

b

Au





Section 01

.

Section 02

.Z

Section 03

.

Section 04

W

.

Section 05

Z

.



12 In-Depth Interviews

4 Related Fields

Automobile · Robotics · E-commerce · Design

k



Product Manager at
XPENG Motors

@Shanghai



CMO at Black Sesame
Technologies

@Shenzhen



Senior Product
Manager at Galbot

@Shanghai



Investment Director
at MIRAE ASSET

@Shenzhen



Senior Operations
Expert at Alibaba

@Hangzhou



Senior Data Specialist
at Prottime

@Shanghai



b

Owner of XIAOMI
SU7 ULTRA

@Shanghai



Owner of XPENG X9

@Shanghai



General Manager of
Styling Center at
Jomec

@Shanghai



Product Designer at
Thermos

@Shanghai



Glass Art student at
Tsinghua University
Academy of Arts & Design

@Beijing



Industrial Design
student at Tongji
University

@Shanghai

u h

Section 01

.

Section 02

.Z

Section 03

.

Section 04

W

.

Section 05

Z

.

u h

o “(W“ | ’(o | ’“(:Z (h —: :b| “((g :b| “

High-level functions (highway NOA, city NOA) are no longer add-ons but decisive factors in car purchases.



v | (o “(Y’ “((a ~| (V “

While the U.S. still holds a lead in foundational AI models, China is rapidly catching up and already stands out globally in commercialization, large-scale adoption, and strong policy support.



Td| “” “ (W (f —

Transition from rule-based decision trees to AI-driven reasoning (VLA, world models, hybrids), improving generalization, prediction, and semantic understanding.

b u h

In 2024, China’s passenger vehicle sales reached 22.89 million units, with nearly 2 million equipped with advanced ADAS, bringing penetration to 8.6%. Consumer demand is shifting from “usable” to “refined,” as L2+ highway assistance and urban NOA increasingly become core factors in purchase decisions. Intelligent driving is rapidly evolving from a differentiating feature to a mainstream standard.

At the same time, Robotaxi and Robobus pilots are expanding in Beijing, Shanghai, Shenzhen, and Wuxi. Applications in logistics delivery, shuttle services, and specialized vehicles are also accelerating, demonstrating the frontier potential of L4 automation.

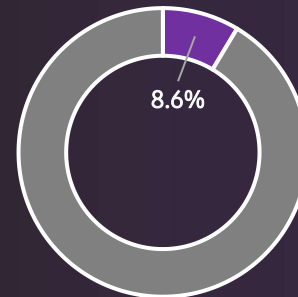
Overall, China’s intelligent driving industry is entering a “commercial deployment phase”: large-scale adoption of L2+ on one side, and scenario-based breakthroughs in L4 on the other. These dual tracks are advancing in parallel, jointly driving the sector’s evolution.

“China’s intelligent driving has entered the commercial deployment phase, marked by rapid tech iteration, faster market penetration, and improving regulations. L2 is widely commercialized, L3 is in pilot and early rollout, while L4 still faces cost and regulatory hurdles.”

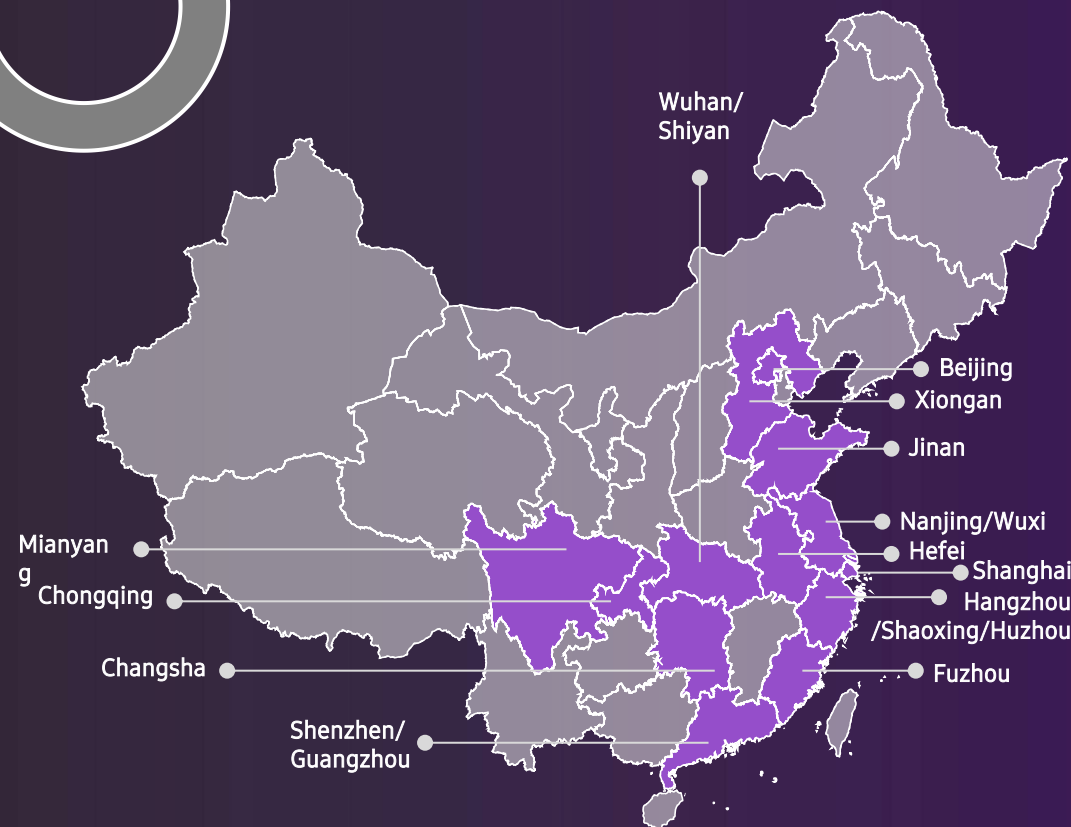


Black Sesame Technologies, CMO

ECEG(TWTm(g | “(j “ “ | (l | “



ECEH(I ~ ~ (O(I ~ | (j



s

u

h

mTY(f“ “ (’(W (T |

f“ “((H



z (T |

f“ “((G



b (T |

小马智行·pony.ai



f“ “((F



v ’ | (T |



XPENG



XIAOMI



NIO



LI AUTO



HUAWEI



BUILD YOUR DREAM

f“ “((E



j | | (T |

阿维塔
AVATR

零跑汽车
LEAPMOTOR

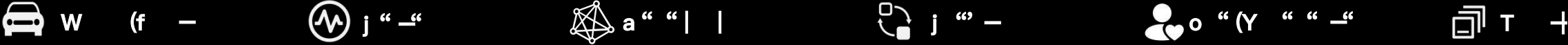
EXEED 星途
EXEED

f“ “((D



w “ (T | —“

Au & Au u h



Rule-based Control
→ Human-like Decisions

2D Signals
→ Multi-dimensional Semantics

Rule Enumeration
→ Data-driven Generalization

Physics-based Forecast
→ Probabilistic Anticipation

Optional Add-on
→ Must-have Companion

Limited Use Cases
→ Diverse Ecosystem

Rule-based, rigid,
long decision chain

2D sensing, isolated data,
heavy rules

Rule enumeration,
weak adaptation to diverse scenes

Physics-based prediction,
unable to infer hidden intent

Add-on feature, low stickiness,
limited interaction

Mainly highway assist
& basic parking

End-to-end learning,
shorter chain, more human-like

Multi-modal fusion,
combining spatial + semantic data

Data-driven by large models,
strong generalization

Probabilistic + semantic reasoning,
human-like anticipation

Becoming standard, high stickiness,
natural interaction

City/Highway NOA, auto-parking,
logistics, Robobus/Robotaxi

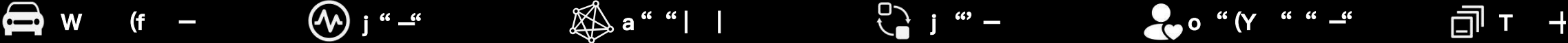
*"In the past, the way to handle generalization—before the era of large models—was to write rules. But at each layer, information would be lost. When humans drive, they don't follow this kind of logic at all."
"Back in 2017–2018, when everyone was working on Robotaxi, the approach was to use LiDAR to detect the vehicle ahead. But humans don't drive like that—for example, you check whether the other car uses a turn signal, or if it has hazard lights on, you keep more distance. This kind of information simply couldn't be processed in the previous dimensions."
"In the past, it was all rule-based, like decision trees. But as you keep adding rules, you find they start to conflict with each other, and the ceiling becomes very hard to break through."*



XPENG Motors, Product Manager of Intelligent Driving

Au u h

Au



Rule-based Control
→ Human-like Decisions

2D Signals
→ Multi-dimensional Semantics

Rule Enumeration
→ Data-driven Generalization

Physics-based Forecast
→ Probabilistic Anticipation

Optional Add-on
→ Must-have Companion

Limited Use Cases
→ Diverse Ecosystem

Rule-based, rigid,
long decision chain

2D sensing, isolated data,
heavy rules

Rule enumeration,
weak adaptation to diverse scenes

Physics-based prediction,
unable to infer hidden intent

Add-on feature, low stickiness,
limited interaction

Mainly highway assist
& basic parking

End-to-end learning,
shorter chain, more human-like

Multi-modal fusion,
combining spatial + semantic data

Data-driven by large models,
strong generalization

Probabilistic + semantic reasoning,
human-like anticipation

Becoming standard, high stickiness,
natural interaction

City/Highway NOA, auto-parking,
logistics, Robobus/Robotaxi

“AI has achieved breakthrough progress in key areas such as driving perception, decision-making and planning, and vehicle control. In particular, the application of large models in vehicles has significantly enhanced user experience in assisted driving and human-vehicle interaction.
“AI has greatly accelerated the popularization of assisted driving functions, enabling rapid adoption and widespread accessibility. Breakthroughs in underlying technologies are the core driving force behind the evolution from assisted driving to higher levels of autonomous driving.”



h u h A

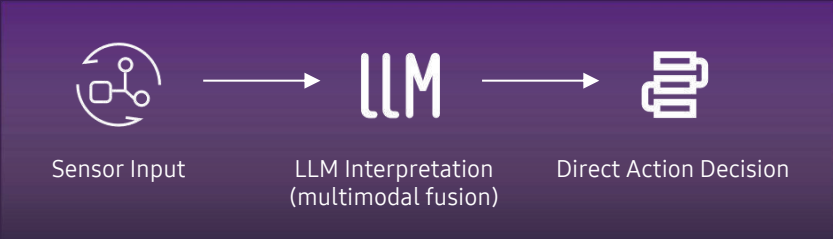
p :f| | “:T–

VLA extends language models to multimodal inputs, enabling natural interaction and semantic-driven driving knowledge.

q ’(g ’“

World Model builds an internal simulation from sensor data to predict future states and guide action planning.

j – “

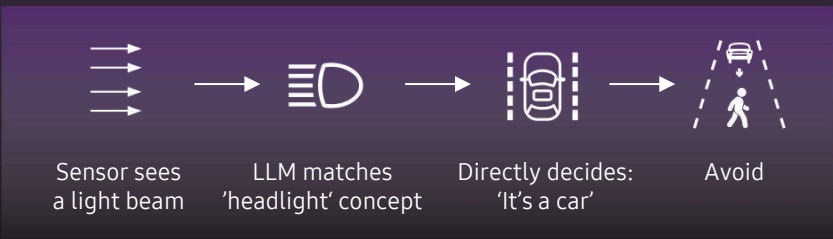


Sensor Input LLM Interpretation (multimodal fusion) Direct Action Decision

m “ ✓

- User-friendly human-machine interaction: Understands natural language
- Strong semantic reasoning: Extracts meaning from environment, combines with experience

