

POWERING FUTURE BUSINESSES WITH ALIBABA CLOUD

Alibaba Group Milestones

With continuous development and innovation, it has formed all-round infrastructure in the digital economy era, from business to finance, to logistics, to cloud computing



Alibaba Group Business



Alibaba Cloud Global Infrastructure. No. 1 in APAC

30

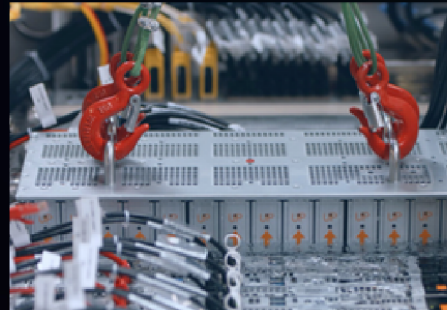
Alibaba Cloud Regions

89

Availability Zones

3200

CDN Nodes



Green Data Centers

Water Cooling, Natural Air Cooling and Liquid Cooling contributes to 1.0 PUE.



Multiple Security Measures

24*7 security monitoring and maintenance system, complies with local data regulations and laws



High Availability

Built with redundant power supplies, networks and connections, higher error tolerance capabilities



Source: alibabacloud.com, as of 2023.8.10

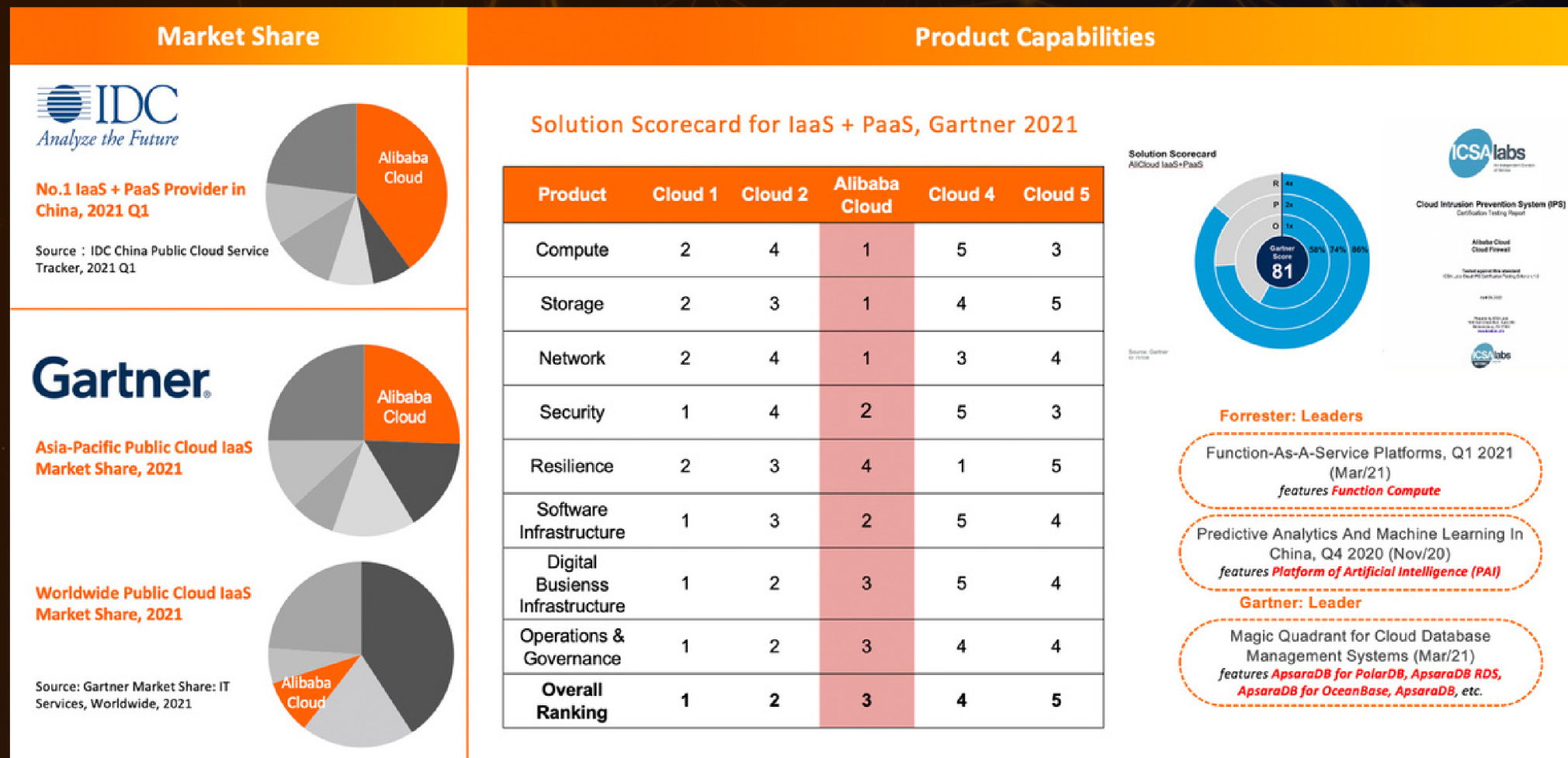
Our Customers Worldwide 4,000,000+





Alibaba Cloud
Worldwide Cloud Services Partner

Global Leader in Cloud Technology



Comprehensive Product Offerings

200+

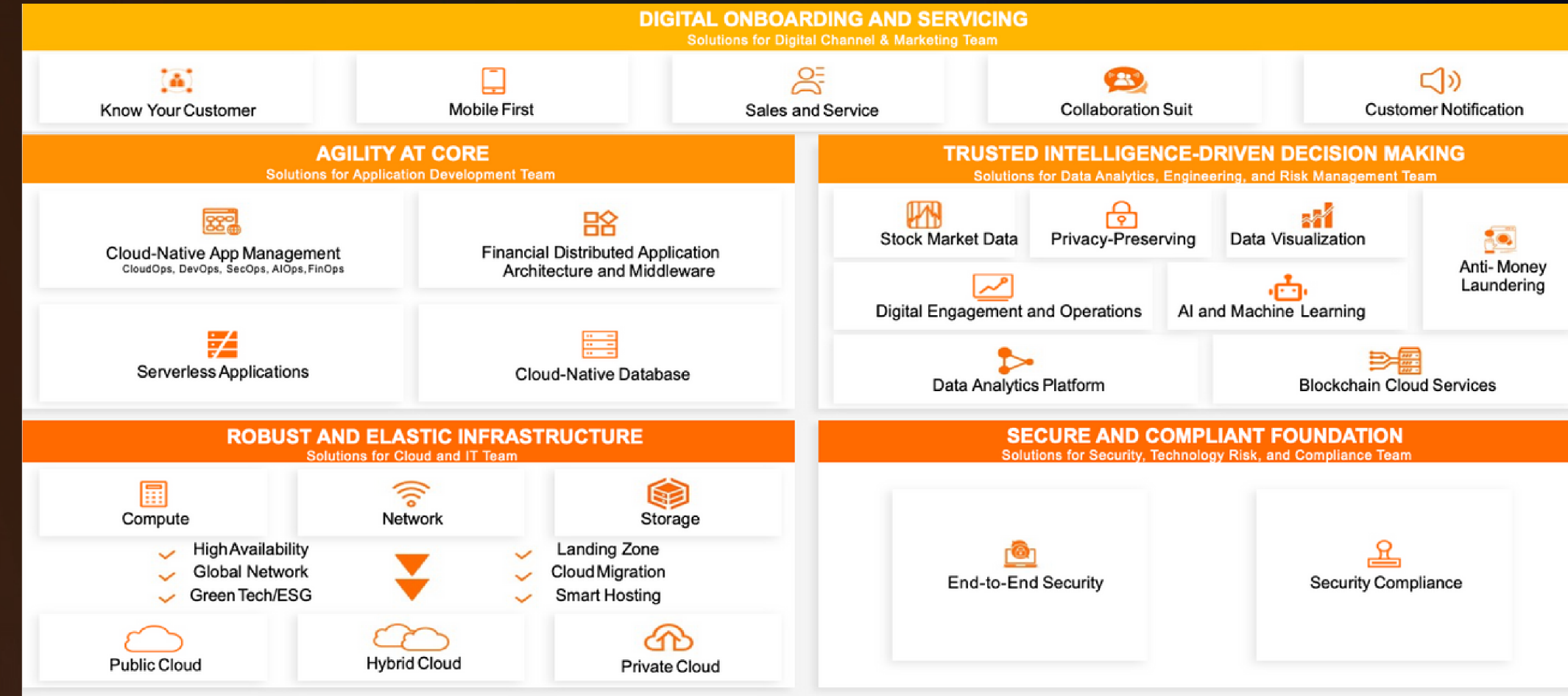
Products in International Market

Industry & Application Engine	<div>Media</div> <div>ApsaraVideo Live</div> <div>ApsaraVideo VOD</div> <div>Media Processing</div> <div>Dynamic CDN</div>	<div>Middleware</div> <div>Enterprise Distributed