



# Die unverzichtbare Rolle von Workstations in der KI-Revolution

Friedrich Stahl  
HP Business Development Data Science & AI

Datum: 10. Oktober 2024

# KI ist in aller Munde



CHATGPT



MIDJOURNEY

*Amazon geht bei generativer  
KI "super aggressiv" vor* **THE WALL STREET JOURNAL.**

*"Microsoft ist zurück." Wie KI  
den fünf Jahrzehnte alten Tech-  
Giganten wieder an die Spitze  
brachte* **The Washington Post**

*Meta stellt Code Llama 70B vor: ein  
leistungsstarkes KI-Modell zur Codegenerierung* **Forbes**

*Google sagt, dass das neue  
KI-Modell Gemini ChatGPT in  
den meisten Tests übertrifft* **The  
Guardian**

*Apple prüft KI-Deals mit  
Nachrichtenverlagen* **The New York Times**

*Mistral gegen Aleph Alpha: Wer  
gewinnt das KI-Rennen in Europa?* **Handelsblatt**

# KI ist überall

## COMPUTER VISION



Medical Imaging



Defect Detection

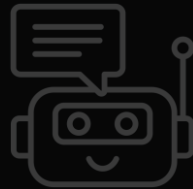


Self-Driving

## NATURAL LANGUAGE PROCESSING



Sentiment Analysis



Chat Bots



Text Summarization

## GENERATIVE AI



Text Generation



Image Generation



Synthetic Data

## RECOMMENDATION ENGINES



Loan Decisions



Personalized Retail



Personalized Experiences

# Die Komponenten von KI-Systemen

Aufbau eines KI-Systems, welches Erkenntnisse generiert



ALGORITHMUS

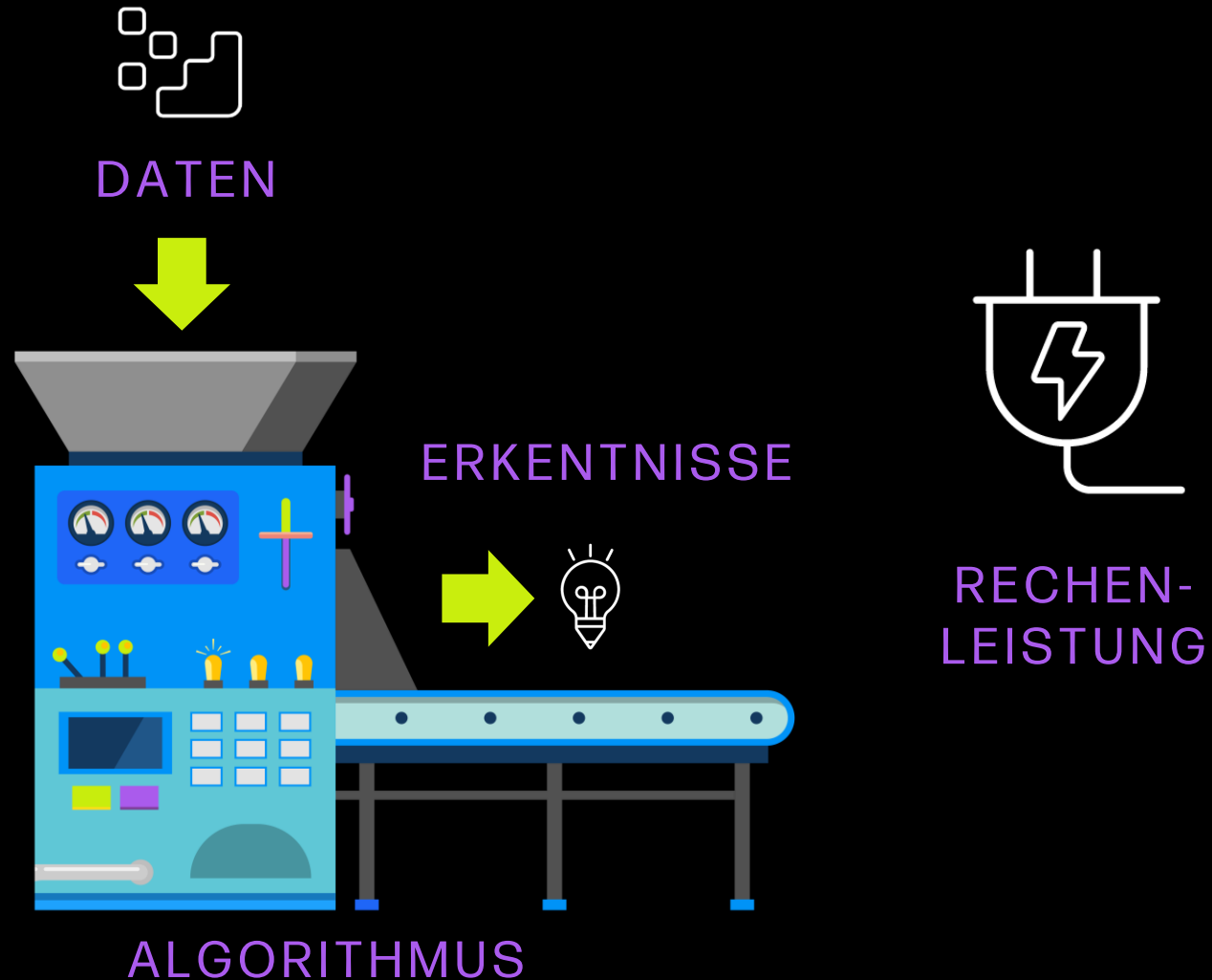


DATEN



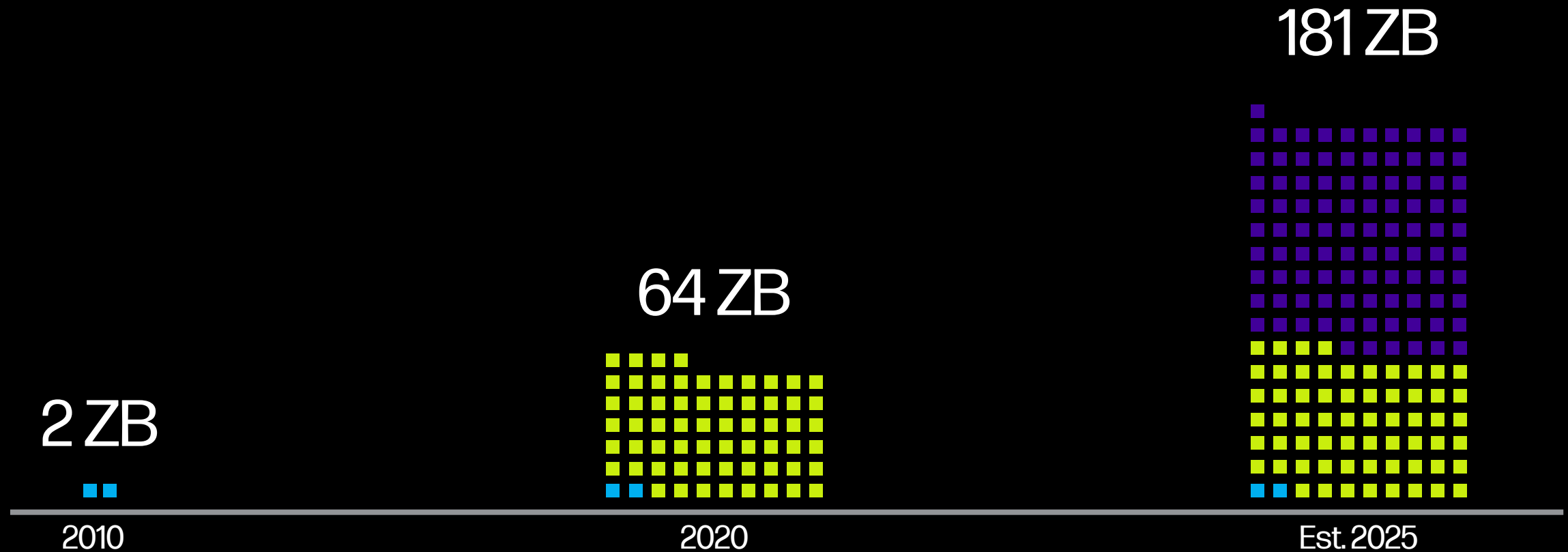
RECHENLEISTUNG

# KI: Die Fabrik der Erkenntnisse



# Explosion der Datengenerierung

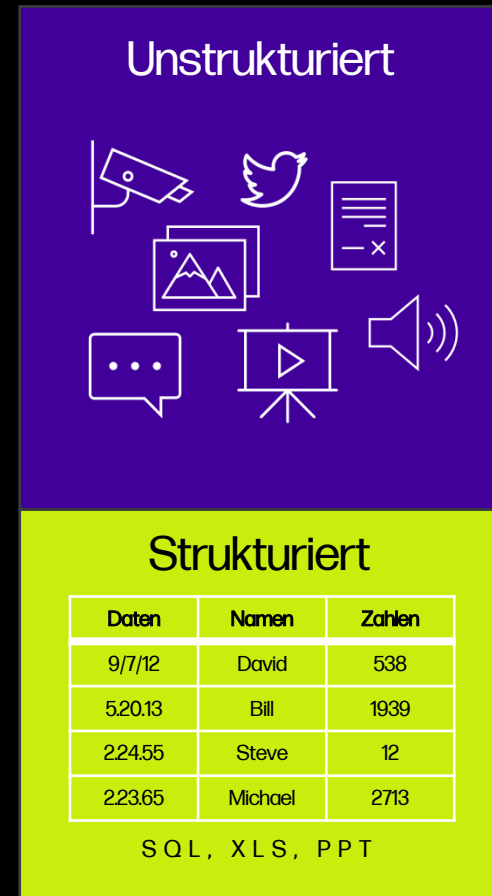
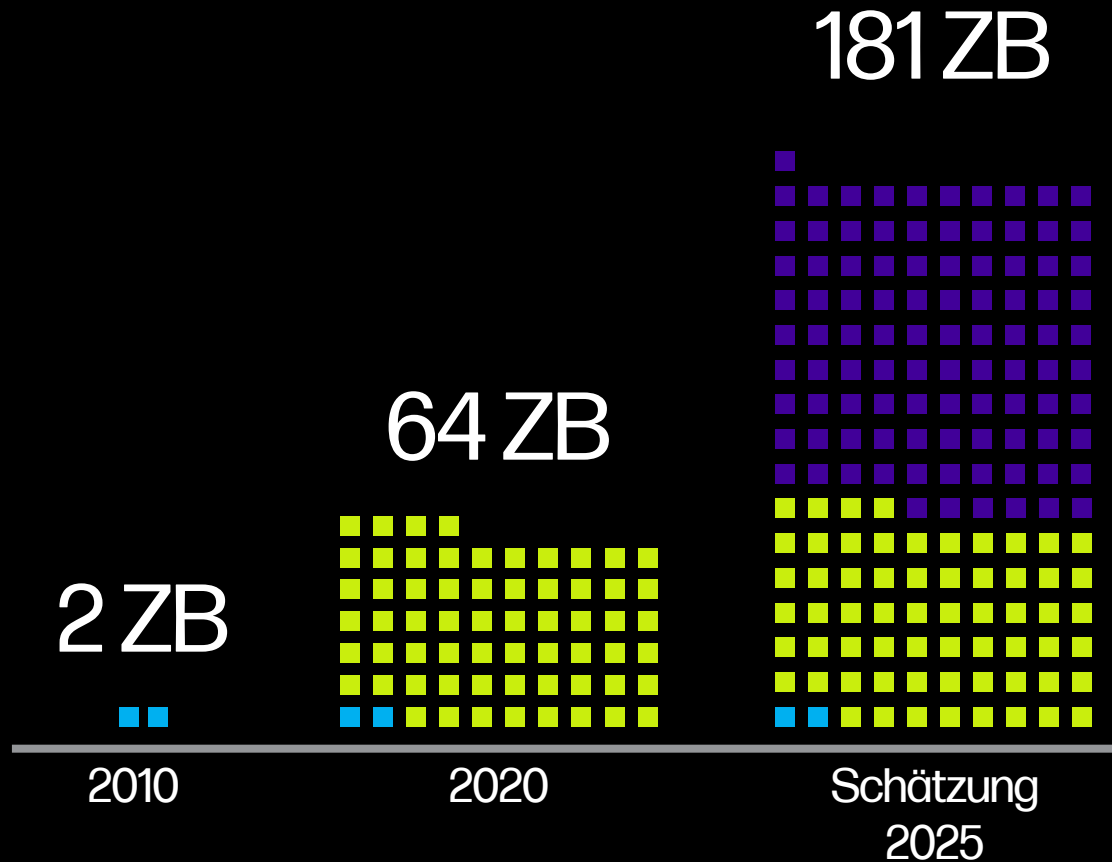
Ein Trend, der sowohl eine Herausforderung als auch eine Chance darstellt