Y | “



Sensor sees a light beam LLM matches 'headlight' concept Directly decides: 'It's a car' Avoid

q “ | “ X

- Black-box nature
- Data dependency
- Weak reasoning depth

j – “

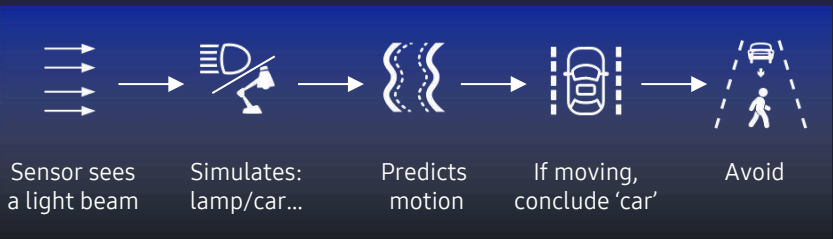


Sensor Input (Mainly Image/Video) World Simulation (predict environment) Action Planning based on simulation

m “ ✓

- Strong spatial reasoning: Uses visual data, simulates instinctive responses
- Better real-time performance: Closer to human driving intuition, instant feedback

Y | “



Sensor sees a light beam Simulates: lamp/car... Predicts motion If moving, conclude 'car' Avoid

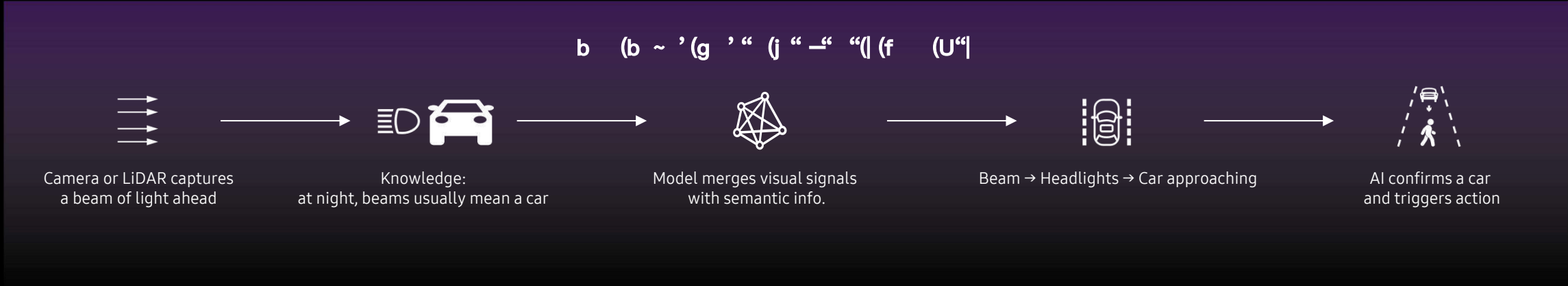
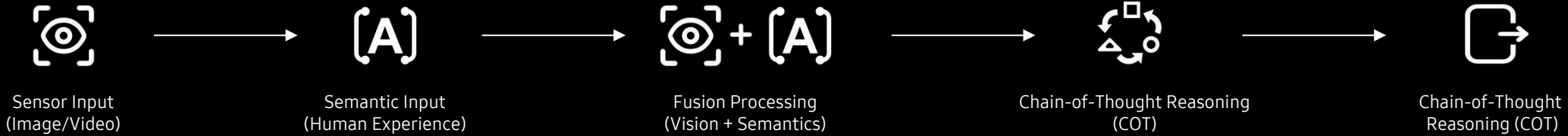
q “ | “ X

- Slow reaction time
- Computationally expensive
- Hard to model reality

y & &

b ~ '(g '“ (

Hybrid models link perception to causal reasoning, enabling AI to “understand why” rather than just “see what”. By combining semantic and experiential inputs, hybrid models enable intelligent driving systems to react quickly while reasoning causally, making decisions more transparent and human-like.



“In the new hybrid training mode, you can input images, videos, and human experiences together, allowing the model to generate results under semantic logic.”
“It not only responds instantly to what it sees, but also draws on prior experience to produce behavior closer to that of humans.”

h s h b u u

h

b “(m| “(p “
Delivers more diverse driving data



I – “(w| |
Improves training coverage, yielding a more robust model



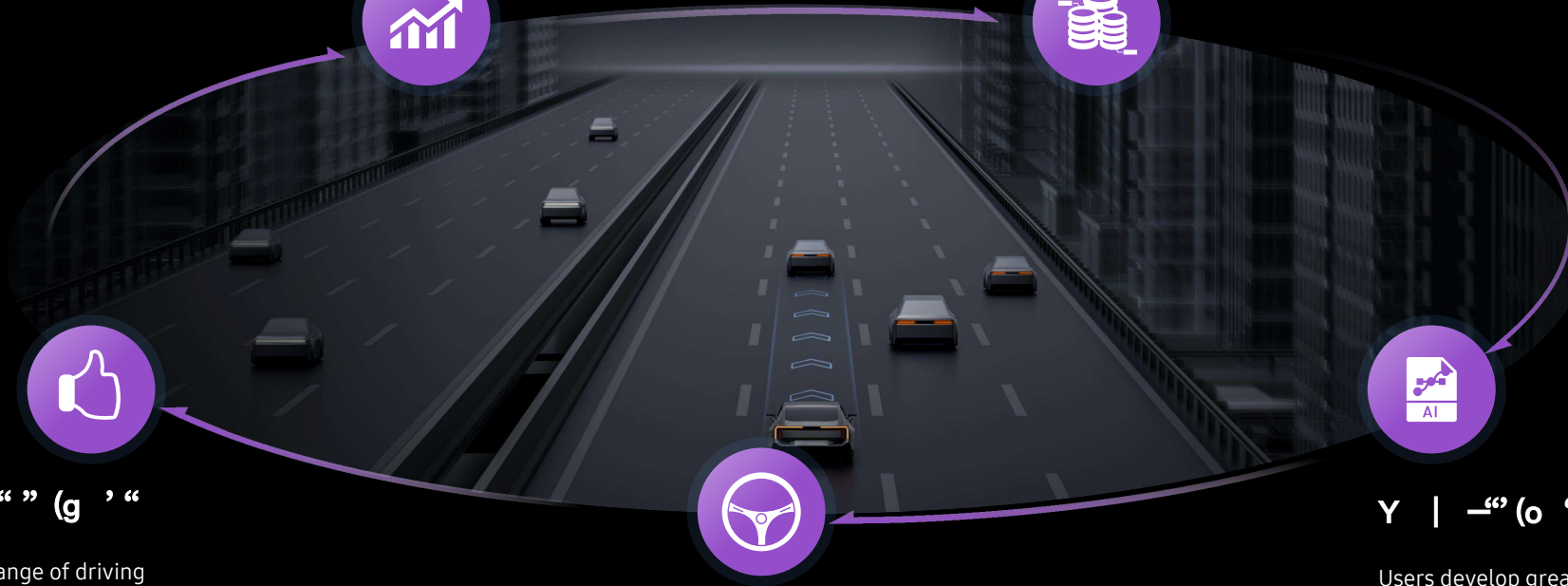
g “(j “”(g ’ “
Model adapts to a wider range of driving scenarios



Y | –“(o “(Y “ “ –“
Users develop greater trust in autonomous driving and show stronger preference for vehicles equipped with it



g “(c “ “ (T :W (Z –
Autonomous driving replaces human drivers in more situations, becoming more human-like and safer



s b b k

Our field study covered leading autonomous driving pilots across three cities, combining Robotaxi, Robobus, and consumer-owned intelligent vehicles.



I ~ | (0(i ~ ~ (Y “ “ — “



Apollo Go (萝卜快跑, Baidu)
China’s largest autonomous ride-hailing platform, operating Robotaxi fleets across multiple cities, including Shenzhen.



Pony.ai (小马智行)
A leading autonomous driving company, focusing on Robotaxi and Robobus services. Strong presence in Shenzhen and Shanghai.



WeRide (文远知行)
A leading global L4 autonomous driving company, operating fully driverless Robotaxi and Robobus services across Guangzhou, Beijing, and Abu Dhabi.



c “ “ :W (i “ (c “ “ (0(Y “ “ —“



XPeng Motors (小鹏汽车)
An EV manufacturer pioneering L2–L4 intelligent driving. We interviewed owners of XPeng X9 for real-world user insights



Xiaomi EV (小米汽车)
Entered the smart EV market with SU7 Ultra. We conducted in-depth interviews with early adopters in Shanghai.

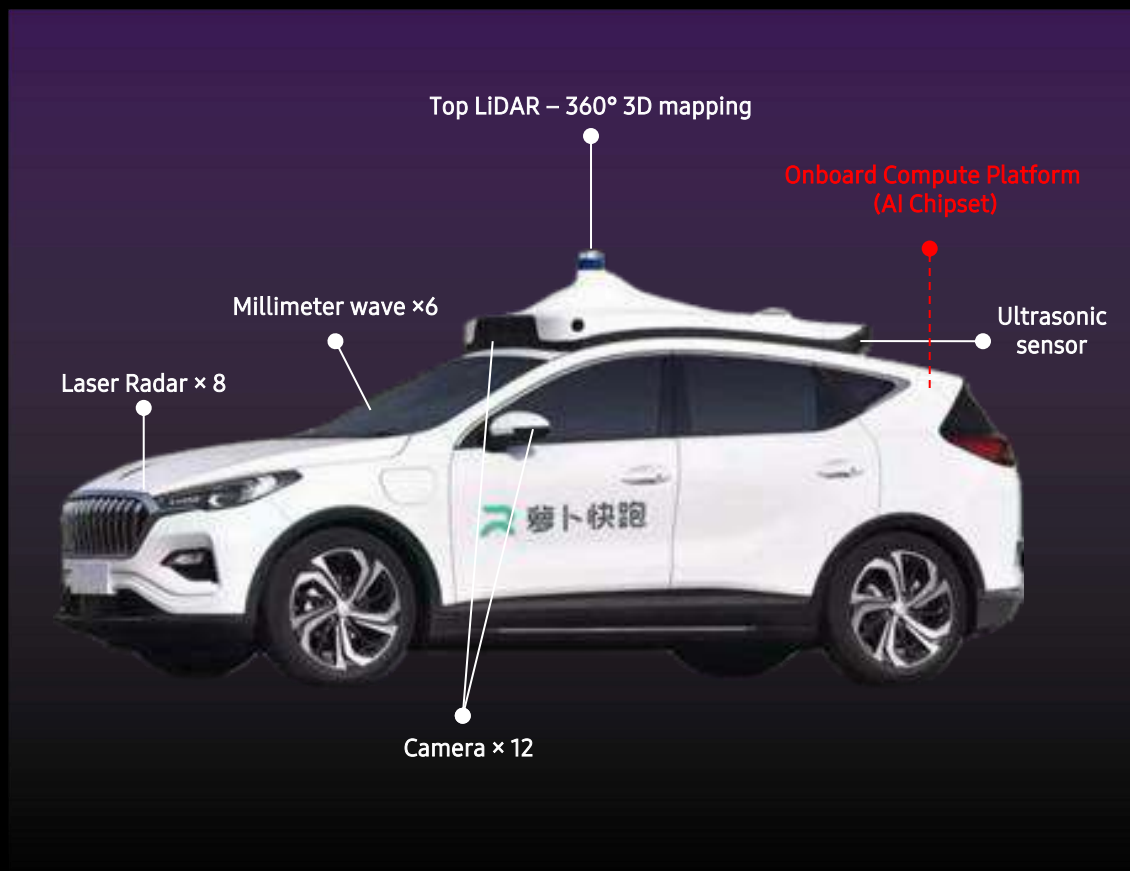


U & A



i ’“(“(’“

c ’“(“(’“



Ride-Hailing App Interface

One-tap booking, clear trip info



In-Vehicle Passenger Screen

Real-time route, speed, and ride control options



Driverless Cockpit

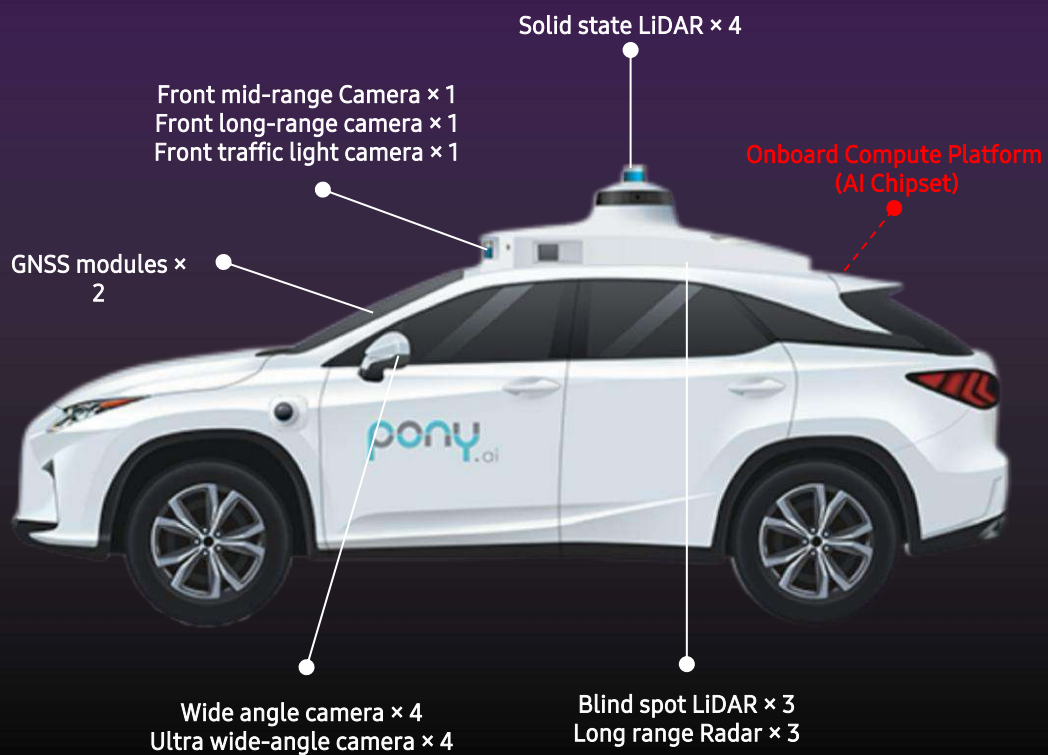
Fully autonomous cabin with no human driver