Application Service</div> <div>Message Queue (RocketMQ/RabbitMQ)</div> <div>Application Configuration Management</div> <div>Application Real-Time Monitoring Service</div>	<div>Management</div> <div>Action Trial</div> <div>Open API Explorer</div> <div>Terraform</div> <div>Cloud Shell</div> <div>Cloud Monitor</div> <div>Cloud Config</div> <div>Resource Access Management</div>		<div>Mobile</div> <div>mPaaS</div> <div>Zoloz eKYC</div> <div>SMS/ChatApp</div> <div>Mobile Testing</div>	<div>Enterprise Service</div> <div>Domains</div> <div>Alibaba Mail</div> <div>Web Hosting</div> <div>API Gateway</div>	<div>Security</div> <div>SDDP</div> <div>Application Security (WAF/Fraud Detection/Content Moderation)</div>		
	Data Intelligence	<div>Database</div> <div>Redis</div> <div>RDS for MySQL</div> <div>RDS for SQL Server</div> <div>RDS for PostgreSQL</div> <div>RDS for ClickHouse</div> <div>MongoDB</div> <div>DTS</div> <div>RDS for MyBase</div> <div>PolarDB PostgreSQL</div> <div>PolarDB MySQL</div>	<div>Big Data</div> <div>DataV</div> <div>Quick BI</div> <div>Data Lake Analytics</div> <div>Dataphin</div> <div>DataWorks</div> <div>DataHub</div> <div>Elastic Search</div> <div>E-MapReduce</div> <div>MaxCompute</div> <div>Realtime Compute</div> <div>Hologres</div> <div>Data Integration</div>		