

**Engineering the Future:
Unlocking Business Value with
Oracle Cloud Infrastructure (OCI)
and
Artificial Intelligence (AI)**



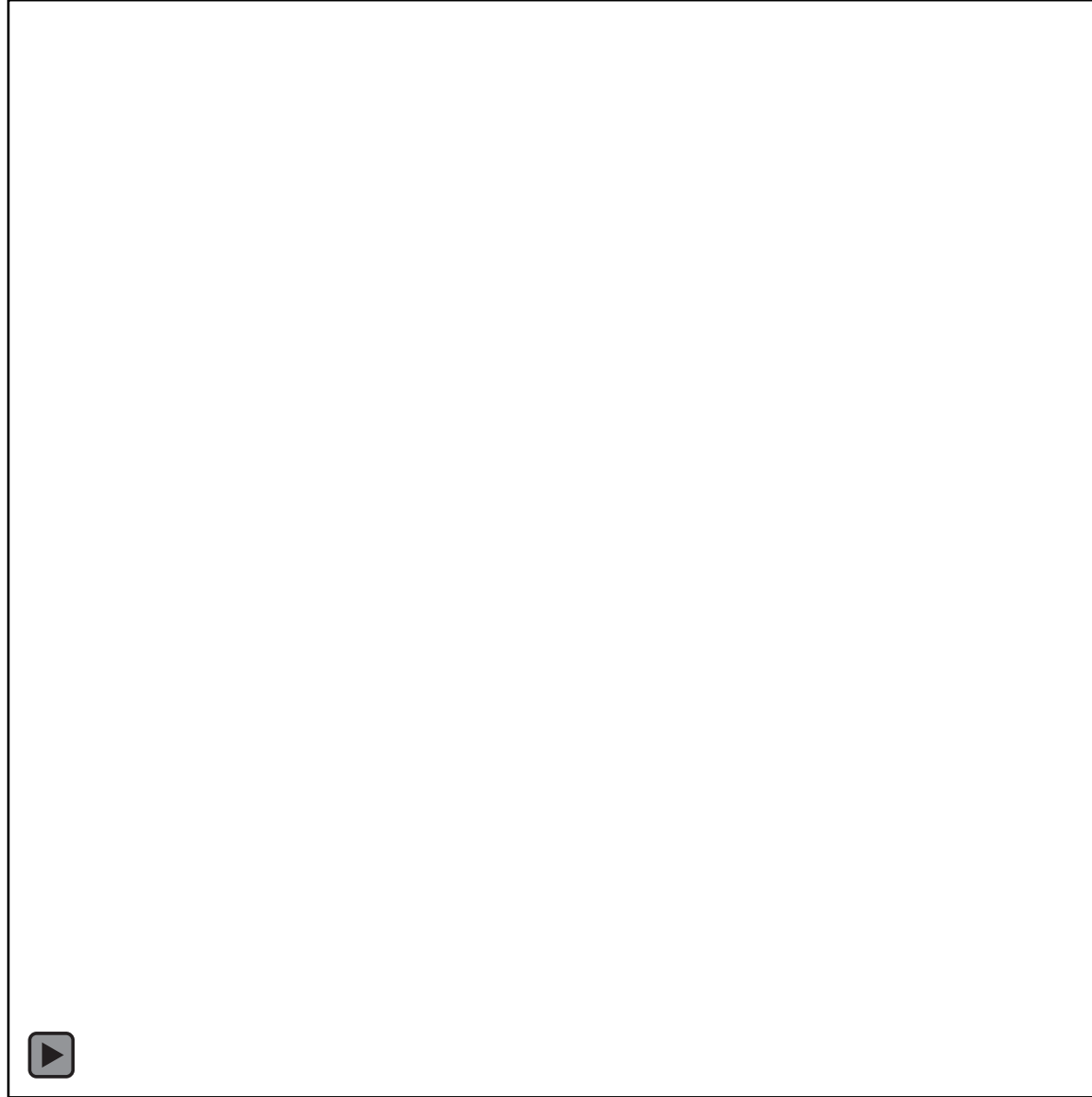
Brett Primo

Oracle Cloud Engineering Leader
South Africa, Kenya and Nigeria

Talking points...90's format 😊

- **Why AI — and OCI — matter more than ever**
- **What business outcomes you and our partners can unlock by using AI on OCI**
- **How Oracle's AI platform and real-world use cases are transforming industries**
- **Q&A**

Talking points...2025 and Beyond...