u

小马智行·pony.ai

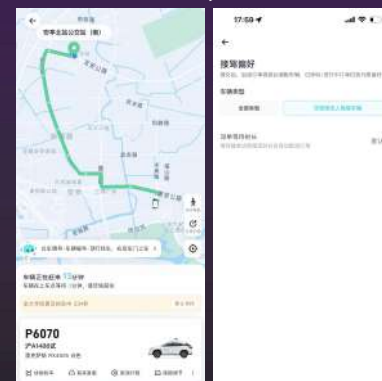
i ’“(“(’“



c ’“(“(’“

Ride-Hailing App Interface

Clear trip info, can choose driverless or safety-driver vehicle



Safety-attended autonomous cabin

Fully autonomous, with remote safety operator ready to take over if needed



In-Vehicle Passenger Screen

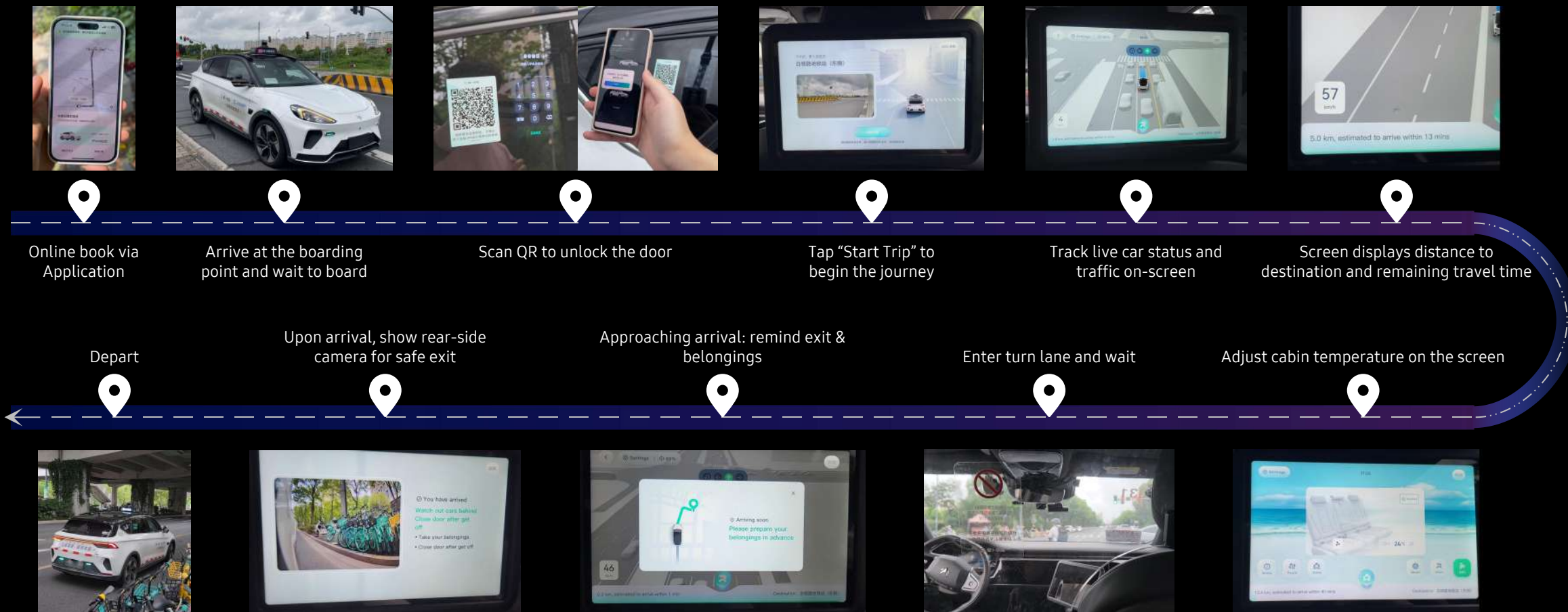
Real-time vehicle status, traffic, and route map



& A



A step-by-step view of Baidu Apollo Robotaxi passenger experience – tested in Jiading, Shanghai on a weekday evening, covering about 10 km with relatively heavy traffic.



A step-by-step view of the Pony.ai Robotaxi passenger experience – tested in Jiading, Shanghai on a weekday evening, covering about 10 km with relatively heavy traffic.



Online book via Application



Arrive at the boarding point and wait to board



Boarding, screen prompts: "Fasten seatbelt before we start"



Enter the last four digits of phone to confirm the trip



Read through the safety instructions



Tap "Start Trip" to begin the journey

Depart

Upon arrival, the vehicle pulls over and reminds you to take your belongings

Enjoy music and other entertainment on the in-vehicle passenger screen

Merge from the slip road into the main carriageway

Pause right turn, yield to oncoming left-turners

Track live car status and traffic on-screen



y

k

U (o(U | ' ,

i ("(| | ' ,

c " | -

T | (o(Y

x
w



- Smooth app booking
- Slower response & long pickup wait & Pick-up limited to designated points
- pre-boarding security verification (QR code scan)



- Smooth app booking with quick acceptance
- Long pickup wait & Pick-up limited to designated points
- In-car security verification (phone number enter)
- Easy onboarding with clear guidance (including seatbelt reminders)



- Smooth start with confident merging



- Clear in-ride information (route, speed...)



- Ride remains stable on most segments



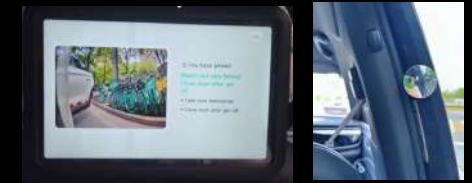
- Misrecognition incident (misidentified a car carrier as moving vehicles, honked unnecessarily, and waited in place.)



- In-car entertainment with music, video, and climate control enhances passenger comfort and engagement



- Accurate curbside drop-off



- Safe exit reminders (in-car screen provides safety reminders, while mirrors and cameras help check the surroundings to ensure safe door opening)



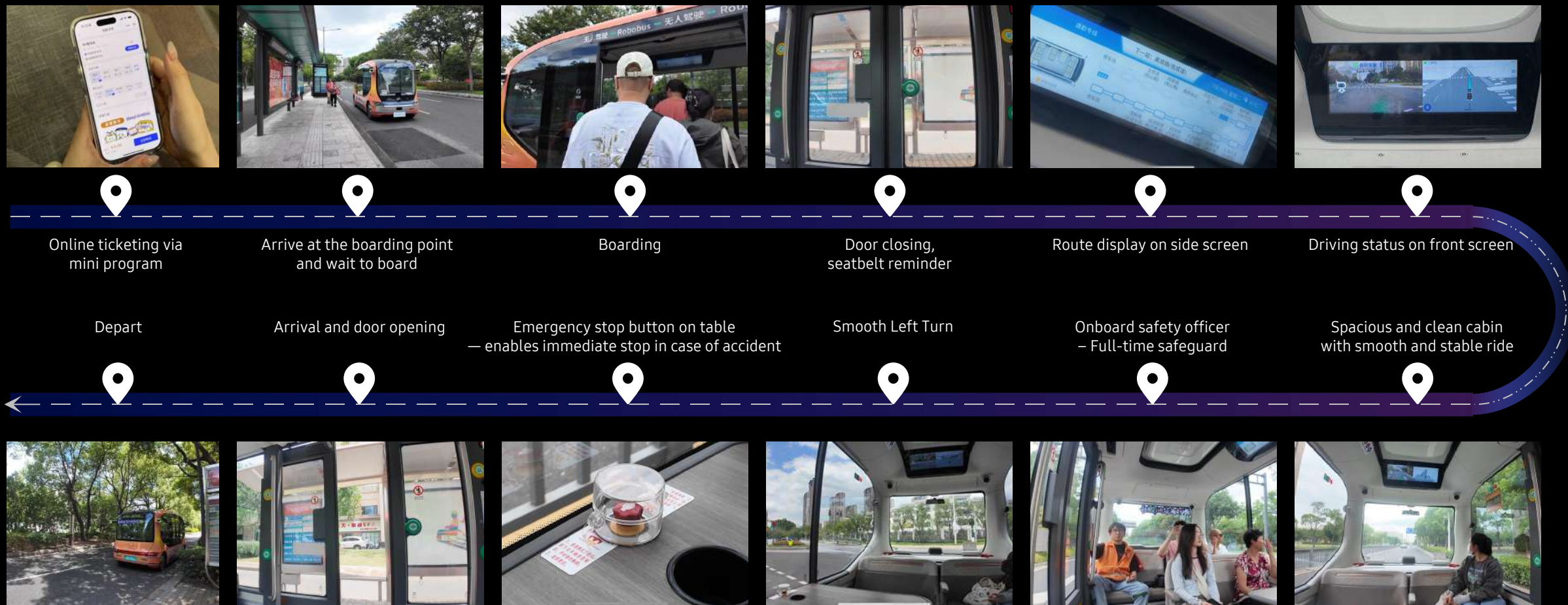
"Curious yet slightly anxious about safety on my first autonomous ride."

"Smooth acceleration and steady driving gradually built my confidence."

"Interactive screens and entertainment features made the journey engaging."

"Safe arrival left me relieved, satisfied, and pleasantly surprised."

A step-by-step view of the WeRide Robobus passenger experience – tested in Wuxi at midday on a weekday, running through a newer district with relatively light traffic.



y

k

U (0(U | '

i ("(| '

T " "

T | (0(Y

- WeChat mini-program: online ticket purchase and live route map
- Service area and routes are limited
- Currently free, ideal for short-distance travel

- Bus arrives on schedule
- Local residents ride it daily as their regular commute

- Silk-smooth turns and lane changes

- Failed to anticipate merging vehicle; sudden hard brake

- Onboard screen shows upcoming stops, live cabin feed, and surrounding traffic

- Safety operator's hand hovers over the emergency-brake button, additional red emergency-brake buttons beside passenger seats

- Glides to a smooth stop at each station and the doors open automatically

"First time taking Robobus feels fresh and fun, and spotting other residents at the stop makes me feel more at ease"

"The smooth turns and lane changes are reassuring, yet an unexpected hard brake can still give you a jolt"

"The cabin displays are comprehensive and smart, both the safety operator and the reachable emergency brake make me feel safer"

"That smooth, safe ride was so cool—I totally want another go!"

u h s y

y

L2+ intelligent driving is gradually solving urban commuting pain points, showing strong adaptability across highways, city roads, and parking. At the same time, smart in-cabin features are reshaping the driving experience.