Jährliche Gesamtdaten, die erstellt und gespeichert werden (ZB<sup>1</sup>)

# Explosion der Datengenerierung

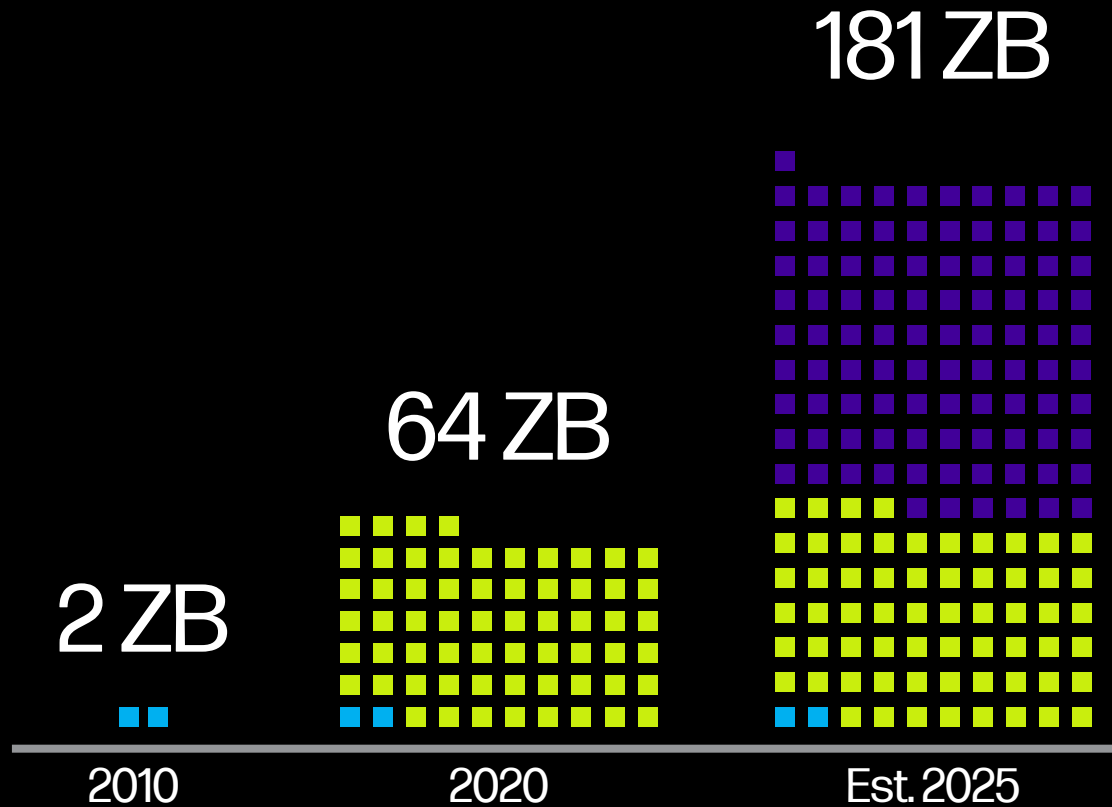
Ein Trend, der sowohl eine Herausforderung als auch eine Chance darstellt