<div>AI</div> <div>Intelligent Service Robot</div> <div>Machine Learning Platform For AI</div>		<div>Private Cloud</div> <div>Apsara Stack Enterprise Edition</div> <div>Apsara Stack Agility Standard Edition</div>	<div>Data Security (KMS/HSM/CloudHSM)</div> <div>Network Security (Anti-DDoS/Secure DCDN/Cloud Firewall)</div>	
		Infrastructure Services	<div>Computing</div> <div>ECS</div> <div>Auto Scaling</div> <div>Container Service</div> <div>Container Service for Kubernetes/ASK</div> <div>Elastic Container Instance</div> <div>Container Registry</div> <div>E-HPC</div> <div>Bare Metal Instance</div> <div>Function Compute</div> <div>Batch Compute</div> <div>Dedicated Host</div>	<div>Networking</div> <div>Virtual Private Cloud</div> <div>Express Connect</div> <div>NAT Gateway</div> <div>Server Load Balance</div> <div>VPN Gateway</div> <div>Cloud Enterprise Network</div> <div>Smart Access Gateway</div> <div>Global Acceleration 2.0</div> <div>Shared Traffic Package</div> <div>Shared Bandwidth</div>		<div>Storage</div> <div>Object Storage Service</div> <div>Block Storage</div> <div>NAS</div> <div>Cloud Storage Gateway</div> <div>Hybrid Backup Recovery</div>		<div>SOFAStack</div> <div>zStack</div> <div>Cloud Box</div>	<div>Security Center</div> <div>Security Management (Managed Security Service)</div>

