechenzentren (Zhang et al., 2021). Sie kann weiter Unterteilt werden in Direkte Luftkühlung und Verdunstungskühlung (Chang et al., 2024). Zudem lässt sich nach ihrer Implementierungsebene in drei Kategorien unterteilen:

- Bei der raumbasierten Kühlung wird der gesamte Serverraum klimatisiert. Die kalte Luft wird typischerweise über ein Plenum unter einem Doppelboden zugeführt und durch perforierte Bodenplatten in den Raum geleitet. Die erwärmte Luft steigt nach oben und wird über Deckenauslässe abgeführt. Diese Methode eignet sich besonders für Rechenzentren mit geringer Gerätedichte und bietet hohe Flexibilität bei der Luftverteilung.
- Die reihenbasierte Kühlung positioniert spezialisierte Kühleinheiten direkt zwischen den Serverreihen, wodurch sich die Luftwege zwischen Kühleinheit und IT-Geräten verkürzen und die Effizienz steigt.
- Bei der rackbasierten Kühlung sind die Kühlsysteme direkt in die Serverracks integriert und führen die Wärme unmittelbar an der Quelle ab, was sich besonders für hochdichte Installationen mit hoher Wärmelast pro Rack eignet (Q. Zhang et al., 2021).

Die Kühlung erfolgt über zwei Haupttypen mechanischer Kühlsysteme. CRAC-Systeme (Computer Room Air Conditioning) funktionieren wie große Klimaanlage mit integrierten Kältemaschinen. Je nach Rückkühlmethode wird die Wärme entweder über einen externen Kondensator an die Umgebungsluft abgegeben (luftgekühlt), mittels Glykol zu einem externen Trockenrückkühler transportiert (glykolgekühlte Variante) oder durch Verdunstung in Kühltürmen abgeführt (wassergekühlt) (Huang et al., 2022). CRAH-Systeme (Computer Room Air Handler) arbeiten dagegen mit Kaltwasser aus einem zentralen Kaltwassersatz, wobei das erwärmte Wasser zwischen CRAH-Einheit und Kaltwassersatz zirkuliert (Habibi Khalaj & Halgamuge, 2017).

Abbildung 3 zeigt eine typische raumbasierte Luftkühlung. Das Grundprinzip ist einfach: Kalte Luft strömt durch die Serverracks, nimmt die von den Komponenten erzeugte Wärme auf und wird als Warmluft wieder abgeführt. Um eine unkontrollierte Vermischung von kalter und warmer Luft zu vermeiden, werden Serverracks in der Hot-Aisle/Cold-Aisle-Konfiguration angeordnet, bei der heiße und kalte Gänge voneinander getrennt sind (2CRSi, o. J.; UBA, 2024). Die CRAC- oder CRAH-Einheiten erzeugen kontinuierlich kalte Luft, die durch das Plenum unter dem Doppelboden strömt. Durch perforierte Bodenplatten im Kaltgang wird die Luft vor die Serverracks geleitet. Angetrieben durch den Druckunterschied zwischen Plenum und Serverraum durchströmt die Kaltluft die Racks und nimmt dabei die Abwärme auf. Die erwärmte Luft wird über den Heißgang zur Decke geführt und über Lüftungsöffnungen zu den Kühleinheiten zurückgeleitet, wo der Kreislauf von neuem beginnt.

Obwohl Luft als primäres Kühlmedium fungiert, spielt Wasser eine wichtige Rolle im Kühlprozess. Der Wasserverbrauch hängt stark von der gewählten Rückkühlmethode ab: Luftgekühlte Kaltwassersätze sind energieintensiver, verbrauchen aber nahezu kein Wasser, während wassergekühlte Kaltwassersätze mit Kühltürmen wasserintensiver, aber energieeffizienter sind (Karimi et al., 2022). Der Wasserverbrauch entsteht hauptsächlich durch Verdunstungskühlung in Kühltürmen, Abschlämmungsprozesse zur Vermeidung von Ablagerungen, Reinigung von Komponenten sowie Luftbefeuchtung zur Einhaltung optimaler Betriebsbedingungen.

Eine zusätzliche Möglichkeit zur Effizienzsteigerung bietet die Freikühlung, bei der kalte Außenluft direkt als Kühlquelle genutzt wird. Diese umfasst sowohl direkte Freikühlung (Außenluft wird direkt in den Serverraum geleitet) als auch indirekte Freikühlung (Wärmeübertragung über Wärmetauscher) (Y. Zhang et al., 2022). Trotz steigender Bedeutung alternativer Technologien bleibt die Luftkühlung aufgrund ihrer weiten Verbreitung und vergleichsweise niedrigen Implementierungskosten ein wichtiger Standard in der Rechenzentrumsbranche.

ADIABATE KÜHLUNG

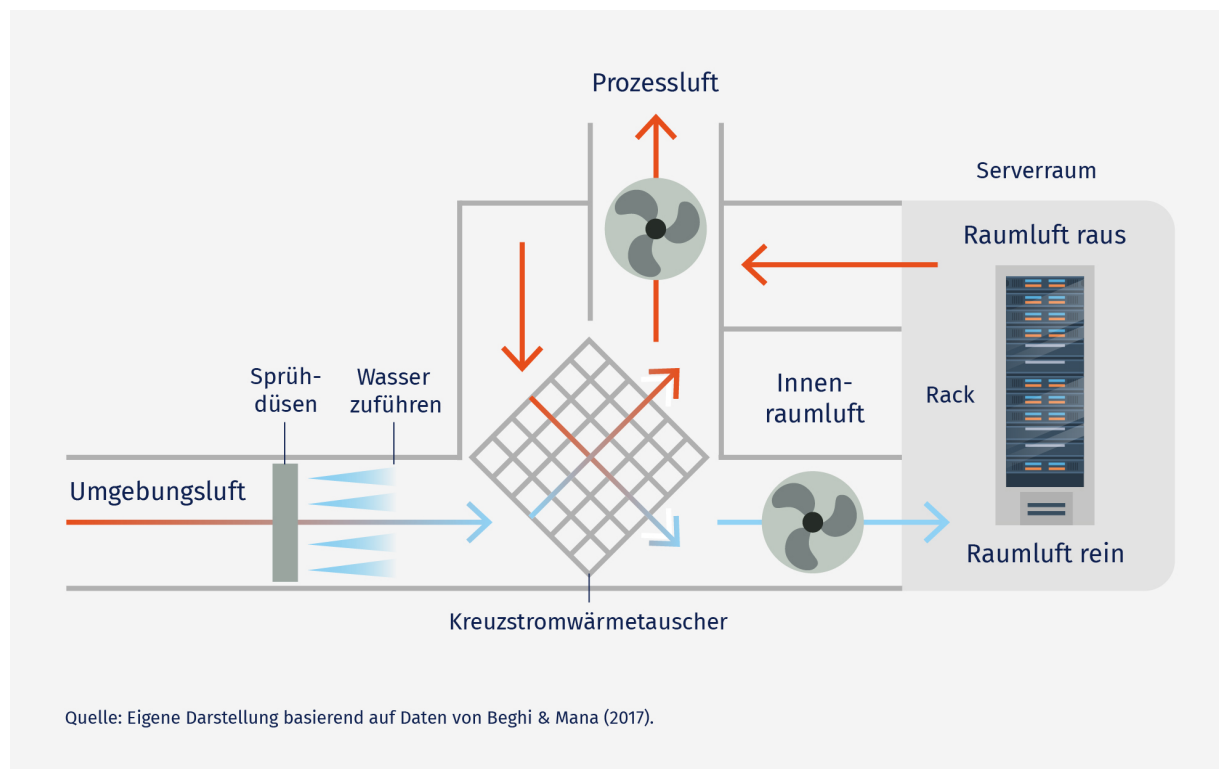


Abbildung 4: Indirekte adiabate Kühlung für Serverräume: Sprühdüsen, Ventilatoren, Wärmetauscher und Serverraum.

Die adiabate Kühlung ist eine Weiterentwicklung klassischer Luftkühlungssysteme, bei der zusätzlich die Verdunstung von Wasser zur Abkühlung der Luft genutzt wird. Sie erweitert den Anwendungsbereich konventioneller Luftkühlung insbesondere in heißen und trockenen Klimazonen. Die adiabate Kühlung nutzt die Verdunstung von Wasser, um die Lufttemperatur zu senken. Dabei wird die Verdunstungsenthalpie genutzt – wenn Wasser verdunstet, entzieht es der umgebenden Luft Wärme und kühlt diese ab. In Rechenzentren erweitert die adiabate Kühlung den Betriebsbereich der Freikühlung: Wenn die Außentemperatur für reine Freikühlung zu hoch ist, kann durch Verdunstung die Lufttemperatur soweit gesenkt werden, dass eine Kühlung ohne mechanische Kälteerzeugung möglich bleibt. In der Praxis werden die Begriffe "adiabate Kühlung" und "Verdunstungskühlung" oft synonym verwendet, da beide auf dem Prinzip der Verdunstung basieren. Der Unterschied liegt hauptsächlich in der Anwendung: Während in Kühltürmen das Prozesswasser direkt gekühlt wird, kühlt die adiabate Kühlung in Rechenzentren primär die Zuluft (Beghi et al., 2017; VISTECH, 2021).

Es gibt zwei Hauptvarianten der adiabaten Kühlung:

1. Indirekte adiabate Kühlung

Bei der indirekten adiabaten Kühlung, wie in Abbildung 4 dargestellt, existieren zwei getrennte Kühlkreisläufe: Einer enthält Luft oder Wasser aus dem Rechenzentrum, der andere gekühlte Außenluft oder ein weiteres flüssiges Kühlmedium. Beide Kreisläufe sind über Wärmetauscher, in diesem Fall ein Kreuzstromwärmetauscher, miteinander verbunden. Die Außenluft wird mit Wasser besprüht, wodurch sie sich erwärmt oder verdunstet und somit Wärme aufnimmt. Anschließend gibt diese Außenluft die aufgenommene Wärme über den Wärmetauscher an die Luft aus dem Rechenzentrum ab (Rüdiger, 2019).

2. Direkte adiabate Kühlung

Hierbei wird Wasser direkt in die Zuluft des Rechenzentrums gesprüht. Durch Verdunstung sinkt die Temperatur der Luft, bevor sie in den Serverraum gelangt. Diese Methode ist besonders effizient, kann jedoch zu einer erhöhten Luftfeuchtigkeit innerhalb des Rechenzentrums führen, was je nach Systemkonfiguration berücksichtigt werden muss (Karimi et al., 2022).

Große Technologieunternehmen setzen zunehmend auf Verdunstungskühlung (adiabate Kühlung) (Microsoft, 2023). Besonders Hyperscale-RZ großer Cloud-Anbieter nutzen diese Technik, da sie eine hohe Kühlleistung bei vergleichsweise geringem Energieaufwand ermöglicht (UBA, 2024). Diese Technologie spart zwar Energie, geht jedoch mit einem erhöhten Wasserverbrauch einher. Derzeit nutzen rund ein Drittel der großen RZ in Deutschland Verdunstungskühlung. Expert*innen gehen davon aus, dass diese Technik künftig verstärkt eingesetzt wird, insbesondere um niedrige PUE-Werte zu erreichen und regulatorische Vorgaben einzuhalten (Bitkom, 2024b).

Der Wasser- und Energieverbrauch adiabater Kühlsysteme variiert stark je nach Systemdesign, Betriebsweise und klimatischen Bedingungen. Moderne adiabate Kühler arbeiten oft im Hybridmodus: Bei kühlen Außentemperaturen funktionieren sie als reine Trockenkühler ohne Wasserverbrauch, bei höheren Temperaturen wird die Verdunstungskühlung zugeschaltet. Dadurch kann der jährliche Wasserverbrauch im Vergleich zu kontinuierlich betriebenen Kühltürmen reduziert werden (Kirvan, 2022). Microsoft berichtet beispielsweise, dass ihre adiabaten Kühlsysteme im Jahresdurchschnitt weniger Wasser verbrauchen als traditionelle Kühlturmsysteme, da sie einen Großteil des Jahres im Trockenkühlmodus betrieben werden können. Die indirekte adiabate Kühlung weist dabei einen geringeren Wasserverbrauch auf als die direkte Variante, da keine Befeuchtung der Serverraumluft erfolgt (Microsoft, 2023).

FLÜSSIGKEITSBASIERTE KÜHLUNG

Flüssigkeitsbasierte Kühlung oder auch Flüssigkeitskühlung (*Liquid Cooling*), dargestellt in Abbildung 5, ist eine weitere Alternative zur Luftkühlung. Die Flüssigkeitskühlung nutzt Wasser oder spezielle dielektrische Flüssigkeiten, um die Abwärme direkt an den entscheidenden Komponenten abzuführen (UBA, 2024).

Die Flüssigkühlungstechnologie verwendet in der Regel Flüssigkeiten als Wärmeübertragungsmedium zur Kühlung von Servern. Sie kombiniert dabei einen Trockenkühler oder einen Kühlturm, um eine passive Kühlung zu erreichen, wodurch sie sich besonders für RZ mit hoher Leistungsdichte eignet. Der Hauptvorteil der Flüssigkühlung liegt im höheren Wärmeübergangskoeffizienten im Vergleich zur Luftkühlung, was zu einer geringeren Temperaturdifferenz zwischen der CPU und dem Kühlmedium führt und somit eine effektivere Rückgewinnung von Abwärme aus DCs und TBSs zur Wiederverwendung ermöglicht (Zhang et al. 2021).

Ein zentraler Vorteil dieser Technologie besteht darin, dass die Wärme unmittelbar dort abgeführt wird, wo sie entsteht, ohne dass Luft als Zwischenmedium erforderlich ist. Da Flüssigkeiten eine deutlich höhere

Wärmekapazität als Luft besitzen, kann die Abwärme effizienter abgeführt werden, wodurch sowohl der Energieverbrauch als auch der Kühlaufwand reduziert werden (Data4, 2024)

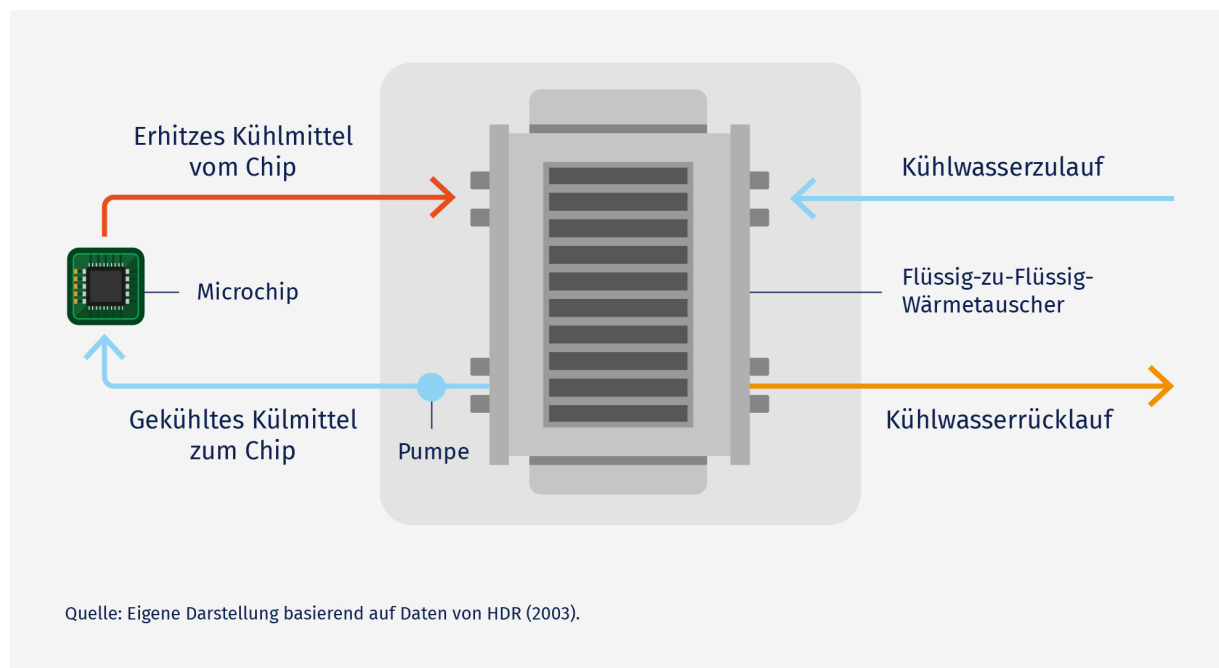


Abbildung 5: Direkte Chipkühlung - Flüssigkeitsbasierte Kühlung.

Es gibt verschiedene Varianten der Flüssigkeitskühlung, die sich hinsichtlich ihrer Effizienz, des Wartungsaufwands und der Integration in bestehende Rechenzentrumsinfrastrukturen unterscheiden. Zwei Varianten (Chang et al. 2024) sind:

- **Indirect liquid cooling:** Indirect liquid cooling is a heat removal process without direct contact between the heat source and liquid coolant. (Habibi Khalaj & Halgamuge, 2017)
- **Direct Liquid Cooling (DLC):** Die Kühlflüssigkeit wird gezielt zu den wärmeerzeugenden Komponenten (CPUs, GPUs) geführt, in der Regel durch sogenannte „Cold Plates“. Diese Methode kann sowohl aktiv als auch passiv betrieben werden, verbessert die Wärmeabfuhr und reduziert den Bedarf an Luftkühlung. DLC kann entweder mit Wasser oder speziellen dielektrischen Flüssigkeiten betrieben werden und umfasst auch die direkte Chipkühlung (Habibi Khalaj & Halgamuge, 2017; HDR, 2023).
- **Immersionenkühlung (Immersion Cooling):** Hierbei werden Server entweder vollständig oder teilweise in eine nicht leitende Flüssigkeit eingetaucht, die die Wärme direkt von den Komponenten aufnimmt. Diese Methode eliminiert den Bedarf an Lüftern, wodurch der Energieverbrauch für Kühlung drastisch reduziert wird. Es gibt zwei Hauptvarianten:
 - *Single-Phase Immersion Cooling:* Die Flüssigkeit bleibt während des gesamten Kühlprozesses im selben Aggregatzustand (flüssig). Die Wärme wird durch Wärmetauscher abgeführt.
 - *Two-Phase Immersion Cooling:* Die Kühlflüssigkeit verdampft bei der Wärmeaufnahme und kondensiert anschließend wieder, um die Wärme noch effizienter abzuführen. Diese Methode bietet eine besonders hohe Kühlleistung, erfordert aber zusätzliche Regelmechanismen.

Abbildung 5 zeigt eine Variante der direkten Chipkühlung, bei der die Wärme in einen Kühlwasserkreislauf abgeführt wird („*Reject Heat to a Cooling Water Loop*“). Diese Methode erfordert einen Flüssigkeitswärmetauscher („*Fluid Heat Exchanger*“), um die Wärme vom wärmeren Direct-to-Chip-Kreislauf in den kälteren Wasserkreislauf zu übertragen (HDR, 2023). Eine Befragung von Bitkom ergab, dass Expert*innen davon ausgehen, dass der Einsatz von Wasser- und Flüssigkeitskühlungen in deutschen RZ weiter an Bedeutung gewinnen wird – insbesondere im Hinblick auf die Einhaltung der PUE-Vorgaben im neuen Energieeffizienzgesetz. Insbesondere Flüssigkeitskühlungskonzepte wie *Immersionskühlung* und DLC werden bei Systemen mit hoher Leistungsdichte zunehmend als essenzielle Technologien betrachtet (Bitkom, 2024b).

Der Wasserverbrauch von Flüssigkeitskühlungssystemen hängt stark vom Systemdesign ab. Geschlossene Systeme, wie *Direct Liquid Cooling* mit *Cold Plates*, können den direkten Wasserverbrauch minimieren, da die Kühlflüssigkeit in einem geschlossenen Kreislauf verbleibt und nur geringfügig nachgefüllt werden muss. Es ist wichtig zu beachten, dass auch geschlossene Flüssigkeitskühlssysteme je nach Rückkühlmethode ebenfalls Wasser verbrauchen können – während Systeme mit Trockenkühlern nahezu wasserfrei arbeiten, benötigen Systeme mit Kühlturm-Rückkühlung durch Verdunstung Wasser. Offene Systeme, insbesondere solche mit Verdunstungskühlung, können hingegen einen hohen Wasserbedarf haben, da das Kühlmedium verdampft und regelmäßig ersetzt werden muss.

Zudem erfordert der Einsatz von Flüssigkeitskühlung eine sorgfältige Planung und Infrastruktur:

- **Leckageschutz:** Systeme mit Wasser als Kühlmittel müssen gegen Leckagen abgesichert sein, um Betriebsausfälle zu vermeiden.
- **Chemische Eigenschaften des Kühlmittels:** Spezielle dielektrische Flüssigkeiten sind oft teurer und können über die Lebensdauer der Systeme hinweg Zusatzkosten verursachen.
- **Integration in bestehende RZ-Infrastrukturen:** Besonders bei Nachrüstungen in bestehenden RZ kann der Umbau auf Flüssigkeitskühlung eine Herausforderung darstellen (Eabel, 2024).

Da moderne KI-Workloads immer mehr Rechenleistung benötigen und herkömmliche Luftkühlssysteme zunehmend an ihre Grenzen stoßen, erwarten Expert*innen, dass Flüssigkeitskühlung in Zukunft eine stärkere Rolle spielen wird (Bitkom, 2024b).

METRIKEN ZUR UMWELTBEWERTUNG VON RECHENZENTREN

Für RZ existieren etablierte Metriken zur Bewertung ihrer Effizienz, darunter die *Power Usage Effectiveness* (PUE) und die *Water Usage Effectiveness* (WUE). Während PUE eine etablierte Metrik darstellt, hat WUE bislang weniger Aufmerksamkeit erhalten. Die PUE ist eine Kennzahl zur Bewertung der Energieeffizienz eines RZ. Sie beschreibt das Verhältnis des gesamten Energieverbrauchs zur tatsächlich von den IT-Geräten genutzten Energie. Ein idealer PUE-Wert von 1,0 würde bedeuten, dass die gesamte zugeführte Energie ausschließlich für IT-Geräte genutzt wird, ohne Verluste durch Kühlung oder andere Infrastruktur (The Green Grid, 2016). In der Praxis liegen die Werte jedoch meist zwischen 1,1 und 2,5, abhängig von Faktoren wie Standort, Klimabedingungen und Kühltechnologie. Seit dem 1. Januar 2025 sind RZ in ganz Europa gemäß der überarbeiteten Energieeffizienz-Richtlinie der Europäischen Union verpflichtet, ihre PUE-Werte gemäß den neuen EU-Vorschriften zu messen und offenzulegen. Im Gegensatz dazu misst die WUE die Wassereffizienz eines RZ, indem sie das Verhältnis des gesamten Wasserverbrauchs zur Energieaufnahme der IT-Geräte in Litern pro Kilowattstunde (l/kWh) angibt (CENELEC, 2022). Allerdings weist diese Metrik einige methodische Einschränkungen auf. Sie erfasst ausschließlich den direkten Wasserverbrauch eines RZ und lässt dabei den Wasserverbrauch unberücksichtigt, der beispielsweise durch die Energieerzeugung oder die Herstellung von Hardware entsteht. Zudem variiert der WUE-Wert erheblich in Abhängigkeit von der verwendeten Kühltechnologie.