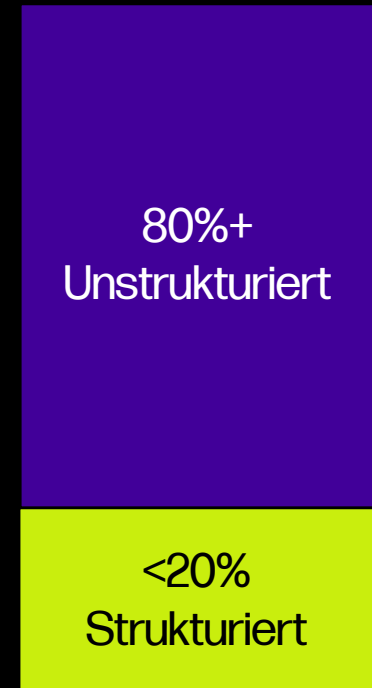


# Explosion der Datengenerierung

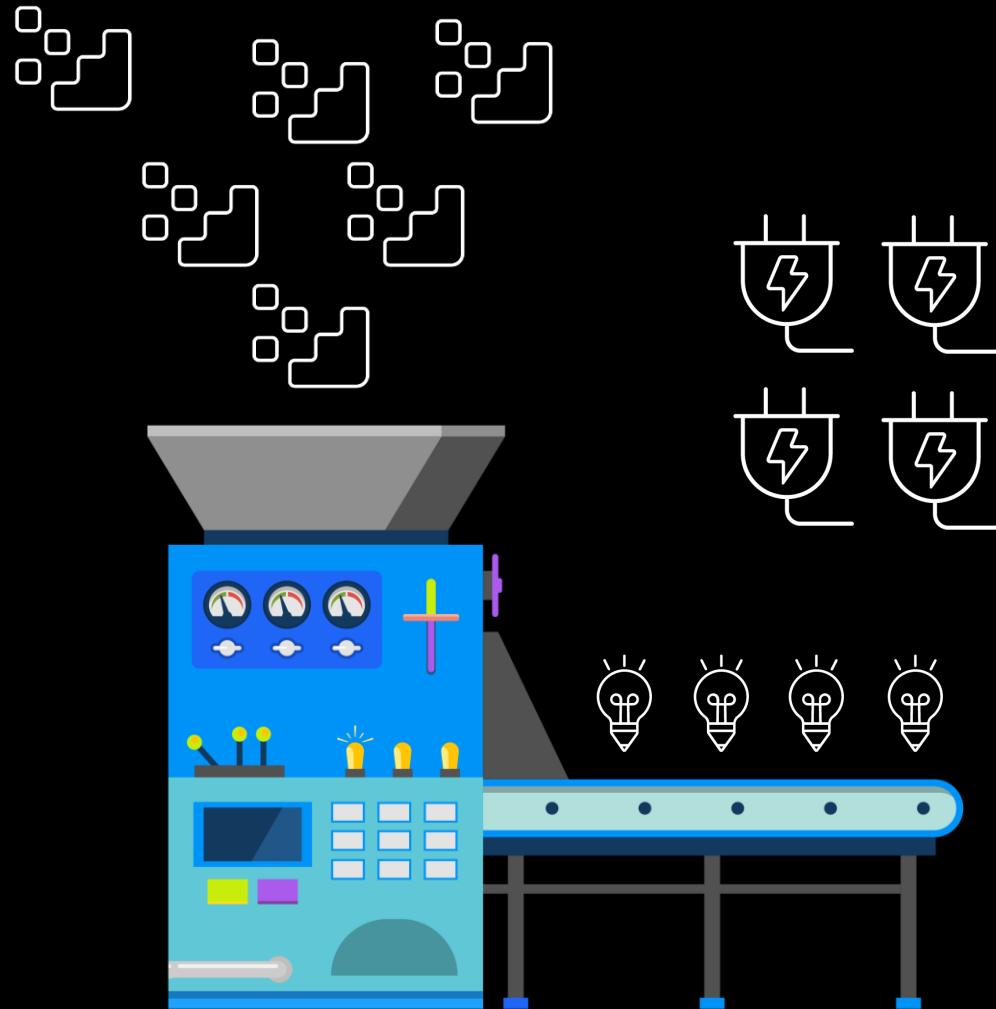
Ein Trend, der sowohl eine Herausforderung als auch eine Chance darstellt



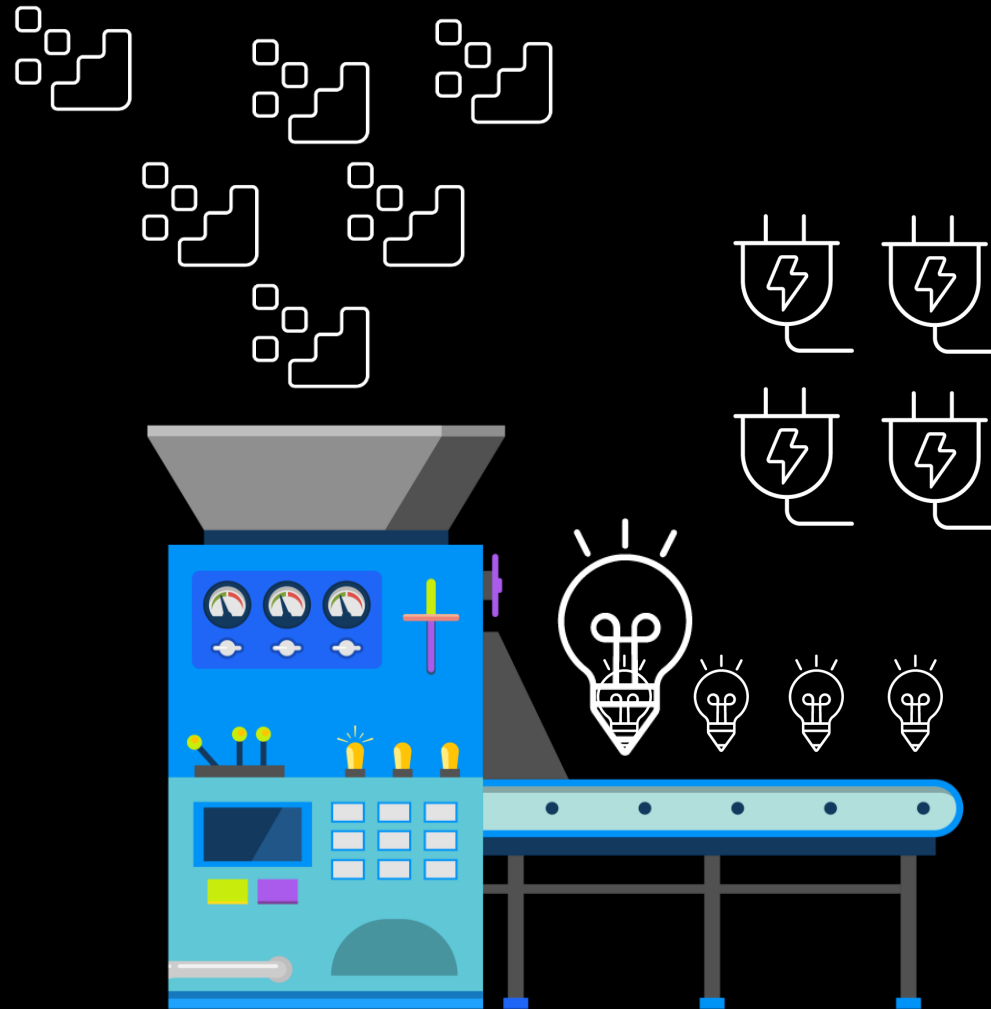
Wachstum nach Datentyp:



# Wie verarbeitet unsere Fabrik diese Datenmenge?



# Wie verarbeitet unsere Fabrik diese Datenmenge?



# Prioritäten bei der Auswahl von Rechenressourcen



## GESCHWINDIGKEIT

Wie schnell kann ich die Berechnungen durchführen?



## AUSGABEN

Wieviel wird mich das kosten?



## SICHERHEIT

Sind meine Daten sicher?