More Than Just Cloud

Financial Services				Retail		Manufacturing		Internet		Culture, Media & Sport		
Banking	Insurance	Securities & Capital Markets	FinTech	New Retail	eCommerce	Automotive	Chemical/ Healthcare/ Others	Gaming	Web3.0 / Metaverse	Media/ Entertainment	Sport/ Culture	
<ul style="list-style-type: none">• Remote Account Opening eKYC• AML• Mobile Application Platform• Mini-Program• Remote Video Service Center• Digital Bank on Cloud• DingTalk Collaborative Platform• Secure Backup• Privacy-Preserving Computing• Oracle Replacement• Teradata Migration• Carbon Credits	<ul style="list-style-type: none">• Remote Account Opening eKYC• AML• Mobile Application Platform• Remote Video Service Center• Privacy-Preserving Computing• Cross-Region Network• SMS Verification & Marketing• Financial Data Platform	<ul style="list-style-type: none">• Cloud Quotation• Remote Account Opening eKYC• AML• DingTalk Collaborative Platform• Trading System on Cloud• Quantitative & High-Frequency Trading (HFT)• Cross-Region Network• Cloud Data Warehouse• Stock Exchange Connectivity	<ul style="list-style-type: none">• Alipay+ Digital Wallet• Blockchain as a Service• Remote Account Opening eKYC• Mobile Application Platform• Quantitative & High-Frequency Trading (HFT)• Cross-Region Network• SMS Verification & Marketing• IPFS E2E solution	<ul style="list-style-type: none">• Cloud POS• Automated Warehouse• Smart Dispatch• Cainiao OWTB Fulfillment System• Consumer Behavior Analytics• Cloud Electronic Shelf Label• Data Mid-End• Business Mid-End• Loyalty & Campaign	<ul style="list-style-type: none">• Image Search• Text Search• AI Recommendation• Cross-Border Online Store Builder• AR/VR Virtual Store• Social Commerce• Digital Ads• eCommerce on Cloud• Consumer Behavior Analytics• Device Security• Mobile Testing• SMS Verification & Marketing• Smart Customer Service	<ul style="list-style-type: none">• Autonomous Driving• Internet of Vehicles• Smart Data Labelling• Big Data Engine & Analytics• AI Modelling & Training Platform• Data Management• IoV Data Platform• Mapping and Local Lifestyle Services• Privacy-Preserving Computing	<ul style="list-style-type: none">• Smart Supply Chain (Forecast, Safety Stock, Replenishment)• Smart Distribution• Production Safety• DICOM on Cloud• Medical Image Processing• Green Manufacturing & ESG• Cainiao OWTB Fulfillment System	<ul style="list-style-type: none">• Game Player Operation & Management• Cloud Gaming• Data Analytics• Game Security• Global Gaming Server• Gaming Acceleration• Cloud Rendering• Cloud Native AI• Real-Person Authentication & eKYC	<ul style="list-style-type: none">• VR/AR/XR• NFT & Blockchain• Digital Avatar• Cloud Rendering• Cloud Native AI• Edge Computing• Content Acceleration	<ul style="list-style-type: none">• User Operation & Marketing• Content Moderation• Media Processing• Live Streaming• VOD & Short Video• RTC (Real-Time Communication)• Narrowband HD• Edge Computing• Content Acceleration• Object Detection• AI Recommendation• AI Beautification SDK• Live Media Gateway• Holographic Interview• High Degree of Freedom Video• Media Assets Management	<ul style="list-style-type: none">• Venue AI Modeling• Venue Panorama Tour• Venue Virtual Navigation• Event Visual Simulation• Event Online Press Conference• Event Dingtalk• Smart Event Guidebook• Staff/Volunteer AI Training• Sport Data Feeds• IOT Smart Device• Digital Ticketing• Fan Zone• Cloud Expo for Sponsorship Activation• Digital Merchandise• Sports AI• Video Assistant Referee	
Common Solutions	Cloud Native				Cloud Migration		Backup & Restore	Security & Compliance		Data Analytics & AI		
	<ul style="list-style-type: none">• Micro Services• Serverless• K8S	<ul style="list-style-type: none">• DevSecOps• FinOps• Application Distribution	<ul style="list-style-type: none">• Zero Trust• Log/Monitor/Trace• Automation	<ul style="list-style-type: none">• Low Code• Multi-Cloud Management	<ul style="list-style-type: none">• Cloud to Cloud• On-Premise to Cloud• Multi-Cloud	<ul style="list-style-type: none">• IT Governance & Landing Zone & CAF	<ul style="list-style-type: none">• SAP on Cloud• VMware on Cloud	<ul style="list-style-type: none">• Backup• Disaster Recovery	<ul style="list-style-type: none">• Security Protection• Continuous Compliance	<ul style="list-style-type: none">• MLPS 2.0 Compliance• Data Security• Privacy Protection	<ul style="list-style-type: none">• AI Competency Center• AI Acceleration	
Ecosystem	Marketplace		Partners & ISV		DAMO		DingTalk		Ant	eCommerce		Amap