FORSCHUNGSBEDARF UND UNTERSUCHUNGSANSATZ

Die theoretischen Grundlagen verdeutlichen die Komplexität des Wasserverbrauchs von KI-Systemen, der sich über den gesamten Lebenszyklus erstreckt. Dabei zeigen sich vielschichtige Zusammenhänge, während ein umfassendes Verständnis des KI-bedingten Wasserverbrauchs noch aussteht. Vor diesem Hintergrund untersucht die vorliegende Studie systematisch die rechenbezogenen Effekte der KI-Nutzung auf die Wassernutzung.

Die Analyse fokussiert sich auf drei zentrale Dimensionen: den direkten Wasserbedarf durch die Kühlung von Rechenzentren, den indirekten Wasserbedarf durch die Energieerzeugung sowie den Wasserverbrauch während der Hardwareproduktion, insbesondere in der Halbleiterindustrie. Dabei wird ein integrierter Ansatz verfolgt, der darauf abzielt, den aktuellen Wissensstand systematisch zu erfassen, Einflussfaktoren auf den Wasserverbrauch zu identifizieren und Best Practices zur Reduktion zu dokumentieren. Gleichzeitig werden technologische, ökonomische und regulatorische Hürden analysiert, die einem nachhaltigen Wasserverbrauch im KI-Sektor entgegenstehen. Das folgende Kapitel stellt die hierfür entwickelte Forschungsmethodik detailliert vor.

3. Methodologie



Abbildung 6: Methodisches Vorgehen der Studie.

Dieses Kapitel beschreibt die methodische Vorgehensweise zur Untersuchung der Forschungsfragen. Zur Beantwortung dieser wurde ein Mixed-Methods-Ansatz gewählt (siehe Abbildung 6). Diese Kombination ermöglicht eine umfassende Analyse aus wissenschaftlicher und praxisnaher Perspektive. Konkret wurden eine systematische Literaturrecherche (SLR), die Analyse von Nachhaltigkeitsberichten führender Technologieunternehmen sowie Experteninterviews und ein Workshop durchgeführt. Durch diese Methodenkombination sollen der aktuelle Forschungsstand systematisch erfasst, praktische Herausforderungen identifiziert und evidenzbasierte Handlungsempfehlungen für eine nachhaltige Wassernutzung im KI-Sektor entwickelt werden.

3.1. SYSTEMATISCHE LITERATURRECHERCHE (SLR)

Die SLR wurde primär in der *ScienceDirect*-Datenbank durchgeführt und durch allgemeine Suchmaschinen ergänzt, um eine breitere Abdeckung zu gewährleisten. Die Recherche erfolgte durch verschiedene Kombinationen der Schlüsselwörter "artificial intelligence", "AI", "water consumption", "water usage", "data center" und "water footprint". Um den technologischen Wandel zu berücksichtigen, wurde der Analysezeitraum auf 2020–2024 festgelegt. Ältere relevante Arbeiten wurden durch das Schneeballverfahren einbezogen.

Aus den 1.104 initial gefundenen Studien wurden zunächst Duplikate entfernt. Anschließend erfolgte eine Sichtung von Titel und Abstract, um irrelevante Studien auszuschließen. Im letzten Schritt wurden die vollständigen Texte gelesen, um relevante Studien zu identifizieren. Studien, die sich nicht explizit mit den First-Order-Effekten von KI auf den Wasserverbrauch befassen, wurden aussortiert, da sie nicht im Zentrum der Untersuchung standen. Dazu zählten unter anderem Arbeiten, die den Fokus auf Nachhaltigkeit durch KI legten, sowie Studien, die sich mit spezifischen Kühlungstechnologien in RZ befassen, ohne eine Meta-Analyse oder eine breitere Betrachtung des Themas vorzunehmen, oder die den Fokus auf den Energieverbrauch statt auf den Wasserverbrauch legten. Zudem wurden Literaturquellen einbezogen, die im Verlauf der Interviews und des Workshops von Expert*innen empfohlen wurden, um eine fundierte und breit angelegte Perspektive zu gewährleisten. Insgesamt wurden 52 Studien als relevant identifiziert und in die detaillierte Analyse einbezogen.

Die SLR basiert ausschließlich auf öffentlich zugänglichen Quellen, sodass nicht veröffentlichte oder proprietäre Unternehmensdaten nicht einbezogen werden konnten. Dies könnte zu einer Unterschätzung bestimmter Aspekte des Wasserverbrauchs führen, insbesondere wenn Unternehmen nur selektiv über ihre Nachhaltigkeitsmaßnahmen berichten. Zudem besteht die Möglichkeit eines Publikationsbias, da umweltbezogene Forschung häufig in spezifischen Fachzeitschriften veröffentlicht wird, wodurch andere Perspektiven möglicherweise unterrepräsentiert sind.

Ein weiteres methodisches Limit besteht darin, dass der Wasserverbrauch von Hardwareplattformen nicht für alle relevanten Komponenten detailliert berücksichtigt werden konnte, da hierzu oft keine spezifischen Daten verfügbar sind. Schließlich erfolgte keine methodische Differenzierung oder Bewertung der Studienqualität, sodass sich methodische Unterschiede zwischen den ausgewerteten Arbeiten auf die Vergleichbarkeit der Ergebnisse auswirken könnten.

3.2. ANALYSE VON NACHHALTIGKEITSBERICHTEN

Um die Berichterstattung zum Wasserverbrauch in der Industrie zu erfassen, wurden die Nachhaltigkeitsberichte der „Magnificent 7“ – der sieben weltweit führenden Technologieunternehmen (*Nvidia*, *Apple*, *Microsoft*, *Amazon*, *Alphabet (Google)*, *Meta Platforms*, *Tesla*) – analysiert. Die Auswahl dieser Unternehmen basierte auf ihrer Marktmacht, ihrer Rolle in der KI-gestützten Rechenzentrumsinfrastruktur und ihrer Relevanz für den globalen IT-Sektor.

Der Fokus lag dabei auf der Dokumentation des Wasserverbrauchs, differenziert nach direktem Wasserverbrauch und Wasserentnahme sowie auf spezifischen Maßnahmen zur Reduktion des Wasserbedarfs in RZ. Die Berichte wurden anhand eines einheitlichen Kategorienschemas ausgewertet und nach dem Umfang der Berichterstattung in drei Kategorien klassifiziert:

1. **Keine Angabe:** Unternehmen nennen keine spezifischen Daten zum Wasserverbrauch.
2. **Teilweise Erwähnung:** Wasserverbrauch wird genannt, aber ohne detaillierte Differenzierung.
3. **Detaillierte Berichterstattung:** Umfangreiche Daten und Maßnahmen zur Wassereinsparung werden beschrieben.

Ein zentrales methodisches Problem besteht in der intransparenten Berichterstattung vieler Unternehmen. Während einige Firmen detaillierte Daten zu ihrem Energieverbrauch bereitstellen, sind Angaben zum Wasserverbrauch oft unvollständig oder uneinheitlich dargestellt. Dies erschwert eine vergleichbare Bewertung zwischen verschiedenen Unternehmen. Zudem besteht die Möglichkeit einer selektiven Berichterstattung, bekannt als „Greenwashing“, bei dem Unternehmen gezielt Maßnahmen betonen, die nur einen begrenzten Einfluss auf die tatsächliche Umweltbilanz haben. Da die Untersuchung ausschließlich auf öffentlich zugänglichen Berichten basiert, konnte keine externe Validierung der angegebenen Daten erfolgen.

3.3. EXPERTENINTERVIEWS UND WORKSHOP

Zur Ergänzung der quantitativen Analyse wurden zehn semi-strukturierte Experteninterviews mit Fachleuten aus Wissenschaft und Praxis durchgeführt. Die Interviews erfolgten in zwei Phasen:

1. **Wissenschaftliche Perspektiven:** Fokus auf First-Order-Effekte von KI auf den Wasserverbrauch, methodische Ansätze und bestehende Forschungslücken.
2. **Praxisperspektiven:** Herausforderungen und Lösungsansätze für eine nachhaltige Wassernutzung in RZ.

Die Teilnehmer*innen wurden anhand folgender Kriterien ausgewählt:

- Fachliche Expertise in Umweltwirkungen digitaler Technologien
- Berufserfahrung in der KI- oder Rechenzentrumsbranche
- Autor*innen relevanter wissenschaftlicher Publikationen oder praktische Erfahrung mit Wassermanagement in IT-Infrastrukturen

Die Erkenntnisse aus der ersten Interviewrunde wurden in einem Expertenworkshop validiert, in dem insbesondere regulatorische, ökonomische und technologische Hürden für eine nachhaltigere Wassernutzung in RZ diskutiert wurden. Die Ergebnisse des Workshops flossen iterativ in die zweite Interviewrunde ein, um praxisnahe Handlungsempfehlungen zu entwickeln.

Die begrenzte Anzahl an Interviews stellt eine methodische Einschränkung dar, da sie möglicherweise nicht alle relevanten Perspektiven ausreichend berücksichtigt. Dies könnte dazu führen, dass bestimmte Herausforderungen oder Lösungsansätze in der Untersuchung unterrepräsentiert bleiben. Zudem besteht ein Selektionsbias, da die Interviewpartner*innen bewusst anhand ihrer Fachkompetenz ausgewählt wurden, wodurch bestimmte Akteursgruppen, wie kleinere Unternehmen oder Organisationen aus Regionen mit spezifischen Wassermanagement-Herausforderungen, möglicherweise nicht adäquat repräsentiert sind. Ein weiteres potenzielles Problem ist die soziale Erwünschtheit in den Antworten der Interviewten – insbesondere im Unternehmenskontext könnten kritische Aspekte weniger offen angesprochen oder betriebliche Herausforderungen verharmlost dargestellt werden.

3.4. DATENANALYSE

Die Datenanalyse erfolgte durch die Integration und systematische Auswertung der Ergebnisse aus der Literaturrecherche, den Experteninterviews und dem Workshop. Ziel war es, zentrale Muster und Themen zu identifizieren, die zur Beantwortung der Forschungsfragen beitragen.

Die im Rahmen der SLR identifizierten Studien wurden hinsichtlich zentraler Themen und First-Order-Effekte von KI auf den Wasserverbrauch analysiert. Die Studien lassen sich in vier Kategorien einteilen:

1. Allgemeine Nachhaltigkeitsstudien, die mehrere Umweltindikatoren untersuchen, darunter Wasser, jedoch nicht ausschließlich.
2. Studien zum direkten Wasserverbrauch von RZ, insbesondere im Zusammenhang mit Kühlungstechnologien.
3. Untersuchungen zum Wasserverbrauch durch den Energiebedarf von KI und RZ.
4. Analysen zum Wasserverbrauch in der Hardwareproduktion, insbesondere bei der Halbleiterfertigung.

Die extrahierten Inhalte dienten als Grundlage für die weiteren Analysen der Interviews und des Workshops. Die qualitativen Daten aus den Experteninterviews und dem Workshop wurden mithilfe der *Thematic Analysis* untersucht. Dabei wurden Hauptthemen und wiederkehrende Muster systematisch identifiziert.

Die Analyse erfolgte in drei Schritten:

1. **Kodierung:** Die transkribierten Interviews und Workshop-Daten wurden mit der Datenanalysesoftware *atlas.ti* codiert. Dabei wurden sowohl induktive (neue, aus den Daten abgeleitete Themen) als auch deduktive (vorab definierte Kategorien aus der Literaturrecherche) Ansätze genutzt.
2. **Themenidentifikation:** Relevante Themen wie Hindernisse für wassersparende Innovationen, bestehende Best Practices und mögliche regulatorische oder technologische Maßnahmen wurden herausgearbeitet.
3. **Datenintegration:** Die Ergebnisse der SLR, Interviews und des Workshops wurden zusammengeführt, um ein umfassendes Bild der Auswirkungen von KI auf den Wasserverbrauch zu erhalten. Offene Fragen und Lücken wurden durch gezielte Zusatzrecherchen und vertiefende Gespräche in einer zweiten Interviewrunde adressiert.

4. Ergebnisse I: Wasserverbrauch durch KI

Forschungsfrage: *Wie beeinflusst die zunehmende Nutzung von KI den globalen Wasserverbrauch und welche Faktoren bestimmen den Wasserverbrauch von KI in verschiedenen Phasen des Lebenszyklus?*

Dieses Kapitel beschreibt die Ergebnisse der SLR. Im Rahmen der Literaturrecherche wurden 52 relevante Studien und Berichte identifiziert, die sich in drei Kategorien unterteilen:

- Studien zu den allgemeinen Umweltauswirkungen von KI, einschließlich der Kategorie Wasser,
- Studien zum Wasserverbrauch in RZ,
- Studien zum Wasserverbrauch zentraler Hardwarekomponenten, die für KI und Software essenziell sind.

Dieses Kapitel bietet eine umfassende Analyse des Wasserverbrauchs im Zusammenhang mit KI-Anwendungen. Kapitel 4.1 gibt einen Überblick über den aktuellen Stand der Wissenschaft und zeigt unterschiedliche Ansätze zur Quantifizierung des Wasserverbrauchs in der Fachliteratur. Kapitel 4.2 beleuchtet die Rolle von RZ und ergänzt die wissenschaftlichen Erkenntnisse durch eine vereinfachte Übersicht über drei weit verbreitete Kühltechnologien. In Kapitel 4.3 liegt der Fokus auf dem "embodied water consumption" – also dem Wasserverbrauch, der durch die Herstellung der für KI erforderlichen Hardware entsteht. Kapitel 4.4 widmet sich der Dokumentation des Wasserverbrauchs in Nachhaltigkeitsberichten von Unternehmen und Organisationen und analysiert, inwieweit dieser Aspekt dort berücksichtigt wird.

4.1. QUANTIFIZIERUNG DES WASSERVERBRAUCHS DURCH KI

ERGEBNISSE DER LITERATURRECHERCHE

Die SLR identifizierte sechs wissenschaftliche Studien, die sich mit dem Wasserverbrauch durch KI befassen. Während einige dieser Arbeiten KI im Kontext allgemeiner Umweltfolgen betrachten und Wasser dabei nur eine untergeordnete Rolle spielt, fokussieren sich andere explizit auf RZ und die mit KI verbundenen wasserbezogenen Auswirkungen. Studien, die sich insbesondere mit dem Wasserverbrauch von RZ befassen, weisen inhaltliche Überschneidungen mit Kapitel 4.1.2 „Wasserverbrauch durch RZ“ auf. Sie werden jedoch in diesem Abschnitt eingeordnet, da sie speziell die durch KI bedingten Effekte analysieren. Insgesamt konnten drei Studien identifiziert werden, die eine Quantifizierung des durch KI verursachten Wasserverbrauchs vorgenommen haben. Tabelle 1 gibt einen Überblick über zentrale Studien.

AUTOR*IN/JAHR/TITEL	FOKUS	ZENTRALE ERKENNTNISSE
Li et al. (2023): Making AI Less Thirsty: Uncovering and Addressing the Secret Water Footprint of AI Models	- Entwickelt eine Methodologie zur Quantifizierung des Wasserverbrauchs von KI-Modellen - Betrachtung von On-site (Kühlung in RZ) und Off-site (Energieerzeugung) Wasserverbrauch	- Schätzt, dass das Training von GPT-3 in Microsofts RZ 700.000 Liter Wasser verbrauchte. - Schätzt, dass der globale Wasserbedarf durch KI im Jahr 2027 zwischen 4,2 und 6,6 Milliarden m³ betragen könnte - Fordert mehr Transparenz und strategisches Wassermanagement, um den Wasserverbrauch großer KI-Modelle zu reduzieren.
Shumba et al. (2024): A Water Efficiency Dataset for African Data Centers	- Analyse des Wasserverbrauchs von RZ 41 afrikanischen Ländern - Kombiniert Wetter- und Stromerzeugungsdaten, um Daten, um die Wassernutzungseffizienz von Rechenzentren zu schätzen. - Nutzt den Datensatz, um den Wasserverbrauch während der Inferenz zweier großer Sprachmodelle (LLaMA-3-70B und GPT-4) in elf ausgewählten afrikanischen Ländern zu schätzen	- Ein 10-seitiger Bericht mit Llama-3-70B verbraucht ca. 0,7 Liter Wasser, während GPT-4 bis zu 60 Liter benötigen kann. - Acht von elf untersuchten afrikanischen Ländern haben einen geringeren Wasserverbrauch als der globale Durchschnitt, da ihre Stromerzeugung weniger wasserintensiv ist. - Einige Länder mit Steppenklima haben hingegen einen überdurchschnittlich hohen Wasserverbrauch, was eine sorgfältige Planung für KI-Implementierungen erforderlich macht.
Zuccon et al. (2023): Beyond CO ₂ Emissions: The Overlooked Impact of Water Consumption of Information Retrieval Models	- Untersuchung des Wasserverbrauchs von neuronalen <i>Information Retrieval</i> - Modellen (IR) - Baut auf der Methode von Li et al. (2023) auf - Untersucht den Einfluss der Wasserqualität auf On-site Wasserverbrauch	- Neuronale IR-Modelle (wie TILDEv2 mit docTquery) verbrauchen bis zu 927 Liter Wasser, während klassische Modelle (BM25) nur 0,01 Liter benötigen. - Der Wasserverbrauch von neuronalen IR -Modellen wird bisher stark unterschätzt. - Der Wasserverbrauch wird von externen Faktoren beeinflusst: Wasserqualität (schlechte Qualität kann den Verbrauch steigern), Tageszeit (bis zu 32 Liter Unterschied zwischen 9 Uhr und 15 Uhr) und

Jahreszeit (23% höherer Verbrauch im Sommer vs. Winter)

Tabelle 1: Zentrale Studien zum Wasserverbrauch durch KI

Die drei Studien nutzen unterschiedliche methodische Ansätze zur Quantifizierung des Wasserverbrauchs von KI. Li et al. (2023) entwickelten eine Methodik zur Schätzung des Wasser-Fußabdrucks von KI-Modellen anhand von Energiebedarf und Kühlanforderungen. In einer Hochrechnung schätzt die Studie, dass der globale Wasserbedarf durch KI bis 2027 auf 4,2–6,6 Milliarden m³ steigen könnte. Diese Schätzung basiert jedoch auf mehreren kritischen Annahmen und ist daher mit erheblichen Unsicherheiten behaftet. Sie sollte eher als eine grobe Größenordnung und Trendindikation denn als präzise Prognose interpretiert werden. Die Autoren selbst weisen darauf hin, dass die Bandbreite der Schätzung die Komplexität und Dynamik der KI-Entwicklung widerspiegelt.

Shumba et al. (2024) untersuchten die WUE in afrikanischen RZ und zeigten, dass der Wasserverbrauch stark von der Wasserintensität der Stromerzeugung abhängt. Regionen mit erneuerbaren Energien verbrauchen weniger Wasser als solche mit fossilen Kraftwerken oder ariden Klimazonen (Shumba et al., 2024). Zuccon et al. (2023) analysierten den Wasserverbrauch von neuronale Informationsretrieval (IR)-Modellen und fanden heraus, dass größere Modelle überproportional viel Wasser benötigen – sowohl direkt für die Kühlung als auch indirekt durch wasserintensive Stromerzeugung und das Faktoren wie Jahreszeit und Tageszeit den Wasserverbrauch in RZ beeinflussen können (Zuccon et al., 2023).

Insgesamt zeigen die Studien, dass der Wasserverbrauch von KI stark variiert – abhängig vom Standort der Rechenzentren, der Art der Kühlung und der genutzten Energiequelle. Deutlich wird jedoch auch, dass bislang nur sehr wenige Studien existieren, die sich gezielt mit dem Wasserverbrauch durch KI befassen und diesen systematisch zu quantifizieren versuchen. Zudem fehlen standardisierte Messmethoden für den gesamten KI-bedingten Wasserverbrauch, insbesondere unter Berücksichtigung der Onsite- und Offsite-Verbräuche von Rechenzentren sowie des sogenannten „embodied water use“, also des Wasserverbrauchs während der Herstellung KI-spezifischer Hardware wie GPUs und Chips. Dies erschwert den Vergleich verschiedener Standorte und Technologien erheblich. Viele Technologieunternehmen veröffentlichen keine detaillierten Verbrauchsdaten, sodass Modellierungen häufig auf Schätzungen beruhen (Shumba et al., 2024; Zuccon et al., 2023). Dadurch sind belastbare Aussagen zum Wasserfußabdruck von KI bislang nur eingeschränkt möglich.