# Arten von Rechenressourcen



WORKSTATION

## VORTEIL

- Datensicherheit
- Flexibilität der physischen Umgebung

## NACHTEIL

- Einschränkungen der Rechenleistung
- Einschränkungen für Benutzer



EDGE

## VORTEIL

- Datensicherheit
- Leistungsstarkes Rechnen mit geringer Latenz

## NACHTEIL

- Einschränkungen der physischen Umgebung
- Bandbreiten-Limitierungen



CLOUD

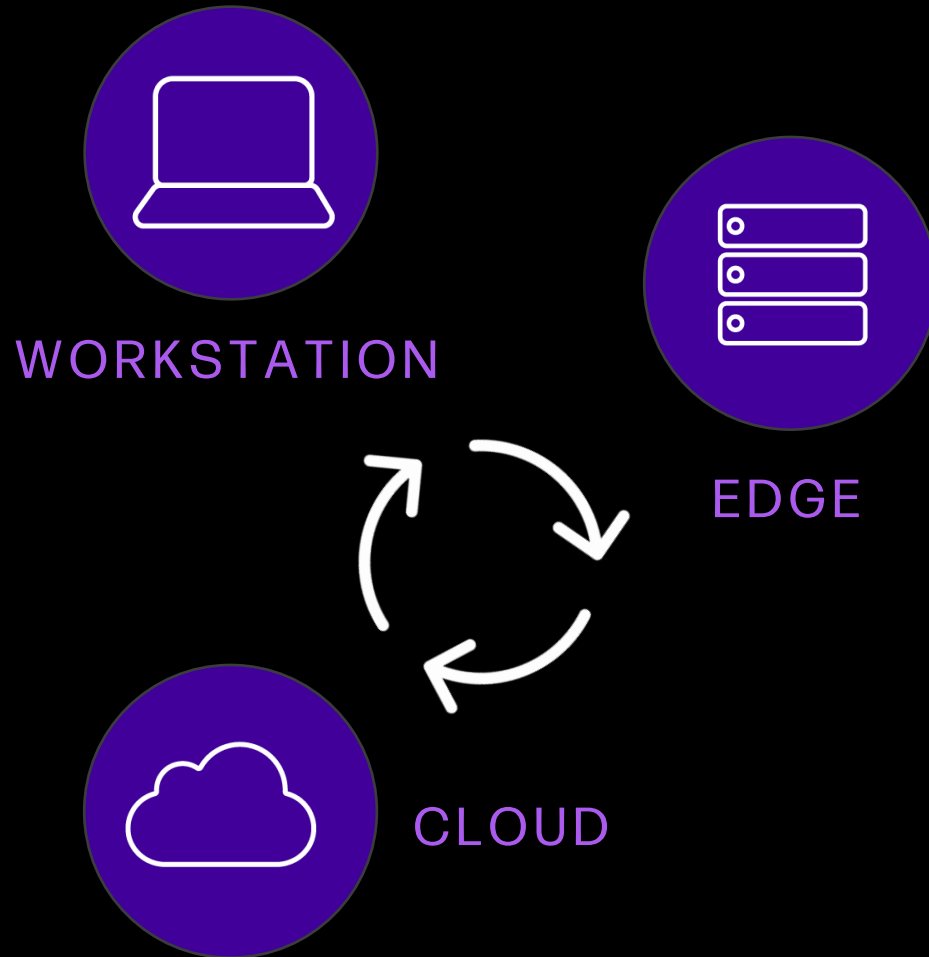
## VORTEIL

- Leistungsstarke Rechenleistung
- Minimale Anschaffungskosten

## NACHTEIL

- Datensicherheitsrisiko
- Wiederkehrende Kosten

# Die Zukunft der Datenverarbeitung



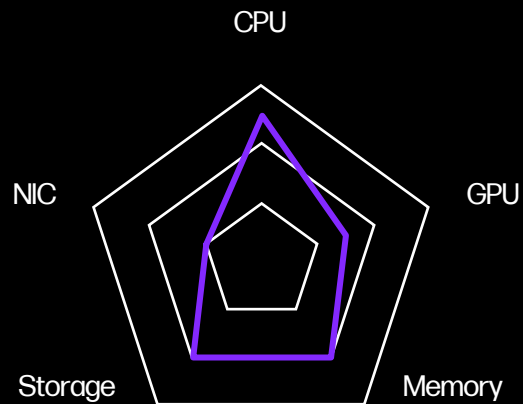
## Hybrides Rechnen

ermöglicht es Data Scientists, die Effizienz jeder Art von Rechenressource zu maximieren und KI-Projekte schnell, kostengünstig und sicher zusammenzuführen.

# Hybrides Rechnen beginnt mit der Hardware

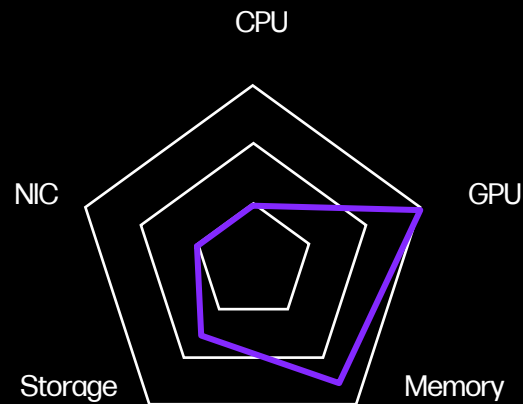
Die Flexibilität von KI-Workstations ermöglicht vollständig optimierte hybride Rechennetzwerke

## DATEN-AUFBEREITUNG



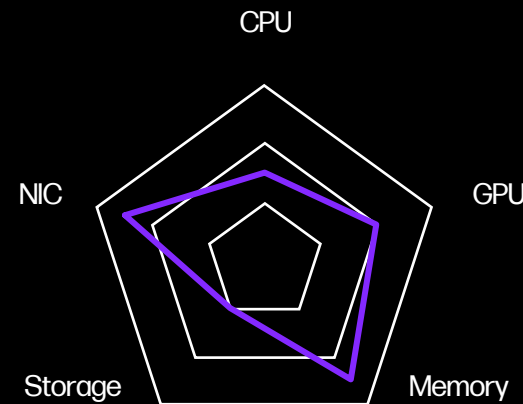
CPU Intensiv

## MODELL-TRAINING



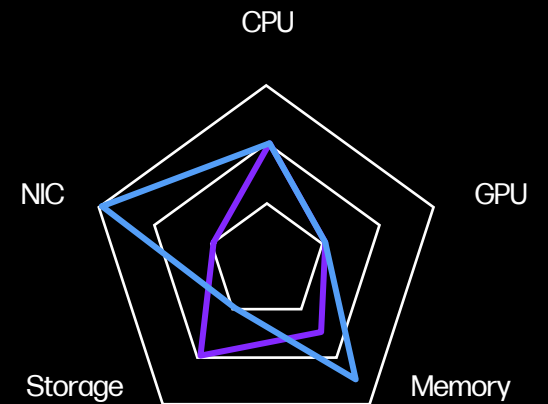
GPU Intensiv

## MODELL-INFERENCING



Memory/ NIC Intensiv

## DATEN ANALYTIK



— Batch — Streaming

↑ Storage Intensiv

↑ NIC/Memory Intensiv

# Hybrides Rechnen beginnt mit der Hardware



ZBook Firefly  
G11 14"