More and more drivers' attitudes are shifting from curiosity to reliance, turning intelligent driving from a "nice-to-have" to a "must-have" in car purchases.

b | (hi T

- Stable, confident lane changes and overtaking, nearly hands-free

V (hi T

- Recognizes traffic lights, pedestrians, lane changes, and U-turns, easing urban driving

T | “(j |

- Covers auto/remote parking and pull-out; effective in narrow spaces and garages



c :V| ~ (c “ “ —“

- Smart scenes with voice AI, adaptive lighting, and personalized comfort

“As someone who has driven with intelligent driving, I honestly feel I rely on it a lot. I definitely wouldn't buy a car without intelligent driving anymore.”

“Once, when I went on a trip, the system had just updated, and NGP couldn't be used on the road. I was so anxious at that moment—I was about to head out for my trip, and without intelligent driving I had to drive manually, which I just couldn't stand.”



Mr. Song, Owner of XPENG X9

y A b k

Highway NOA is among the most mature features of intelligent driving, handling lane changes, overtaking, and highway ramps with stable and smooth performance that feels highly human-like. On most segments, it is nearly hands-free, with manual takeover only required in construction zones or highly complex scenarios.

b | (Y (l |



b | (f| “(V | “



“Overall, when I use intelligent driving, highway NOA is definitely the feature I rely on most. Elevated roads are the scenarios that fit me best. So if the experience weren’t good, I simply wouldn’t be using it every day.”



Mr. Cheng, Owner of XIAOMI SU7 ULTRA

“If I were to rate highway intelligent driving, I’d give it at least 9.5 out of 10. On highways, there’s basically no need for manual takeovers—it’s extremely rare. The only special case is during road construction. Other than that, I almost never need to intervene.”



Mr. Song, Owner of XPENG X9

b A S

City NOA can now recognize traffic lights, handle lane changes, pedestrian avoidance, and U-turns, significantly easing the burden of urban commuting. However, in cases such as aggressive cut-ins or sudden complex traffic conditions, it may still misjudge or brake abruptly, requiring manual takeover. It is best described as “usable but not flawless.”

m (o:

f“((“ “ | (’ | —“

I ((“ “ | (’ | —“



XIAOMI SU7 ULTRA



XIAOMI SU7 ULTRA



XIAOMI SU7 ULTRA

“City NOA has been a process of going from nothing to something—my previous car didn’t even have urban intelligent driving. The good part is that it does bring more convenience: for example, if I want to take a sip of water, the system can briefly take over. Just having this function is already an advantage. But at the same time, City NOA still has quite a few shortcomings—for instance, sometimes it misidentifies lanes or unexpectedly takes the wrong route.”



Mr. Song, Owner of XPENG X9

“I use City NOA less often, as I don’t fully trust it. But in the few times I have tried, the experience was generally fine—whether waiting at traffic lights or changing lanes, it worked okay. Still, there are issues. For example, sometimes it fails to clearly recognize the lane markings, ends up drifting over the line, which can be quite dangerous.”



Mr. Cheng, Owner of XIAOMI SU7 ULTRA

A

b

Parking features include in-car auto parking, remote parking, and auto pull-out. These significantly reduce driving stress in tight spaces and underground garages. Overall, the functions are practical and convenient, though scanning and execution can sometimes be slightly slow.

c: + (| | + |



T (:



I “ “(|



"Today I was taking my daughter to summer camp, and she had a suitcase. My parking spot was a bit far back, and the trunk couldn't open, so I had the car pull itself out of the space in advance. That way, when I got to the lot, I could load her things and just drive off directly. But there are drawbacks—for instance, when cars are parked on both sides and the space is tight, using the auto-parking function makes it move very slowly."



Mr. Song, Owner of XPENG X9

u b u b

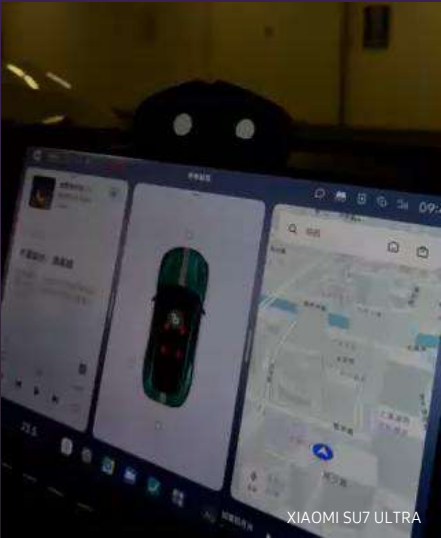
In-cabin intelligence offers customizable features, such as lighting modes, child entertainment mode, enhanced voice interaction, and even personalized greetings. These functions enhance comfort and strengthen emotional connection between driver and vehicle, positioning it at the forefront of AI-powered cabin design.

c “ (~ “ ((



XIAOMI SU7 ULTRA

c:† (“ “ (Tc



XIAOMI SU7 ULTRA

T ’ : “ “(~ “ (



XIAOMI SU7 ULTRA

Y “ | “ (’ “



XPENG X9

“Anything you want to control can be done through Xiaomi’s built-in voice assistant, Xiao Ai, using voice commands. It even comes with facial expressions, which add emotional value—for example, you can have Xiao Ai compliment the passenger in the front seat. Sometimes I also ask Xiao Ai questions, like how much a Lexus ES300 costs. The car is also equipped with ambient lighting, where you can choose the color and let the lights pulse along with the music volume.”



Mr. Cheng, Owner of XIAOMI SU7 ULTRA

k u

Section 01

.

Section 02

.Z

Section 03

.

Section 04

W

.

Section 05

Z

.

k h

z (w“ ((m+ | ~“(T +

Most companies have moved beyond single-purpose lab demos, now targeting L2/L3 levels — single-task or cross-task execution in semi-open environments. The focus is shifting from “can it work” to “can it scale.”



n “(— “(—| “ “(“ (~| | — (“ “| | (| ’(—

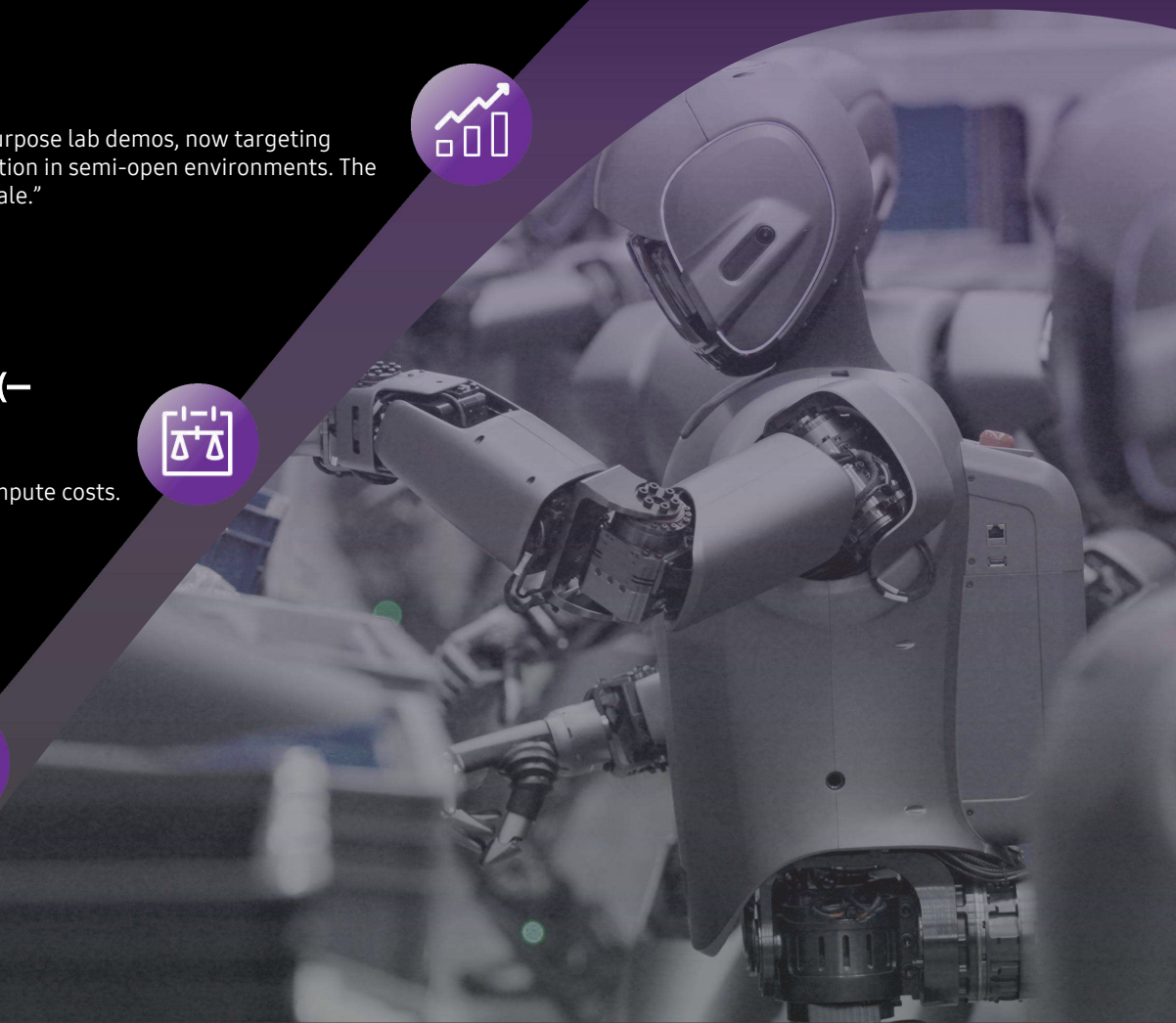
Layered models: Fast to deploy, but brittle in dynamic settings.

VLA & World Models: Greater adaptability, but constrained by data and compute costs.



v | (m (g “ (| ’(v “— |

With strong policy support, capital inflows, and a full industry chain, China is rapidly moving from follower to global leader in embodied intelligence, driven by fast commercialization in service, logistics, and manufacturing.

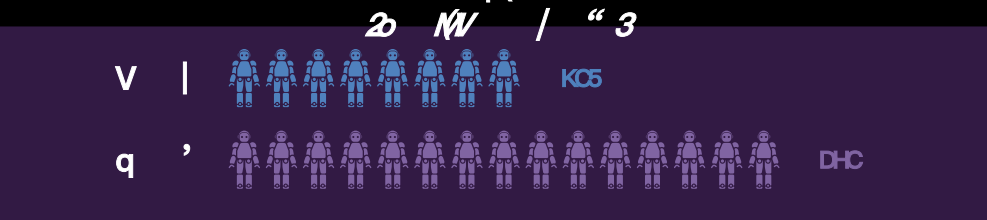


b k

z (n “ ((V “ — |

By 2024, about 150 firms worldwide—over 80 in China—are developing general-purpose humanoids. The sector is accelerating toward commercialization, with 2025 seen as the first mass-production year; several firms have already announced mass-production plans ranging from several thousand to tens of thousands of units.

h ~“(’(b | ’(l ~ (V | “ ((“(q ’B/



j “—“(ECEH(j ’ — (p “(’(f “|’ (V “ “(b | ’:l ~ (V | “



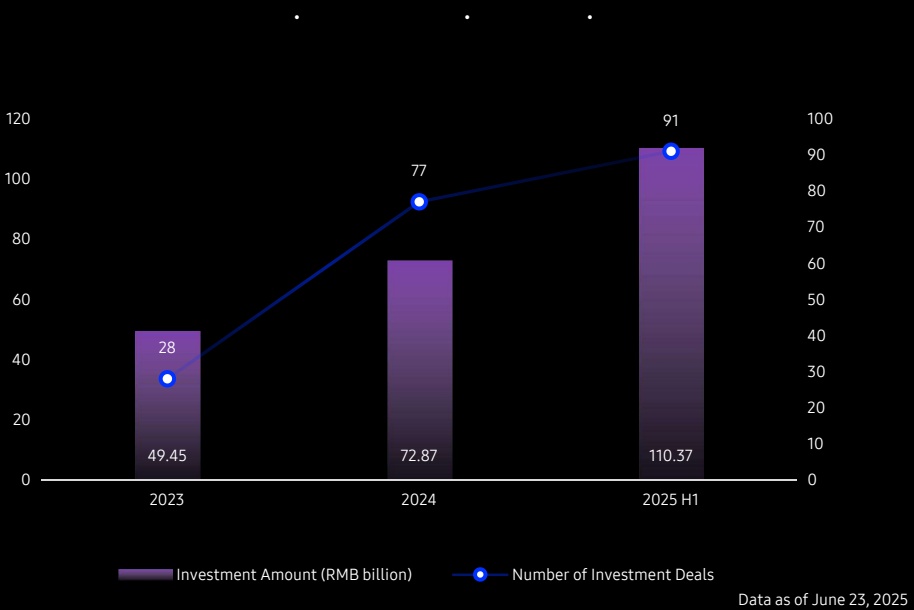
“2025 will be the first year of commercialization for embodied intelligence. On one hand, government policies are providing guidance; on the other hand, leading enterprises like Zhiyuan have set KPIs of 500 million yuan in revenue, and they are determined to achieve them.”