EXPERTENINTERVIEWS UND WORKSHOP

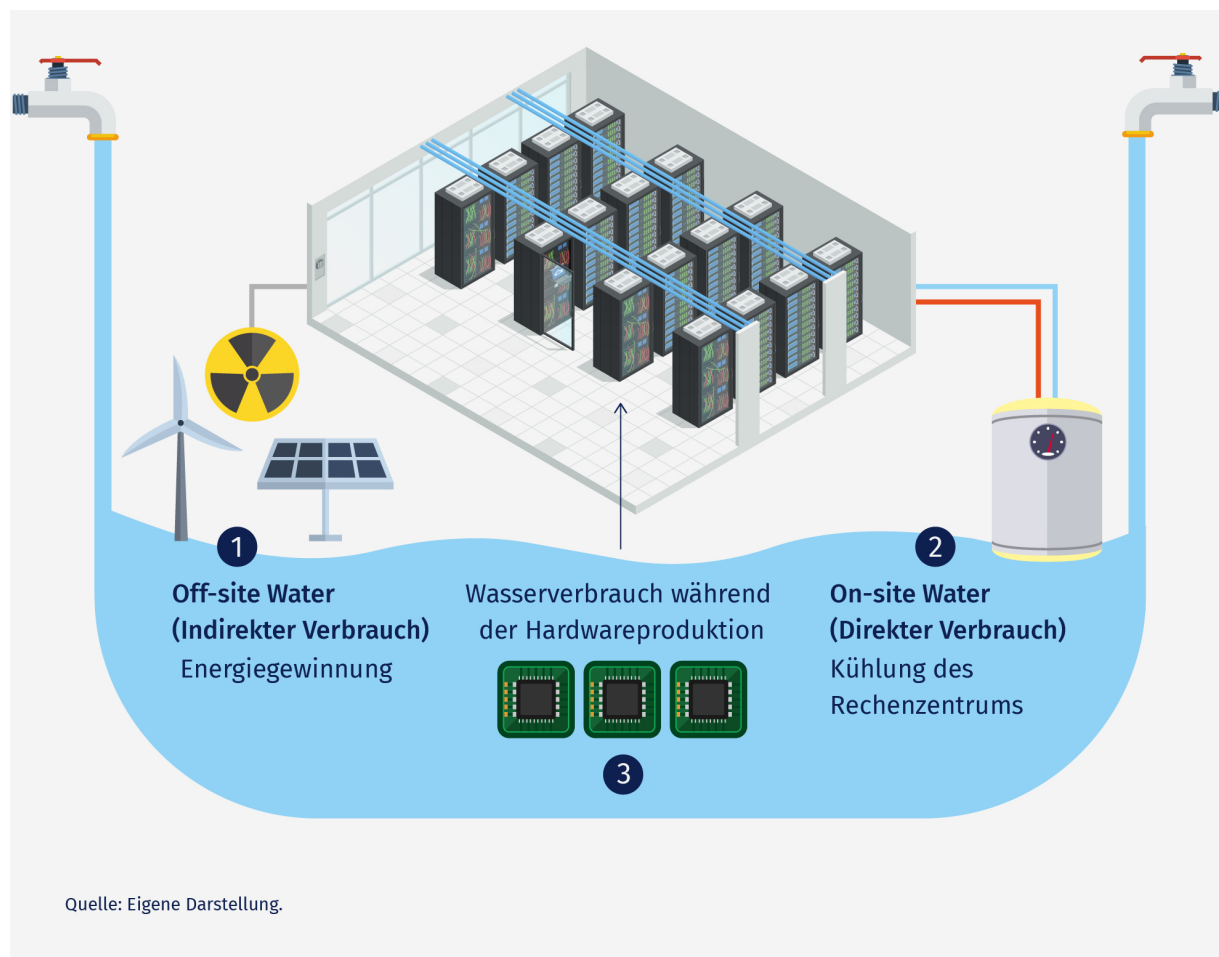


Abbildung 7: Direkter und indirekter Wasserverbrauch in Rechenzentren.

Die in den Studien untersuchten Effekte von KI auf den Wasserverbrauch wurden auch von den Expert*innen bestätigt (siehe Abbildung 7): (1) Direkter Wasserverbrauch durch Kühlprozesse in RZ, (2) Wasserverbrauch durch die Stromerzeugung für den Betrieb der RZ, (3) Wasserverbrauch bei der Hardwareproduktion, insbesondere der Herstellung von Computerchips.

Zudem betonen Expert*innen, dass der Fokus ursprünglich auf dem extrem ressourcenintensiven Training großer Modelle lag. Mittlerweile zeigt sich jedoch, dass die kontinuierliche Nutzung (*Inference*) einen wesentlich größeren und langfristigeren Faktor beim Wasserverbrauch darstellt. Auch die räumliche Flexibilität von Training und Nutzung beeinflusst den Wasserverbrauch erheblich: Während das Training ortsunabhängig ist und daher in wasserreichen Regionen stattfinden könnte, muss die Rechenleistung für die Nutzung häufig nahe an den Endnutzenden erfolgen. Darüber hinaus berichten Expert*innen von Untersuchungen in wasserarmen Regionen, in denen der Bau großer RZ bestehende Wasserknappheiten verschärfte und zu Konflikten über die Verteilung von Wasser zwischen Industrie, Landwirtschaft und der lokalen Bevölkerung führte.

Wie stark der KI-Lebenszyklus den Wasserverbrauch tatsächlich beeinflusst, wird unter den Expert*innen unterschiedlich bewertet. Während einige die potenziellen Gefahren insbesondere durch den steigenden Wasserbedarf von RZ betonen, weisen andere darauf hin, dass der On-site Wasserverbrauch stark vom jeweiligen Design und insbesondere von der verwendeten Kühltechnologie abhängt und in einigen Fällen nur minimal ausfallen kann. Mehrere Expert*innen heben jedoch hervor, dass es schwierig ist, hierzu belastbare Aussagen zu treffen, da die Datengrundlage und die Transparenz in diesem Bereich oft unzureichend sind. Insbesondere fehlen detaillierte Daten darüber, welcher Anteil des Wasserverbrauchs in RZ direkt auf KI-Anwendungen zurückzuführen ist. Ein ganzheitlicher Lebenszyklusansatz ist daher entscheidend, um die langfristigen Wasserauswirkungen von KI realistisch zu bewerten und Strategien für eine nachhaltigere Nutzung zu entwickeln.

4.2. WASSERVERBRAUCH DURCH RECHENZENTREN

ERGEBNISSE DER LITERATURRECHERCHE

Die meisten der als relevant identifizierten Studien befassen sich mit dem Wasserverbrauch von RZ. Dabei fokussieren sich die Arbeiten vorwiegend auf den direkten Wasserverbrauch durch die Kühlung des RZ und/oder auf den indirekten Wasserverbrauch, der durch die Stromerzeugung. Tabelle 2 gibt eine Übersicht über die als besonders relevant identifizierten Studien.

AUTOR*IN/JAHR/TITEL	FOKUS DER STUDIE	ZENTRALE ERKENNTNISSE
Mytton (2021): <i>Data Centre Water Consumption</i>	• Untersuchung des direkten und indirekten Wasserverbrauchs von RZ. • Untersucht den Wasserverbrauch von RZ sowie die Berichterstattung der Branche, z. B. anhand des WUE	• Der größte Anteil am Wasserfußabdruck eines Rechenzentrums entfällt auf die Stromerzeugung. • Weniger als ein Drittel der RZ-Betreiber misst den Wasserverbrauch systematisch. • Die Rechenzentrumsbranche leidet unter einem Mangel an Transparenz.

Karimi et al. (2022): <i>Water-energy tradeoffs in data centers: A case study in hot-arid climates</i>	• Untersucht zweier RZ hinsichtlich ihres Energie- und Wasserverbrauchs anhand des PUE und WUE • Regionale Begrenzung auf aride Klimazonen (Phönix, USA). • Vergleicht Kühlung der RZ: RZ1 verwendet luftgekühlte Kaltwassersätze (Chillers), RZ2 nutzt wassergekühlte Kaltwassersätze und zusätzlich Verdunstungskühlung.	• Die luftgekühlten Chiller von RZ1 sind energieintensiver als das wasserbasierte Kühlsystem von RZ2; jedoch beträgt der Wasserfußabdruck von RZ1 nur etwa ein Viertel desjenigen von RZ2. • Der PUE von RZ1 unterliegt stärkeren jahreszeitlichen Schwankungen, während die WUE in RZ2 saisonabhängig deutlich stärker variiert. • Luftgekühlte Chiller mit direkter Verdunstungskühlung weisen im untersuchten Klima von Phoenix den geringsten kombinierten Energie- und Wasserfußabdruck aller betrachteten Systeme auf.
Lei et al. (2025): <i>Water Footprint of Data Center Workloads: A Review of Key Determinants</i>	• Untersuchung der Einflussfaktoren auf den Wasserverbrauch pro Workload. • Entwicklung von Leitlinien für Wassereinsparstrategien auf Workload-Ebene • Systematische Analyse der Variabilität des Wasserverbrauchs zwischen verschiedenen Workloads	• Extreme Variabilität: Wasserverbrauch pro Workload kann um mehr als das 10.000-fache schwanken • Ranking der Haupteinflussfaktoren: Servereffizienz > Stromnetz- Wasserfaktor > Serverauslastung > Kühlsystem • Optimierung erfordert standortspezifische Kombinationen - keine universelle Lösung für alle RZ

Tabelle 2: Ausgewählte Studien zum Wasserverbrauch von Rechenzentren

Die betrachteten Studien unterscheiden sich deutlich hinsichtlich ihres Erkenntnisinteresses, der gewählten Systemgrenzen, der untersuchten RZ-Typen sowie der methodischen Herangehensweise. Ein Teil der Literatur quantifiziert den Wasserverbrauch bzw. Wasserfußabdruck mit unterschiedlichen methodischen Ansätzen, zugrunde liegenden Annahmen und regionalen Schwerpunkten (Farfan & Lohrmann, 2023; Jerléus et al., 2024; Ristic et al., 2015). Erste Schätzungen zum globalen Wasserfußabdruck von RZ stammen von Ristic et al. (2015). Unter Berücksichtigung des direkten, kühlungsbedingten Wasserverbrauchs durch HVAC-Systeme sowie des indirekten, energiebedingten Wasserverbrauchs schätzt die Studie den globalen Wasserfußabdruck von RZ auf 767 bis 147.810 Millionen Kubikmeter Wasser. Die Autoren merken jedoch an, dass diese Schätzungen mit hohen Unsicherheiten behaftet sind. Farfan & Lohrmann (2023) liefern eine der wenigen identifizierten Studien mit Fokus auf Europa. Sie schätzen, dass der jährliche Wasserverbrauch digitaler Dienste im Jahr 2030 zwischen 273,4 und 820,1 Mio. m³ betragen wird – bei einem gleichzeitigen Strombedarf von 56,3 bis 169 TWh.

Weitere Studien betrachten die ökologische Nachhaltigkeit von ‚grünen‘ RZ anhand von Best Practices und Use Cases wobei Wasser als eine von mehreren Umweltkategorien analysiert wird (Al Kez et al., 2022; Monserrate, 2022; Pegus et al., 2016; Sharma et al., 2017). Die meisten identifizierten Studien fokussieren sich jedoch auf den Wasserverbrauch von verschiedenen RZ-Typen und analysieren konkrete RZ-Designs und Kühlsysteme hinsichtlich ihrer Energie- und Wassereffizienz. Dabei variiert der Betrachtungshorizont: Während manche Studien den direkten Wasserverbrauch (on-site) untersuchen (Sharma et al., 2017), beziehen andere zusätzlich den indirekten Wasserverbrauch durch Stromerzeugung mit ein (Lei et al., 2025; Lei & Masanet, 2022; Shehabi et al., 2016; Siddik et al., 2021).

Ein weiterer Untersuchungsstrang widmet sich den geografischen und infrastrukturellen Einflussfaktoren. Diese Studien zeigen, wie Standortwahl, Klimazone und lokale Netzinfrastuktur die Energie- und Wassereffizienz von Rechenzentren maßgeblich beeinflussen können (Cáceres et al., 2024; Islam et al., 2017; Lei et al., 2023; Silva-Llanca et al., 2023). Silva et al. (2023) untersuchen die Einsparpotenziale durch luftbasierte Freikühlung von RZ unter verschiedenen klimatischen Bedingungen und findet, dass Besonders Standorte mit niedrigeren Temperaturen ($< 20^{\circ}\text{C}$) und hoher Luftfeuchtigkeit ($> 60\%$) besonders geeignet sind, da hier der Einsatz von Chiller und/oder Luftbefeuchter minimal ausfällt.

Auch Karimi et al. (2022) untersuchten den Energie- und Wasserverbrauch zweier RZ und fokussiert sich dabei auf RZ im Großraum Phoenix, die unterschiedliche Kühlsysteme nutzen. Beide Rechenzentren setzen CRAH-Einheiten zur Kühlung der Serverräume ein, wobei die Abwärme an durch Chiller erzeugtes Kaltwasser abgegeben wird. Während RZ1 luftgekühlte Kaltwassersätze verwendet, nutzt RZ2 eine zentrale wassergekühlte Anlage mit Kühltürmen sowie teilweise eine hybride Verdunstungskühlung über sogenannte Rooftop Units (RTUs). Dabei handelt es sich um vorkonfigurierte Dachklimageräte mit direkter adiabater Verdunstungskühlung, die Umgebungsluft durch Wasserverdunstung abkühlen und in die Raumluft einbringen – besonders effizient bei heiß-trockenen Klimabedingungen. Die Analyse der Effizienzkennzahlen PUE, $\text{WUE}_{\text{onsite}}$ und $\text{WUE}_{\text{source}}$ zeigt, dass RZ1 zwar einen höheren Energieverbrauch (durchschnittlich 13 % höhere PUE) aufweist, aber gleichzeitig einen deutlich geringeren Wasserverbrauch hat – die WUE liegt 66 % unter der von RZ2. Darüber hinaus wurde festgestellt, dass die saisonalen Schwankungen den PUE in RZ1 stärker beeinflussen, während in RZ2 vor allem der Wasserverbrauch in heißen Monaten deutlich ansteigt. Die Studie hebt zudem hervor, dass der gezielte Einsatz von freier Kühlung und adiabater Verdunstungskühlung – bei geeigneten Außenbedingungen etwa 40 % des Jahres möglich – ein großes Potenzial zur gleichzeitigen Reduktion von Energie- und Wasserverbrauch bietet. Zhang et al. (2025) untersuchen den Einsatz von thermischer Energiespeicherung im Untergrund (Reservoir Thermal Energy Storage, RTES) zur energie- und wassersparenden Kühlung von Rechenzentren. Die Ergebnisse zeigen, dass RTES für Hyperscale- und Krypto-Rechenzentren technisch umsetzbar und über 20 Jahre nachhaltig betrieben werden kann. Besonders vorteilhaft ist, dass im Winter günstig erzeugte Kälte gespeichert und im Sommer genutzt wird – dies senkt nicht nur die Kühlkosten deutlich (11,9 \$/MWh bzw. 14,8 \$/MWh), sondern reduziert auch den Wasserverbrauch erheblich durch geringeren Bedarf an Kühltürmen (Y. Zhang et al., 2025).

Eine differenzierte Betrachtung des wasserbezogenen Ressourcenverbrauchs pro Recheneinheit (Workload) liefern Lei et al. Diese Studie analysiert die Einflussfaktoren auf den wasserbezogenen Ressourcenverbrauch pro Recheneinheit (Workload) in Rechenzentren, gemessen in Litern pro Workload, um Strategien zur Wassereinsparung zu unterstützen. Die Ergebnisse zeigen, dass der wasserbezogene Ressourcenverbrauch pro Workload um mehr als das 10.000-Fache variieren kann. Diese Unterschiede werden durch mehr als 1000-fache Unterschiede im Wasserverbrauch pro Kilowattstunde Serverstrom sowie durch etwa 10-fache Unterschiede in der Effizienz der Serverauslastung verursacht. Die wichtigsten Einflussfaktoren sind: Servereffizienz, wasserbedingte Eigenschaften des Stromnetzes, Serverauslastung, Art des Kühlsystems, Infrastruktureffizienz, Klimazone, Anteil inaktiver Server und der Erneuerungszyklus der Server. Der verstärkte Einsatz erneuerbarer

Energien (z. B. Wind oder Photovoltaik), der Einsatz hocheffizienter Server und Technologien wie Virtualisierung zur Erhöhung der Serverauslastung können auf nationaler Ebene wirksam zur Reduzierung des Wasserverbrauchs beitragen und gleichzeitig die Dienstleistungsleistung maximieren. Es gibt jedoch kein universelles Rezept zur Minimierung des Wasserverbrauchs – vielmehr hängt ein optimaler Ressourceneinsatz von einer sorgfältig abgestimmten Kombination dieser Faktoren ab.

Insgesamt macht die Analyse der Fachliteratur verdeutlicht zudem, dass der Wasserverbrauch von RZ das Ergebnis eines komplexen Zusammenspiels aus Kühlsystem, Energiequelle, klimatischen Standortbedingungen und betrieblicher Ausgestaltung ist. Es bestehende bereit viele Ansätze, die den direkten und/oder indirekten Wasserverbrauch von RZ senken können. Ein zentrales Problem bleibt die mangelnde Transparenz von Seiten der Industrie. Mytton (2021), Lei et al. (2025) und Masanet et al. (2024) kritisieren die bislang unzureichende systematische Erfassung und Offenlegung wasserbezogener Daten, was die Vergleichbarkeit und Entwicklung belastbarer Standards erschwerten. Die Literatur legt nahe, dass der Wasserverbrauch von Rechenzentren nicht isoliert betrachtet werden kann. Vielmehr ist er eng mit Energieeffizienz, Kühlsystemwahl und Standortfaktoren verknüpft. Im Kontext zunehmender Rechenlasten – etwa durch KI-Anwendungen mit hoher Energiedichte – gewinnt die Frage nach einer nachhaltigen Wasserbewirtschaftung in Rechenzentren weiter an Bedeutung (Lei et al., 2025; Y. Zhang et al., 2022).

EXPERTENINTERVIEWS UND WORKSHOP

Expert*innen betonen die Bedeutung von Kühltechnologie, Standortwahl und Energiequelle für den Wasserverbrauch von RZ. Während Luftkühlung traditionell weit verbreitet ist, gewinnt Flüssigkeitskühlung – insbesondere direkte Flüssigkeitskühlung und Immersionskühlung – zunehmend an Relevanz, insbesondere bei Hochleistungsrechenzentren mit hoher Wärmedichte, etwa im Kontext von KI-Anwendungen. Diese Systeme ermöglichen eine effizientere Wärmeabfuhr. Durch geschlossene Wasserkreisläufe und den Einsatz freier Kühlung ist der direkte Wasserverbrauch in der Regel gering. Einige Expert*innen hoben hervor, dass der Wasserverlust im Kühlkreislauf (z. B. durch Nachfüllen) meist vernachlässigbar ist, während der Hauptanteil des Wasserverbrauchs bei der Rückkühlung anfällt – insbesondere in den Sommermonaten. Gleichzeitig wurde betont, dass in vielen RZ in Deutschland sowie in Zentral- und Nordeuropa der direkte Wasserverbrauch häufig relativ niedrig ist, da die klimatischen Bedingungen eine effiziente Kühlung auch ohne wasserintensive Systeme ermöglichen.

Als vielversprechend wurden verschiedene hybride Kühlsysteme beschrieben, die bei niedrigen Außentemperaturen mit trockener Rückkühlung arbeiten und erst bei großer Hitze auf Verdunstungskühlung umstellen. Weitere Ansätze zur Reduktion des Ressourcenverbrauchs umfassen adiabate Kühlung sowie Hochtemperaturkühlung. Die steigende Nachfrage nach KI-Anwendungen wurde mehrfach als Treiber für wasserintensivere Kühllösungen benannt, da Luftkühlung zunehmend an ihre Leistungsgrenzen stößt. Mehrere Expert*innen wiesen auf den Zielkonflikt zwischen Wasser- und Energieeffizienz hin: Während wasserbasierte Kühlsysteme den Stromverbrauch deutlich senken können, führen sie potenziell zu einem erhöhten Wasserbedarf. Die Einschätzungen zur Relevanz des Wasserverbrauchs durch RZ fielen unterschiedlich aus: Einige – mit globaler Perspektive – bewerten den direkten und indirekten Wasserverbrauch als bedeutend, insbesondere angesichts der wachsenden Nachfrage nach Rechenleistung durch KI. Andere betonten, dass viele RZ bereits über Systeme verfügen, bei denen der direkte Wasserverbrauch gering ist – abhängig von Design und Technologie. Eine ganzheitliche Betrachtung sei jedoch in jedem Fall notwendig, um die Umweltauswirkungen angemessen zu erfassen.

Ein häufig angesprochenes Problem war die mangelnde Transparenz beim Wasserverbrauch: Viele Betreiber veröffentlichen keine entsprechenden Daten, was Vergleiche sowie die Entwicklung regulatorischer Maßnahmen

erheblich erschwert. Mit wachsendem Kühlbedarf, insbesondere durch datenintensive Anwendungen, steigt laut Expert*innen der Handlungsdruck. Wiederholt thematisiert wurde auch die Standortwahl, die bislang überwiegend nach ökonomischen Kriterien – etwa der Nähe zu Netzknoten – erfolgt. Dabei bleiben standortbedingte Ressourcenvorteile, wie die effiziente Nutzung von Abwärme, teilweise wenig. Letztlich ist eine bessere Balance zwischen Energie- und Wassereffizienz entscheidend, um den wachsenden Kühlbedarf von RZ mit ökologischen und wirtschaftlichen Anforderungen in Einklang zu bringen.

4.3. WASSERVERBRAUCH DURCH HARDWAREPRODUKTION

Ergebnisse der Literaturrecherche

Die Produktion von Hardware, insbesondere von Halbleitern, stellt einen der ressourcenintensivsten Prozesse in der IT-Industrie dar. Die SLR hat acht relevante Studien identifiziert, die sich mit dem Wasserverbrauch der Chipfertigung befassen. Tabelle 3 gibt eine Übersicht der zentralen Ergebnisse vier der als am wichtigsten eingestuften Studien. Während einige Arbeiten einen allgemeinen Fokus auf die Umweltauswirkungen der Halbleiterproduktion legen und den Wasserverbrauch nur am Rande thematisieren, gibt es auch Studien, die explizit den Wasserfußabdruck der Chipfertigung analysieren und quantifizieren.

AUSGEWÄHLTE STUDIEN

AUTOR*IN, JAHR, TITEL	FOKUS DER STUDIE	ZENTRALE ERKENNTNISSE
Ruberti (2023): The chip manufacturing industry: Environmental impacts and eco-efficiency analysis	• Untersucht die Umweltwirkungen und die Öko-Effizienz der weltweit führenden Halbleiterfabriken. • Berechnung und Analyse verschiedener zentraler Umweltleistungsindikatoren, basierend auf den Daten aus den Corporate-Social-Responsibility-(CSR)-Berichten der Unternehmen.	• Höherer Umsatz pro produziertem Wafer steht oft mit erhöhtem Ressourcenverbrauch in Zusammenhang. • Größere Unternehmen und fortschrittlichere Produkte führen nicht zwangsläufig zu geringeren Umweltbelastungen pro produzierter Einheit. • Unterschiedliche Fertigungstechnologien führen zu stark variierenden Wasserverbräuchen.
Wang et al. (2023): Water strategies and practices for sustainable development in the semiconductor industry	• Analysiert die Nachhaltigkeitsberichte von 24 Halbleiterunternehmen und untersucht deren Wasserstrategien und -praktiken	• Wasserrecycling, Wiederverwendung und Wiederaufbereitung sind die zentralen Strategien für eine nachhaltige Halbleiterproduktion. • Reverse Osmosis spielt eine zentrale Rolle zur Wasseraufbereitung und Rückgewinnung.