ZBook Firefly  
G11 16"



ZBook Power  
G11 16"



ZBook Studio  
G11 16"



ZBook Fury  
G11 16"

## Workflows

- Modellerstellung
- Remote-Data-Science-Verbindung mit Z Desktop über HP Anyware
- Explorative Datenanalyse
- Informatik und Data Science in der Ausbildung
- Datenvisualisierung
- Mittlere Datenanalyse
- Light Data Science (ML-Inferenz)
- Umfangreiche Datenanalyse
- Medium Data Science (ML-Inferenz und -Training; DL-Inferenz)

## Daten

Kleinere Datensätze/strukturierte Daten

Größere Datensätze/unstrukturierte Daten

## Modelle

Modellerstellung

ML Inferenz

ML Inferenz/Training & DL Inferenz



# Hybrides Rechnen beginnt mit der Hardware



Entry				High-end					
Z1 Tower G9	Z2 Mini G9	Z2 SFF G9	Z2 Tower G9	Z4 Rack G5	Z4 G5	Z6 G5	Z6 G5 A	Z8 G5	Z8 Fury G5
Step into professional performance	Powers large workflows. Fits a tiny space.	Full size Z tower power at half the size.	Your new benchmark for entry workstation power.	Compactly rackable. Uniquely powerful.	Our best-selling workhorse, now more powerful.	Unleash extreme graphics performance	Maximized Cores. Astounding Graphics. Unbound Potential.	Take on processor-intensive workloads. Relentlessly.	Extreme performance. Infinite possibilities.
<b>Up to</b>									
<b>Intel® Core™ 1</b>				<b>Intel® Xeon™ &amp; AMD® Threadripper Pro™</b>					
<ul style="list-style-type: none"> <li>Intel Core i9</li> <li>NVIDIA T1000<sup>*,2,3</sup></li> <li>128GB of RAM</li> </ul>	<ul style="list-style-type: none"> <li>Intel Core i9, K-series</li> <li>NVIDIA RTX A2000<sup>2,3</sup></li> <li>64GB of RAM</li> </ul>	<ul style="list-style-type: none"> <li>Intel Core i9 K-series</li> <li>NVIDIA A4000<sup>2,3</sup></li> <li>128GB of RAM</li> </ul>	<ul style="list-style-type: none"> <li>Intel Core i9 K-series</li> <li>NVIDIA 5000 Ada<sup>2,3</sup></li> <li>128 GB of RAM</li> </ul>	<ul style="list-style-type: none"> <li>Intel Xeon 24 core</li> <li>1 high-end GPU up to NVIDIA 6000 Ada<sup>2,3</sup></li> <li>256GB of RAM</li> </ul>	<ul style="list-style-type: none"> <li>Intel Xeon 24 core in a single CPU W9</li> <li>2 high-end GPUs up to NVIDIA 6000 Ada<sup>2,3</sup></li> <li>512 GB of RAM</li> </ul>	<ul style="list-style-type: none"> <li>Intel Xeon 36 core in a single CPU W9</li> <li>3 high-end GPUs up to NVIDIA 6000 Ada<sup>2,3</sup></li> <li>1 TB of RAM</li> </ul>	<ul style="list-style-type: none"> <li>AMD Ryzen™ Threadripper™ PRO 96 core</li> <li>3 high-end GPUs up to NVIDIA 6000 Ada<sup>2,3</sup></li> <li>1 TB of RAM</li> </ul>	<ul style="list-style-type: none"> <li>Intel Xeon 64 core with 2X scalable</li> <li>2 high-end GPUs up to NVIDIA 6000 Ada<sup>2,3</sup></li> <li>1 TB of RAM</li> </ul>	<ul style="list-style-type: none"> <li>Intel Xeon 56-cores / single CPU, W9</li> <li>4 high-end GPUs up to NVIDIA 6000 Ada<sup>2,3</sup></li> <li>2 TB of RAM</li> </ul>

# Fallbeispiel Einsatz von Workstations für KI

DELIFLOR

Chrysanten



Delifor ist der größte Züchter von Chrysanthemen in den Niederlanden.

Die Niederlande sind für über 40 % des weltweiten Handels mit Schnittblumen verantwortlich.

Der Markt für Blumen und Zierpflanzen erwirtschaftet jährlich 53 Milliarden Euro

## Die Herausforderung:

Genauere Vorhersage der Blütezeit für Blumen, die international verschickt werden, mithilfe von Bildanalysen zur Blütenklassifizierung.