Galbot, Senior Product Manager

j —(0(V| | (U (V | (Y ~ ’ “(c “ “ —“(n| “ ”

Since 2024, local governments have rolled out policies for humanoid robots, while investment has surged: by June 23, 2025, 91 funding events had been recorded—already surpassing 2024’s 77—with a total of RMB 11.04 billion.



“This year, financing in the embodied robotics industry during the first half has already exceeded the total for last year, showing extremely high market enthusiasm.”



MIRAE ASSET, Investment Director

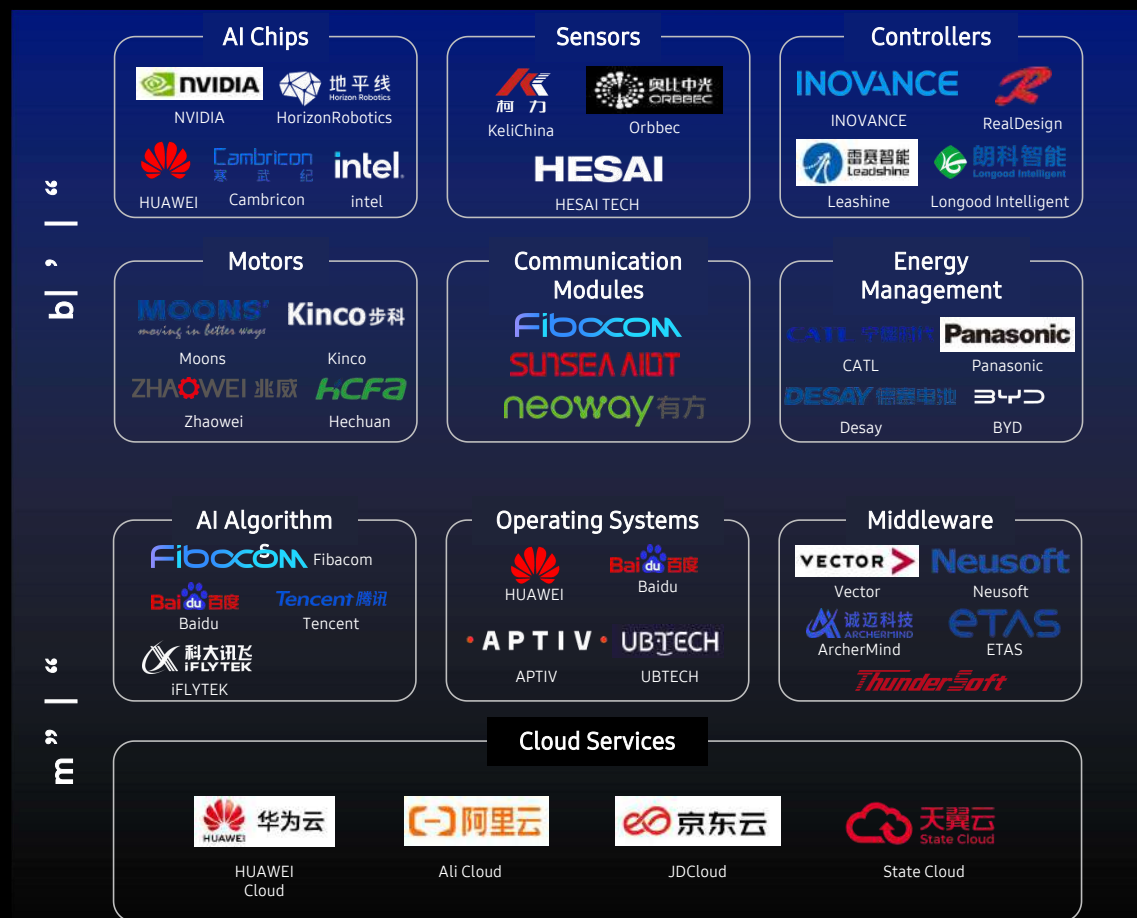
b k u b

From the perspective of the industrial ecosystem, China has already established a full industry chain in the field of embodied intelligence, covering core components, software algorithms, and complete-machine applications.

o “ | NZ ’ | “ | (b | ’ | “(0(m ” | “

Y ~ ’ “ :c “ “ —j ’ —

w “ | NT + (m— |



Au A k

Leveraging a robust Artificial General Intelligence (AGI) central system markedly enhances model generalization, endowing robots with genuine cognitive and decision-making capabilities and thereby substantially expanding the breadth and depth of their application domains.

U|

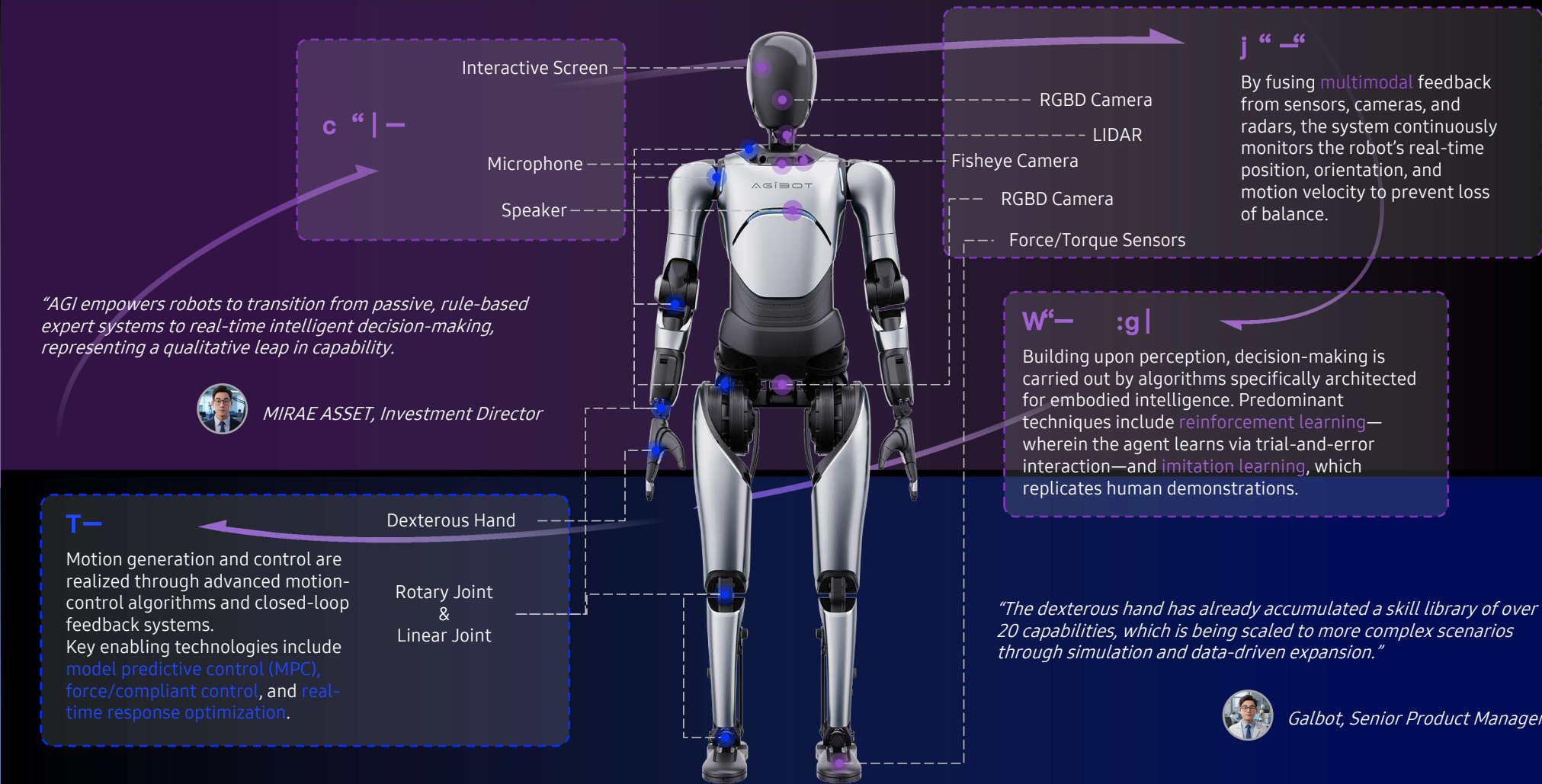
Responsible for perception, comprehension, and planning—primarily driven by large language models and vision–language–action foundation models.

v“ ~“

Responsible for converting decisions into concrete actions, i.e., motor control and motion generation.

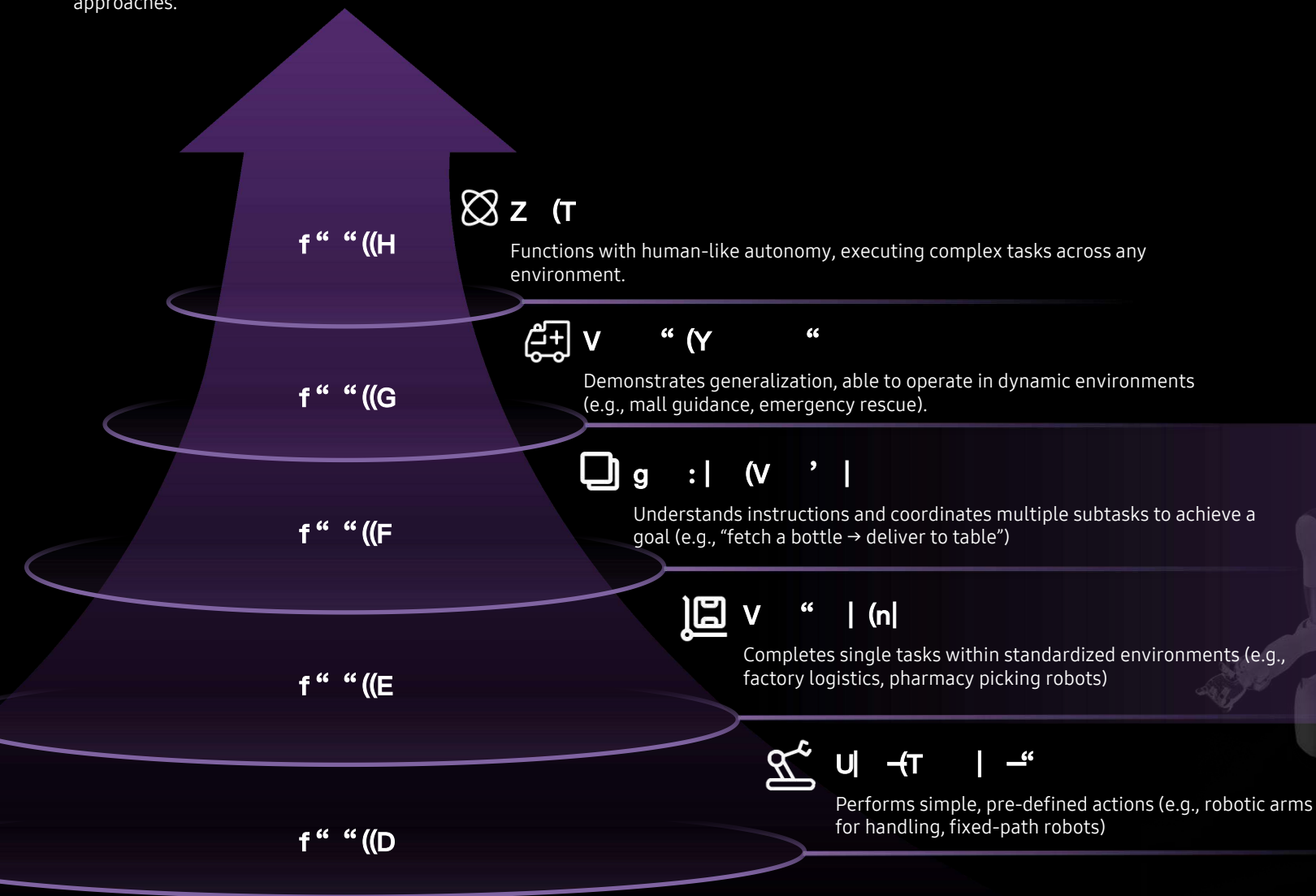
U ’

Responsible for the physical execution of actions, supported by the robot’s embodiment.