	• Ein zukünftiges Wasserzyklus-System für Halbleiter-Industrieparks wird vorgeschlagen, welches Wasserquelle, Nutzung, Recycling und Abwasserentsorgung integriert. • Der globale Wasserverbrauch für die Halbleiterproduktion betrug $7,51 \times 10^8$ m³ in 2021.
Frost & Hua (2019): Quantifying spatiotemporal impacts of the interaction of water scarcity and water use by the global semiconductor manufacturing industry	• Umfassende Analyse zum globalen Wasserverbrauch der Halbleiterindustrie. • Berücksichtigt sowohl den direkten Wasserverbrauch in der Herstellung als auch den strombedingten Wasserverbrauch. • Fast die Hälfte der Halbleiterfertigung findet in Regionen mit überdurchschnittlicher Wasserknappheit statt. • Besonders hoher Wasserverbrauch für die Reinigung der Wafer mit UPW. • Investitionen in wassersparende Technologien können den Wasserverbrauch deutlich senken. • Die wirksamste Maßnahme zur Reduzierung des gesamten Wasserbedarfs in der Chipproduktion ist die Senkung des Stromverbrauchs in den Fertigungsanlagen.
Lepawsky (2024): Climate change induced water stress and future semiconductor supply chain risk	• Diese Studie kombiniert Standortdaten von Halbleiterfertigungsanlagen mit Daten zu spezifischen Kunden-Lieferanten-Netzwerken sowie Daten zum globalen Wasserstressrisiko unter drei Klimaszenarien für die Jahre 2030 und 2040. • 40 % der Produktionsstätten werden sich 2030 und 2040 in Regionen mit hohem oder extrem hohem Wasserstress befinden. • 24–40 % der derzeit im Bau befindlichen und mehr als 40% der neu angekündigten Produktionsstätten werden ebenfalls in Gebieten mit hohem oder extrem hohem Wasserstress liegen. • Besonders betroffen sind Taiwan und der Südwesten der USA. • Die Wasserstressrisiken könnten sich durch die globalen Lieferketten ausbreiten und Auswirkungen auf die gesamte Halbleiterproduktion haben

Tabelle 3: Ausgewählte Studien zum Wasserverbrauch in der KI-Hardwareproduktion.

Die Studien zeigen, dass Wasserverbrauch ein relevanter Faktor in der weltweiter Halbleiterproduktion ist, wobei erhebliche Unterschiede im Wasserverbrauch bestehen.

Der Wasserverbrauch in der Fertigung von Halbleitern ist besonders hoch bei komplexen Komponenten wie High Performance Computing-Chips und GPUs – insbesondere durch den Einsatz großer Mengen an Ultrapure Water (Frost & Hua, 2019). Laut Wang et al. (2023) verstärkt lokale Wasserknappheit den Druck auf Halbleiterunternehmen, ihre Wassernutzung effizienter zu gestalten. Gleichzeitig zeigt die Forschung von Lepawsky (2024), dass Recyclingstrategien und der Einsatz von nicht-trinkbarem Wasser als Lösungen für den hohen Wasserverbrauch an Bedeutung gewinnen.

Wang et al. (2023) und Ruberti et al. (2023) zeigen, in ihren Studien, dass insbesondere die Reinigung der Siliziumwafer¹ und die Herstellung von *Ultrapure Water* (UPW) – Wasser, das nahezu frei von Verunreinigungen und gelösten Substanzen ist – zentrale Einflussfaktoren für den Wasserverbrauch in der Halbleiterproduktion sind (Ruberti, 2023; Wang et al., 2023). Gleichzeitig existieren erhebliche regionale Unterschiede im Wasserverbrauch, da sich die Produktionsmethoden und die Verfügbarkeit von Wasserressourcen stark unterscheiden (Lepawsky, 2024). Außerdem sind Schätzungen und Zukunftsprognosen aufgrund ihrer unterschiedlichen methodischen Ansätze und stetigen Wandel der Industriebranche nur schwer vergleichbar sind.

Einige Unternehmen setzen bereits auf wassersparende Maßnahmen, indem sie geschlossene Wasserkreisläufe und Recyclingtechnologien nutzen. Diese Systeme ermöglichen eine Wiederverwendung des Wassers mit minimalen Nachfüllungen. Besonders große Hersteller wie TSMC und Samsung setzen verstärkt auf recyceltes Wasser, um den Wasserverbrauch zu senken und gleichzeitig ihre Produktionskapazitäten aufrechtzuerhalten (Wang et al., 2023). Darüber hinaus werden innovative Wasseraufbereitungsverfahren entwickelt, die eine nachhaltigere Nutzung der Ressource ermöglichen.

Eine direkte Vergleichbarkeit der Studienergebnisse ist aufgrund regionaler Gegebenheiten, technologischer Entwicklungen und methodischer Ansätze nicht möglich. Wichtige Einflussfaktoren sind:

- **Regionale Unterschiede:** Die Wasserversorgung, Recyclingkapazitäten und klimatischen Bedingungen beeinflussen den Wasserverbrauch erheblich (Lepawsky, 2024).
- **Technologische Entwicklungen:** Neuere Fertigungstechnologien wie 5-nm- oder 3-nm-Chips benötigen deutlich mehr Wasser als ältere Generationen (Ruberti, 2023).
- **Datenverfügbarkeit:** Viele Unternehmen veröffentlichen keine detaillierten Angaben zu ihrem Wasserverbrauch, wodurch Studien oft auf Modellrechnungen oder Industrieb Berichte angewiesen sind (Wang et al., 2023).
- **Indirekter Wasserverbrauch:** Neben der direkten Nutzung für die Wafer-Produktion wird der Wasserbedarf für vorgelagerte Prozesse wie Rohstoffgewinnung oder Transport häufig nicht berücksichtigt.

EXPERTENINTERVIEWS UND WORKSHOP

Die Expert*innen betonten ebenfalls die zentrale Rolle der Hardwareproduktion für den gesamten Wasserverbrauch der digitalen Infrastruktur. Neben der wasserintensiven Halbleiterfertigung erfordert auch die allgemeine Herstellung von Servern und Rechenzentrumskomponenten erhebliche Wassermengen, insbesondere für Kühlung und Reinigung während der Produktion. Ein weiteres Problem stellt die mangelnde Transparenz des Wasserverbrauchs in der Hardwareproduktion dar. Verfügbare Daten sind schwer vergleichbar, da sie nicht nach einheitlichen Methoden erhoben und zentral erfasst werden. Zudem handelt es sich um ein globales Phänomen, das durch internationale Lieferketten zusätzlich erschwert wird. Ein weiterer wichtiger

¹ Wafer sind das bearbeitete Halbleiterausgangsmaterial für die Chipproduktion in der Halbleitertechnologie. Wafer werden durch Sägen gezogener Einkristallstäbe aus z.B. Silicium und nachfolgender Oberflächenbehandlung wie Polieren oder Ätzpolieren (Ätzverfahren) hergestellt.

Aspekt, der von den Expert*innen hervorgehoben wurde, ist die geografische Lage der Produktionsstätten. Die Halbleiterfertigung findet zu großen Teilen nicht in Europa, sondern in Taiwan, Südkorea, Japan, den USA und China statt. In Regionen, die bereits unter hohem Wasserstress leiden, kann das zu zusätzlichen Herausforderungen führen.

4.4. BERICHTERSTATTUNG ZU DEM WASSERVERBRAUCH AUS DER PRAXIS - NACHHALTIGKEITSBERICHTE

Die Auswertung der Nachhaltigkeitsberichte hat gezeigt, dass nahezu alle der betrachteten Unternehmen Angaben zum Wasserverbrauch machen, wobei die meisten konkrete Zahlen vorlegen. Allerdings bleibt die Differenzierung zwischen direkten und indirekten Verbräuchen (Scope 1–3) sowie zwischen „Water Withdrawal“ und „Water Consumption“ oft unberücksichtigt. Vergleichswerte zu Vorjahren oder Branchenstandards bieten viele, aber nicht alle, der untersuchten Berichte. Zudem zeigt sich, dass die Transparenz und Nachvollziehbarkeit der Angaben, insbesondere in Bezug auf die Herleitung der Zahlen, oft begrenzt sind.

Ziele zur Reduktion des Wasserverbrauchs haben fast alle Unternehmen definiert, wobei viele anstreben, bis 2030 „wasserpositiv“ zu werden. Microsoft setzte dieses Ziel als erstes der betrachteten Unternehmen im Jahr 2020 (Microsoft, 2022). Langfristige Trends und Basisjahre, die die Zielerreichung messbar machen könnten, werden jedoch nur selten berücksichtigt. Strategien zur Wasserreduktion, wie Wasserrecycling oder innovative Technologien, werden in einer wachsenden Zahl von Berichten aufgeführt – ein deutlicher Anstieg im Vergleich zu 2020. Positiv hervorzuheben ist zudem das Engagement vieler Unternehmen in Projekten zur Förderung eines nachhaltigen Wassermanagements, etwa durch Investitionen in regionale Wasserprojekte oder Partnerschaften zur Verbesserung der Wasserversorgung und -qualität in besonders betroffenen Gebieten.

Ein weiterer Fokus der Analyse lag auf der Berichterstattung zum Wasserverbrauch in RZ. Der Umfang der Berichterstattung zu diesem Thema ist begrenzt: Ein beträchtlicher Teil der Unternehmen erwähnt den Wasserverbrauch ihrer RZ gar nicht. Angaben zur WUE fehlen in vielen Berichten, haben jedoch im untersuchten Zeitraum leicht zugenommen. Nur wenige Unternehmen nennen Ziele zur Verbesserung der Wassereffizienz in ihren RZ, und oft fehlen konkrete Angaben zur Quantifizierung dieser Ziele. Ein auffälliges Defizit zeigt sich ebenfalls bei der Berichterstattung über die Umweltwirkungen von KI. Nur ein kleiner Teil der Berichte behandelt dieses Thema, und auch dort mangelt es an konkreten Maßnahmen, Zielen oder Indikatoren. Stattdessen wird KI vor allem im Kontext ihrer potenziellen Vorteile für die Entwicklung nachhaltiger Lösungen diskutiert, ohne die eigenen Umweltwirkungen wie Wasser- und Energieverbrauch zu thematisieren.

Insgesamt zeigt die Analyse der Nachhaltigkeitsberichte, dass Wassermanagement in den Nachhaltigkeitsberichten der „Magnificent 7“ zunehmend thematisiert wird. Dennoch bestehen weiterhin erhebliche Unterschiede in der Tiefe und Transparenz der Berichterstattung, insbesondere im Hinblick auf RZ, KI-Umweltwirkungen und die langfristige Kontextualisierung von Wasserzielen.

5. Ergebnisse II: Einflussfaktoren und Best Practices

Forschungsfrage: *Welche Möglichkeiten gibt es, den direkten Wasserfußabdruck von KI und insbesondere von Rechenzentren zu reduzieren?*

Der Wasserverbrauch durch KI wird von verschiedenen Faktoren beeinflusst, die in enger Wechselwirkung zueinanderstehen. Im Expertenworkshop wurden Einflussfaktoren identifiziert und zunächst in acht Kategorien unterteilt: *Art der Kühlung, Praktiken der Hardware- und Softwaregestaltung, Energie-Wasser-Nexus, Standortfaktoren, Rohstoffe und Herstellung, Politische Rahmenfaktoren, Ökonomische Dynamiken und sozioökonomische Rahmenfaktoren.*

Ursprünglich wurden sozioökonomische Rahmenfaktoren nicht als Hauptgegenstand der Untersuchung betrachtet. Da sie jedoch sowohl im Workshop als auch in Experteninterviews als zentrale Einflussfaktoren hervorgehoben wurden, sind sie in die weitere Analyse eingeflossen. Basierend auf den Erkenntnissen aus dem Workshop und den Experteninterviews wurden die Einflussfaktoren in vier übergeordnete Kategorien zusammengeführt (Abbildung 8):

1. **Praktiken der Softwaregestaltung:** Aspekte der Auswahl, Nutzung und Entwicklung von KI-Software mit Auswirkungen auf den Ressourcenverbrauch.
2. **Praktiken der Hardwaregestaltung:** Betrachtung des gesamten Lebenszyklus von Hardwarekomponenten.
3. **Rechenzentrumsdesign und -betrieb:** Einfluss der Kühltechnologie, der Energieversorgung und des Energie-Wasser-Nexus auf den Gesamtverbrauch sowie die Optimierung der Hard-Software-Interaktion im RZ.
4. **Standortfaktoren:** Berücksichtigung geografischer Gegebenheiten sowie politischer und sozioökonomischer Rahmenfaktoren.



Abbildung 8: Einflussfaktoren und Best Practices.

In den Experteninterviews wurden zahlreiche Best Practices identifiziert, die zur Reduzierung des Wasserverbrauchs im KI-Sektor beitragen können. Diese lassen sich den genannten Kategorien zuordnen. Dabei wird kein Anspruch auf Vollständigkeit erhoben. Vielmehr dienen sie als Orientierung für mögliche Maßnahmen und zukünftige Entwicklungen.

5.1. PRAKTIKEN DER SOFTWAREGESTALTUNG

Ein zentraler Aspekt, den die Expert*innen mehrfach betonten, ist die energie- und ressourceneffiziente Gestaltung von Software und KI-Modellen. Diese beeinflusst maßgeblich die Ressourcennutzung von RZ, da nicht alle Berechnungsprozesse denselben Energie- und Wasserverbrauch verursachen. Faktoren wie Modellgröße, Programmiersprache und Algorithmenineffizienz spielen eine entscheidende Rolle für den ökologischen Fußabdruck von KI-Anwendungen.

Obwohl zahlreiche Best Practices existieren, erschweren fehlende Vergleichbarkeit und standardisierte Benchmarks eine umfassende Quantifizierung ihrer Effekte. Durch gezielte Softwareoptimierung lassen sich dennoch erhebliche Einsparungen erzielen, etwa durch parallele Programmierung oder eine effizientere Nutzung von Mehrkernprozessoren. Die Expert*innen empfehlen daher verstärkt den Einsatz intelligenter Optimierungsstrategien wie kontinuierliches und föderiertes Lernen, um KI-Anwendungen nachhaltiger zu gestalten und den Ressourcenverbrauch langfristig zu reduzieren.

ENERGIE- UND RESSOURCENEFFIZIENTE PROGRAMMIERPRAKTIKEN (GREEN CODING)

Green Coding umfasst Methoden zur Optimierung der Energie- und Ressourceneffizienz von Software, etwa durch den Wechsel zu ressourcenschonenden Programmiersprachen oder die Reduzierung unnötig großer Datenmengen. Ziel ist es, überflüssige Rechenprozesse zu vermeiden und den Gesamtverbrauch an Energie und Wasser zu senken. Da der Ressourcenbedarf von RZ direkt mit der Rechenleistung und Speicheranforderung von Softwareanwendungen zusammenhängt, könnte eine stärkere Berücksichtigung von Energie- und Ressourceneffizienz als Entwicklungsindikator erheblich zur Nachhaltigkeit beitragen. Besonders in der KI-Entwicklung ist eine enge Abstimmung zwischen Software und Hardware essenziell, um Rechenkapazitäten effizient zu nutzen und unnötige Lasten zu vermeiden. Während verschiedene *Green Coding*-Praktiken zunehmend erforscht werden, sind sie in der Praxis noch nicht weit verbreitet.

Ein zentraler Ansatz dabei ist die Vermeidung von *Software Bloat*, also die übermäßige Aufblähung von Software durch ineffizienten Code und unnötige Funktionen. Diese unnötige Komplexität führt zu einem erhöhten Rechen-, Energie- und Wasserverbrauch. Durch schlankere Softwarearchitekturen, modulare Entwicklung und regelmäßige Code-Überprüfung können ineffiziente Berechnungen reduziert und der Betrieb ressourcenschonender gestaltet werden. Während solche Optimierungen bereits in der Embedded-Software-Entwicklung gängige Praxis sind, steht die KI-Entwicklung häufig vor der Herausforderung, zwischen Effizienz und Genauigkeit abzuwägen. Eine stärkere Integration von Nachhaltigkeitskriterien in die Softwareentwicklung könnte dazu beitragen, den ökologischen Fußabdruck digitaler Anwendungen zu verringern.

RESSOURCENSCHONENDE OPTIMIERUNG VON KI-MODELLEN

Durch gezielte Optimierungstechniken lässt sich der Rechenaufwand von KI-Modellen potenziell verringern – was indirekt auch Energie- und Wasserverbrauch reduzieren kann, insbesondere über energieintensive RZ. Ein Ansatz ist das **Approximate Computing**, bei dem Rechenprozesse gezielt vereinfacht werden, etwa durch geringere Genauigkeit oder frühzeitigen Abbruch von Trainingsläufen. Diese Technik bietet vor allem bei rechenintensiven Anwendungen wie ML und generativer KI theoretisch Einsparpotenziale, ist bislang jedoch vor allem in der Forschung verbreitet. Ein weiterer Ansatz ist die Modellkompression: Hierbei werden durch Methoden wie *Pruning* oder *Knowledge Distillation*-Modelle so angepasst, dass sie mit weniger Parametern und Rechenaufwand auskommen – oft mit nur geringen Einbußen bei der Modellleistung. Solche Verfahren finden bereits Anwendung bei großen Technologiekonzernen und in der Forschung, sind jedoch noch nicht flächendeckend in der Praxis etabliert. Eine breitere Nutzung könnte langfristig dazu beitragen, den Ressourcenverbrauch von KI-Systemen zu reduzieren – vorausgesetzt, sie werden zielgerichtet und kontextabhängig eingesetzt.

OPTIMIERUNG DURCH PARETO-ANALYSEN

Pareto-Analysen ermöglichen es, KI-Modelle nicht nur anhand ihrer reinen Leistung, sondern gleichzeitig im Hinblick auf Kriterien wie Energie- und Wassereffizienz zu bewerten. Dabei werden Modelle systematisch hinsichtlich mehrerer Zielgrößen analysiert, sodass nicht ausschließlich Genauigkeit oder Rechenleistung, sondern auch der Ressourcenverbrauch in die Optimierung einfließt. Auf diese Weise lassen sich Modelle identifizieren, die Effizienzgewinne ermöglichen können.

Solche Ansätze werden bereits in einzelnen Bereichen der Software- und Hardwareentwicklung eingesetzt, etwa im Kontext von Modellkomprimierung oder Architekturdesign. In der breiten KI-Entwicklung sind sie jedoch bislang noch nicht fest verankert. Eine stärkere Integration von Multi-Objective-Optimierungen wie der Pareto-Analyse in den Entwicklungsprozess könnte dazu beitragen, nachhaltigere KI-Modelle zu fördern – die sowohl leistungsfähig als auch ressourcenschonender sind.

EINSATZ SPEZIALISierter, KLEINERER KI-MODELLE ZUR RESSOURCENSCHONUNG

Anstatt große, generalisierte KI-Modelle zu nutzen, können kleinere, auf spezifische Aufgaben optimierte Modelle den Rechenaufwand erheblich reduzieren. Diese benötigen weniger Speicher, Rechenleistung und Energie, was sich potenziell auch positiv auf den Wasserverbrauch in RZ auswirken kann. Da task-spezifische Modelle nur die für eine bestimmte Anwendung notwendigen Berechnungen ausführen, können sie in vielen Fällen effizienter sein als breit angelegte, ressourcenintensive Modelle.

Obwohl große KI-Modelle derzeit viele Anwendungen dominieren, fehlt es an standardisierten Benchmarks, um den Energieverbrauch und die Umweltbilanz verschiedener Modellgrößen systematisch zu vergleichen. Erste Analysen deuten jedoch darauf hin, dass kleinere Modelle für spezifische Aufgaben ähnliche Ergebnisse liefern können, während sie gleichzeitig einen geringeren Ressourcenverbrauch aufweisen. Eine verstärkte Forschung und standardisierte Effizienzmetriken könnten dazu beitragen, den gezielten Einsatz spezialisierter Modelle zu fördern und damit eine nachhaltigere Nutzung von KI zu unterstützen.

NUTZUNG VON OPEN-SOURCE-MODELLEN ZUR EFFIZIENZSTEIGERUNG

Open-Source-Modelle bieten die Möglichkeit, bereits optimierte Algorithmen weiterzuverwenden und an spezifische Anforderungen anzupassen, anstatt neue Modelle von Grund auf zu entwickeln. Durch die Wiederverwendung bestehender Lösungen kann der Rechenaufwand für das Training neuer Modelle verringert werden, was potenziell zur Reduktion des Energie- und Wasserverbrauchs beiträgt. Ein wesentlicher Vorteil von Open-Source-Modellen liegt im Transfer Learning, bei den vortrainierten Modellen für neue Aufgaben angepasst werden, anstatt sie vollständig neu zu trainieren. Dies kann den Ressourcenaufwand erheblich reduzieren.

Dennoch setzen viele Unternehmen weiterhin auf proprietäre Modelle, die mit hohem Rechenaufwand trainiert werden. Eine verstärkte Nutzung von Open-Source-Ansätzen könnte zu mehr Effizienz und Ressourcenschonung beitragen, insbesondere wenn bestehende Modelle gezielt weiterentwickelt anstatt von Grund auf neu trainiert werden. Der tatsächliche Einfluss auf den ökologischen Fußabdruck der KI-Entwicklung und insbesondere auf den Wasserverbrauch ist jedoch nicht abschließend geklärt. Weitere Forschung wäre notwendig, um die ökologischen Vorteile dieser Ansätze genauer zu quantifizieren.

5.2. PRAKTIKEN DER HARDWAREGESTALTUNG

Die Gestaltung und Nutzung von Hardware in RZ spielt eine entscheidende Rolle für den Wasserverbrauch in der digitalen Infrastruktur. Neben dem direkten Wasserverbrauch durch Kühlprozesse beeinflussen auch die Herstellung, Lebensdauer und das Recycling von Hardware maßgeblich den ökologischen Fußabdruck. Ergänzend zu den in Kapitel 4.3 analysierten Maßnahmen der Halbleiterindustrie zur Reduzierung des Wasserverbrauchs während der Produktion zeigen die Experteninterviews, dass insbesondere die steigende Nutzung spezialisierter KI-Hardware wie GPUs und TPUs sowie die geringe Recyclingquote von IT-Komponenten erhebliche Herausforderungen darstellen. Eine effizientere Nutzung, längere Lebensdauer und eine stärkere Kreislaufwirtschaft könnten den Wasserverbrauch erheblich reduzieren, erfordern jedoch gezielte Maßnahmen und wirtschaftliche Anreize.