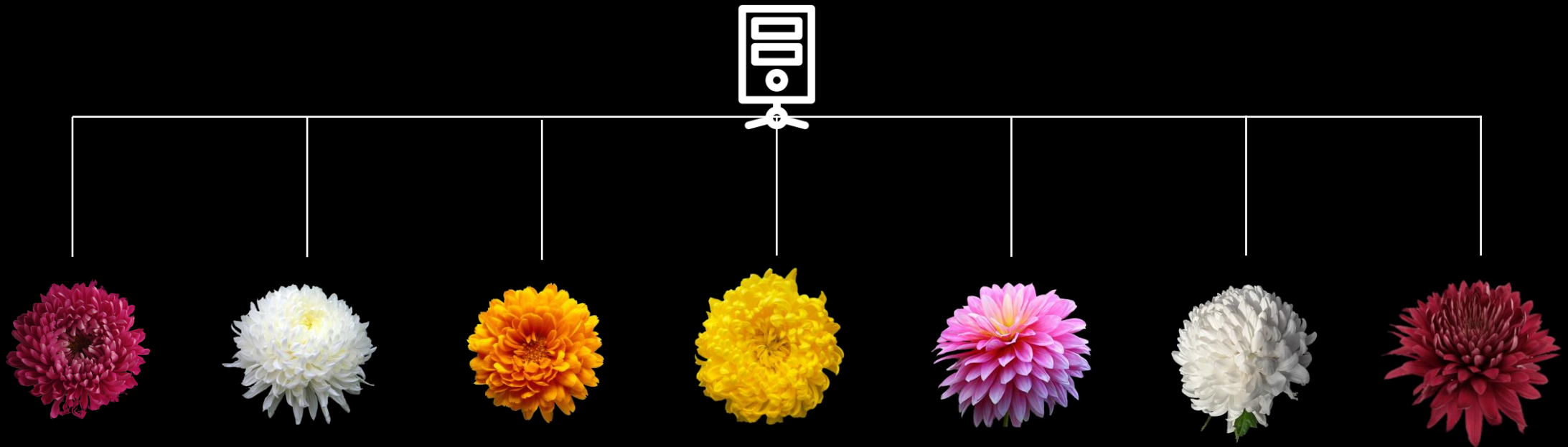
# Vor Einsatz von HP Workstations

Analyse von 5000+ Bilder über einen einzigen Server

DELÍFLOR

Chrysanten

Analysezeit:  
45 Minuten



# Mit HP Workstations

DELÍFLOR

Chrysanten

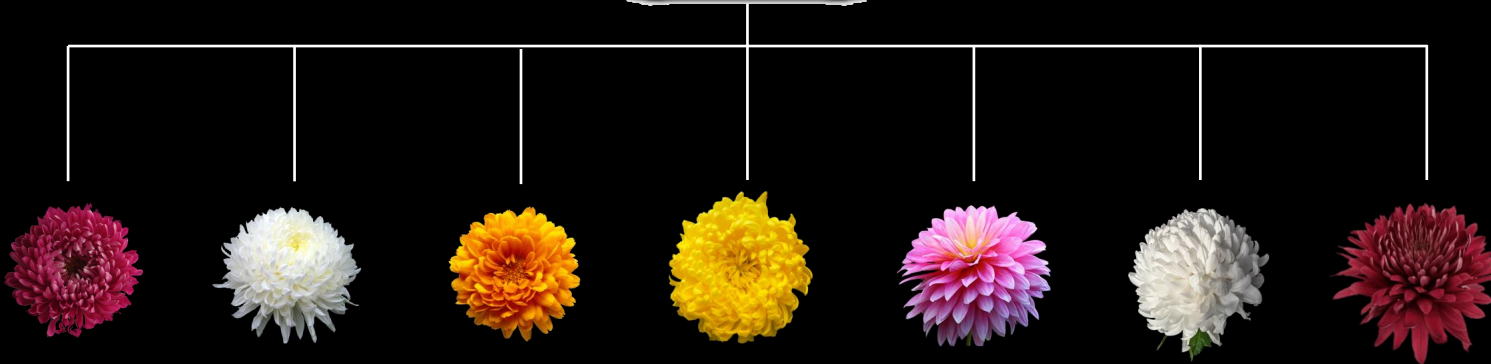
## Z8 G5 Desktop



"Mit dem Z by HP mit NVIDIA-GPUs läuft eine Epoche in nur wenigen Minuten, etwa 6x schneller. Das gibt uns die Möglichkeit, mit der Geschwindigkeit unserer Ideen zu arbeiten und die Kontinuität des Denkens zu wahren."

**Carolina Urrea Hernandez**

Researcher



# Erschließung neuer Workflows

Verwendung molekularer Marker zur Vorhersage der hexasomischen Vererbung, Optimierung der Kreuzung, um die besten Blüten zu erzeugen.

"Das sind komplexe Daten, mit denen wir noch vor wenigen Jahren nicht hätten arbeiten können. Die Leistungsfähigkeit von Z by HP ermöglichte es uns, diese Modelle zu betreiben."



# Software bringt KI-Projekte zusammen

Die Geräte von Z by HP werden mit der gesamten Software geliefert, die Sie für die Durchführung von Data-Science-Projekten benötigen, und zwar direkt nach dem Auspacken

## Software Tools

Holen Sie sich die beliebtesten Tools zum Analysieren von Daten und Erstellen leistungsstarker Modelle. Profitieren Sie von einer nahtlosen Verwaltung von Paketaktualisierungen und Abhängigkeiten, damit Ihre Umgebung immer einsatzbereit ist.



## Dienstprogramme für Entwickler

Erstellen und implementieren Sie ganz einfach Ihre Data-Science-Anwendungen, während Sie Ihre Data-Science-Tools und -Modelle verwalten.



## Verbessertes Cloud-Erlebnis

Interagieren Sie nahtlos mit Ihrer Haupt-Cloud-Umgebung über den Cloud-Befehlszeilenzugriff.



## Ubuntu. Zertifiziert und vorinstalliert.

Wir arbeiten eng mit Canonical zusammen, um die neueste Version von Ubuntu OS ausgiebig zu testen und zu zertifizieren, damit sie optimal funktioniert.



## Microsoft WSL 2 vorinstalliert

Erleben Sie das Beste von Windows und Ubuntu mit dem Microsoft Windows-Subsystem für Linux 2. WSL 2 ist eine einfache und schnelle Möglichkeit, Linux direkt unter Windows auszuführen.

# HP Anyware

Nur HP bietet ein High-Definition-Computing-Erlebnis für sichere digitale Arbeitsplätze



Flexibilität bei der  
Bereitstellung der IT-  
Infrastruktur



Starke Sicherheit und  
Datenschutz



Virtuelle  
Kollaborationsfähigkeit



Erstklassige Remoting-  
Leistung zur Steigerung  
der Produktivität

# AI Studio setzt die Puzzleteile zusammen

Eine Plattform zur Verbindung von Menschen, Daten, Tools und Rechenleistung, um die Modellentwicklung zu beschleunigen



## ZENTRALISIEREN

Alle Ihre Daten, Tools, Teammitglieder und Projekte auf einer Plattform.



## ZUSAMMENARBEITEN

Greifen Sie einfach auf Daten, Vorlagen und Experimente zu und teilen Sie sie mit Ihrem Team.



## BESCHLEUNIGEN

Beschleunigen Sie die Bereitstellung von KI-Modellen durch betriebliche Effizienz.



