

Thesenpapier zur Künstlichen Intelligenz und ihrer Infrastruktur in Deutschland

Schwerpunkte: Trägerschaft · Nutzen · Recht · Akzeptanz · Kernprozesse · Betreibermodell

DOKUMENTTYP	Thesenpapier zur KI-Infrastruktur in Deutschland
ZIELBILD	Belastbare Standortbestimmung für Betreiber, Kommunen, Investoren und Politik
ERSTELLT VON	Helmut Kohl Erstellt mit Hilfe mehrerer KI-Modelle
STAND	22. September 2026 — Fassung 1.0 Redaktionsschluss: 22. September 2026

Das vorliegende Papier ordnet fünf Thesen zur Künstlichen Intelligenz und zu der Infrastruktur, die sie in Deutschland braucht. Es behandelt die Trägerschaft des Ausbaus, den bereits messbaren Nutzen, den wachsenden Abstand zwischen Technik und Recht, die Akzeptanz als Standortfrage und den bevorstehenden Sprung der KI-Nutzung von den Unterstützungsfunktionen in die Kernprozesse der Unternehmen. Ein Ausblickskapitel behandelt die Frage, wer welche Schicht eines verteilten Rechenzentrumsbetriebs halten sollte. Quellengrundlage sind Bitkom und Borderstep, das Statistische Bundesamt, die Übertragungsnetzbetreiber, Gesetzentwürfe des Bundes sowie dokumentierte kommunale Verfahren.

Inhaltsverzeichnis

1 Executive Summary	3
2 Anlass, Methodik und Quellenlage	6
3 These 1: Ein weltveränderndes Ereignis — erstmals ohne den Staat am Steuer	8
4 These 2: Der Nutzen ist größer als die Debatte	10
5 These 3: Das Recht läuft der Technik hinterher	13
6 These 4: Akzeptanz entscheidet — und für manche Stadt ist das Rechenzentrum die Antwort	16
7 These 5: Von den Supportaufgaben in die Kernprozesse — die These mit der größten Bedeutung für Unternehmen	20
8 Ausblick: Wer welche Schicht hält	23
9 Prüffragen und Handlungsempfehlungen	28
Anhang A: Quellenverzeichnis	29
Anhang B: Verzeichnisse	32

1 Executive Summary

Künstliche Intelligenz verändert die Welt. Das ist keine Prognose mehr, sondern eine Feststellung — und sie fällt in eine deutsche Debatte, die fast ausschließlich über Risiken geführt wird. Die Risiken sind real. Sie sind aber nicht die ganze Rechnung, und sie sind nicht der Grund, warum der Ausbau der nötigen Infrastruktur hier langsamer läuft als anderswo. Dieses Papier ordnet fünf Thesen dazu.

These 1: Ein weltveränderndes Vorhaben — erstmals ohne den Staat am Steuer

Weltverändernde Vorhaben wie das Atombombenprogramm oder die Mondlandung waren staatlich beauftragt, staatlich finanziert und staatlich kontrolliert. Bei der Künstlichen Intelligenz treibt die Privatwirtschaft, und die staatliche Seite arbeitet in Zeiträumen, die zur Geschwindigkeit der Entwicklung nicht passen — sie ist nicht unwillig, sie ist zu langsam. Eine wichtige Ausnahme vom westlichen Muster bildet China mit stärker staatlich gelenkter Industrie- und Infrastrukturpolitik. Für Deutschland ändert das die Frage, die ein Standort beantworten muss: nicht welche Förderung es gibt, sondern ob Genehmigung und Netzanschluss verlässlich sind. Der Staat ist in dieser Konstellation nicht Auftraggeber, sondern Bedingung. Wie ernst die Lage eingeschätzt wird, zeigt die Bitkom-Erhebung vom 14. September 2026: 91 Prozent der Unternehmen halten KI für die wichtigste Zukunftstechnologie, 56 Prozent meinen zugleich, Deutschland habe den Anschluss verloren, und 34 Prozent sorgen sich um die eigene Existenz — nach 23 Prozent im Vorjahr.^[2]