EFFIZIENTE HARDWARE & OPTIMIERTE NUTZUNG

Eine zentrale Strategie zur Reduktion des Wasserverbrauchs in RZ ist die Entwicklung und Nutzung energieeffizienter Hardware. Moderne Chips und Server mit optimierter Energieeffizienz senken nicht nur den Stromverbrauch, sondern verringern auch die Kühlanforderungen, was direkt den Wasserverbrauch beeinflusst. Expert*innen betonten zudem die Bedeutung einer besseren Auslastung bestehender Hardware. Durch Virtualisierung und die verstärkte Nutzung von Mehrkernprozessoren kann die benötigte Serverkapazität

optimiert werden, wodurch sich der Bedarf an neuer Hardware reduziert. Diese Maßnahmen sind in vielen Unternehmen bereits etabliert, jedoch besteht weiterhin Verbesserungspotenzial, insbesondere in Hinblick auf KI-Workloads, die hohe Rechenleistung erfordern.

VERLÄNGERUNG DER HARDWARE-LEBENSDAUER

Die nachhaltige Nutzung bestehender Hardware stellt eine weitere Best Practice dar, um den Ressourcenverbrauch langfristig zu senken. Eine verbesserte Wartung und gezielte Aufrüstung einzelner Komponenten anstelle des kompletten Hardware-Austauschs tragen dazu bei, die Lebensdauer von Servern und Speicherlösungen zu verlängern. In der Praxis fehlt jedoch oft der wirtschaftliche Anreiz für eine langfristige Nutzung, da Herstellerzyklen und Innovationsdruck häufig zu einem schnellen Austausch führen.

RECYCLING & KREISLAUFWIRTSCHAFT

Ein weiteres großes Potenzial zur Reduktion des Wasserverbrauchs liegt in der Verbesserung der Recycling- und Wiederverwendungsraten von IT-Hardware. Die Expert*innen betonten, dass der verstärkte Einsatz recycelter Materialien in der Chip- und Serverproduktion sowie die Entwicklung von Programmen zur Wiederverwendung von IT-Hardware den Ressourcenverbrauch erheblich senken könnte. Der Übergang zu einer Kreislaufwirtschaft in der IT-Branche ist jedoch noch nicht weit verbreitet und erfordert regulatorische Maßnahmen sowie wirtschaftliche Anreize.

5.3. RECHENZENTRUMSDSIGN & NUTZUNG

Wie bereits in Kapitel 4 beschrieben, gibt es verschiedene Rechenzentrums- und Kühlungsdesigns, die maßgeblich den Energie- und Wasserverbrauch beeinflussen. Neben den bereits erläuterten Maßnahmen zur Anpassung der Kühlung an die klimatischen Bedingungen des Standorts sowie ressourcenschonenden Kühlsystemen wie der adiabaten Kühlung wurden in den Experteninterviews weitere Best Practices identifiziert. Diese zusätzlichen Ansätze bieten weiteres Potenzial zur Reduzierung des Wasser- und Energieverbrauchs und tragen zur nachhaltigen Optimierung des Rechenzentrumsbetriebs bei.

HOCHTEMPERATUR-KÜHLUNG

Hochtemperatur-Kühlung reduziert den Wasserverbrauch, indem sie den Kühlkreislauf mit höheren Eintrittstemperaturen (bis zu 50 °C oder mehr) betreibt. Dadurch kann in vielen Klimazonen auf energie- und wasserintensive Kältemaschinen oder Verdunstungskühlung verzichtet werden. Besonders in Regionen mit hohen Außentemperaturen ermöglicht diese Methode eine nahezu verdunstungsfreie Kühlung. Hochtemperatur-Flüssigkeitskühlung in Servern kann zudem die Effizienz steigern und die Rückkühlung vereinfachen.

Diese Technologie wird bereits in Hochleistungs-RZ und spezialisierten KI-RZ eingesetzt, ist jedoch noch nicht weit verbreitet. Eine Herausforderung ist die Anpassung der Hardware an höhere Betriebstemperaturen, da viele IT-Komponenten derzeit auf niedrigere Temperaturen ausgelegt sind. Fortschritte in der Material- und Chiptechnologie könnten jedoch die Akzeptanz dieser Methode in Zukunft weiter erhöhen.

HYBRIDE KÜHLVERFAHREN

Hybride Kühlsysteme kombinieren Luft-, adiabate und Flüssigkeitskühlung, um je nach Last und Umweltbedingungen die effizienteste Methode zu nutzen und auf diese Weise Wasser und Energie zu sparen. In

kühleren Monaten oder bei moderaten Workloads wird Luftkühlung bevorzugt, während bei hoher Belastung oder warmen Außentemperaturen Flüssigkeits- oder adiabate Kühlung zugeschaltet wird. Intelligente Steuerungssysteme ermöglichen eine dynamische Anpassung an wechselnde Bedingungen. Hybride Kühlung ist bereits in Hyperscale-RZ verbreitet, allerdings erfordert die Implementierung eine komplexere Infrastruktur und sorgfältige Planung, da nicht alle Standorte oder Hardware-Designs für den flexiblen Wechsel zwischen verschiedenen Kühlmethoden optimiert sind.

OPTIMIERUNG DER SERVER- UND HARDWARE-AUSLASTUNG UND MESSUNG DES ENERGIEVERBRAUCHS IN RZ

Durch eine gezielte Anpassung der Hardware-Betriebspunkte kann der Energieverbrauch gesenkt werden, ohne die Leistung signifikant zu beeinträchtigen. Standardmäßig sollten Server und GPUs nicht auf Höchstleistung laufen, sondern sich dynamisch an Workloads anpassen. Vergleichbar mit einem Auto, das nicht im „Sportmodus“ startet, kann eine energieeffiziente Voreinstellung große Einsparungen bringen. Diese Strategie ist bereits in *Supercomputing*- und *Hyperscale*-RZ verbreitet, aber nicht durchgängig implementiert. Automatisierte Systeme zur Workload-Optimierung könnten weitere Einsparungen ermöglichen, indem sie den Stromverbrauch und die Prozessorleistung in Echtzeit anpassen, was Energie- und Wassereffizienz verbessert. Zudem gibt es bereits Tools, die den Energieverbrauch von RZ optimieren, indem sie den Stromverbrauch kontinuierlich überwachen und den Betriebspunkt dynamisch anpassen. In Hochleistungs-RZ konnten durch solche Systeme bereits Energieeinsparungen von bis zu 20 % erzielt werden.

NUTZUNG VON FREE COOLING DURCH KLIMATISCHE BEDINGUNGEN

In Regionen mit kühlen Temperaturen kann *Free Cooling*, also die Nutzung der Umgebungsluft zur Kühlung, den Energie- und Wasserverbrauch erheblich senken. In Ländern wie Deutschland, Skandinavien oder Island kann *Free Cooling* bis zu 100 % der Kühlanforderungen abdecken, solange Außentemperaturen niedrig genug sind. Die Methode ermöglicht erhebliche Einsparungen, da wasserbasierte Kühlsysteme reduziert oder sogar ganz vermieden werden können.

Der Einsatz von *Free Cooling* ist jedoch stark standortabhängig und eignet sich nicht für tropische oder heiße Klimazonen. Zudem erfordert diese Methode spezielle Infrastrukturen, um eine effiziente Wärmerückführung zu gewährleisten. Unternehmen müssen daher abwägen, ob *Free Cooling* an ihren geplanten Standorten wirtschaftlich und technologisch sinnvoll realisierbar ist.

5.4. STANDORTFAKTOREN

Ein zentraler Einflussfaktor, den die Expert*innen im Workshop hervorhoben, war die Wahl des Standorts für RZ. Neben klimatischen Bedingungen spielen dabei der Flächenbedarf, die Nähe zu Energie- und Wasserquellen sowie regionale Möglichkeiten zur Nutzung von Abwärme eine wichtige Rolle. Laut den Expert*innen wird die Standortentscheidung jedoch primär von der Anbindung an das Energienetz und die Nähe zu Internetknotenpunkten bestimmt, während ökologische Faktoren bisher nur eine untergeordnete Rolle spielen.

STANDORTWAHL & NUTZUNG ALTERNATIVER WASSERQUELLEN

Die Wahl des Standorts beeinflusst maßgeblich den Wasserverbrauch eines RZ. Eine nachhaltige Strategie umfasst die Platzierung von RZ, deren Kühltechnologie große Mengen an Wasser benötigt, in wasserreichen Gebieten, um die Belastung lokaler Trinkwasserressourcen zu vermeiden. Zudem kann die Nutzung von aufbereitetem Abwasser oder Oberflächenwasser für Kühlzwecke eine sinnvolle Alternative darstellen, da auf diese Weise die Abhängigkeit von Trinkwasserquellen reduziert und eine effizientere Wassernutzung durch Integration in bestehende Wasserkreisläufe ermöglicht wird. Dieses Konzept gewinnt zunehmend an Bedeutung,

insbesondere in Regionen mit akuter Wasserknappheit wie Arizona oder Kalifornien. Dennoch bestehen Herausforderungen, insbesondere durch regulatorische Vorgaben zur Wasserqualität oder durch logistische Hürden bei der Bereitstellung alternativer Wasserquellen in großem Maßstab. Eine erfolgreiche Implementierung erfordert daher enge Zusammenarbeit mit lokalen Behörden und Betreiber*innen der Infrastruktur.

ABWÄRMENUTZUNG IN FERNWÄRMENETZEN & INDUSTRIEN

RZ erzeugen erhebliche Mengen an Abwärme, die für die Beheizung von Gebäuden oder industrielle Prozesse genutzt werden könnte. Integrierte Lösungen ermöglichen es, Server-Abwärme direkt in Fernwärmenetze einzuspeisen oder für Prozesse mit hohem Wärmebedarf, wie in der Lebensmittel- oder Chemieindustrie, zu nutzen. Dadurch könnte ein erheblicher Teil der bislang ungenutzten Wärmeenergie effizient weiterverwendet werden.

Während Abwärmenutzung in einigen Regionen, etwa Skandinavien oder den Niederlanden, bereits etabliert ist, bestehen in vielen Ländern noch technische und regulatorische Hürden. Ein Problem ist die relativ niedrige Temperatur der Abwärme, die oft nicht direkt genutzt werden kann. Die steigende gesetzliche Verpflichtung zur Abwärmenutzung, wie im deutschen Energieeffizienzgesetz festgelegt, könnte jedoch dazu beitragen, diese Praktik in größerem Umfang zu implementieren und wirtschaftlich rentabel zu machen.

DEZENTRALISIERTE RECHENZENTRUMSSTRUKTUREN ZUR RESSOURCENVERTEILUNG

Anstatt große RZ an wenigen Standorten zu bündeln, könnten kleinere, geografisch verteilte RZ helfen, Ressourcen effizienter zu nutzen und regionale Engpässe zu vermeiden. Eine dezentrale Infrastruktur ermöglicht eine bessere Verteilung der Wasser- und Energiebelastung auf verschiedene Regionen, was das Risiko lokaler Überlastungen verringert. Zudem kann die Integration von Abwärmenutzung in lokale Netzwerke erleichtert werden, da kleinere RZ oft näher an potenziellen Wärmeabnehmern liegen.

Diese Strategie wird insbesondere für KI-RZ diskutiert, bei denen die physische Nähe zum Nutzer/ zur Nutzerin nicht immer erforderlich ist. In vielen Fällen könnten Trainingsmodelle von KI in abgelegenen, energieeffizienten Standorten laufen, während nur das *Inferencing* näher am Nutzer erfolgt. Allerdings kann die Verlagerung von Rechenkapazitäten mit zusätzlichen Herausforderungen verbunden sein, etwa höheren Netzwerklatenzen oder erhöhtem Verwaltungsaufwand für verteilte Infrastrukturen.

6. Ergebnis III: Hindernisse und Herausforderungen

Forschungsfrage: Welche technologischen, ökonomischen und regulatorischen Hindernisse stehen der Umsetzung wassersparender Lösungen und Innovationen entgegen?

Die Entwicklung ressourcenschonender KI-Technologien und der energie- und ressourceneffiziente Betrieb von RZ stehen vor zahlreichen Herausforderungen. Diese betreffen sowohl technische als auch strukturelle Aspekte und erstrecken sich über verschiedene Dimensionen – von politischen Rahmenbedingungen über wirtschaftliche Anreize bis hin zu geografischen Gegebenheiten und Marktdynamiken im KI-Sektor.

Unzureichende Regulierungen, fehlende Transparenz und das Fehlen einheitlicher Standards führen häufig zu einer ineffizienten Nutzung von Ressourcen und behindern notwendige Fortschritte. Gleichzeitig prägen wirtschaftliche Interessen und die steigende Nachfrage nach leistungsfähigeren KI-Modellen die Entwicklung der digitalen Infrastruktur, was ressourcenschonende Praktiken erschwert.

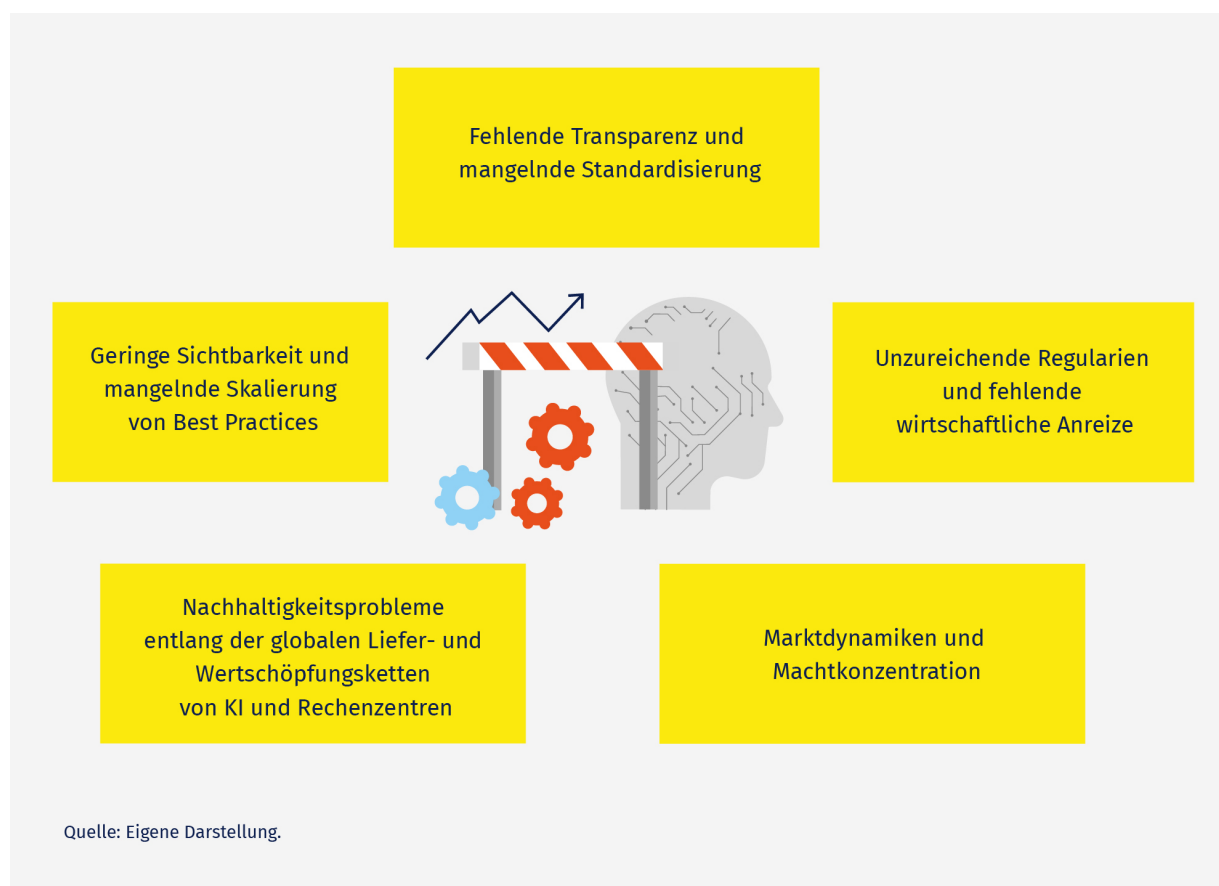


Abbildung 9: Hindernisse und Herausforderungen für die ressourcenschonende KI und Rechenzentren.

Um diese komplexen Herausforderungen systematisch zu erfassen, wurden im Expertenworkshop sowie in den Interviews zentrale Problemfelder identifiziert. Basierend auf diesen Erkenntnissen wurden fünf zentrale Kategorien gebildet, die die wesentlichen Hindernisse für eine ressourceneffizientere Gestaltung von KI und RZ zusammenfassen:

1. **Fehlende Transparenz und mangelnde Standardisierung**

Unzureichende Datenerhebung und fehlende Vergleichbarkeit von Umweltkennzahlen erschweren die Bewertung und Optimierung des Ressourcenverbrauchs von KI und RZ. Einheitliche Metriken und verpflichtende Offenlegungen könnten hier Abhilfe schaffen.

2. **Geringe Sichtbarkeit und mangelnde Skalierung von Best Practices**

Ressourcenschonende Lösungen sind oft nur in einzelnen Unternehmen oder Institutionen implementiert und werden nicht systematisch verbreitet. Dies verhindert eine breitere Adaption und damit eine großflächige Reduktion des Ressourcenverbrauchs.

3. **Unzureichende Regularien und fehlende wirtschaftliche Anreize**

Fehlende oder unzureichende gesetzliche Vorgaben sowie mangelnde wirtschaftliche Anreize behindern Investitionen in ressourcenschonende Technologien. Eine gezielte Regulierung sowie Anreize für nachhaltige Investitionen könnten diesen Trend umkehren.

4. **Marktdynamiken und Machtkonzentration**

Wenige große Akteure dominieren den Markt für KI-Modelle und RZ. Dies erschwert den Zugang zu alternativen, ressourcenschonenden Lösungen und begrenzt Innovationspotenziale, insbesondere für kleinere Unternehmen und Forschungseinrichtungen.

5. **Nachhaltigkeitsprobleme entlang der globalen Liefer- und Wertschöpfungsketten von KI und Rechenzentren**

Die Produktion und Nutzung von KI-Hardware ist durch komplexe, global vernetzte Lieferketten geprägt, die mit erheblichen Umweltbelastungen verbunden sind. Fehlende Transparenz und mangelnde Standards erschweren eine nachhaltige Steuerung dieser Prozesse.

Dieses Kapitel analysiert die genannten Herausforderungen im Detail und zeigt auf, welche strukturellen und politischen Rahmenbedingungen notwendig sind, um langfristig eine ressourcenschonende Entwicklung im KI- und Rechenzentrumssektor zu ermöglichen.

6.1. FEHLENDE TRANSPARENZ UND MANGELNDE STANDARDISIERUNG

Unzureichende Datenerhebung und fehlende Transparenz

Ein wesentliches Hindernis für nachhaltige KI und RZ ist der Mangel an präzisen, vergleichbaren Daten über den gesamten Lebenszyklus – von der Hardwareproduktion über den Betrieb bis zur Entsorgung. Besonders kritisch ist die fehlende Transparenz beim Wasser- und Energieverbrauch in der Halbleiterproduktion. Unternehmen, die RZ betreiben, veröffentlichen oft nur aggregierte Zahlen, die weder regionale Unterschiede noch spezifische Verbrauchsmuster einzelner RZ offenlegen.

Während einige Betreiber neben dem PUE auch den WUE als Kennzahl für den Wasserverbrauch veröffentlichen, bleiben diese Werte oft unvollständig oder nicht vergleichbar. Zudem fehlen etablierte Methoden zur Berechnung des anteiligen Verbrauchs für unterschiedliche KI-Anwendungen oder Hardwareplattformen, was eine gezielte Optimierung erschwert.

Obwohl es Initiativen zur Definition von Messgrößen wie Liter pro Kilowattstunde oder spezifischen Wasserverbrauchskennzahlen gibt, fehlt eine zentrale Instanz zur systematischen Sammlung und Prüfung dieser Daten. Dies behindert nicht nur die wissenschaftliche Forschung, sondern auch die Einführung nachhaltiger Praktiken durch Unternehmen und Regulierungsbehörden. Ohne verpflichtende Offenlegung bleibt es schwierig, Benchmarks für ressourcenschonende KI-Modelle und RZ zu entwickeln sowie geeignete politische Maßnahmen abzuleiten.

Fehlende Standards und Bewertungsmethoden

Neben der unzureichenden Datenerhebung fehlen verbindliche Standards und Benchmarks zur Bewertung des Energie- und Ressourcenverbrauchs von KI-Modellen und RZ. Während für Supercomputer etablierte Benchmarks wie der Green500-Index² existieren, fehlen vergleichbare Systeme für KI-Anwendungen. Bestehende Benchmarks, wie der Linpack-Benchmark, auf dem sowohl der TOP500- als auch der Green500-Index basieren, sind nicht optimal zur Bewertung von KI-Workloads geeignet, da sie ursprünglich für lineare Gleichungssysteme entwickelt wurden (top500.org, o. J.). Trotz gewisser Überschneidungen, etwa bei Matrixmultiplikationen, fehlt ein spezifischer Maßstab, um die Energieeffizienz verschiedener KI-Hardwarearchitekturen objektiv zu bewerten. Deshalb wird die Entwicklung neuer Benchmarks gefordert, die eine international vergleichbare Bewertung der Energieeffizienz von KI-Systemen ermöglichen.

Fehlende Standards erschweren zudem eine objektive Einordnung des Ressourcenverbrauchs:

- **Unklare Effizienzbewertung:** Generative KI-Modelle für Bild- und Videogenerierung verbrauchen enorme Rechenkapazitäten, ohne dass klar ist, ob größere Modelle tatsächlich leistungsfähiger sind oder ob kleinere Modelle denselben Zweck mit geringerem Ressourcenaufwand erfüllen könnten.
- **Fehlende Vergleichbarkeit:** Unternehmen bewerten ihre KI-Modelle nach eigenen Maßstäben, was Transparenz erschwert und Greenwashing begünstigt.

Die Wassernutzung in RZ wird durch bestehende Metriken nur unzureichend abgedeckt. Der WUE erfasst ausschließlich den laufenden Betrieb, nicht jedoch den Wasserverbrauch für die Herstellung oder das Recycling von Hardware. Auch unterscheiden sich die Anforderungen zwischen großen Hyperscalern und kleineren, lokal betriebenen RZ erheblich, was eine Standardisierung erschwert. Eine einheitliche Lösung, die für alle Betreiber praktikabel ist, bleibt eine große Herausforderung.

6.2. GERINGE SICHTBARKEIT UND MANGELNDE SKALIERUNG VON BEST PRACTICES

Obwohl es entlang der gesamten KI-Lieferkette bereits zahlreiche Best Practices gibt, fehlt es an Sichtbarkeit, Validierung und systematischer Verbreitung dieser Ansätze. In der Debatte um die Reduktion des Ressourcenverbrauchs von KI und RZ wird oft versäumt, bestehende Maßnahmen zu erfassen, zu bewerten und gezielt weiterzuentwickeln. Trotz vorhandener Lösungen bleiben viele effiziente Praktiken ungenutzt, da sie nicht ausreichend dokumentiert oder skaliert werden.