A

With progress concentrated at L2/L3, embodied robots already demonstrate practical value. Unlocking L4/L5 autonomy presents both the greatest challenge and the biggest opportunity, hinging on advances in learning approaches.



Their capabilities are built on three dominant technical routes: **Layered, VLA, and World Models**

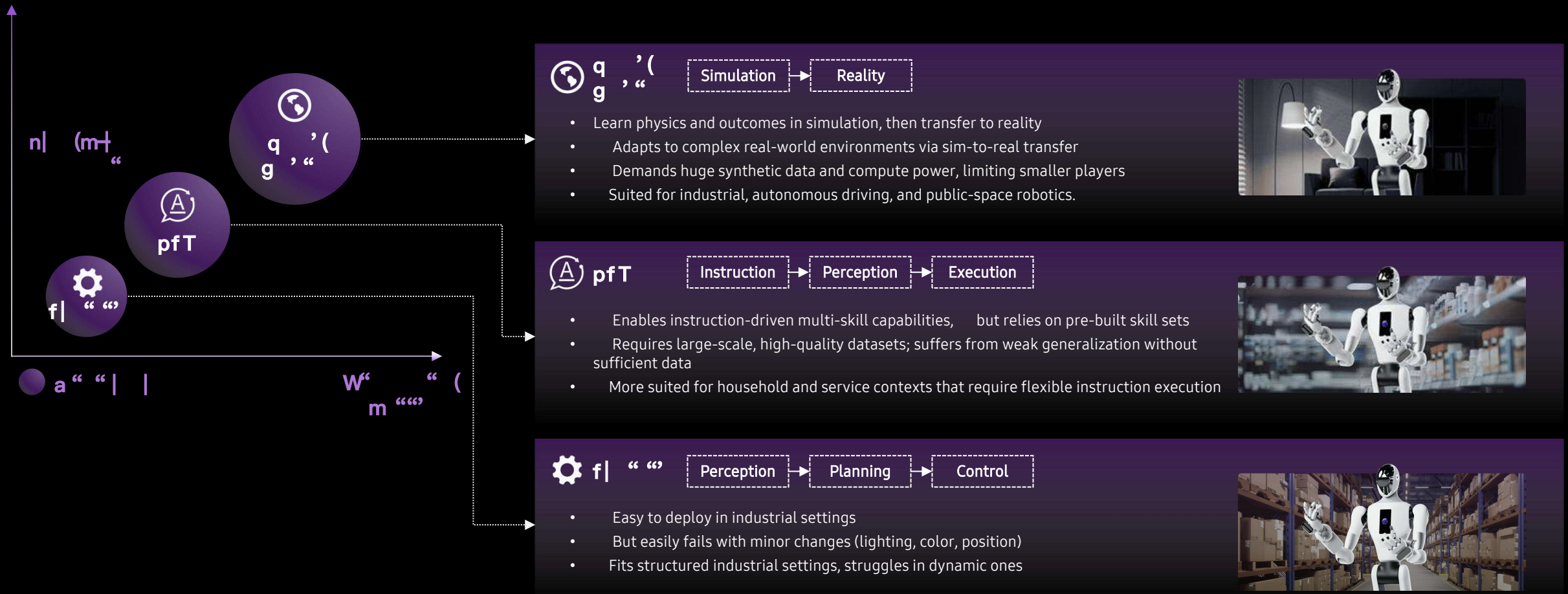
"In industrial scenarios, most tasks are still handled using layered models, which can be trained quickly but have limited generalization. Meanwhile, current research is primarily focused on VLA and world models."



Galbot, Senior Product Manager

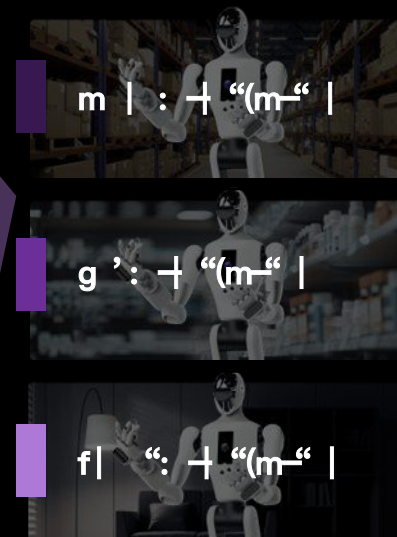
A h

- **Layered** - Clear structure, Fast deployment, Weak generalization, Limited scope
- **VLA** - Multi-modal alignment, Skill library expansion, Data-intensive, Flexible commands
- **World Models** - Simulation learning, Strong generalization, High cost, Complex environments



A A b u

Different technical routes correspond to different application scenarios, which vary by task complexity and the degree of human interaction.



"Small scenes are like cooking or fetching items; medium ones include material handling or unmanned pharmacies; large ones are household service or eldercare — these complex environments are still under exploration."



Galbot, Senior Product Manager

"Scenarios involving close human interaction require extreme caution, as safety and complexity are much higher than in pure industrial tasks."



MIRAE ASSET, Investment Director

k

k

Relatively standardized, low-interaction tasks in controlled settings.

Fast to deploy, highly efficient, but easily disrupted by changes in object size, shape, or environment.

q | “ “(m



U (n | |



j | —(



Z | ~ —(m | —



b

Multi-step tasks with medium complexity, still requiring limited human interaction.

Requires higher task coordination (pick → carry → deliver). Efficiency proven in pilots, but generalization to diverse layouts remains a challenge.

((o T | “(j -
“ (| | -“)

v “ “ -“(m “(“ -

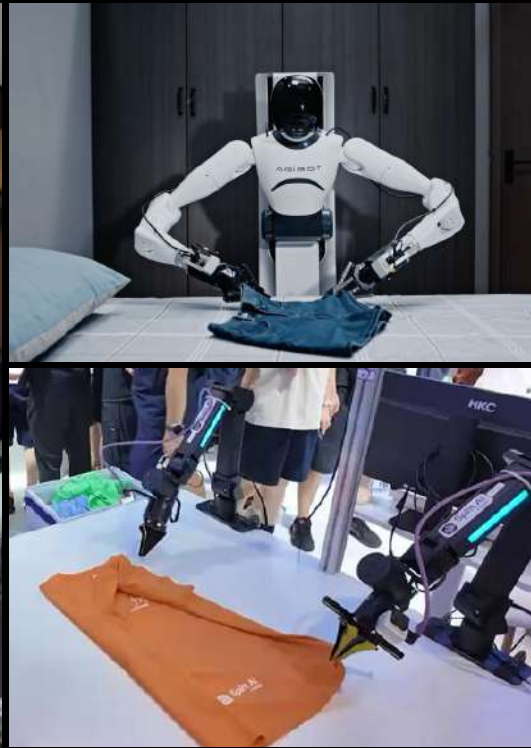


b y A

Dynamic, complex environments requiring both advanced generalization and intensive human interaction.
Highest potential value, but also the hardest to achieve. Needs robust perception–decision–execution pipelines plus strong safety assurances.

b “ ’(m“ —“

| “ | ~ |



o ’ “ | “ (
 (Marine research and exploration/security and rescue...)



b k & kby UBTECH

Boasts a full-stack technology portfolio spanning hardware drivers to AI algorithms, with its core competitive moat rooted in the vertical integration of three proprietary modules. However, its software capabilities remain comparatively weak—particularly in machine vision, speech recognition, and algorithmic platforms—which require substantial reinforcement.

q | “ (mE



UBTECH's next-generation industrial humanoid robot, Walker S2, enables fully autonomous battery swapping—no human intervention or shutdown required—completing a rapid swap in just three minutes and delivering 7×24-hour continuous operation.

q | “ (mD



UBTECH's humanoid robot Walker S1 entered the ZEEKR 5G Smart Factory in 2024 to participate in quality-inspection tasks and other operations.



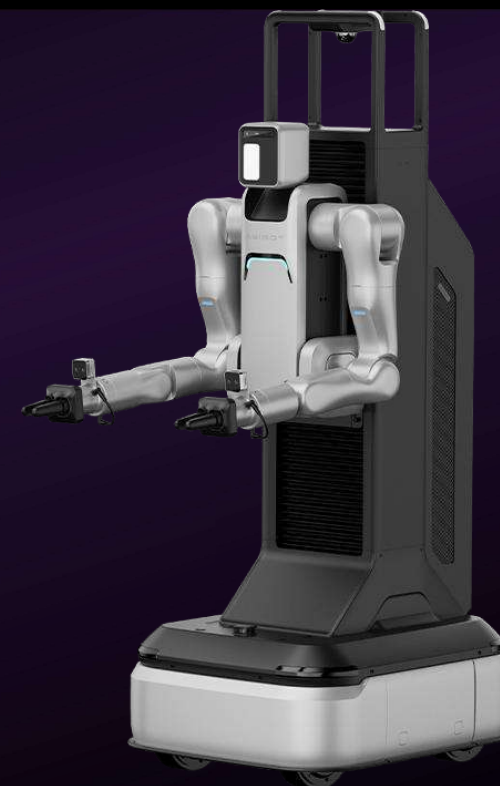
UBTECH's humanoid robot Walker S1 began on-site trials at BYD plants in 2025, operating in concert with L4 autonomous logistics vehicles, unmanned forklifts, industrial mobile robots, and intelligent manufacturing management systems.

b k

A/ &



Pursues a “full-stack self-research + multi-scenario deployment” strategy, underpinned by proprietary hardware and its GO-1 large model, yielding high generalization capability. Nonetheless, elevated hardware-development costs and stringent scenario-adaptation requirements continue to pose significant commercialization challenges.



In 2025, Zhiyuan's Expedition A2-W has entered routine industrial operations on the loading line at Fulprecision's manufacturing plant.

s | t “ (TE:q

b k / A & GALBOT

Maintains a technological edge in embodied large models and end-to-end control—e.g., GroceryVLA and GraspVLA—delivering strong generalization. The company has forged broad commercial partnerships and established a presence across retail, manufacturing, urban services, pharmacies, and other application scenarios.



In 2024, Galbot announced a joint launch with Meituan Pharmacy of the Galbot Smart Pharmacy solution, in which Galbot operates the pharmacy completely unattended.



In August 2025, Galbot began serving customers at the “Galaxy Capsule” convenience store in Beijing and launched its “Ten Cities, One Hundred Stores” expansion program.



At Mercedes-Benz and ZEEKR plants, Galbot handles sunroof transfers and tote-box shuttling; on BYD’s production line, Galbot coordinates with autonomous vehicles for integrated logistics.

k b **u**

Section 01

.

Section 02

.Z

Section 03

.

Section 04

W

.

Section 05

Z

.

k b u



Z (g| | ((c “ “ ((n | ’(c – “(a

AI brings intelligent decision-making to marketing, pricing, and customer service. While big players still dominate, SMEs now gain greater room to compete, making the e-commerce ecosystem more inclusive.



V “(m ”((Z (j –“:W “ ((p| “:W “

Consumers are moving beyond pure price sensitivity toward experience, personalization, and trust. AI tools like smart Q&A and virtual try-on align with this shift, making intelligent consumption a new norm.



Y (c’ (f –((n | ’(c “ “ (g| – |

AI is gradually reshaping the “people–goods–place” framework, moving from static rules to adaptive, data-driven decision-making. While still in progress, this trend points to a future where real-time demand forecasting and logistics optimization become the industry norm.