These 2: Der Nutzen ist größer als die Debatte

Der Nutzen lässt sich beispielhaft an zwei Feldern zeigen, die jeden betreffen: der Präzisionsonkologie und der Verkehrssicherheit. In der Onkologie hat im August 2026 erstmals eine individualisierte Krebstherapie eine Phase-III-Studie mit 1.137 Patienten bestanden; welche Merkmale eines einzelnen Tumors überhaupt angreifbar sind, berechnen dabei trainierte Vorhersagemodelle.^[34] Eine Phase-II-Studie von BioNTech wurde Ende August beendet — dort wurde derselbe Wirkstoff allein gegeben. Die Kombinationsstudie IMcode003 mit demselben Wirkstoff läuft unverändert weiter, und auch die erfolgreiche Studie war eine Kombination. Gescheitert ist also die Monotherapie, nicht das Verfahren.^[35] Im Straßenverkehr starben 2025 in Deutschland 2.832 Menschen, 62 mehr als im Vorjahr; das Bundesziel von 1.537 für 2030 rückt weg statt näher. Die Hauptursachen sind Geschwindigkeit, Abstand und Ablenkung^[13] — genau die Größen, die Assistenz- und Automatisierungssysteme adressieren. Entscheidend für dieses Papier ist die Folgerung: Beide Felder erhöhen den Bedarf an gesicherter, leistungsfähiger Rechenkapazität — im einen Fall vor allem wegen Datenmenge und Datenschutz, im anderen wegen der Kopplung an einen physischen Prozess. Wo diese Anforderungen greifen, wird der Standort der Rechenleistung zur Geschäftsfrage; wo sie nicht greifen, bleibt er beliebig. Genau diese Unterscheidung trägt den Rest des Papiers.

These 3: Das Recht läuft der Technik hinterher, und der Abstand wächst

Der Engpass ist der Netzanschluss. Nach der Begründung des Referentenentwurfs zur Reform des Netzanschlussverfahrens liegen auf Verteil- und Übertragungsnetzebene Anträge über rund 400 Gigawatt vor, von denen nur ein Bruchteil realisierbar ist;^[16] das bisherige Windhundprinzip wird darin selbst als nicht mehr zeitgemäß bezeichnet. Die Praxis hat im April 2026 auf ein Reifegradverfahren umgestellt, das zugehörige Gesetz befindet sich im September noch im Verfahren, und die verbindliche Auskunft darüber, ob an einem Standort überhaupt Kapazität frei ist, soll erst 2028 verfügbar sein.^[17] Dasselbe Muster zeigt sich auf der Anwenderseite: 66 Prozent der Unternehmen halten den AI Act für nachteilig, nach 56 Prozent im Vorjahr^[2] — die Ablehnung wächst in dem Maß, in dem man sich mit der Regelung befasst. Das ist kein Argument gegen Regulierung, sondern eines für Fristen: Ein Genehmigungsverfahren, das länger dauert als der Bau, trifft die Standortentscheidung, ohne dass sie irgendwo getroffen worden wäre.

These 4: Akzeptanz entscheidet — und für manche Stadt ist das Rechenzentrum die Antwort

Es sind zwei verschiedene Lagen. Im Ballungsraum geht es um Verdrängung: Vier hessische Verfahren — Groß-Gerau, Maintal, Birstein, Babenhausen — zeigen die Spannweite kommunaler Konflikte,^[38] von der Ablehnung im Stadtparlament über den Rückzug eines Investors bis zum politisch fortgeführten, aber umstrittenen Vorhaben. In keinem der Fälle war die Technik der Streitpunkt; in einem zog sich ein Investor mit bis zu drei Milliarden Euro zurück, weil ein Aufstellungsbeschluss fehlte. Zugleich ist Frankfurt mit rund 997 MW und vier Prozent Leerstand physisch am Anschlag;^[5] Wartezeiten für größere Netzanschlüsse können dort über zehn Jahre betragen. In einer Stadt mit 30.000 bis 70.000 Einwohnern stellt sich die Frage umgekehrt, und sie lautet nicht „Wollen wir ein Rechenzentrum?“, sondern „Womit werden wir für jüngere Menschen und Familien wieder attraktiv?“. Dort geht es um Abwanderung und Alterung: Nach einer RWI-Auswertung aus dem Jahr 2019 entfielen 43 Prozent aller Binnenumzüge auf die 18- bis 29-Jährigen^[7], und wer geht, gründet seine Familie anderswo. Dieselbe Auswertung zeigt, dass die Umzugsentscheidung unter fünfzig deutlich an den wirtschaftlichen Möglichkeiten hängt — qualifizierte, gut bezahlte Arbeit vor Ort ist damit ein wesentlicher Hebel dagegen, wenn auch nicht der einzige. Ein Rechenzentrum kann genau das liefern, dazu über lange Zeiträume erhebliche Gewerbesteuerereinnahmen, Abwärme für das kommunale Wärmenetz und Rechenleistung für Kliniken und Industrie in der Region. Es ist damit ein Anlass zur Standortentwicklung, aber kein Ersatz für sie: Ohne Betreuungsplätze, Wohnraum und Anbindung entstehen Pendler statt Einwohner. Diese Rechnung stellt niemand von allein auf; sie zu erklären ist Aufgabe des Betreibers, und zwar bevor der Aufstellungsbeschluss auf der Tagesordnung steht.