Für eine breitere Implementierung ressourcenschonender Technologien müssen bewährte Lösungen transparenter gemacht und durch Standards gestützt werden. Dies betrifft insbesondere effiziente Kühltechnologien in RZ, die oft über das notwendige Maß hinaus betrieben werden. Eine präzisere Steuerung könnte den Wasser- und Energieverbrauch erheblich senken, wird jedoch aufgrund fehlender Transparenz und standardisierter Vorgaben nur selten umgesetzt.

Zudem werden bestehende Workload-Optimierungsstrategien nicht konsequent genutzt. Unnötige Serverkapazitäten bleiben oft aktiviert, da es an automatisierten Managementsystemen fehlt, die ungenutzte Hardware abschalten oder Workloads effizienter verteilen. Ein weiteres Hindernis ist das ungenutzte Innovationspotenzial im Softwarebereich. Während erhebliche Investitionen in energieeffiziente Hardware

² Der TOP500-Index listet die weltweit leistungsstärksten Supercomputer basierend auf dem Linpack-Benchmark, während der Green500-Index diese Systeme nach ihrer Energieeffizienz bewertet, also der Rechenleistung pro Watt Energieverbrauch. Der Linpack-Benchmark misst die Rechenleistung eines Systems bei der Lösung dichter linearer Gleichungssysteme

fließen, erhalten ressourcenschonende Softwarelösungen und alternative Betriebsstrategien zu wenig Aufmerksamkeit und sind nicht ausreichend sichtbar.

6.3. UNZUREICHENDE REGULARIEN UND FEHLENDE WIRTSCHAFTLICHE ANREIZE

Fehlende Regulierung und politische Rahmenbedingungen

Die Regulierung von RZ konzentriert sich bislang hauptsächlich auf Energieeffizienz, während der Wasserverbrauch oft unberücksichtigt bleibt. Zwar existieren in der Europäischen Union Vorgaben für nachhaltige RZ, doch es fehlen einheitliche Richtlinien zur Begrenzung des Wasserverbrauchs und zur verpflichtenden Veröffentlichung von Daten zur Wassernutzung sowie der Menge und Qualität von abgeleitetem Wasser. Dies erschwert eine effektive Regulierung sowie die Vergleichbarkeit nachhaltiger Maßnahmen. Ein weiteres Problem ist die internationale Uneinheitlichkeit der Regulierung. Während einige Länder wie Singapur bereits strikte Vorgaben für den Energieverbrauch von RZ eingeführt haben, bleibt der Wasserverbrauch global gesehen relativ unterreguliert. Dies ermöglicht es Unternehmen, Standorte gezielt nach Regionen mit laxeren Vorschriften zu wählen.

Aus Sicht von RZ-Betreibern wird auch der bürokratische Aufwand als Hindernis wahrgenommen. Viele bestehende Vorschriften gelten als schwer verständlich oder ineffektiv, da sie hohe Meldepflichten mit sich bringen, aber keine klaren Handlungsempfehlungen oder wirtschaftlichen Anreize bieten. Ein weiteres Problem ist die mangelnde behördliche Koordination: Zuständigkeiten für Wasserverbrauch, Stromversorgung und digitale Infrastruktur sind häufig auf verschiedene Ministerien oder Behörden verteilt, was eine kohärente Regulierung erschwert.

Hohe Investitionskosten und wirtschaftliche Hemmnisse

Nachhaltige Infrastrukturmaßnahmen für RZ erfordern häufig hohe Investitionen. Flüssigkeitsgekühlte Server oder alternative Kühlmittel wie Ammoniak oder Propan sind zwar ressourcenschonender, aber teurer in der Anschaffung als herkömmliche Luftkühlungssysteme. Aufgrund fehlender wirtschaftlicher Anreize scheuen viele Unternehmen diese zusätzlichen Kosten.

In Deutschland kommt hinzu, dass die hohen Energiepreise den Betrieb nachhaltiger RZ verteuern. Während RZ in kühleren Klimazonen mit nachhaltiger Energieversorgung grundsätzlich effizienter betrieben werden könnten, erfolgt die Standortwahl häufig aus wirtschaftlichen oder steuerlichen Erwägungen – etwa aufgrund der Nähe zu großen Internetknotenpunkten.

Ein weiteres Problem ergibt sich aus unzureichenden Anreizen für nachhaltige Infrastrukturmaßnahmen. Während einige Länder gezielt auf Förderungen setzen, fehlen in vielen Regionen wirtschaftliche Vorteile für energieeffiziente oder wassersparende RZ. Gleichzeitig werden Maßnahmen zur Nutzung von Abwärme durch zu starre Vorgaben erschwert. So erfordern einige Regularien eine jährliche Mindestquote für Abwärmenutzung, ohne saisonale Unterschiede im Wärmebedarf zu berücksichtigen.

6.4. MARKTDYNAMIKEN UND MACHTKONZENTRATION

Steigende Nachfrage nach immer leistungsfähigeren KI-Modellen

Ein möglicher Hemmschuh für eine nachhaltige Entwicklung im KI-Sektor ist die wachsende Nachfrage nach immer leistungsfähigeren und komplexeren KI-Modellen. Insbesondere große Sprach- und Bildverarbeitungsmodelle erfordern erhebliche Rechenkapazitäten, was potenziell zu einem steigenden

Energie- und Wasserverbrauch in RZ führen kann. Zugleich wird die Effizienz von KI-Modellen nicht immer ausreichend berücksichtigt. Größere Modelle gelten oft als leistungsfähiger, obwohl spezialisierte, kleinere Modelle in einigen Anwendungsfällen denselben Zweck mit geringerem Ressourcenaufwand erfüllen könnten. Die Marketingstrategien großer Technologieunternehmen setzen in der Regel auf das kontinuierliche Wachstum von KI-Modellen, während deren Ressourcenverbrauch bislang nur bedingt systematisch hinterfragt wird.

Marktmacht weniger großer Unternehmen und ungleicher Wettbewerb

Der KI-Markt wird derzeit von einigen großen Akteuren wie Google, Microsoft und Amazon dominiert, die sowohl die Entwicklung von KI-Modellen als auch die Bereitstellung der zugrunde liegenden Recheninfrastruktur steuern. Diese Machtkonzentration kann es kleineren Unternehmen und Forschungseinrichtungen erschweren, sich mit alternativen Ansätzen am Markt zu etablieren oder sich den Zugang zu essenziellen Ressourcen wie Rechenleistung, Landflächen für den Bau von RZ oder Investitionskapital zu sichern. Ein potenzielles Problem besteht im fehlenden Wettbewerb um nachhaltigere Alternativen. Da viele KI-Unternehmen auf die Cloud-Dienste der großen Anbieter angewiesen sind, gibt es bislang nur begrenzte wirtschaftliche Anreize, energie- und wassereffizientere Lösungen zu entwickeln oder in größerem Umfang zu skalieren. Zudem kann die Marktdominanz großer Technologieunternehmen Innovationen im Bereich nachhaltiger KI erschweren. Kleinere Anbieter und Forschungseinrichtungen haben nicht immer die finanziellen oder infrastrukturellen Mittel, um ressourcenschonendere Alternativen zu entwickeln und zu testen. Dadurch besteht die Gefahr, dass bestehende Marktstrukturen verstärkt werden, während nachhaltige Lösungen nicht im gleichen Maße gefördert oder skaliert werden.

6.5. NACHHALTIGKEITSPROBLEME ENTLANG DER GLOBALEN LIEFER- UND WERTSCHÖPFUNGSKETTEN VON KI UND RECHENZENTREN

Internationale Abhängigkeiten und eingeschränkte Einflussmöglichkeiten

Die globale Struktur der KI-Industrie führt zu erheblichen Abhängigkeiten und erschwert regionale Nachhaltigkeitsmaßnahmen. Viele der leistungsfähigsten generativen KI-Modelle, wie ChatGPT, werden in den USA entwickelt und betrieben, wodurch ihre Recheninfrastruktur außerhalb Europas liegt. Dies schränkt den Einfluss auf ökologische Standards erheblich ein. Zudem werden KI-Modelle oft auf weltweit verteilten RZ ausgeführt, wodurch es schwierig ist, lokale Nachhaltigkeitsvorgaben durchzusetzen. Ähnlich verhält es sich mit der Hardwareproduktion: Auch die Lieferketten für KI-Chips sind global verflochten, mit dominanten Standorten wie Taiwan und den USA, während die Produktionskapazitäten in Europa begrenzt sind. Diese internationale Struktur verringert die Transparenz der Prozesse und reduziert die Möglichkeit, nachhaltige Praktiken entlang der gesamten Lieferkette zu etablieren.

Hoher Wasserverbrauch bei der Halbleiterproduktion

Die Herstellung von Halbleitern, insbesondere von Hochleistungs-GPUs und spezialisierten KI-Chips, ist extrem wasserintensiv. Millionen Liter Wasser werden pro Produktionszyklus für die Reinigung, das Ätzen und die Kühlung von Wafern benötigt. Angesichts der steigenden Nachfrage nach KI-Hardware verschärft sich diese Problematik weiter.

Besonders problematisch ist die Produktion in wasserarmen Regionen wie Taiwan oder Arizona, wo die Halbleiterindustrie direkt mit der lokalen Wasserversorgung der Bevölkerung und der Landwirtschaft konkurriert. Gleichzeitig fehlt es an transparenten Berichten zum Wasserverbrauch führender Hersteller wie Nvidia oder TSMC, was die Regulierung und eine nachhaltige Steuerung erschwert. Da es bisher keine

international verbindlichen Vorgaben zur Begrenzung des Wasserverbrauchs in der Chipproduktion gibt, bleibt dieses Problem ungelöst.

Kurze Lebensdauer und fehlende Anreize für nachhaltige Hardware-Nutzung

Ein weiteres Nachhaltigkeitsproblem ist die kurze Lebensdauer von IT-Hardware. Insbesondere Server und Speicherlösungen wie SSDs haben oft nur eine Nutzungsdauer von wenigen Jahren, bevor sie ersetzt werden. Dies führt zu einem hohen Austauschbedarf und steigert den Ressourcenverbrauch.

Zudem werden Hardwarekomponenten oft vollständig ersetzt, anstatt gezielt aufgerüstet oder recycelt zu werden. Wirtschaftliche Anreize zur Verlängerung der Lebensdauer fehlen weitgehend. Während Unternehmen zunehmend in energieeffiziente Technologien investieren, bleibt der Wasserverbrauch in der Hardwareproduktion weitgehend unbeachtet. Wasser ist in vielen Regionen vergleichsweise günstig, sodass kaum wirtschaftlicher Druck besteht, wassersparende Maßnahmen zu priorisieren.

Regulatorische Vorgaben oder steuerliche Anreize für eine nachhaltigere Beschaffung und Nutzung von IT-Infrastruktur könnten dazu beitragen, den Ressourcenverbrauch zu reduzieren. Derzeit fehlt es jedoch an klaren Rahmenbedingungen, die Unternehmen dazu verpflichten oder motivieren, langlebige und ressourcenschonende Hardware einzusetzen.

7. Handlungsempfehlungen für ressourceneffiziente KI und Rechenzentren

Die vorliegende Studie untersucht die Umweltauswirkungen, insbesondere den Wasserverbrauch, der durch die zunehmende Entwicklung und Nutzung von KI entsteht. Dabei wurden zentrale Einflussfaktoren identifiziert, die den Wasserverbrauch bestimmen, sowie Best Practices zur Reduzierung dieser Belastung analysiert. Zudem wurden Barrieren und Herausforderungen herausgearbeitet, die einer nachhaltigeren Gestaltung entgegenstehen.

Basierend auf den Studienergebnissen wurde eine Reihe von Handlungsempfehlungen formuliert (siehe Abbildung 10), die in sieben Bereiche unterteilt sind: Einführung verbindlicher Transparenzvorgaben und Etablierung von Standards, Forschung und Innovationsförderung für ressourcen- und energieeffiziente KI-Technologien, Förderung von energie- und ressourceneffizienter Rechenzentrumsinfrastruktur und Betriebsweisen, Schaffung von Anreizen für energie- und ressourcenschonende Investitionen, Sichtbarmachung und Verbreitung von Best Practices, Förderung einer ganzheitlichen Betrachtung des direkten und indirekten Energie- und Ressourcenverbrauchs durch KI, Förderung ressourcenschonender Hardware und Stärkung der Kreislaufwirtschaft



Abbildung 10: Handlungsempfehlungen für ressourceneffiziente KI und Rechenzentren.

7.1. EINFÜHRUNG VERBINDLICHER TRANSPARENZVORGABEN UND ETABLIERUNG VON STANDARDS

Verpflichtende Offenlegung des Wasser- und Energieverbrauchs von Rechenzentren

Eine zentrale Empfehlung ist die Einführung verbindlicher Transparenzvorgaben für den Wasser- und Energieverbrauch von RZ und KI-Infrastrukturen. Zusätzlich zu Angaben zum verbrauchten Wasser, sollten auch die Wasserqualität des genutzten und wieder abgeleiteten Wassers Berücksichtigung finden. Derzeit fehlen einheitliche Monitoring- und Reporting-Systeme, wodurch viele Unternehmen entweder selbst nur eine eingeschränkte Übersicht über ihren Ressourcenverbrauch haben oder die offengelegten Daten nicht ausreichend methodisch erfasst sind, um eine fundierte Einordnung und Vergleichbarkeit zu ermöglichen. Um diese Lücke zu schließen, sollten Unternehmen neben den Energieverbräuchen auch den Wasser- und weiteren Ressourcenverbrauch anhand standardisierter Methoden erfassen und offenlegen. Dies erfordert die Entwicklung einheitlicher Metriken, die eine präzisere Bewertung des Ressourcenverbrauchs entlang des gesamten Lebenszyklus von KI-Modellen und RZ ermöglichen.

Ein zentraler Schritt wäre die Zusammenarbeit zwischen Forschungseinrichtungen, politischen Akteuren und Industrieverbänden, um verbindliche Standards für nachhaltige KI-Modelle zu entwickeln. Diese sollten nicht nur den Betrieb von RZ, sondern auch die vorgelagerten Produktionsprozesse und die Endnutzung von KI-Systemen berücksichtigen. Besonders relevant ist dies für die Halbleiterproduktion, deren Wasserverbrauch bislang kaum dokumentiert ist. Unternehmen sollten zudem verpflichtend angeben, in welchem Umfang verwendetes Wasser recycelt oder rückgeführt wird, um die tatsächlichen Umweltauswirkungen realistisch bewerten zu können.

Integration von Transparenzvorgaben in Nachhaltigkeitsbewertungen

Um die erfassten Verbrauchsdaten gezielt für nachhaltige Maßnahmen zu nutzen, sollten umfassendere Nachhaltigkeitsindikatoren für KI-Anwendungen entwickelt werden. Diese sollten neben dem Energieverbrauch auch weitere relevante Umweltwirkungen wie den Wasserverbrauch systematisch berücksichtigen. Ein weiterentwickeltes Bewertungssystem für KI-Modelle könnte dazu beitragen, Prioritäten für nachhaltige Investitionen zu setzen. Neue Benchmarks sollten daher nicht nur den Wasser- und Energieverbrauch im Rechenzentrumsbetrieb, sondern auch die Umweltauswirkungen der Halbleiterproduktion sowie den langfristigen gesellschaftlichen Nutzen einzelner KI-Anwendungen einbeziehen.

Regulatorische und politische Maßnahmen zur Verbesserung der Transparenz

Die Einführung verbindlicher Offenlegungspflichten könnte durch regulatorische Maßnahmen auf nationaler und internationaler Ebene unterstützt werden. Ähnlich wie bestehende Vorgaben zur CO₂-Bilanzierung könnte eine gesetzliche Verpflichtung zur Wasserberichterstattung für RZ und KI-Anwendungen eingeführt werden. Derzeit sind viele Umweltverpflichtungen freiwillig und unterliegen keiner unabhängigen Prüfung. Dies ermöglicht es Unternehmen, durch selektive Berichterstattung *Greenwashing* zu betreiben oder kritische Daten zurückzuhalten. Ein weiteres Problem ist die fehlende Verfügbarkeit regionaler Verbrauchsdaten. Derzeit existieren nur grobe Durchschnittswerte für den Wasserverbrauch großer Cloud-Anbieter, ohne dass ersichtlich ist, welche Standorte besonders hohe Ressourcenbeanspruchungen haben. Eine differenziertere, standortspezifische Erfassung würde es jedoch ermöglichen, gezielte Maßnahmen für wasserarme Regionen zu ergreifen und nachhaltigere Standortentscheidungen zu treffen.

7.2. FORSCHUNG UND INNOVATIONSFÖRDERUNG FÜR RESSOURCEN- UND ENERGIEEFFIZIENTE KI-TECHNOLOGIEN

Weiterentwicklung nachhaltiger Speichertechnologien

Ein zentraler Forschungsbereich für eine ressourcenschonendere KI-Entwicklung liegt in der Optimierung und Weiterentwicklung von Speichertechnologien. Insbesondere synthetische DNA-Speicherung und optische Speichermedien wie 5D-Kristalle, Glas und Keramiken bieten Ansätze zur Verbesserung der Speicherdichte und Energieeffizienz. Auch der verstärkte Forschungsfokus auf langlebigere SSDs könnte dazu beitragen, den Materialverbrauch in RZ zu reduzieren. SSDs unterliegen heute einem hohen Austauschzyklus, was zu steigenden Umweltbelastungen durch die Rohstoffgewinnung und Entsorgung führt. Forschungsförderung sollte daher gezielt auf die Entwicklung langlebiger und ressourcenschonender Speicherlösungen ausgerichtet werden.

Erforschung alternativer Kühltechnologien

Da Kühlung einen erheblichen Anteil des Wasser- und Energieverbrauchs von RZ ausmacht, ist die Entwicklung innovativer Kühlverfahren essenziell. Während Flüssigkeitskühlung zunehmend etabliert wird, gibt es noch unerschlossene Potenziale, beispielsweise durch den Einsatz von Phasenkühlung, wasserfreien Kühlsystemen oder geothermischen Lösungen. Besondere Aufmerksamkeit sollte dabei der Skalierbarkeit und den ökologischen Auswirkungen neuer Kühltechnologien gewidmet werden. Auch Ansätze wie Unterwasser-RZ bieten mögliche Effizienzgewinne, doch es fehlt bislang an umfassenden Studien zu deren langfristigen Auswirkungen auf marine Ökosysteme. Eine verantwortungsbewusste Forschung sollte daher nicht nur technische Machbarkeit, sondern auch potenzielle Umweltfolgen systematisch untersuchen.

Förderung des Wissenstransfers zwischen Wissenschaft, Industrie und Politik

Neben der reinen Technologieentwicklung bedarf es auch einer besseren Verzahnung zwischen Forschung, Industrie und politischen Entscheidungsträger*innen. Im Expertenworkshop wurde angemerkt, dass Fördermittel für energieeffizientes Rechnen in der Vergangenheit zeitweise stark rückläufig waren, aber aktuell wieder an Bedeutung gewinnen. Ein nachhaltiger, langfristig gesicherter Förderschwerpunkt ist daher entscheidend, um technologische Innovationen voranzutreiben.

Auch Plattformen zur Dissemination von Forschungsergebnissen könnten dabei helfen, Wissen aus der Wissenschaft schneller in die Praxis zu überführen. Besonders wichtig ist es, die Ergebnisse nicht nur auf globale Skalierung auszurichten, sondern auch auf kleinere und mittelständische Rechenzentrumsbetreiber, die oft nicht über dieselben Innovationskapazitäten wie große Anbieter verfügen.

7.3. FÖRDERUNG VON ENERGIE- UND RESSOURCENEFFIZIENTER RECHENZENTRUMSINFRASTRUKTUR UND BETRIEBSWEISEN

Optimierung des Energie-Wasser-Nexus durch systematische Standortplanung

Die Standortwahl von RZ hat maßgeblichen Einfluss auf ihren Wasser- und Energieverbrauch. Bisher werden ökologische Faktoren wie die regionale Wasserverfügbarkeit oder die Integration in nachhaltige Energiequellen oft nicht ausreichend berücksichtigt. In Regionen mit Wasserknappheit kann der Betrieb von wasserintensiven RZ erhebliche Umweltprobleme verursachen, während Standorte mit nachhaltiger Energie- und Wasserversorgung ungenutzt bleiben. Steuerliche Anreize und Subventionen könnten die Ansiedlung an Standorten mit nachhaltiger Energieversorgung fördern und damit die Umweltbelastung reduzieren.

Förderung intelligenter Lastverteilung und Betriebspunkt-Optimierung

Viele RZ betreiben ihre Server durchgehend auf maximaler Leistung, obwohl dies nicht immer erforderlich ist. Durch intelligente Steuerungssysteme könnten erhebliche Mengen an Energie und Wasser eingespart werden. Eine dynamische Lastverteilung, die sich an die tatsächliche Nutzung anpasst, würde ermöglichen, dass Rechenkapazitäten effizienter genutzt und unnötige Energieverbräuche vermieden werden. Hierzu sollten verstärkt *Workload-Management-Tools* entwickelt und regulatorisch verankert werden, um Cloud-Anbieter zur Bereitstellung energieeffizienter Lösungen zu verpflichten. Zudem sollten adaptive Betriebsmodi implementiert werden, die es ermöglichen, dass Server sich je nach Nutzungslast automatisch anpassen, um Überlastung zu vermeiden und unnötigen Stromverbrauch zu reduzieren. Eine optimierte Laststeuerung würde nicht nur die Betriebskosten senken, sondern auch die Umweltbelastung durch RZ erheblich verringern.

7.4. SCHAFFUNG VON ANREIZEN FÜR ENERGIE- UND RESSOURCENSCHONENDE INVESTITIONEN

Politische und wirtschaftliche Anreize für Investitionen in ressourcenschonende Rechenzentren

Der hohe Strompreis in einigen Regionen, insbesondere in Deutschland, kann dazu führen, dass Unternehmen ihre RZ in Länder mit niedrigeren Energiekosten verlagern. Dies hat nicht nur wirtschaftliche Konsequenzen, sondern hemmt auch Innovationen im Bereich energie- und wassereffizienter IT-Infrastrukturen. Um diesem Trend entgegenzuwirken, könnten gezielte wirtschaftliche Anreize geschaffen werden, die es für Unternehmen attraktiver machen, energie- und ressourceneffiziente RZ in Deutschland und anderen Regionen mit nachhaltigen Energiequellen zu betreiben. Eine Möglichkeit wäre die Einführung von Strompreisvorteilen für RZ, die gezielt in energie- und ressourceneffiziente Technologien investieren. Gleichzeitig könnten gezielte Subventionen und Innovationsprogramme dazu beitragen, die Forschung und Entwicklung neuer ressourcenschonender Infrastrukturen zu beschleunigen. Eine weitere Möglichkeit wäre die Einführung staatlicher Förderprogramme, die speziell auf nachhaltige IT-Infrastrukturen zugeschnitten sind. Ebenso könnten regulatorische Rahmenbedingungen so angepasst werden, dass nachhaltige RZ bevorzugt behandelt werden, beispielsweise durch vereinfachte Genehmigungsprozesse oder reduzierte Netzentgelte für Unternehmen, die nachweislich ressourcenschonend arbeiten. Nur durch gezielte wirtschaftliche Anreize können nachhaltige Investitionen gefördert und der Abwanderung von RZ in weniger regulierte Märkte entgegengewirkt werden.