Digital-Native Capabilities Built for Financial Services



Digital Transformation Journey for Retailers



CLOUD FIRST, SECURITY FIRST

Scalable Infrastructure & Global Connectivity
Adopting cloud native technology to easily scale up or down the infrastructure base on business needs

Protect Personal Data
Securing and protecting end customers' personal data and meeting different regional data residency requirements



EXPAND ONLINE FOOTPRINTS

Build Digital Touchpoint
Establishing retailers' own online touchpoint, start to build private traffic

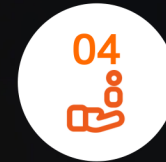
Direct-to-Consumer
Directly reach end customers' via digital channels, helping retailers to build better brand image and customer service



OMNI-CHANNEL OPERATION

All Channels as One
Omni-channel platform helps retailers to manage SKUs, inventory and orders across all sales channels

Online-Merge-Offline Experience
Provide customer a seamless experience across online and offline



SMARTER CUSTOMER ENGAGEMENT

Know Your Customer
Use retailers' direct-to-customer channels and data analytics to effectively segment their customers

Engage Your Customer
Find the right time, reach to the right customer, offer them the right things, through the right channel

Maintain Your Customer
Enhance the ability of tracking and analyzing the full lifecycle of your customer journey



SMARTER SUPPLY CHAIN

Digital Supply Chain
Holistic solutions seamlessly covering your warehousing and logistics business by engaging OWTB (OMS, WMS, TMS, BMS)

Smart Warehouse
Comprehensively automate your warehouse and DC by engaging the integrated hardware & software solutions powered by Cainiao

Smart Resources Dispatch
Dispatch your resources in an efficient and economic way by leveraging big data and AI technology



Generative AI : The Next Inflection Point For Digital Transformation

Driven by Increased Computational Power, Data Growth, Model Advancements

58%

58% of Chinese Companies Have
Used Artificial Intelligence
Technology

40%

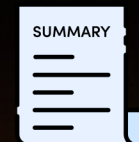
Over 40% of State-owned
Enterprises Have
AI Technology R&D Institutions

67%

67% of IT leaders surveyed said
they have prioritized generative AI
for their business within the next
18 months.