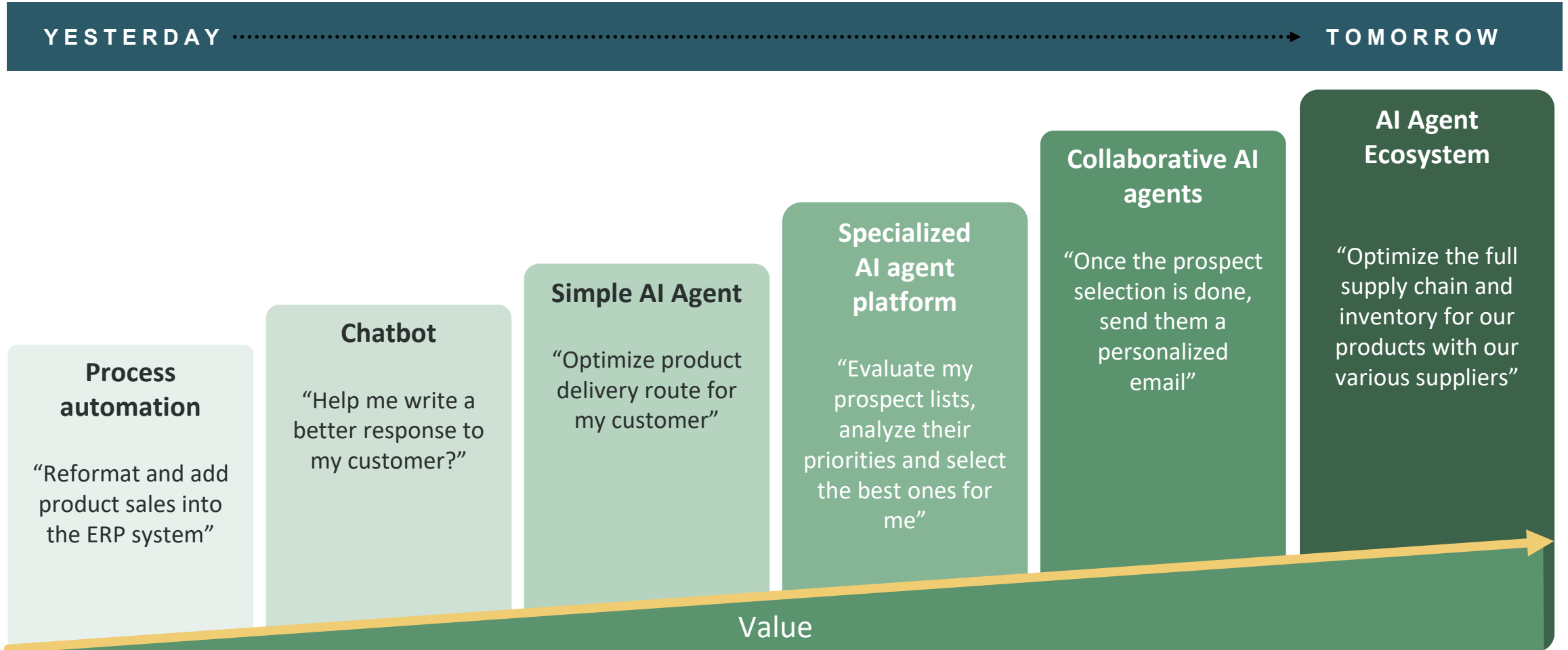


The background of the slide is a photograph of an industrial manufacturing environment. A large, orange robotic arm is positioned on the left, with its end effector (a welding torch) focused on a metal component on a conveyor belt. Bright sparks are visible at the point of contact. In the foreground on the right, a person's hand is holding a tablet computer. The tablet screen displays a technical diagram of a robotic arm, likely a simulation or a control interface. The overall scene is dimly lit, with the primary light source being the sparks from the welding process.

In 10 years, AI
could **double the pace**
of productivity growth in
the US.

—Goldman Sachs

The shift from static AI models to autonomous, AI-driven actions



Enterprise AI adoption challenges

Most enterprises are under pressure to adopt AI, but they're running into challenges:

- It's expensive and complex to build from scratch
- Prebuilt solutions lack flexibility or domain relevance
- Security and compliance concerns require attention
- Teams are unsure where to start or how to scale

Oracle's approach: We meet you where you are...

CREATE

The performance and power you need to build AI and machine learning models from the ground up. Get more price-performance with Oracle's complete AI stack

OPERATE

Oracle provides a platform and tools so you can build AI enhanced applications and focus on solving your specific business needs. Customize with your data, without needing a team of data scientists

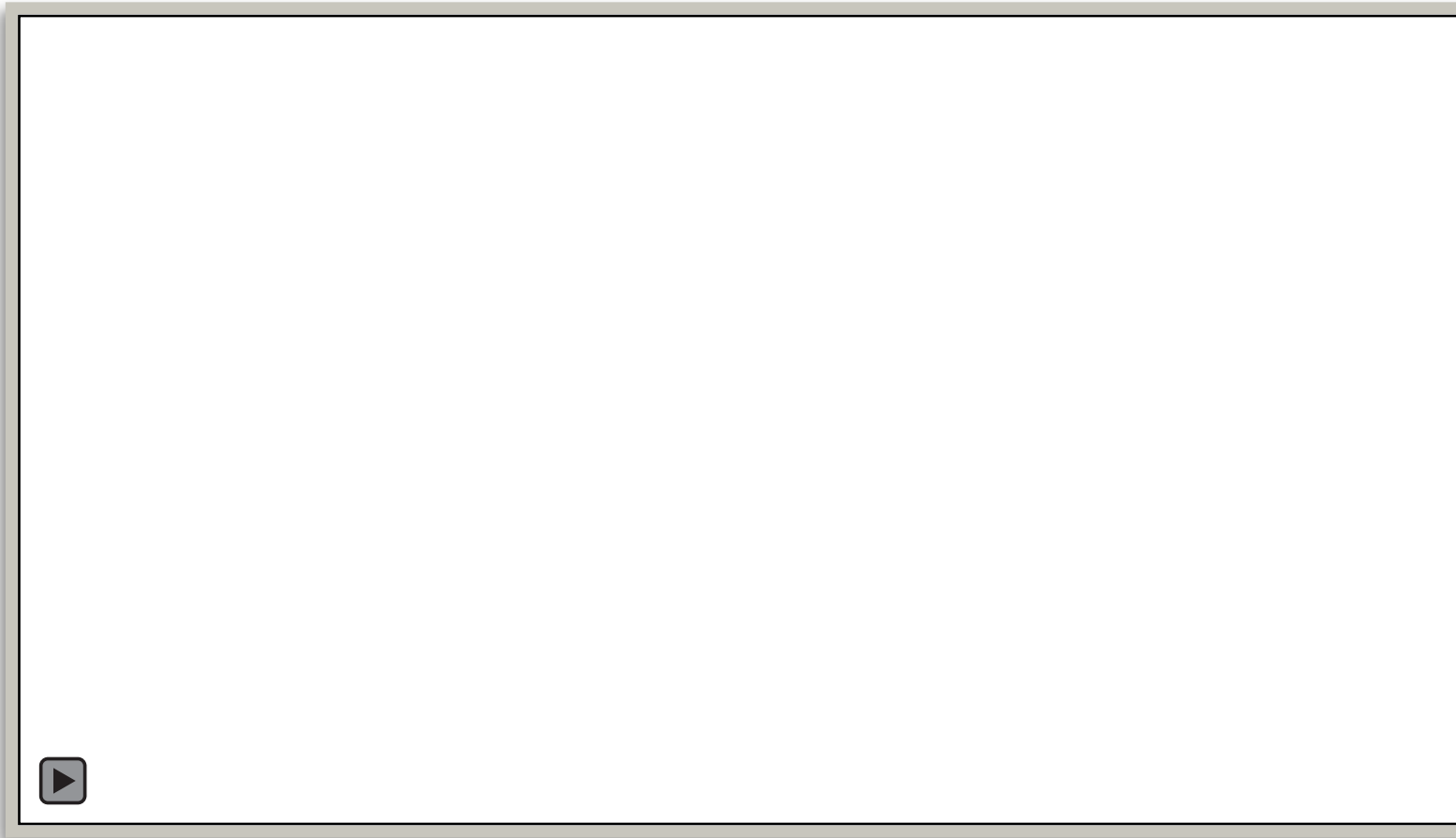
ACTIVATE

AI embedded in tools and services you already use, providing value from day one. No AI expertise, configuration, or complexity needed.