Infrastrukturkosten und langfristige Investitionssicherheit

Der Bau und Betrieb nachhaltiger RZ erfordert erhebliche Investitionen, insbesondere wenn energieeffiziente Kühltechnologien oder nachhaltige Stromquellen integriert werden sollen. Diese hohen Anfangsinvestitionen stellen für viele Unternehmen eine Hürde dar, sodass oftmals kostengünstigere, aber ressourcenintensive Lösungen bevorzugt werden. Um langfristige Investitionen in nachhaltige RZ zu fördern, bedarf es politischer Maßnahmen, die klare Planungssicherheit für Unternehmen schaffen. Dazu gehört die Einführung langfristig stabiler Fördermechanismen sowie regulatorischer Rahmenbedingungen, die nachhaltige Standorte begünstigen. Gleichzeitig sollten umfassende Leitlinien entwickelt werden, die Unternehmen dabei unterstützen, nachhaltige Entscheidungen zu treffen und langfristig von ressourcenschonenden Maßnahmen zu profitieren.

7.5. SICHTBARMACHUNG UND VERBREITUNG VON BEST PRACTICES

Nutzung bestehender Best Practices als Vorbild für nachhaltige Rechenzentren

Bereits heute gibt es zahlreiche RZ, insbesondere im Bereich öffentlicher Auftraggeber, universitärer Forschungseinrichtungen und Hochleistungs-RZ, die mit hohen Effizienzstandards arbeiten und nachhaltige Betriebsweisen erfolgreich umsetzen. Diese Best Practices bieten wertvolle Erfahrungswerte und könnten als Leitlinien für andere Unternehmen und Institutionen dienen. Insbesondere mittelständische Unternehmen stehen häufig vor Herausforderungen, wenn es darum geht, nachhaltige Maßnahmen wirtschaftlich und technisch umzusetzen. Ein strukturiertes Best-Practice-Programm mit Fallstudien, Leitfäden und konkreten Ansprechpartner*innen könnte dazu beitragen, erprobte Lösungen sichtbarer zu machen und deren Übertragbarkeit auf andere Kontexte zu erleichtern.

Förderung einer aktiven Wissens- und Erfahrungsteilung

Um die Verbreitung nachhaltiger Best Practices zu beschleunigen, sollte ein verstärkter Austausch zwischen Vorreiter*innen der Branche, politischen Akteuren und Unternehmen gefördert werden. Derzeit bleiben nachhaltige Praktiken oft auf einzelne Institutionen beschränkt, da es an systematischen Plattformen und Mechanismen für den Wissenstransfer fehlt. Eine gezielte Öffentlichkeitsarbeit und die Bereitstellung von strukturierten Informationsquellen könnten dazu beitragen, das Bewusstsein für bereits umgesetzte Lösungen zu schärfen. Neben klassischen Publikationen wie Leitfäden und Whitepapers wären auch interaktive Formate sinnvoll, beispielsweise Workshops, Exkursionen zu nachhaltigen RZ oder Vernetzungsveranstaltungen zwischen etablierten und neuen Akteuren der Branche.

Verankerung von Best Practices in politischen und wirtschaftlichen Strategien

Damit Best Practices eine breitere Wirkung entfalten, sollten sie aktiv in politische und wirtschaftliche Entscheidungsprozesse integriert werden. Nachhaltige RZ könnten gezielt gefördert werden, indem Best Practices als Grundlage für gesetzliche Vorgaben und Förderprogramme herangezogen werden. Insbesondere im öffentlichen Sektor sollten nachhaltige Kriterien verstärkt in Vergabeprozesse für IT-Infrastruktur einfließen. Dies könnte durch verbindliche Nachhaltigkeitsvorgaben bei Ausschreibungen geschehen, wodurch Unternehmen, die ressourcenschonende Technologien einsetzen, bevorzugt berücksichtigt werden. Gleichzeitig sollten staatliche Fördermittel gezielt in Technologien fließen, die nachweislich zu einer Reduzierung des Wasser- und Energieverbrauchs beitragen.

Sensibilisierung der Gesellschaft für nachhaltige IT-Infrastrukturen

Ein wesentlicher Hemmfaktor für die Umsetzung nachhaltiger Maßnahmen in RZ ist die geringe öffentliche Wahrnehmung der Thematik. Während Energieeffizienz in vielen anderen Branchen zunehmend als relevantes Nachhaltigkeitsthema diskutiert wird, bleibt der ökologische Fußabdruck digitaler Infrastrukturen oft im Hintergrund. Dies führt dazu, dass viele Unternehmen wenig Anreize haben, in ressourcenschonende Technologien zu investieren, da es weder wirtschaftliche Vorteile noch gesellschaftlichen Druck gibt. Eine stärkere Sensibilisierung für die ökologischen Auswirkungen von RZ könnte dazu beitragen, nachhaltige IT-Infrastrukturen als Standard zu etablieren. Öffentlichkeitskampagnen, Berichterstattung über Vorzeigeprojekte sowie eine verstärkte Integration der Thematik in politische Debatten könnten dazu beitragen, die gesellschaftliche Relevanz nachhaltiger RZ zu erhöhen.

7.6. FÖRDERUNG EINER GANZHEITLICHEN BETRACHTUNG DES DIREKTEN UND INDIREKTEN ENERGIE- UND RESSOURCENVERBRAUCHS DURCH KI

Die ökologische Bewertung von KI darf sich nicht nur auf einzelne Umweltfaktoren wie Energie- oder Wasserverbrauch beschränken, sondern muss den gesamten Lebenszyklus von KI-Systemen berücksichtigen. Dazu zählen nicht nur der Betrieb von RZ, sondern auch die vorgelagerte Hardwareproduktion, der Ressourcenverbrauch entlang globaler Lieferketten und die Entsorgung von Geräten. Der Bau neuer Hyperscale-RZ führt zudem zu Flächenversiegelung und weiteren Umweltfolgen, die bislang oft vernachlässigt werden. Ein isolierter Fokus auf bestimmte Umweltaspekte birgt Risiken durch negative Trade-offs. So kann eine erhöhte Energieeffizienz zu einem höheren Wasserverbrauch führen, während wassersparende Kühltechnologien oft mehr Strom benötigen. Besonders problematisch ist die Herstellung von Hochleistungschips, da sie große Mengen an seltenen Rohstoffen erfordert, deren Abbau mit ökologischen und sozialen Folgen verbunden ist. Eine nachhaltige Strategie muss daher Umwelt- und Sozialfaktoren gemeinsam betrachten.

Indirekte Auswirkungen und Rebound-Effekte von KI

Neben den direkten Umweltauswirkungen müssen auch indirekte Effekte berücksichtigt werden. KI kann zwar zur Effizienzsteigerung in manchen Bereichen beitragen, kann aber auch Rebound-Effekte erzeugen, wenn ressourcenschonende Prozesse durch eine wachsende Nachfrage nach datenintensiven Anwendungen kompensiert werden. Eine nachhaltige Entwicklung von KI sollte daher nicht nur den Ressourcenverbrauch einzelner Systeme, sondern deren langfristige gesellschaftliche und wirtschaftliche Wechselwirkungen einbeziehen.

Soziale Verantwortung und globale Lieferketten

Die Produktion von KI-Hardware ist eng mit internationalen Lieferketten verknüpft und stark von globalen Machtverhältnissen geprägt. Während Hochleistungs-RZ meist in wirtschaftlich attraktiven Regionen angesiedelt werden, entstehen Umweltbelastungen durch Rohstoffgewinnung, Chipproduktion und Entsorgung entlang der gesamten Wertschöpfungskette. Oft profitieren lokale Gemeinschaften wirtschaftlich nur begrenzt von diesen Entwicklungen. Nachhaltige Strategien müssen daher soziale und ökologische Kriterien gleichermaßen berücksichtigen, um faire und verantwortungsvolle Produktions- und Nutzungsbedingungen zu gewährleisten.

Eine umfassende Regulierung sollte nicht nur Mindeststandards für den Wasser- und Energieverbrauch von RZ setzen, sondern auch Anforderungen an nachhaltige Rohstoffbeschaffung, Kreislaufwirtschaft und eine gerechte Verteilung der Wertschöpfung entlang globaler Lieferketten formulieren. Nur durch einen solchen ganzheitlichen und interdisziplinären Ansatz lassen sich langfristig tragfähige Lösungen für eine ressourcenschonende und sozial gerechte Entwicklung von KI-Infrastrukturen schaffen.

7.7. FÖRDERUNG RESSOURCENSCHONENDER HARDWARE UND STÄRKUNG DER KREISLAUFWIRTSCHAFT

Die Herstellung von Hochleistungschips und Serverkomponenten für KI ist extrem ressourcenintensiv und erfordert große Mengen an Wasser, Energie und seltenen Rohstoffen. Viele dieser Materialien sind begrenzt verfügbar und werden unter umweltschädlichen sowie sozial problematischen Bedingungen abgebaut. Aktuell gibt es nur begrenzte Mechanismen zur Rohstoffrückgewinnung oder zur Verlängerung der Hardware-Lebensdauer durch nachhaltiges Design.

Ein zentrales Problem ist die begrenzte Nutzungsdauer von KI-Hardware, insbesondere von Speichermedien und Recheneinheiten. SSDs und Chips werden oft bereits nach wenigen Jahren ausgetauscht, was den Ressourcenverbrauch und Elektroschrott weiter erhöht. Zudem führt der rasante Fortschritt in der KI-Entwicklung dazu, dass ältere Hardwaregenerationen frühzeitig obsolet werden. Notwendig sind daher neue Designstandards für langlebigere, reparaturfreundlichere und recycelbare Hardware, kombiniert mit wirtschaftlichen Anreizen für Unternehmen, die ressourcenschonende Hardware nutzen und weiterentwickeln.

Neben der Hardwareproduktion sollte auch die Standortwahl von RZ nachhaltiger gestaltet werden. Der Bau neuer RZ erfordert große Mengen an Baumaterialien wie Beton und Stahl, deren Herstellung hohe CO₂-Emissionen verursacht. Dennoch entstehen neue RZ häufig isoliert in Industriegebieten, anstatt vorhandene Infrastrukturen zu nutzen. Eine gezieltere Standortpolitik könnte dazu beitragen, Materialeinsatz und Umweltbelastungen zu reduzieren.

Ein weiteres Problem ist die mangelnde Transparenz der Hardwareproduktion, insbesondere in der Halbleiterindustrie. Marktführer wie Nvidia veröffentlichen kaum Informationen über den Wasser- und Energieverbrauch ihrer Chipproduktion. Ohne verbindliche Offenlegungspflichten bleibt ein großer Teil des Ressourcenverbrauchs unkontrolliert. Daher sollten verpflichtende Berichterstattungsstandards eingeführt

werden, die den gesamten Lebenszyklus von Hardware – von der Rohstoffgewinnung über die Produktion bis hin zur Entsorgung – erfassen.

Um die Kreislaufwirtschaft in der KI-Hardware zu stärken, sollten gezielte Maßnahmen zur Förderung von Recycling-Technologien und Second-Life-Nutzung entwickelt werden. RZ könnten verpflichtet werden, ausgemusterte Hardware durch Refurbishing-Programme weiterzuverwenden oder durch Rücknahme- und Wiederverwertungsprogramme nachhaltig zu nutzen. Eine systematische Kreislaufwirtschaft könnte nicht nur den Materialverbrauch senken, sondern auch wirtschaftliche Vorteile schaffen, indem Kosten für neue Hardware reduziert und Wertschöpfungspotenziale in der Wiederverwertung erschlossen werden.

Eine nachhaltigere Hardware-Nutzung erfordert eine Kombination aus regulativen Maßnahmen, wirtschaftlichen Anreizen und mehr Transparenz in den Lieferketten. Nur durch eine konsequente Förderung ressourcenschonender Hardware und eine Stärkung der Kreislaufwirtschaft kann die Umweltbelastung der KI-Industrie langfristig reduziert werden.

8. Diskussion & Fazit

8.1. KRITISCHE REFLEXION DER ERGEBNISSE & LIMITATIONEN

Die vorliegende Studie analysiert den Wasserverbrauch von RZ im Kontext von KI und identifiziert zentrale Einflussfaktoren, Herausforderungen sowie Handlungsempfehlungen. Trotz einer breit angelegten Methodik, die eine systematische Literaturrecherche, Experteninterviews und einen Workshop kombiniert, unterliegt die Untersuchung verschiedenen Limitationen, die die Aussagekraft und Übertragbarkeit der Ergebnisse einschränken.

Begrenzte Datenlage und Verfügbarkeit von Primärdaten

Ein zentrales Problem der Studie ist die eingeschränkte Verfügbarkeit detaillierter Primärdaten zum Wasserverbrauch von RZ. Während einige Unternehmen Nachhaltigkeitsberichte veröffentlichen, fehlt es an standardisierten Messmethoden und einheitlichen Metriken, um den Wasserverbrauch verlässlich zu quantifizieren. Zudem sind viele Daten nur auf aggregierter Ebene verfügbar und lassen keine detaillierte Differenzierung nach Kühltechnologien oder regionalen Bedingungen zu. Da die Studie auf öffentlich zugänglichen Daten sowie Expertenaussagen basiert, können Verzerrungen durch selektive Berichterstattung nicht ausgeschlossen werden.

Methodische Herausforderungen bei der Datenerhebung

Die Kombination aus Literaturrecherche, Experteninterviews und Workshop-Methodik bietet eine breite Perspektive auf das Forschungsthema, weist jedoch auch methodische Grenzen auf:

- **Systematische Literaturrecherche:** Der Forschungsstand zur Quantifizierung des Wasserverbrauchs von KI-RZ ist begrenzt, weshalb einige Fragestellungen nicht durch belastbare empirische Studien beantwortet werden konnten. Zudem variieren die verwendeten Methoden zur Schätzung des Wasserverbrauchs zwischen den identifizierten Quellen, wodurch keine direkte Vergleichbarkeit gewährleistet ist.
- **Experteninterviews:** Die Interviews deckten verschiedene Perspektiven aus Wissenschaft und Praxis ab, waren jedoch in ihrer Anzahl begrenzt. Dies schränkt die Generalisierbarkeit der Aussagen ein. Zudem basieren die Expertenmeinungen auf individuellen Erfahrungen und Einschätzungen, was zu einer gewissen Subjektivität führt.
- **Expertenworkshop:** Der Workshop fand in Präsenz statt, wodurch ausschließlich deutsche Expert*innen ihre Einschätzungen einbringen konnten. Dies führte zu einer begrenzten geografischen Perspektive und möglicherweise zu einer Verzerrung in der Betrachtung internationaler Best Practices. Zudem war es aufgrund der zeitlichen Begrenzung des Workshops nicht möglich, alle relevanten Themen in ausreichender Tiefe zu behandeln oder eine umfassende Abdeckung internationaler Entwicklungen sicherzustellen.

Eingeschränkte Übertragbarkeit der Ergebnisse

Die Studienergebnisse sind stark kontextabhängig und lassen sich nicht uneingeschränkt auf alle RZ oder geografischen Regionen übertragen. Der Wasserverbrauch hängt von zahlreichen Faktoren ab, darunter Standortbedingungen, Energiequellen und spezifische Hardwareanforderungen. Während allgemeine Trends und Einflussfaktoren identifiziert wurden, sind präzise Hochrechnungen für den globalen Wasserverbrauch von KI-gestützten RZ aufgrund der heterogenen Datenlage nicht möglich. Auch konnte die Studie nicht sämtliche Best Practices umfassend analysieren. Insbesondere fehlte es an gezielten empirischen Untersuchungen zur Verbreitung neuer technischer Innovationen, die den Wasserverbrauch von RZ beeinflussen könnten.

Dynamik der KI-Entwicklung und technologische Veränderungen

Ein weiterer wesentlicher Punkt ist die rasante Entwicklung der KI-Technologien. Das Feld der KI-Nutzung und -Entwicklung wächst und verändert sich kontinuierlich, sodass einige der aktuellen Erkenntnisse innerhalb kurzer Zeit durch neue technische Fortschritte oder veränderte regulatorische Rahmenbedingungen überholt sein könnten. Da die Studie auf bestehenden Entwicklungen und bereits veröffentlichten Daten basiert, konnten zukünftige technologische Innovationen und deren mögliche Auswirkungen auf den Wasserverbrauch nicht vollständig berücksichtigt werden.

Trotz dieser Limitationen liefert die Studie wertvolle Erkenntnisse über die zentralen Einflussfaktoren des Wasserverbrauchs von KI-RZ. Sie verdeutlicht, dass eine Kombination aus technischen, wirtschaftlichen und politischen Maßnahmen erforderlich ist, um den Wasserverbrauch nachhaltig zu reduzieren. Gleichzeitig zeigen die begrenzte Datenlage und die methodische Herangehensweise, dass weiterführende Forschung notwendig ist. Insbesondere sind gezielte empirische Studien erforderlich, um die tatsächliche Verbreitung neuer Technologien und deren Einfluss auf den Wasserverbrauch genauer zu untersuchen. Zudem wäre eine stärkere internationale Perspektive wünschenswert, um regionale Unterschiede in der Wassernutzung besser zu erfassen und eine fundiertere Grundlage für politische und wirtschaftliche Entscheidungen zu schaffen.

8.2. FAZIT & AUSBLICK

Die Ergebnisse dieser Studie zeigen, dass der Wasserverbrauch von KI-RZ eine zunehmend relevante Herausforderung darstellt. Gleichzeitig existieren bereits zahlreiche technologische Lösungen, die den Wasserverbrauch innerhalb von RZ reduzieren können. Allerdings reicht es nicht aus, einzelne Maßnahmen zu betrachten – eine ganzheitliche Perspektive ist erforderlich, die den gesamten Lebenszyklus von KI und der zugehörigen Infrastruktur berücksichtigt. Dies schließt nicht nur den direkten Wasserverbrauch in RZ ein, sondern auch den Wasserbedarf für die Stromerzeugung und insbesondere die ressourcenintensive Herstellung von Hardware. Ein zentrales Problem bleibt der Mangel an öffentlich zugänglichen und transparent erhobenen Daten, die eine belastbare Quantifizierung der Wassernutzung und eine systematische Bewertung von Best Practices ermöglichen würden.

Handlungsfelder zur Reduktion des Wasserverbrauchs in KI-Rechenzentren

Auf Basis der Analyse wurden mehrere zentrale Handlungsfelder identifiziert, die eine nachhaltige Entwicklung der Rechenzentrumsbranche fördern können:

- **Einführung verbindlicher Transparenzvorgaben und Etablierung von Standards:** Für eine nachhaltige KI-Infrastruktur sind verbindliche Transparenzvorgaben zum Wasser- und Energieverbrauch von RZ erforderlich. Standardisierte Monitoring-Systeme und neue Metriken sollten die Umweltbilanz vergleichbar machen und neben Energie- auch Wasserverbräuche erfassen. Ergänzend dazu braucht es Nachhaltigkeitsindikatoren, die ökologischen Auswirkungen und den gesellschaftlichen Nutzen von KI-Modellen berücksichtigen.
- **Forschung und Innovationsförderung für ressourcen- und energieeffiziente KI-Technologien:** Neue Speichertechnologien, effiziente Kühlverfahren und nachhaltige Rechenarchitekturen bieten Potenzial zur Reduktion des Ressourcenverbrauchs. Zudem sollte die Umweltbilanz generativer KI-Modelle intensiver erforscht werden. Eine stärkere Vernetzung von Wissenschaft, Industrie und Politik kann den Transfer nachhaltiger Technologien beschleunigen.

- **Förderung ressourcenschonender Rechenzentren:** Die Standortwahl beeinflusst den Wasser- und Energieverbrauch erheblich. Steuerliche Anreize und regulatorische Vorgaben könnten nachhaltige Standorte fördern. Effiziente Lastverteilung, adaptive Betriebsmodi und alternative Kühltechnologien sollten verstärkt eingesetzt werden, um Umweltbelastungen zu minimieren.
- **Anreize für nachhaltige Investitionen:** Fehlende wirtschaftliche Anreize hemmen Investitionen in ressourcenschonende RZ. Förderprogramme, steuerliche Vergünstigungen und regulatorische Anreize könnten nachhaltige Infrastrukturmaßnahmen attraktiver machen und den Umstieg auf umweltschonende Technologien beschleunigen.
- **Verbreitung von Best Practices:** Nachhaltige Lösungen sind vorhanden, werden jedoch zu wenig verbreitet. Ein strukturiertes Best-Practice-Programm könnte erfolgreiche Modelle sichtbarer machen und Unternehmen bei der Umsetzung unterstützen. Öffentlichkeitsarbeit und regulatorische Vorgaben sollten nachhaltige Betriebsweisen stärker verankern.
- **Ganzheitliche Betrachtung der Umweltwirkungen von KI:** Die Umweltbewertung von KI muss den gesamten Lebenszyklus einbeziehen – von der Hardwareproduktion bis zur Entsorgung. Rebound-Effekte und negative Trade-offs zwischen Energie- und Wasserverbrauch müssen vermieden werden. Eine umfassende Regulierung sollte ökologische und soziale Kriterien gleichermaßen berücksichtigen.
- **Förderung ressourcenschonender Hardware und Kreislaufwirtschaft:** KI-Hardware ist ressourcenintensiv und oft nur kurz nutzbar. Langlebigere, reparaturfreundliche Designs sowie verpflichtende Transparenzvorgaben für Produktionsprozesse könnten Umweltbelastungen senken. Zudem sollten Recycling- und Wiederverwertungsprogramme gezielt gefördert werden.