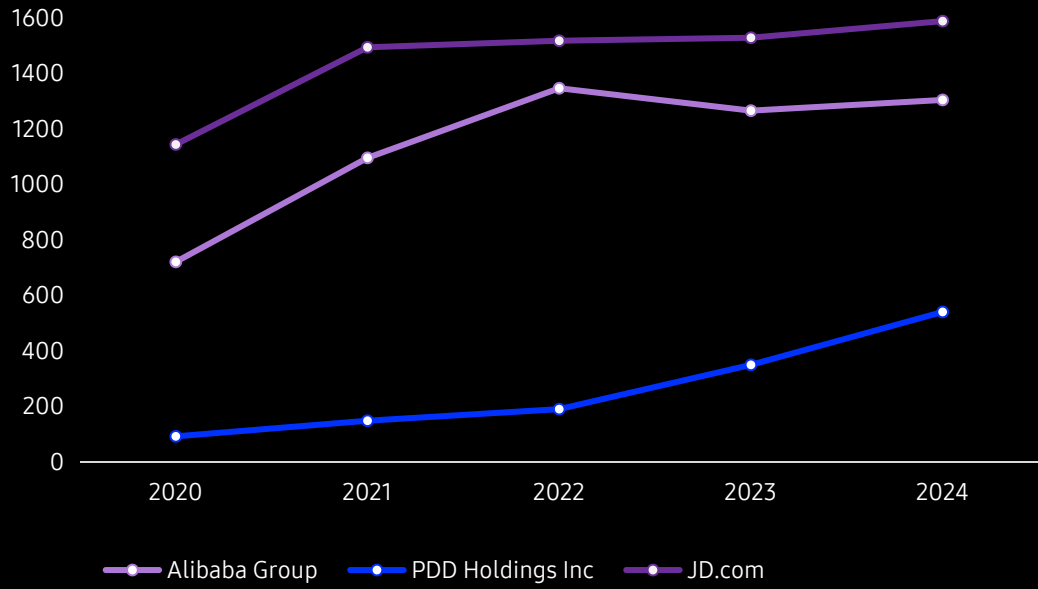
b k

China has remained the world's largest online retail market for 12 consecutive years. In 2024, online retail sales reached 15.5 trillion yuan (US\$2.16 trillion), according to the Ministry of Commerce (Jan 24, 2025).

I (m+ “W“ | ’ (TcW “ (i “ |

With transaction volumes and order counts surging, leading platforms face growing pressure on operations and management. Manual order processing can no longer keep pace, making the integration of AI technologies essential for efficiency and scalability.

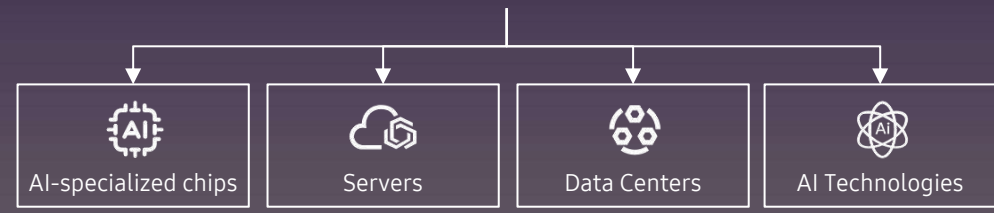
i “ | (l “ “ “(8EOEC EOEH
2b Nm(U 3



T ~|~| (T ((Tq5(V ’ (m | “

Alibaba will invest RMB 380B (US\$53B) in cloud computing and AI hardware infrastructure over FY2025–FY2027. The investment exceeds Alibaba's total infrastructure spend of the past decade. Average annual investment: >RMB 120B.

FKOU(c “ “ Ne“ (Z – (T “|



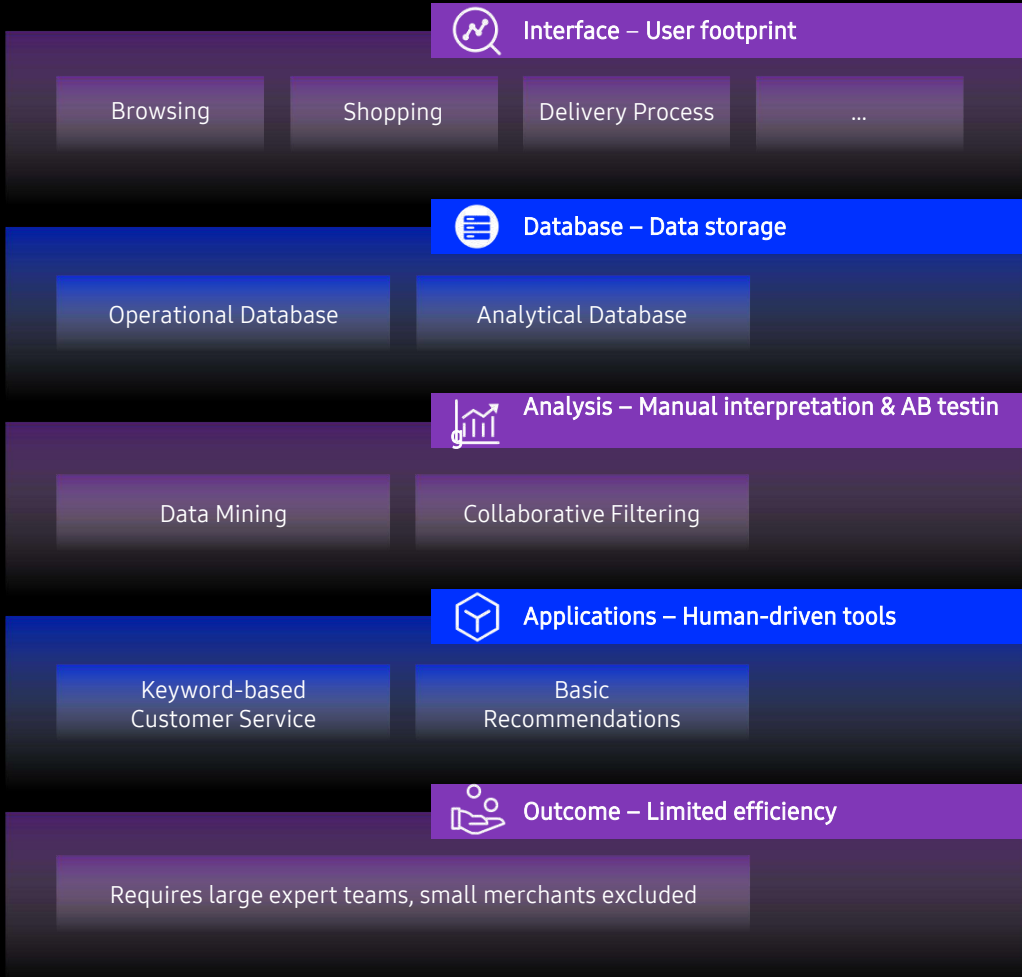
“Around 2023, when the new CEO took over, Alibaba raised the strategic slogan: User First, AI Driven. This shows that AI has been elevated to the very core of the group’s strategy. Jack Ma now only closely follows two businesses: Instant Retail and Cloud + AI. The AI budget itself is RMB 380 billion over three years.”

Alibaba, Senior Operations Expert

h h & Au

E-commerce relied on big data and manual analysis, which demanded large expert teams. This gave big brands an edge, while small merchants struggled to keep up.

k b



“Personalized recommendation (‘thousand faces for thousand users’) already existed back in 2016, but it wasn’t called AI then — it was mainly big data statistics plus manual experience.”

“In terms of pricing and promotion strategies, decisions used to be made based on data and human experience — for example, setting the price as 10% off the lowest price in the past 90 days.”



Alibaba, Senior Operations Expert

“Every day we have to deal with a large number of products, such as analyzing which items sold best in the past six months. In the past, we had to manually clean the data and then organize it, which made the process highly inefficient.

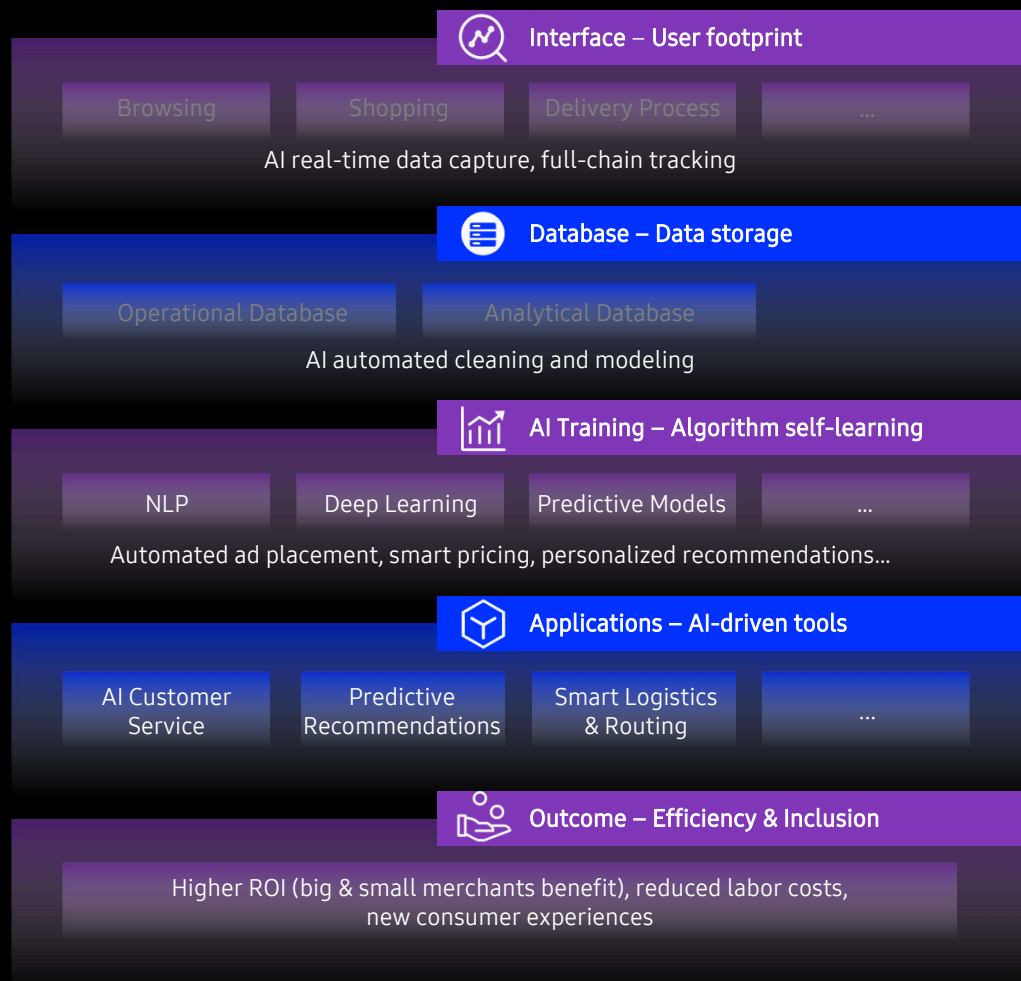
“I focus on consumer segmentation for brands and develop related strategies. Previously, these strategies had to be tested through paid media campaigns, with ad results used to refine the approach. Big data was the main basis for brand communication.”



Prottime, Senior Data Specialist

A u k b Au

AI reshapes e-commerce through predictive algorithms, personalized recommendations, and automation. It reduces reliance on large teams, improves efficiency, and makes advanced tools accessible to all merchants.



"Now with AI, even small merchants can at least move from 2 to 5 — it's truly inclusive...In the past, operations relied heavily on manual teams. Now AI empowers small merchants to compete, no longer crushed by the scale advantage of big players."