Business Impact:

- ✓ **Faster time to value** with AI that's ready on day one
- ✓ **Better performance** for demanding workloads
- ✓ **Control** over security, compliance, and cost

OCI Generative AI service



ORACLE®

Available in a fully managed service with open source and proprietary models:

 Meta

 cohere

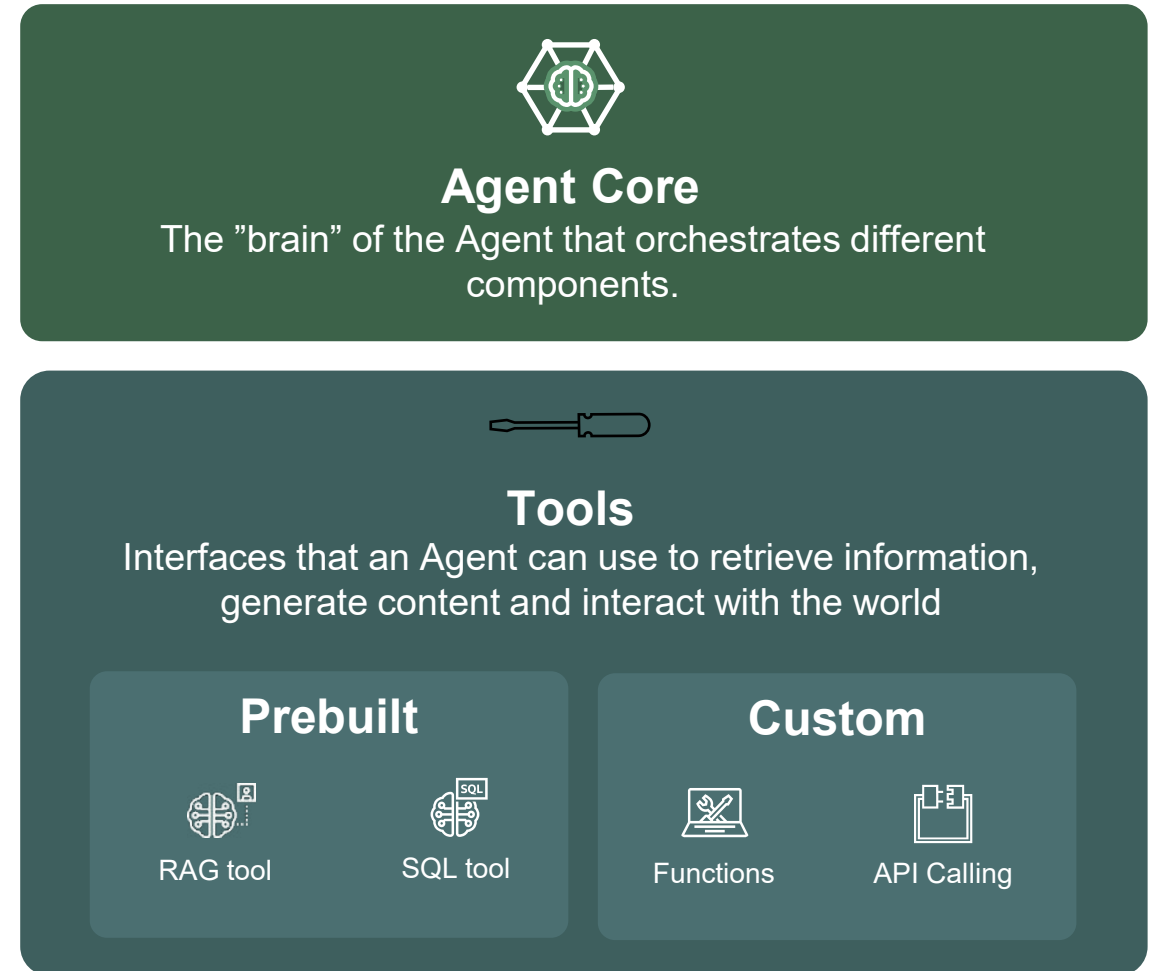




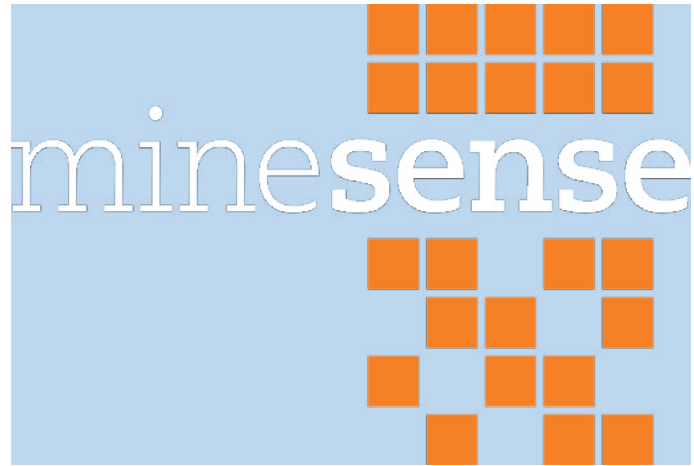
 OpenAI

Accelerate AI agent development with a complete platform

- ✓ Simplified development of AI Agents for faster time to market
- ✓ Customization and flexibility to address your unique business needs
- ✓ Build a foundation for Agentic AI and access tools needed for the next frontier of AI innovation