Ausblick

Angesichts der rasanten Entwicklung von KI-Technologien und der damit verbundenen Infrastruktur ist davon auszugehen, dass der Wasserverbrauch von RZ weiter steigen wird. Um nachhaltige Lösungen zu entwickeln, sind daher weiterführende Forschungen erforderlich. Insbesondere sollten folgende Aspekte in zukünftigen Studien näher betrachtet werden:

- **Genauere Erfassung und Berichterstattung des Wasserverbrauchs:** Internationale Standards zur Messung und Offenlegung des Wasserverbrauchs von RZ könnten dazu beitragen, belastbare Vergleichswerte zu schaffen und Unternehmen zur Einführung wassersparender Technologien zu motivieren.
- **Systematische Bewertung neuer Technologien:** Der Einsatz innovativer Kühltechnologien und effizienterer Hardware-Architekturen sollte gezielt erforscht werden, um das Potenzial zur Wassereinsparung besser quantifizieren zu können.
- **Berücksichtigung der sozioökonomischen Auswirkungen der Rechenzentrumsentwicklung:** Während die vorliegende Studie primär technische und wirtschaftliche Aspekte untersucht hat, wurden in den Experteninterviews und dem Workshop auch soziale und ethische Fragen aufgeworfen. Die gesellschaftlichen und ökologischen Auswirkungen von RZ in verschiedenen Regionen sollten systematisch untersucht werden. Dies schließt die Wechselwirkungen zwischen wirtschaftlichen Interessen, Wasserknappheit und lokaler Widerstandsfähigkeit gegenüber Klimawandel ein.

Insgesamt zeigt die Studie, dass der Wasserverbrauch von KI-RZ ein zunehmend wichtiges Thema ist, das technische, regulatorische und gesellschaftliche Maßnahmen erfordert. Während bereits Lösungen existieren, ist eine umfassendere Betrachtung des gesamten Lebenszyklus von KI erforderlich, um nachhaltige Strategien für die Zukunft zu entwickeln.

Literaturverzeichnis

- 2CRSi. (o. J.). *Air Cooling System: Why Is It Essential for Data Centers?* <https://2crsi.com/air-cooling>
- Al Kez, D., Foley, A. M., Laverty, D., Del Rio, D. F., & Sovacool, B. (2022). Exploring the sustainability challenges facing digitalization and internet data centers. *Journal of Cleaner Production*, 371, 133633. <https://doi.org/10.1016/j.jclepro.2022.133633>
- Beghi, A., Mana, G. D., Lionello, M., Rampazzo, M., & Sisti, E. (2017). Energy-efficient operation of an indirect adiabatic cooling system for data centers. *2017 American Control Conference (ACC)*, 1984–1989. <https://doi.org/10.23919/ACC.2017.7963243>
- Berthelot, A., Caron, E., Jay, M., & Lefèvre, L. (2024). *Estimating the environmental impact of Generative-AI services using an LCA-based methodology*. 1. <https://inria.hal.science/hal-04346102>
- Bieser, J. C. T., & Hilty, L. M. (2018). *Indirect Effects of the Digital Transformation on Environmental Sustainability: Methodological Challenges in Assessing the Greenhouse Gas Abatement Potential of ICT*. 68–53. <https://doi.org/10.29007/lx7q>
- Bitkom. (2024a). KI und Nachhaltigkeit: Chancen und Herausforderungen für eine nachhaltige Digitalisierung. *Bitkom e. V.*
- Bitkom. (2024b). *Rechenzentren in Deutschland*. <https://www.bitkom.org/Bitkom/Publikationen/Studie-Rechenzentren-in-Deutschland>
- Cáceres, C. R., Sandberg, M., & Sotoca, A. (2024). Planning data center locations in Swedish municipalities. A comparative case study of Luleå and Stockholm. *Cities*, 150, 105063. <https://doi.org/10.1016/j.cities.2024.105063>
- CENELEC. (2022). *EN 50600-4-9:2022 Information technology—Data centre facilities and infrastructures—Part 4-9: Water Usage Effectiveness*. <https://standards.iteh.ai/catalog/standards/clc/a6277500-02b4-49de-9647-318d6918e9f1/en-50600-4-9-2022>
- Chang, Q., Huang, Y., Liu, K., Xu, X., Zhao, Y., & Pan, S. (2024). Optimization Control Strategies and Evaluation Metrics of Cooling Systems in Data Centers: A Review. *Sustainability*, 16(16), 7222. <https://doi.org/10.3390/su16167222>
- Copley, M. (2022, August 30). Data centers, backbone of the digital economy, face water scarcity and climate risk. *npr*. <https://www.npr.org/2022/08/30/1119938708/data-centers-backbone-of-the-digital-economy-face-water-scarcity-and-climate-ris>

Data4. (2024, August 10). *Liquid cooling in data centers: A revolution in energy efficiency*.

<https://www.data4group.com/en/datacenter-dictionary/liquid-cooling-in-data-centers-a-revolution-in-energy-efficiency/>

Eabel. (2024, September 5). Flüssigkeitskühlung vs. Luftkühlung im Rechenzentrum. *E-abel*.

<https://www.eabel.com/de/flussigkeitskuhlung-vs-luftkuhlung-im-rechenzentrum/>

Farfan, J., & Lohrmann, A. (2023). Gone with the clouds: Estimating the electricity and water footprint of digital data services in Europe. *Energy Conversion and Management*, 290, 117225.

<https://doi.org/10.1016/j.enconman.2023.117225>

Frost, K., & Hua, I. (2019). “Quantifying spatiotemporal impacts of the interaction of water scarcity and water use by the global semiconductor manufacturing industry”. *Water Resources and Industry*, 22, 100115.

<https://doi.org/10.1016/j.wri.2019.100115>

Google. (2023). *2023 Environmental Report*.

Gröger, J., Köhler, A., Naumann, S., Filler, A., Guldner, A., Kern, E., Lorenz, H., & Maksimov, Y. (2018). *Entwicklung und Anwendung von Bewertungsgrundlagen für ressourceneffiziente Software unter Berücksichtigung bestehender Methodik*.

Habibi Khalaj, A., & Halgamuge, S. K. (2017). A Review on efficient thermal management of air- and liquid-cooled data centers: From chip to the cooling system. *Applied Energy*, 205, 1165–1188.

<https://doi.org/10.1016/j.apenergy.2017.08.037>

HDR. (2023, März 24). *Direct-To-Chip Liquid Cooling*. <https://www.hdrinc.com/insights/direct-chip-liquid-cooling>

He, C., Liu, Z., Wu, J., Pan, X., Fang, Z., Li, J., & Bryan, B. A. (2021). Future global urban water scarcity and potential solutions. *Nature Communications*, 12(1). Crossref. <https://doi.org/10.1038/s41467-021-25026-3>

Hernandez, D., & Brown, T. B. (2020). *Measuring the Algorithmic Efficiency of Neural Networks* (Version 1). arXiv.

<https://doi.org/10.48550/ARXIV.2005.04305>

Hilty, L. M., & Aebischer, B. (2015). ICT for Sustainability: An Emerging Research Field. In L. M. Hilty & B. Aebischer (Hrsg.), *ICT Innovations for Sustainability* (Bd. 310, S. 3–36). Springer International Publishing.

https://doi.org/10.1007/978-3-319-09228-7_1

Huang, B., Zheng, Z., Lu, G., & Zhai, X. (2022). Design and experimental investigation of a PCM based cooling storage unit for emergency cooling in data center. *Energy and Buildings*, 259, 111871.

<https://doi.org/10.1016/j.enbuild.2022.111871>

- IPCC. (2023). *Climate Change 2022 – Impacts, Adaptation and Vulnerability: Working Group II Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (1. Aufl.). Cambridge University Press. <https://doi.org/10.1017/9781009325844>
- Islam, M. A., Ren, S., Quan, G., Shakir, M. Z., & Vasilakos, A. V. (2017). Water-Constrained Geographic Load Balancing in Data Centers. *IEEE Transactions on Cloud Computing*, 5(2), 208–220. IEEE Transactions on Cloud Computing. <https://doi.org/10.1109/TCC.2015.2453982>
- Jerléus, K., Ibrahim, M. A., & Augustsson, A. (2024). Environmental footprints of the data center service sector in Sweden. *Heliyon*, 10(11), e31290. <https://doi.org/10.1016/j.heliyon.2024.e31290>
- Kaack, L. H., Donti, P. L., Strubell, E., Kamiya, G., Creutzig, F., & Rolnick, D. (2022). Aligning artificial intelligence with climate change mitigation. *Nature Climate Change*, 12(6), 518–527. <https://doi.org/10.1038/s41558-022-01377-7>
- Karimi, L., Yacuel, L., Johnson, J. D., Ashby, J., Green, M., Renner, M., Bergman, A., Norwood, R., & Hickenbottom, K. L. (2022). Water-energy tradeoffs in data centers: A case study in hot-arid climates. *Resources, Conservation and Recycling*, 181, 106194. <https://doi.org/10.1016/j.resconrec.2022.106194>
- Kirvan, P. (2022). Adiabatic cooling. *TechTarget*. <https://www.techtarget.com/whatis/definition/adiabatic-cooling>
- Langley, H. (2023, Juli 25). Google's water use is soaring. AI is only going to make it worse. *Business Insider*. <https://www.businessinsider.com/google-water-use-soaring-ai-make-it-worse-data-centers-2023-7>
- Lei, N., Lu, J., Cheng, Z., Cao, Z., Shehabi, A., & Masanet, E. (2023). Geospatial assessment of water footprints for hyperscale data centers in the United States. *Journal of Physics: Conference Series*, 2600(17), 172003. <https://doi.org/10.1088/1742-6596/2600/17/172003>
- Lei, N., Lu, J., Shehabi, A., & Masanet, E. (2025). The water use of data center workloads: A review and assessment of key determinants. *Resources, Conservation and Recycling*, 219, 108310. <https://doi.org/10.1016/j.resconrec.2025.108310>
- 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. *Resources, Conservation and Recycling*, 182, 106323. <https://doi.org/10.1016/j.resconrec.2022.106323>

- Leiserson, C. E., Thompson, N. C., Emer, J. S., Kuszmaul, B. C., Lampson, B. W., Sanchez, D., & Schardl, T. B. (2020). There's plenty of room at the Top: What will drive computer performance after Moore's law? *Science*, 368(6495), eaam9744. <https://doi.org/10.1126/science.aam9744>
- Lepawsky, J. (2024). Climate change induced water stress and future semiconductor supply chain risk. *iScience*, 27(2), 108791. Crossref. <https://doi.org/10.1016/j.isci.2024.108791>
- Li, P., Yang, J., Islam, M. A., & Ren, S. (2023). *Making AI Less „Thirsty“: Uncovering and Addressing the Secret Water Footprint of AI Models* (arXiv:2304.03271). arXiv. <http://arxiv.org/abs/2304.03271>
- Ligozat, A.-L., Lefèvre, J., Bugeau, A., & Combaz, J. (2022). *Unraveling the Hidden Environmental Impacts of AI Solutions for Environment* (arXiv:2110.11822). arXiv. <http://arxiv.org/abs/2110.11822>
- Livingstone, G. (2023, Juli 11). 'It's pillage': Thirsty Uruguayans decry Google's plan to exploit water supply. *The Guardian*. <https://www.theguardian.com/world/2023/jul/11/uruguay-drought-water-google-data-center>
- Luccioni, A. S., & Hernandez-Garcia, A. (2023). *Counting Carbon: A Survey of Factors Influencing the Emissions of Machine Learning* (arXiv:2302.08476). arXiv. <http://arxiv.org/abs/2302.08476>
- Luccioni, A. S., Jernite, Y., & Strubell, E. (2024). Power Hungry Processing: Watts Driving the Cost of AI Deployment? *The 2024 ACM Conference on Fairness, Accountability, and Transparency*, 85–99. <https://doi.org/10.1145/3630106.3658542>
- Luccioni, A. S., Viguier, S., & Ligozat, A.-L. (2022). *Estimating the Carbon Footprint of BLOOM, a 176B Parameter Language Model* (arXiv:2211.02001). arXiv. <http://arxiv.org/abs/2211.02001>
- McKinsey. (2024, Oktober 29). *AI power: Expanding data center capacity to meet growing demand*. <https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/ai-power-expanding-data-center-capacity-to-meet-growing-demand#/>
- Microsoft. (2022). 2022 Environmental Sustainability Report. *Global Sustainability*.
- Microsoft. (2023). *Modern datacenter cooling*. https://datacenters.microsoft.com/wp-content/uploads/2023/05/Azure_Modern-Datacenter-Cooling_Infographic.pdf
- Monserrate, S. G. (2022). The Cloud Is Material: On the Environmental Impacts of Computation and Data Storage. *MIT Case Studies in Social and Ethical Responsibilities of Computing*, Winter 2022. <https://doi.org/10.21428/2c646de5.031d4553>

- Munia, H. A., Guillaume, J. H. A., Wada, Y., Veldkamp, T., Virkki, V., & Kummu, M. (2020). Future Transboundary Water Stress and Its Drivers Under Climate Change: A Global Study. *Earth's Future*, 8(7). Crossref. <https://doi.org/10.1029/2019ef001321>
- Mytton, D. (2021). Author Correction: Data centre water consumption. *Npj Clean Water*, 4(1), 1–1. <https://doi.org/10.1038/s41545-021-00140-3>
- 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. <https://doi.org/10.3390/en17153615>
- Pegus, P., II, Varghese, B., Guo, T., Irwin, D., Shenoy, P., Mahanti, A., Culbert, J., Goodhue, J., & Hill, C. (2016, März 12). Analyzing the Efficiency of a Green University Data Center. *Proceedings of the 7th ACM/SPEC on International Conference on Performance Engineering*. Crossref. <https://doi.org/10.1145/2851553.2851557>
- Reig, P. (2013). *What's the Difference Between Water Use and Water Consumption?* <https://www.wri.org/insights/whats-difference-between-water-use-and-water-consumption>
- Ristic, B., Madani, K., & Makuch, Z. (2015). The Water Footprint of Data Centers. *Sustainability*, 7(8), 11260–11284. <https://doi.org/10.3390/su70811260>
- Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., Ross, A. S., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, A., Luccioni, A. S., Maharaj, T., Sherwin, E. D., Mukkavilli, S. K., Kording, K. P., Gomes, C. P., Ng, A. Y., Hassabis, D., Platt, J. C., ... Bengio, Y. (2023). Tackling Climate Change with Machine Learning. *ACM Computing Surveys*, 55(2), 1–96. <https://doi.org/10.1145/3485128>
- Ruberti, M. (2023). The chip manufacturing industry: Environmental impacts and eco-efficiency analysis. *Science of The Total Environment*, 858, 159873. Crossref. <https://doi.org/10.1016/j.scitotenv.2022.159873>
- Rüdiger, A. (2019, Juli 23). Was ist und wie funktioniert Adiabatik? *Data Center Insider*. <https://www.datacenter-insider.de/was-ist-und-wie-funktioniert-adiabatik-a-845438/>
- Sharma, P., Pegus II, P., Irwin, D., Shenoy, P., Goodhue, J., & Culbert, J. (2017). Design and Operational Analysis of a Green Data Center. *IEEE Internet Computing*, 21(4), 16–24. Crossref. <https://doi.org/10.1109/mic.2017.2911421>
- Shehabi, A., Smith, S., Sartor, D., Brown, R., Herrlin, M., Koomey, J., Masanet, E., Horner, N., Azevedo, I., & Lintner, W. (2016). *United States Data Center Energy Usage Report* (LBNL--1005775, 1372902; S. LBNL--1005775, 1372902). <https://doi.org/10.2172/1372902>

- Shumba, N., Tshekiso, O., Li, P., Fanti, G., & Ren, S. (2024). *A Water Efficiency Dataset for African Data Centers* (Version 2). arXiv. <https://doi.org/10.48550/ARXIV.2412.03716>
- Siddik, M. A. B., Shehabi, A., & Marston, L. (2021). The environmental footprint of data centers in the United States. *Environmental Research Letters*, 16(6), 064017. <https://doi.org/10.1088/1748-9326/abfba1>
- Silva-Llanca, L., Ponce, C., Bermúdez, E., Martínez, D., Díaz, A. J., & Aguirre, F. (2023). Improving energy and water consumption of a data center via air free-cooling economization: The effect weather on its performance. *Energy Conversion and Management*, 292, 117344. <https://doi.org/10.1016/j.enconman.2023.117344>
- Statista. (2024). *Anzahl der Hyperscale-Rechenzentren weltweit von 2015 bis 2023*. <https://www.statista.com/statistics/633826/worldwide-hyperscale-data-center-numbers>
- Strubell, E., Ganesh, A., & McCallum, A. (2019). *Energy and Policy Considerations for Deep Learning in NLP* (arXiv:1906.02243). arXiv. <http://arxiv.org/abs/1906.02243>
- Tagesschau. (2023). *KI und Energieverbrauch: Die unsichtbare Umweltbelastung von Künstlicher Intelligenz*. <https://www.tagesschau.de/wissen/klima/ki-energieverbrauch-100.html>
- The Green Grid. (2016). *ISO/IEC 30134-2:2016 Information technology—Data centres—Key performance indicators—Part 2: Power usage effectiveness (PUE)*. https://en.wikipedia.org/wiki/Power_usage_effectiveness
- top500.org. (o. J.). *About*. <https://top500.org/project/>
- UBA. (2024). *Rechenzentrumsklimatisierung*. <https://www.umweltbundesamt.de/themen/klima-energie/fluorierte-treibhausgase-fckw/natuerliche-kaeltemittel-in-stationaeren-anlagen/anwendungen/rechenzentrumsklimatisierung>
- United Nations. (2023). *Imminent risk of a global water crisis, warns the UN World Water Development Report 2023*. UNESCO. <https://www.unesco.org/en/articles/imminent-risk-global-water-crisis-warns-un-world-water-development-report-2023>
- United Nations. (2024). *The United Nations World Water Development Report 2024: Water for Prosperity and Peace*. UNESCO.
- Verma, P., & Tan, S. (2024, September 18). A bottle of water per email: The hidden environmental costs of using AI chatbots. *The Washington Post*. <https://www.washingtonpost.com/technology/2024/09/18/energy-ai-use-electricity-water-data-centers/>

- VISTECH. (2021). *Adiabatic Coolers Explained*. <https://www.vistechcooling.co.uk/knowledge-centre/articles/adiabatic-coolers-explained/#:~:text=In%20summary%2C%20Evaporative%20cooling%2C%20as,air%20passing%20into%20the%20cooler.>
- Wang, Q., Huang, N., Cai, H., Chen, X., & Wu, Y. (2023). Water strategies and practices for sustainable development in the semi conductor industry. *Water Cycle*, 4, 12–16. Crossref. <https://doi.org/10.1016/j.watcyc.2022.12.001>
- WRI. (2023). *25 Countries, Housing One-quarter of the Population, Face Extremely High Water Stress*. World Resource Institut. <https://www.wri.org/insights/highest-water-stressed-countries>
- WRI. (n.y.). *Water Risk Atlas*. https://www.wri.org/applications/aqueduct/water-risk-atlas/#/?advanced=false&basemap=hydro&indicator=1b4f2592-09fd-4ac4-afcd-5a0a9a63617b&lat=47.42808726171425&lng=6.020507812500001&mapMode=view&month=5&opacity=0.5&ponderation=DEF&predefined=false&projection=absolute&scenario=business_as_usual&scope=future&threshold&timeScale=annual&year=2030&zoom=5
- Zhang, J., Zhang, R., Schmidt, G., & Guo, B. (2019). *RP-1755—Impact of Gaseous Contamination and High Humidity on the Reliable Operation of Information Technology Equipment in Data Centers*. https://store accuristech.com/standards/rp-1755-impact-of-gaseous-contamination-and-high-humidity-on-the-reliable-operation-of-information-technology-equipment-in-data-centers?product_id=2089942
- Zhang, Q., Meng, Z., Hong, X., Zhan, Y., Liu, J., Dong, J., Bai, T., Niu, J., & Deen, M. J. (2021). A survey on data center cooling systems: Technology, power consumption modeling and control strategy optimization. *Journal of Systems Architecture*, 119, 102253. <https://doi.org/10.1016/j.sysarc.2021.102253>
- Zhang, Y., Peng, P., Oh, H., Jin, W., Winick, J., Sickinger, D., Atkinson, T., Sartor, D., Lei, N., Sauter, I., & Dobson, P. F. (2025). Addressing Data Center Cooling Needs Through the Use of Reservoir Thermal Energy Storage Systems. *Journal of Clean Energy and Energy Storage*, 02, 2550002. <https://doi.org/10.1142/S2811034X25500029>
- Zhang, Y., Yanqi Zhao, Siyuan Dai, Binjian Nie, Hongkun Ma, Jianming Li, Qi Miao, Yi Jin, Linghua Tan, & Yulong Ding. (2022). Cooling technologies for data centres and telecommunication base stations – A comprehensive review. *Journal of Cleaner Production*.

Zheng, Z., Huang, B., Lu, G., & Zhai, X. (2022). Design and optimization of an air-based phase change cold storage unit through cascaded construction for emergency cooling in IDC. *Energy*, 241, 122897.

<https://doi.org/10.1016/j.energy.2021.122897>

Zuccon, G., Scells, H., & Zhuang, S. (2023). Beyond CO2 Emissions: The Overlooked Impact of Water Consumption of Information Retrieval Models. *Proceedings of the 2023 ACM SIGIR International Conference on Theory of Information Retrieval*, 283–289. <https://doi.org/10.1145/3578337.3605121>

Danksagung

Für die wertvollen Beiträge zur inhaltlichen Gestaltung dieser Studie gilt unser Dank allen Teilnehmenden der Interviews und des Workshops. Ihre vielfältigen Perspektiven und Anregungen haben maßgeblich zur Qualität der Ergebnisse beigetragen.

Besonderer Dank gebührt iRights.Lab GmbH für die Unterstützung bei der Planung und Durchführung des Workshops sowie dem Grafikladen für die professionelle Umsetzung der Grafiken und die ansprechende Gestaltung der Studie.

Ebenso danken wir dem Bundesministerium für Forschung, Technologie und Raumfahrt für die Initiierung der Studie und die Bereitstellung der notwendigen finanziellen Mittel.

Diese Arbeit wäre ohne diese Unterstützung in dieser Form nicht möglich gewesen.

GESELLSCHAFT FÜR INFORMATIK E. V. (GI)

Geschäftsstelle Bonn

Ahrstr. 45
53175 Bonn
Tel.: +49 228 302-145
Fax: +49 228 302-167
E-Mail: bonn@gi.de

Geschäftsstelle Berlin

Weydingerstr. 14-16
10178 Berlin
Tel.: +49 30 7261 566-15
Fax: +49 30 7261 566-19
E-Mail: berlin@gi.de

gs@gi.de
www.gi.de

[@mas.to/@informatik](https://www.instagram.com/mas.to/@informatik)
[in/company/gesellschaft-fuer-informatik](https://www.linkedin.com/company/gesellschaft-fuer-informatik)