Text Generation



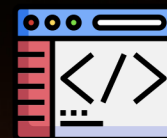
Text Summarization



Q&A



Search



Code Generation



Image Generation



Audio Generation



Video Generation



Communications

AI Chatbot , Q&A , Search



Retail

Product and Pricing Analysis,
Marketing Content Generation



Supply Chain

Logistics Planning, Demand
& Production Forecasting



Media & Gaming

Video & Image Generation,
Content upscaling



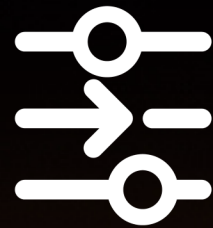
Financial Services Industries

Risk Management,
Fraud Detection



Healthcare & Life Sciences

Drug discovery, medical research,
personalized patient treatment



Publicly Available Models

Do not contain Enterprise Data and possesses outdated Knowledge Base



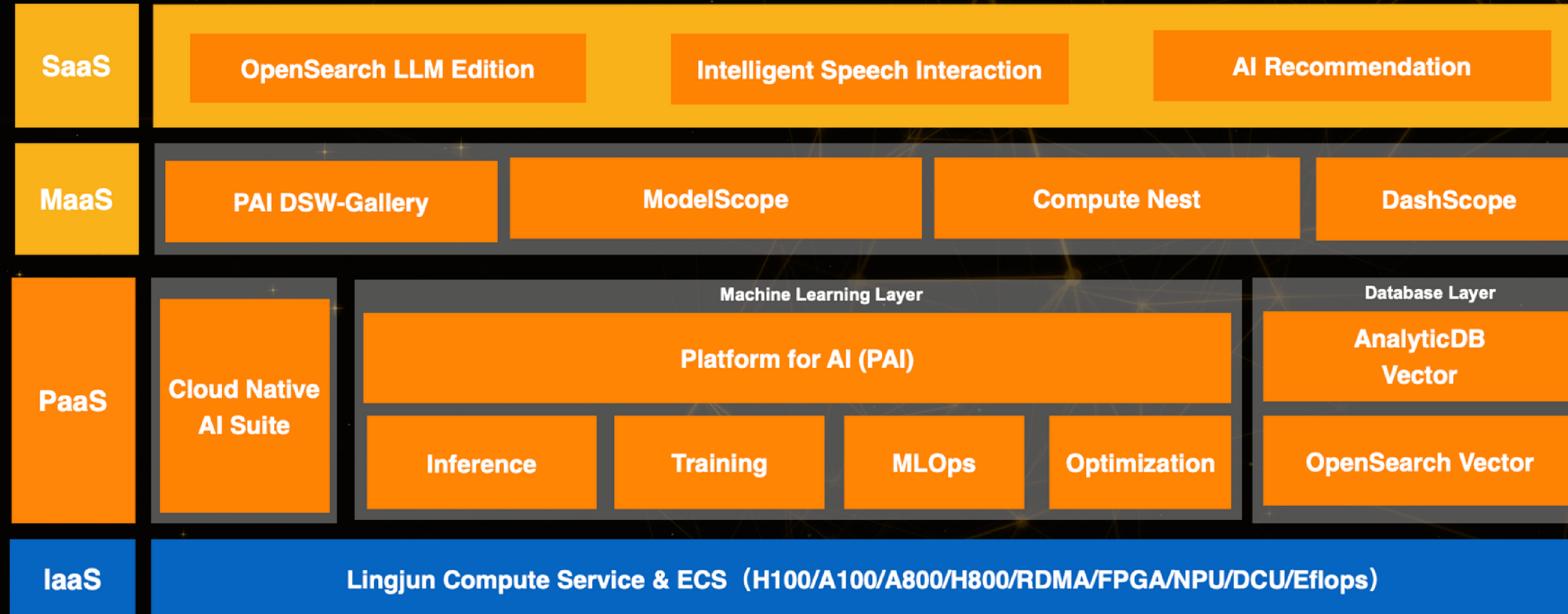
Unfamiliar with Generative AI Technology