OCI AI Agent Platform

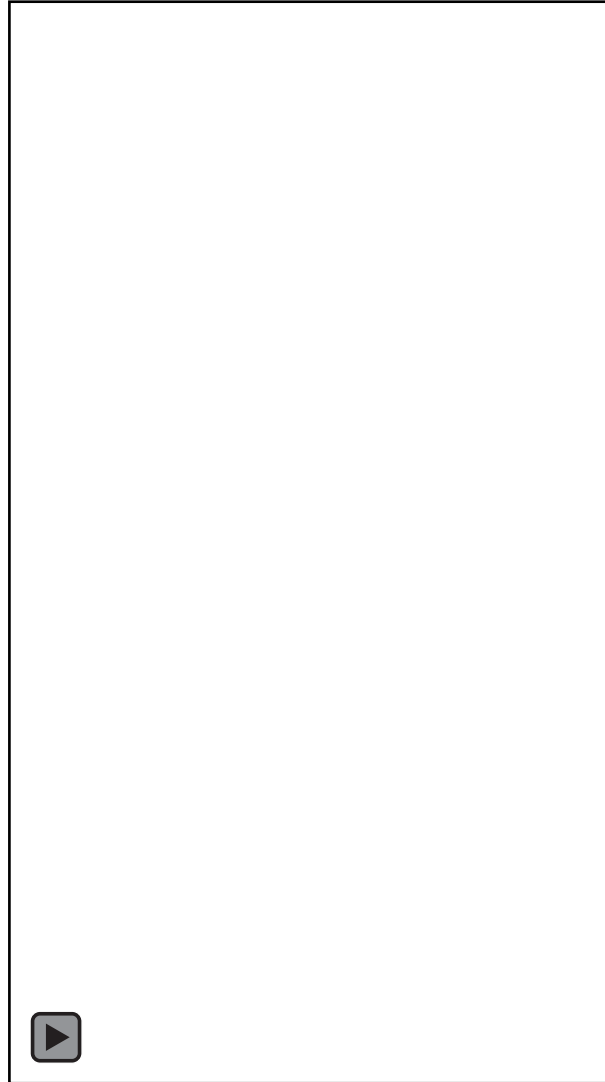


**Improved engineering
productivity for equipment
issues by 5X**

OCI AI Agent Platform

**Lowered DBA
workloads by 65%**

Oracle Autonomous Data Warehouse



ORACLE
Fusion Applications

Agents in the apps you're already using...!

50+

AI Agents across the
Fusion Application Suite



Designed to help you streamline operations, improve decision-making, and reduce manual workloads

Functional Agents

Act as digital assistants for different personas within the enterprise and perform domain-specific tasks.

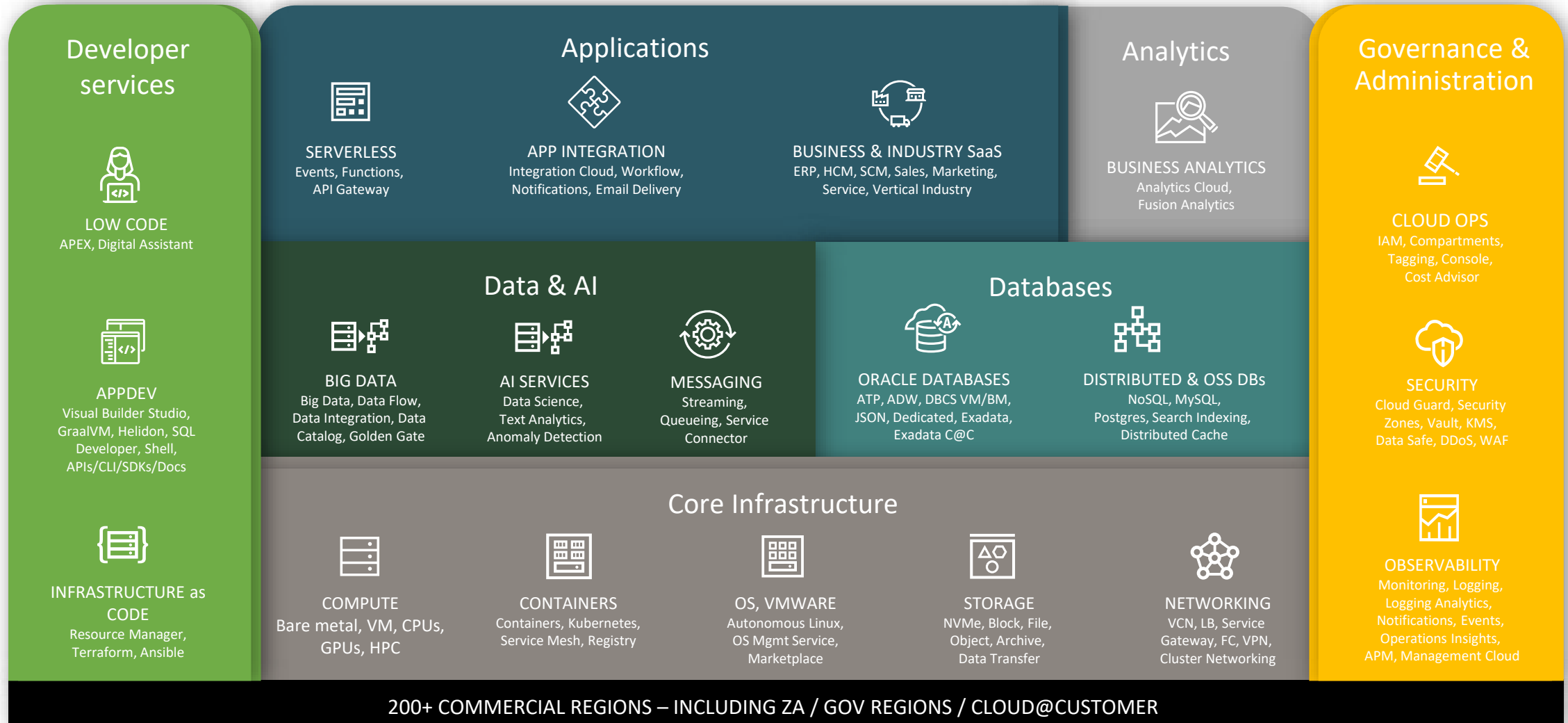
Supervisory Agents

Manage other agents, overseeing complex workflows and making decisions on whether human intervention is needed.

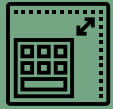
Utility Agents

Perform routine, low-risk tasks such as retrieving data, sending notifications, or running reports.

Oracle Cloud Infrastructure Portfolio: Powered by Supercluster



Those same OCI Superclusters power your AI infrastructure...!