These 5: Von den Supportaufgaben in die Kernprozesse — die These mit der größten Bedeutung für Unternehmen

57 Prozent der deutschen Unternehmen ab 20 Beschäftigten setzen KI ein, nach 20 Prozent im Jahr 2024. Kein einziges gibt an, das Potenzial vollständig zu nutzen; 59 Prozent sagen „überhaupt nicht“, und dieser Anteil ist über alle Größenklassen hinweg fast gleich.^[2] Der Grund steht in der Verteilung der Einsatzfelder: KI sitzt im Kundenkontakt (72 Prozent) und im Marketing (54 Prozent), also in den Unterstützungsfunktionen. In den Kernprozessen liegt der Einsatz deutlich niedriger: Produktion 22 Prozent, Forschung und Entwicklung 16, Management 7. Das wird sich ändern: Wissensmanagement hat sich binnen eines Jahres verdoppelt, der Einsatz in der IT-Abteilung mehr als vervierfacht, und 11 Prozent setzen bereits KI-Agenten ein. Dieser Schritt ist keine Forschung mehr, sondern Einführungsarbeit — und er läuft schneller ab als der Aufbau der Infrastruktur darunter. Welche Kostenblöcke Unternehmen besonders häufig als groß bezeichnen, zeigt in dieselbe Richtung: Infrastruktur 51 Prozent, Lizenzen 21, Tokenverbrauch 8.^[2]

Ausblick: Wer welche Schicht hält

Wo Latenz, Datenresidenz oder Datenvolumen Nähe verlangen, entsteht ein Zielsegment, das dieses Papier mit einer Größenordnung von etwa fünf bis fünfzig Megawatt ansetzt — eine Arbeitshypothese, keine Marktstatistik, aber eine Größe, die Stadtwerke, Regionalbetreiber und bestehende Colocation-Häuser bauen können. Edge-Rechenzentrum und Inferenzbetrieb — also der laufende Betrieb fertiger trainierter Modelle, im Unterschied zum einmaligen Training, bei dem das Modell erst entsteht — können dabei am Standort zusammenfallen, ohne in derselben Gesellschaft zu liegen: Ein Gebäude ist ein Immobiliengeschäft mit zwanzig Jahren Nutzungsdauer, eine Flotte von Beschleunigern ein Technologiegeschäft mit drei bis vier Jahren Abschreibung. Vier Schichten lassen sich unterscheiden — Energie, AssetCo, OpCo, ServCo —, und die einschlägigen Regelwerke treffen unterschiedliche Rollen: das Energieeffizienzgesetz und das KRITIS-Dachgesetz den Betreiber der Anlage, das NIS2-Umsetzungsgesetz die Einrichtung, der AI Act Anbieter und Betreiber von KI-Systemen. Welche Gesellschaft welche Rolle einnimmt, entscheidet das Vertragsmodell. Verteilte Standorte können dabei geografische Redundanz schaffen, die ein einzelner Standort nicht bieten kann — sie ergänzt die technische Redundanz vor Ort, ersetzt sie aber nicht. Was in Deutschland fehlt, ist die mittlere Schicht: ein Betreiber, der eine Flotte verteilter Inferenzknoten hält, vermascht und auslastet. Ohne sie kann die kommunale Schicht ihr Kapital kaum einsetzen.

Kennzahlen dieses Papiers auf einen Blick

Kennzahl	Wert	Quelle und Stand
Unternehmen mit KI-Einsatz (ab 20 Beschäftigte)	57 %	Bitkom, 14.09.2026 ^[2]
davon mit vollständig ausgeschöpftem Potenzial	0 %	Bitkom, 14.09.2026 ^[2]
KI-Einsatz im Kundenkontakt	72 %	Bitkom, 14.09.2026 ^[2]
KI-Einsatz im Management	7 %	Bitkom, 14.09.2026 ^[2]
Installierte IT-Anschlussleistung Deutschland 2025	2.980 MW	Borderstep/Bitkom, 11/2025 ^[3]
Edge-Anschlussleistung Deutschland 2025	rd. 240 MW	Borderstep/Bitkom, 11/2025 ^[3]
Beantragte Anschlussleistung Verteil- und Übertragungsnetz	rd. 400 GW	Referentenentwurf Netzanschluss ^[16]
Verkehrstote Deutschland 2025	2.832	Statistisches Bundesamt, 07/2026 ^[13]
Registrierte Einrichtungen nach NIS2	19.058	Bundesregierung und BSI, Stand 31.07.2026 ^[25]

Tabelle 1: Leitkennzahlen des Papiers

Quelle: siehe Anhang A; Zuordnung je Zeile in der Spalte „Quelle und Stand“