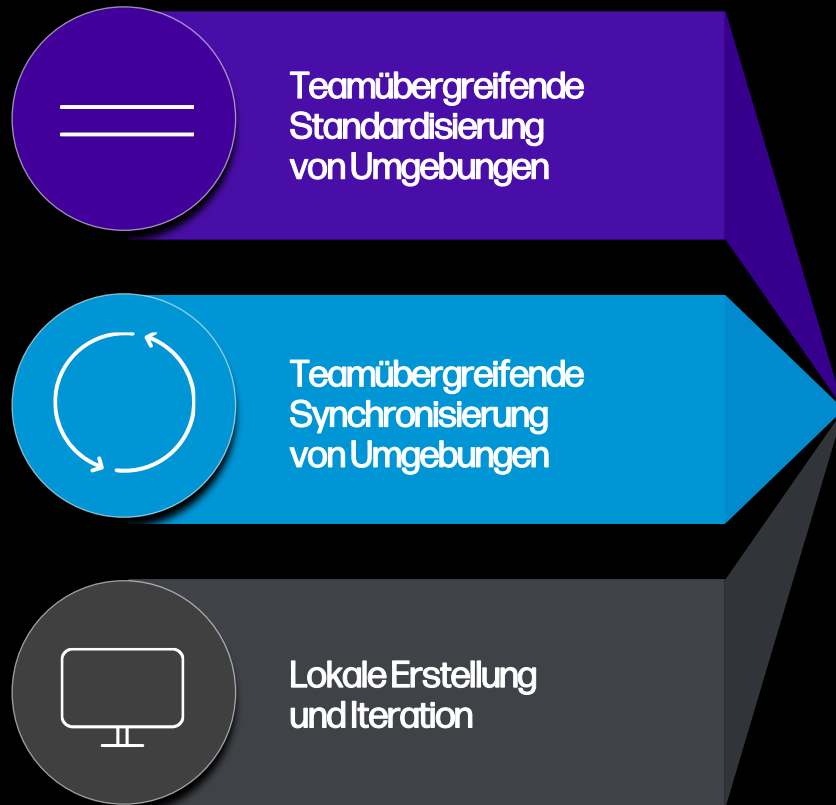
# Transformieren Sie Ihren Workflow

Egal, wo die Arbeit beginnt, AI Studio bringt alles zusammen



# Transformieren Sie Ihren Workflow

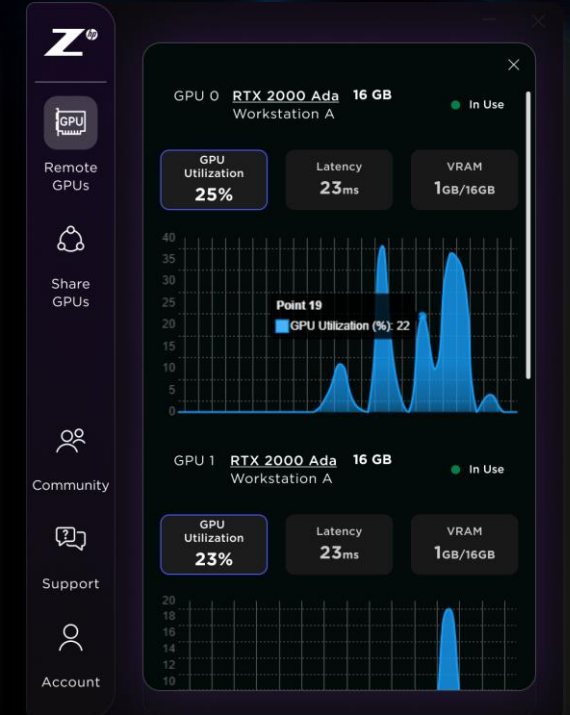
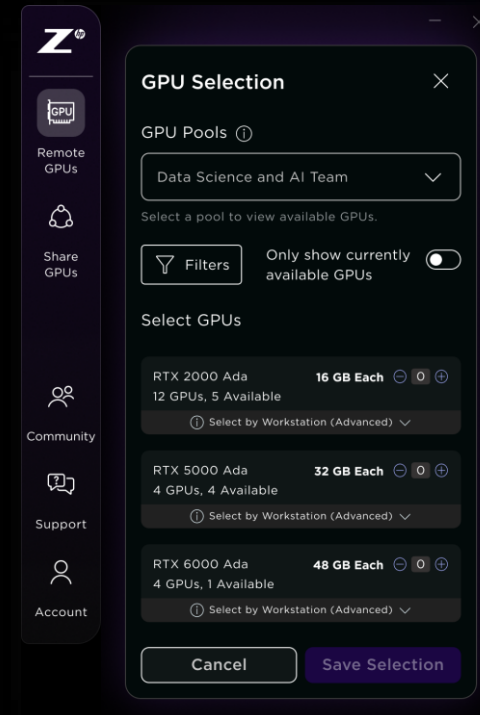
Egal, wo die Arbeit beginnt, AI Studio bringt alles zusammen



# Z by HP Boost

Weltweit erste Workstation-Lösung für On-Demand-GPUs, die die Workstation in eine gemeinsam nutzbare Ressource verwandelt

- **On-Demand-Leistungssteigerung** – nutzen Sie Ihre vorhandenen GPU-Potentiale Ihrer Z by HP Workstation
- **Beschleunigen Sie KI-Innovationen** – liefern Sie Leistung, wenn Sie sie benötigen, um die Modellentwicklung zu beschleunigen
- **GPU-Zugriff – Produktivität zu jeder Zeit** – bewältigen Sie komplexe KI-Workflows von nahezu überall aus mit GPU-Zugriff über Standard-Netzwerkprotokolle
- **Sicheres Trainieren von KI-Modellen** lokal oder per VPN/Verschlüsselung in einer einzelnen Benutzersitzung



# Herzlichen Dank für die Aufmerksamkeit!

Ansprechpartner bei HP für das Thema KI

E-Mail: [friedrich.stahl@hp.com](mailto:friedrich.stahl@hp.com)

LinkedIn: <https://www.linkedin.com/in/friedrichstahl/>

XING: [https://www.xing.com/profile/Friedrich\\_Stahl2](https://www.xing.com/profile/Friedrich_Stahl2)



Tester gesucht für HP AI Studio!