ONLY

Hyperscaler with bare metal GPUs for AI training



Purpose-built RoCEv2 networking at

3.2 Tb/s



High performance file storage with

100x

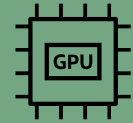
performance of standard file storage



131,072

GPUs

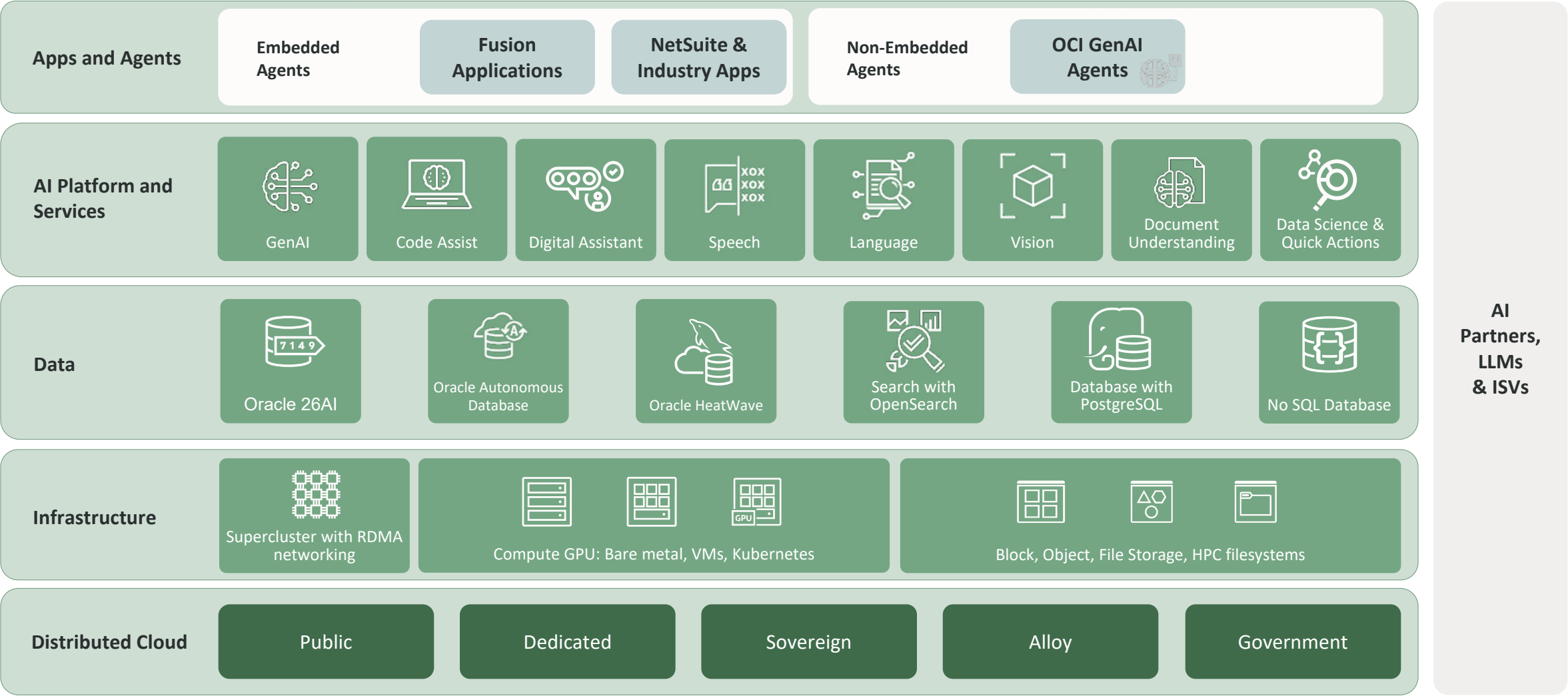
in one OCI Supercluster



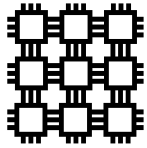
15

Varieties of GPU Shape options available, from 24 GB to 13.5 TB of GPU cluster memory

Comprehensive AI services, data platform, and AI infrastructure



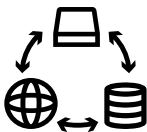
Faster, scalable AI infrastructure to help you get your products to market



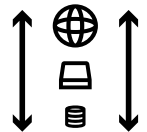
RDMA (remote direct memory access)
for AI/ML workloads at scale



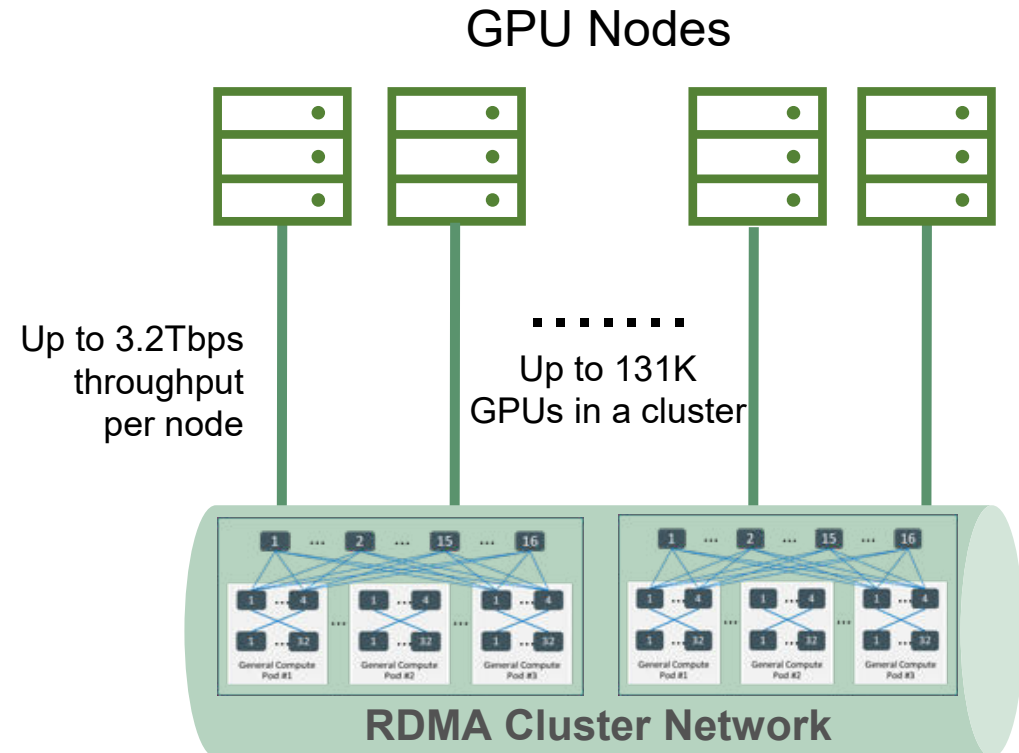
Dedicated, non-blocking networking
with no oversubscription