Expensive, time-consuming and requires deep expertise



Inaccurate & Slow Response

LLM might not be optimized and have slow inference and response speeds





Cost Effective

Leverage on economics of scale and flexible pricing models to train and run FMs for Generative AI



Scalability

Scaling capabilities of the cloud enables agility and resource optimization



Security

Data remains within your own environment and remains private and confidential



Easy to Use

Low/No code solutions allow customers to quickly and easily adopt AIGC solutions



[首页](#)
[模型库](#)
[数据集](#)
[创空间](#)
[文档中心](#)
[动态](#)
[讨论区](#)
[GitHub](#)

[登录 / 注册](#)

热门专题 汇聚热门行业和技术资讯，包括技术趋势、应用场景、模型推荐、实践案例等

[LLM专题页](#)
[智能通行](#)
[基础视觉模型及...](#)
[信息抽取](#)
[零样本学习](#)
[全部专题 >](#)

全部模型 [按任务](#) [按组织](#)

热门任务

[文本生成](#)
[说话人确认](#)
[文本生成视频](#)

计算机视觉 [展开更多](#)

[视觉检测跟踪](#)
[光学字符识别](#)
[人脸人体](#)

[视觉分类](#)
[视觉编辑](#)
[视觉分割](#)

自然语言处理 [展开更多](#)

[文本分类](#)
[文本生成](#)
[分词](#)

[命名实体识别](#)
[翻译](#)
[文本摘要](#)

语音 [展开更多](#)

[语音识别](#)
[逆文本正则化](#)
[语音合成](#)

[支持模型体验](#)
[支持模型训练](#)
[支持模型部署](#)

[综合排序](#)

姜子牙2-13B-Chat

Ziya2-13B-Chat

[文本生成](#)
[transformer](#)
[封神榜](#)

Ziya2-13B-Chat是基于Ziya2-13B-Base的对话模型，在30万高质量的通用指令微调数据以及40万知识增强的指令微调数据上进行了有监督训练，并且在数万条高质量人类偏好数据训练的奖励模型上进行了全参数的人类反馈强化学习训练。

[封神榜](#)
2023.10.20 更新
599
11

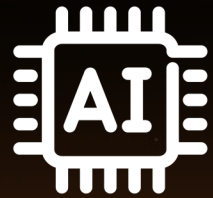
通义千问-14B-Chat

Qwen-14B-Chat

[qwen](#)

通义千问-14B (Qwen-14B) 是阿里云研发的通义千问大模型系列的14亿参数规模的模型。Qwen-14B是基于GPT-4架构的大语言模型，在预训练和微调数据上进行了全参数训练。

One-stop AI-powered conversational search solution for effortless adoption of Generative AI



Extends power of
Large Language Models
on your **own knowledge base**
with high accuracy (>70%)



**Easy Maintenance &
Integration** with
APIs and GUI



Data Security : Data Storage,
Processing occurs within own
environment



Flexibility : Able to switch
between supported models
with ease. Supports
Parameter Optimization and
Manual Intervention



Lower Costs :
No need to maintain/finetune
own models

OpenSearch LLM-Based Convers

OpenSearch / Instance Management

Instance Management

Create Instance

Initiative Alert

application name

Search by application name

Filter by Tag

	Application Name/Instance ID	Tag	Status	Storage Usage (GB)	PVs	Number of Documents	Actions
<input type="checkbox"/>	test_instance ops-sg-vfp3g0zju01		✓ Normal	0/1	14	0	Manage Q&A Test ⋮
<input type="checkbox"/>	llama2test2 ops-sg-dza3dqe6q01		✓ Normal	0/1	0	4	Manage Q&A Test ⋮
<input type="checkbox"/>	llama2test ops-sg-ug83c7is602		✓ Normal	0/1	0	25	Manage Q&A Test ⋮

☐

Edit Tag

Remove Tags



Thank You