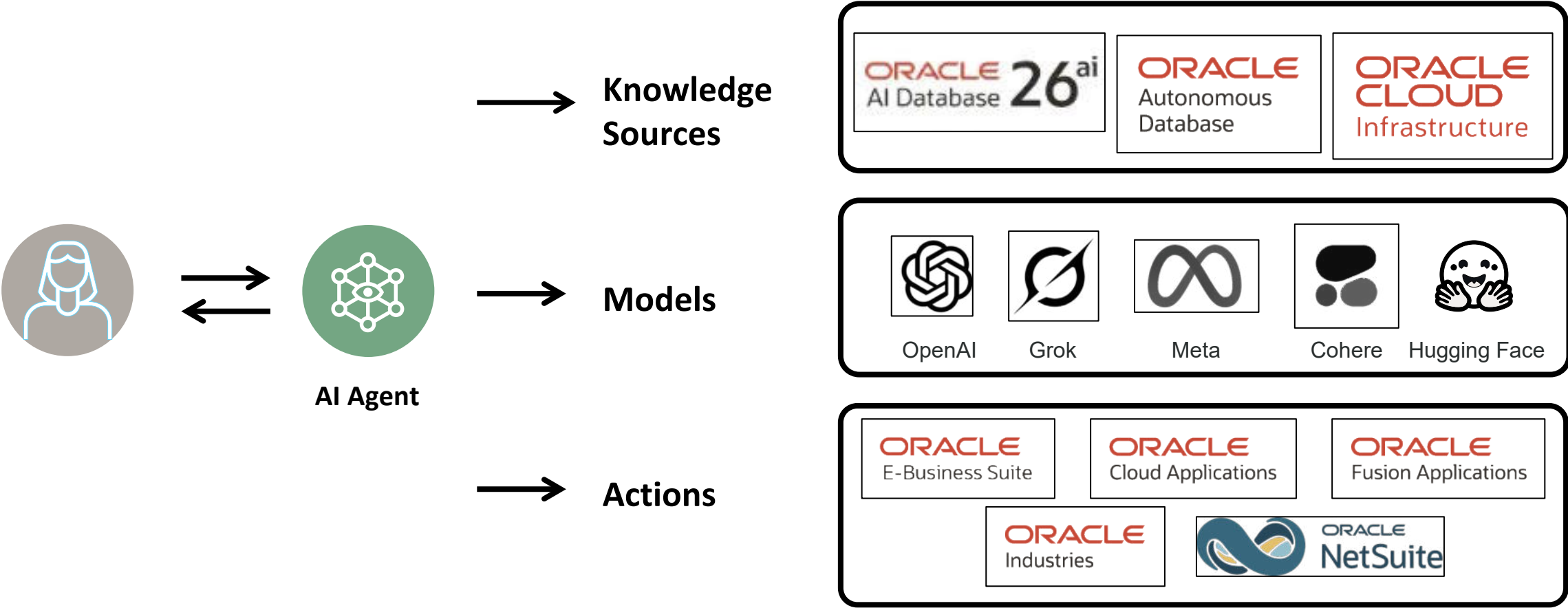
Ultra-low latency and high throughput



Features for Quality of service (QoS)
and intelligent traffic distribution



Oracle is uniquely positioned to help you achieve your AI Aspirations



Deploy AI services effortlessly with easy-to-use packaged features



Generative AI

Fully managed service for using generative AI with customizable models from Meta and Cohere via API.



Language

Extract insights from text data through natural language processing capabilities like classification and entity recognition.



Speech

Convert audio into text and vice versa, enabling speech-to-text and text-to-speech applications.



Vision

Analyze images and videos to extract information and insights using computer vision techniques.



Document Understanding

Extract structured data from documents, automating processes like invoice processing and expensing.



Digital Assistant

Build conversational AI experiences for customers and employees using voice and text interactions.

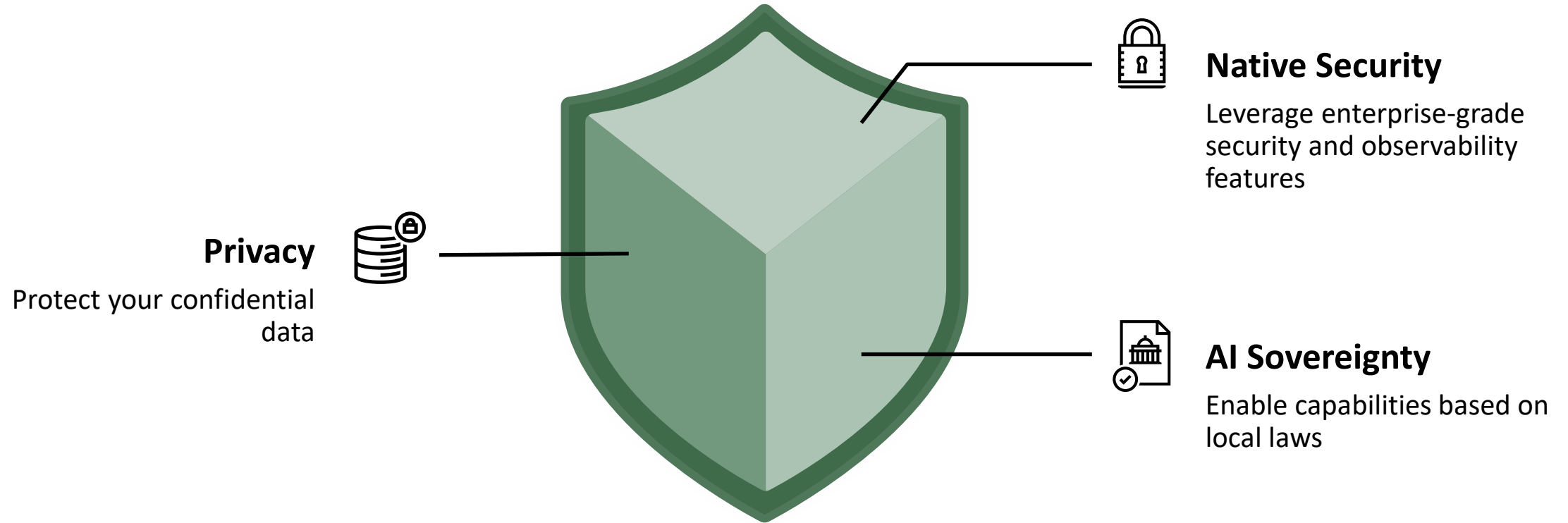
New



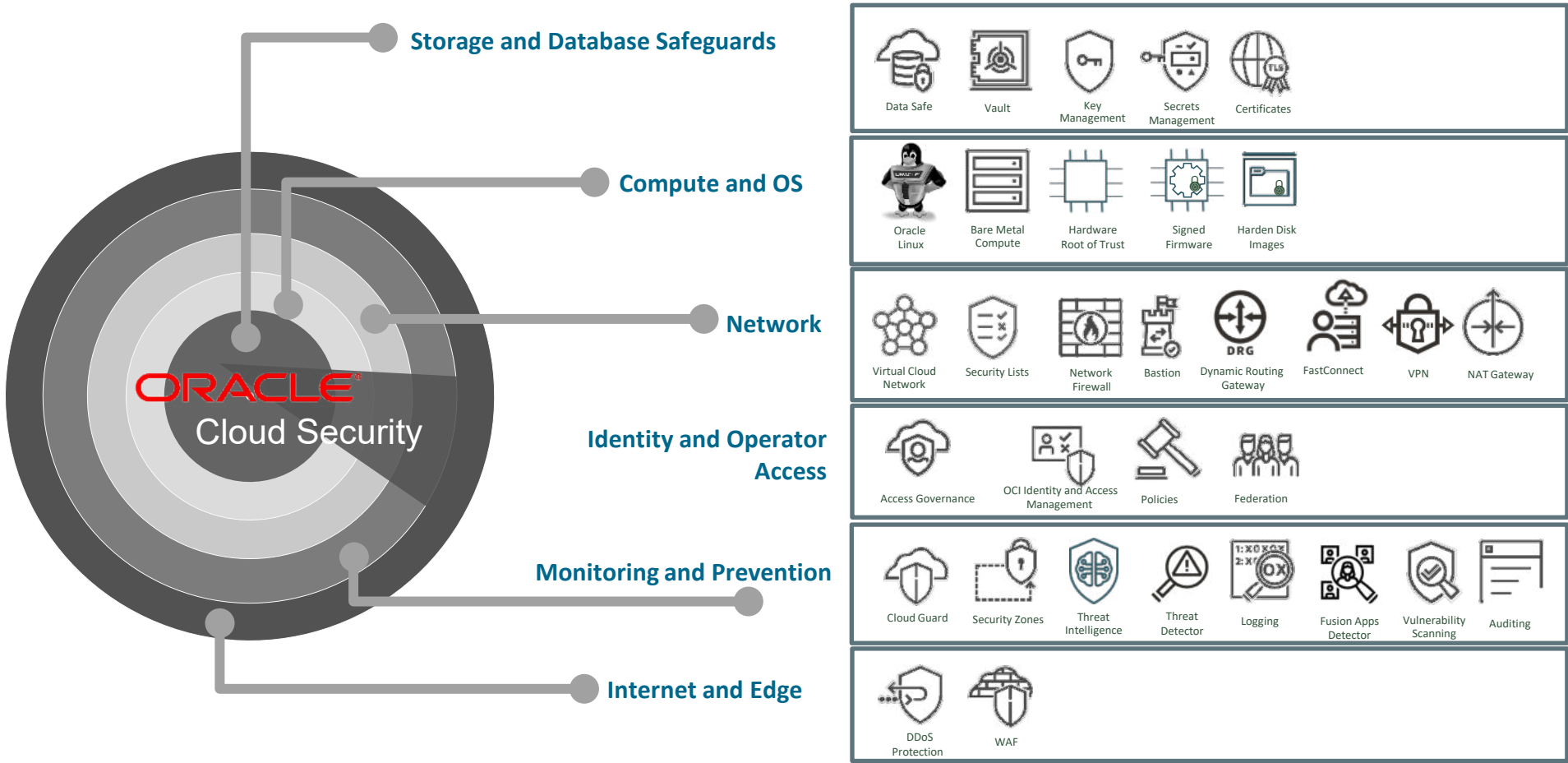
Generative AI Agents

Fully managed service to build and deploy AI agents capable of securely connecting to your data sources.

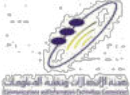
Effective security is multi-faceted: Particularly in context of AI



Effective OCI Security Palette is proactive, adaptive and context aware



Over 80 compliance programs across regions and industries:

REGIONAL			GOVERNMENT			INDUSTRY		
								
GDPR [EU]	PIPEDA [Canada]	ENS [Spain]	DoD DISA SRG IL5	JAB P-ATO	CJIS	HIPAA	PCI DSS – Level 1	HITRUST CSF
								
BSI C5 [Germany]	ISMS [Korea]	NISC [Japan]	EU Model Clauses	LGPD	VPAT-Section 508	TISAX	FINMA	BACEN
								
CITC [Saudi Arabia]	Cyber Essentials Plus [UK]	IRAP [Australia]	Canada Protected B	G-Cloud 12	NIST	EBA	GxP	FIEC

GLOBAL


SOC 1 : SOC 2 : SOC 3

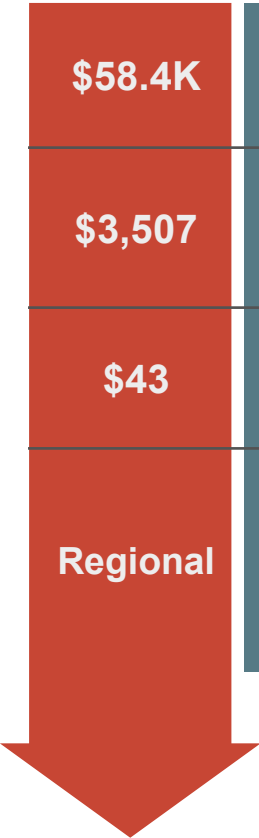

9001 : 27001 : 27017 :
27018 : 27701: 20000-1


Level 2

Oracle's unique approach to AI **for your organization**

- Performance and value for AI infrastructure
- Generative AI across every layer of Oracle's tech stack
- Tight control over data security and compliance

	Global	US East			London			Tokyo			Brazil		
	ORACLE	aws	Microsoft Azure	Google Cloud	aws	Microsoft Azure	Google Cloud	aws	Microsoft Azure	Google Cloud	aws	Microsoft Azure	Google Cloud
GPU instance (NVIDIA H100, Monthly)	\$58.4K	1.5x*	1.6x*	1.1x	1.6x	1.5x	N/A	1.5x	1.8x	N/A	2.1x	N/A	N/A
Kubernetes Cluster (192 vCPU, 1536 GB RAM, Monthly)	\$3,507	2x	2x	2x	3x	3x	2x	3x	3x	2x	4x	4x	3x
Block Storage (1x1TB, 15K IOPS, 125 MB/s, Monthly, balanced performance)	\$43	5x	5x	4x	6x	6x	5x	6x	6x	5x	9x	9x	6x
Public Bandwidth Transferred Out (50 TB, Monthly)	Regional	13x	10x	10x	13x	10x	10x	5x	4x	4x	7x	5x	5x



Price Performant regardless of your chosen region...

* San Jose and California regions as of May 2025



Uber uses OCI for 14 million predictions/second

- Uber modernized its application tier and **AI infrastructure** with thousands of microservices
- Dozens of AI models run on OCI for **inference**
- AI Infrastructure **speeds responses** to trip requests, reducing latency and optimizing routes
- Can easily **scale AI infrastructure** with future growth

“Collaborating with Oracle has allowed us to innovate faster while managing our infrastructure costs. With OCI, our products can run on best-of-breed infrastructure that is designed to support multicloud environments and can scale to support profitable growth.”

**Kamran Zargahi, Senior Director,
Tech Strategy and Cloud Engineering, Uber**



Uber

[Read the press release](#)

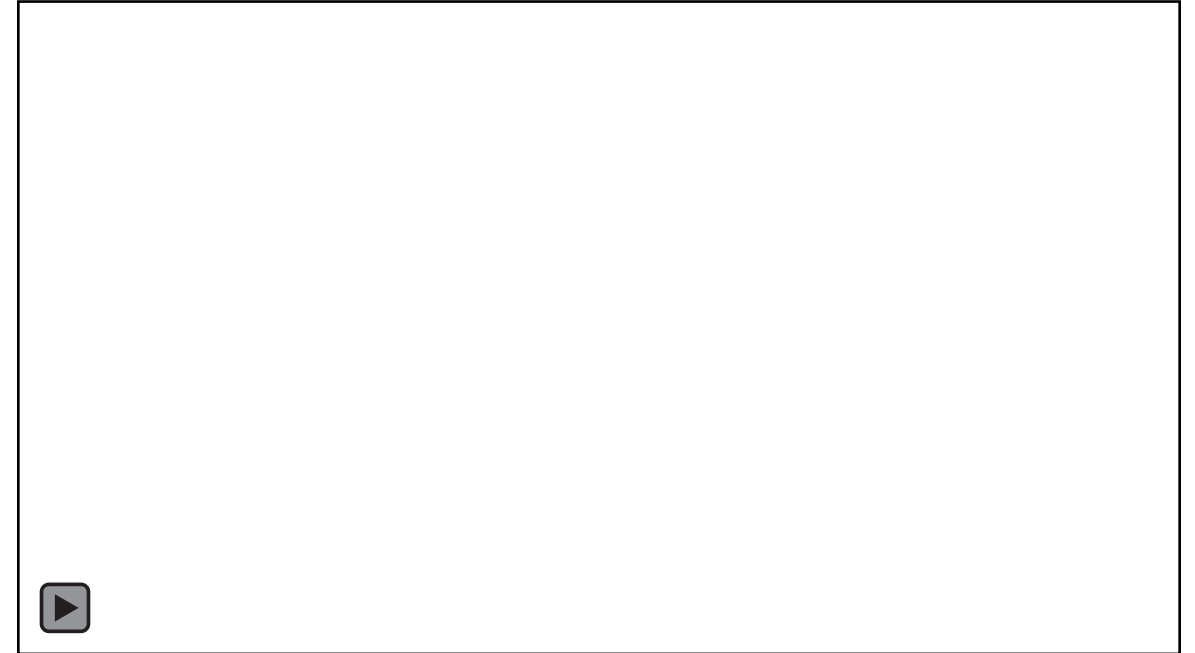
Zoom Uses OCI to Power its AI-first Platform in Saudi Arabia

- Zoom AI Companion uses OCI unique GPU offering for inferencing, data, and AI sovereignty capabilities
- Zoom optimized its AI models to run on efficient OCI GPU shapes with top-tier performance while staying compliant with Saudi Arabian regulations

“By optimizing AI Companion to operate efficiently with **GPU shapes in a local OCI region**, we’re enabling Saudi companies to take full advantage of AI without facing constraints.”

Velchamy Sankarlingam, President of Product and Engineering, Zoom

[Read the press release](#)



Miranda Nash, GVP, Oracle and XD Wang, CTO, Zoom



Protecting communities with AI-driven climate resilience solutions

- RSS Hydro models and maps extreme weather and climate risks using OCI, **NVIDIA GPUs**, and **NVIDIA Omniverse**
- Uses Omniverse 3D visualization for increasing disaster preparedness and impact awareness

Examples: 2020 Kumamoto flooding; simulation of extreme rainstorm events and sea-level rise

Up to 6x the speed

For fine-tuning and training on NVIDIA L40S on OCI vs. on-prem workstation

Up to 2.5x the speed

For inference on NVIDIA L40S on OCI vs. on-prem workstation



RSS-Hydro



“After only a few months of using OCI’s Oracle Database 23ai and Generative AI, our AI agent could answer 80% of end-user questions.”

Jan-Peter Meyer
CTO, Brain4Data



Business Challenge

Brain4Data is a German startup focused on combining AI and RPA for small and midsize businesses to make recommendations for their employees or customers. To provide specific information and context that a human employee would have, Brain4Data needed to develop a GenAI agent that uses Large Language Model (LLM) capabilities, such as text generation, based on the specific knowledge scattered throughout a customer’s documents, databases, and ERP systems to fine-tune the agent’s responses.

Results

Brain4data chose to develop its technology on Oracle Database because of the data simplicity it provides. The company was already using Oracle Database as a document repository, and recognized how the native AI Vector Search capabilities of Oracle Database 23ai simplified the semantic search portion of RAG.

- ✓ Developed GenAI-powered agent Fred, who handles 80% of end-user requests
- ✓ Create a rich graphical user experience tailored for each customer with Oracle APEX
- ✓ Won the European Seal of Excellence, a prestigious honor awarded by the European Innovation Council (EIC) of the European Union for their GenAI innovation

[Read story](#)





Munich Re HealthTech improves insurance analytics with AI

“Oracle’s AI innovation enhances decision-making, efficiency, and accessibility across global markets, making this a game-changer in insurance portfolio management. Actuaries, underwriters, and executives can query data in natural language, receiving real-time insights, KPI analysis, and predictive recommendations without IT intervention.”

Antonis Roussos

Head of Analytics, Munich Re HealthTech



Business Challenge:

Munich Re HealthTech (MRHT) provides solutions for health insurance companies and third-party administrators to assess portfolio and policy performance. As demand grew across multiple countries with increasing data volumes and data sources, the company needed to meet customer’s data residency concerns and rapidly deliver more advanced analytics.

Results:

Migrated to Oracle Autonomous Database 23ai to take advantage of low-latency data access, top performance, and low code development

- ✓ Built AI chatbot with Oracle AI Vector Search, Select AI, APEX and OCI GenAI to search its knowledge base for documentation and advanced data analysis
- ✓ 90% of free text questions accurately answered with next -level efficiency and insight in seconds and reduced IT dependency and development costs
- ✓ Build analytical dashboards from 15 days to 20 minutes while also reducing human error
- ✓ Oracle Globally Distributed Database allowed data access multiple regions in a single instance for faster reporting times and to address data residency
- ✓ Better portfolio management, stronger compliance, and real -time insights for optimized pricing, risk assessment, and underwriting strategies

[Read story](#), [Technical Case Study](#), [Watch Video](#)



Q&A