Alibaba, Senior Operations Expert

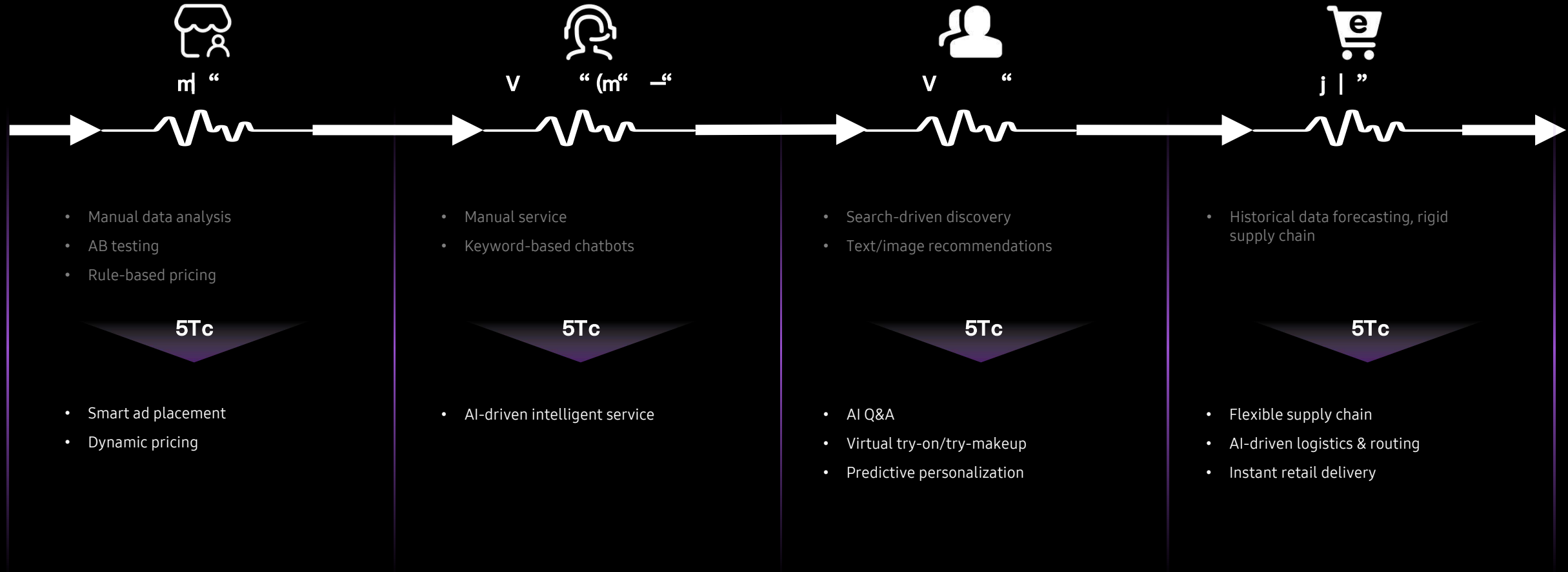
"With the rise of AI, however, advertising logic has shifted: by simply setting campaign objectives (conversion, acquisition, etc.), the platform's AI automatically handles targeting. Although we cannot always be sure if the targeting is fully precise, the overall ROI generally outperforms manual selection."



Protime, Senior Data Specialist

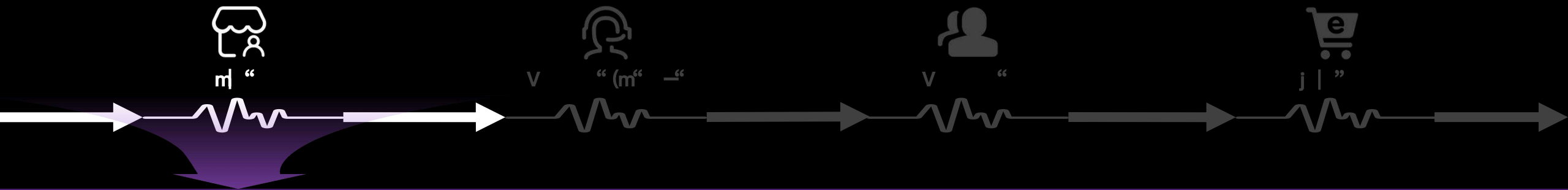
AUA A k b b

AI is reshaping the entire e-commerce value chain — from merchants and consumers to service touchpoints and platform operations.
It is no longer an add-on — it has become the core engine driving efficiency, inclusiveness, and new experiences across the e-commerce chain.



Au & u

Leading Chinese e-commerce platforms leverage AI marketing tools to automate ad targeting, optimize creatives, and empower merchants of all sizes.



ET ~|~|(q| | | |

ET ~|~|(k | | | |



Keyword promotion with cross-channel placement capability

Merchants set the cost they are willing to pay per opportunity

Merchants set an average daily budget
Note: Daily budget ≥ RMB 200 or at least 1× the opportunity bid



After 30 days of continuous placement, up to 3 countries can be set for opportunity de-prioritization

All store products are included by default; up to 200 can be excluded

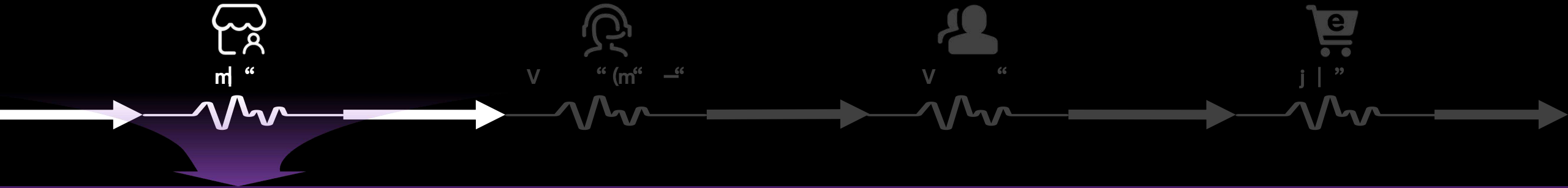
The system recommends costs based on sub-industry and acquisition ability; merchants choose within the range

- b (c(q (| AI-driven platform automating audience targeting, creative optimization, and real-time bidding
- U“ “” | Lower cost, higher ROI, inclusive for SMEs
- o “(V| “(SME merchants boosting exposure with AI-optimized ads during Double 11

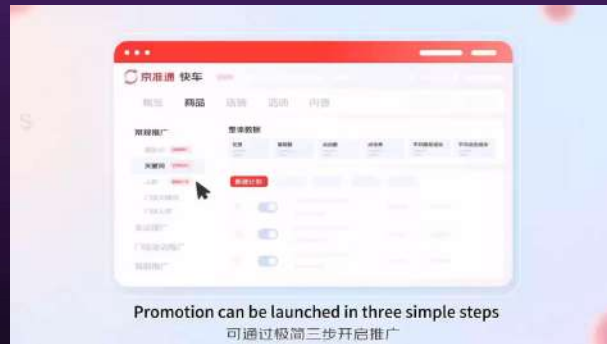
- b (c(q (| A simplified tool linking paid and natural traffic, merchants set budget and ROI goals, while large AI models optimize delivery
- U“ “” | Easy to use, wide traffic coverage, predictable ROI
- o “(V| “(Merchants achieving explosive growth by running full-site promotion during peak campaigns

Au & u

Leading Chinese e-commerce platforms leverage AI marketing tools to automate ad targeting, optimize creatives, and empower merchants of all sizes.



dW(d (



- **b (c(q** | JD's official advertising platform covering search, display, and recommendation ads; supports smart bidding and audience targeting
- **U“ ””** | Broad coverage across JD ecosystem, data-driven precision, strong ROI
- **o “(V| “(** Brands launching new products with precision targeting during JD promotions

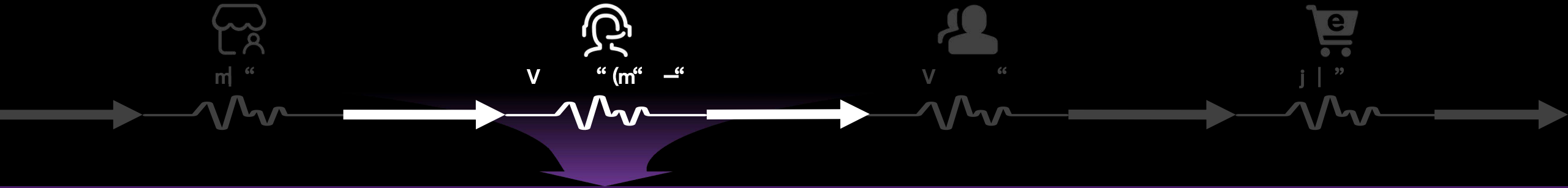
抖音i –((Y “(k | – |



- **b (c(q** | AI ad platform for Douyin e-commerce, combining short-video traffic, livestreaming, and smart bidding
- **U“ ””** | AI smart bidding, creative optimization, traffic monetization
- **o “(V| “(** Douyin merchants using Qianchuan to link viral videos with instant sales

Au b s u

AI-driven customer service tools reduce human workload, improve response speed, and create personalized, round-the-clock consumer experiences.



T ~|~| (n| |



- Q: Are there any major promotions going on recently?
- A: The most recent one is the September Super 88 campaign ...

- **b (c(q** | Store-level AI assistant for Taobao/Tmall merchants. Covers FAQs, order inquiries, logistics tracking, and after-sales service.
- **U“ “”** | Reduces human workload, provides 24/7 service, improves response speed.
- **o “(V| “(** | During Double 11, SMEs used Taoxiaomi to auto-reply thousands of logistics queries per minute.

dW (m | (V “ (m“ —“

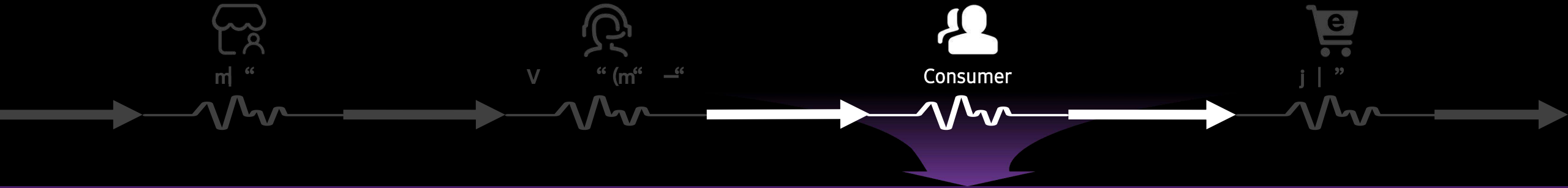


- Report damaged item to customer service → Directly Guided to after-sales service window

- **b (c(q** | JD's AI-driven system integrates voice/text recognition, logistics order tracking, and escalation.
- **U“ “”** | Seamless pre-sales & after-sales, strong logistics integration, high accuracy.
- **o “(V| “(** | JD handled 90%+ refund/return queries automatically during shopping festivals.

Au b s

AI is reshaping the consumer journey by moving beyond search-driven discovery to personalized, interactive experiences. From AI-powered Q&A and virtual try-on to predictive recommendations, consumers now enjoy smarter, more immersive shopping.



ET ~|~| (n| ~| (T

In the dialogue box, it anticipates your questions according to your recent buying preferences

Option to select AI assistants with various voice styles



Pick any two products and have the AI provide an intelligent comparison to guide your choice

- **b (c(q** | Taobao's AI Q&A assistant built on large language models. Provides intelligent search, product comparisons, and conversational shopping guidance.
- **U“ ”** | Makes product discovery easier, reduces information overload.
- **o “(V| “(** Customer asking “What’s the best phone under 6000 RMB?” receive structured comparisons with buying suggestions.

ET ~|~| (p | (n :i (n :g| “

Option to select AI assistants with various voice styles

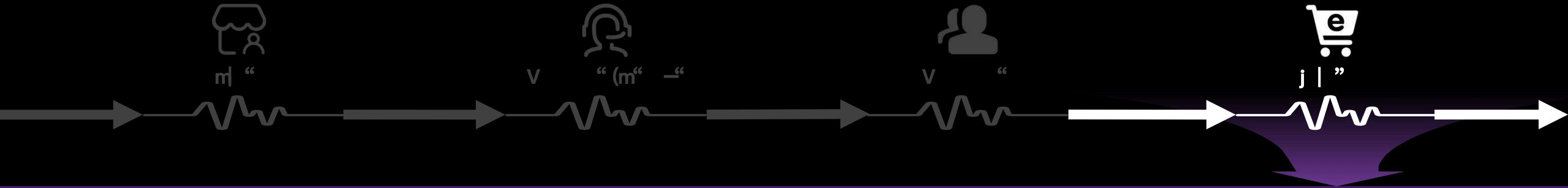


Customer tries on makeup with AR in Tom Ford's official Tmall flagship store, experimenting with the new eyeshadow collection

- **b (c(q** (| AI-powered AR technology integrated into Tmall/Taobao apps. Supports clothing try-on, cosmetic testing, and accessory matching via mobile camera.
- **U“ ”** (| Reduces return rates and enhances immersive and interactive shopping experience.
- **o “(V| “(** Tmall's beauty brands use AI try-makeup tools to boost conversion rates during new product launches.

Au s s h

AI is transforming e-commerce platforms from static, historical-data forecasting to dynamic, real-time operations. By integrating flexible supply chains, AI-driven logistics, and instant retail delivery, platforms can better match supply with demand, reduce inefficiencies, and create seamless consumer experiences.

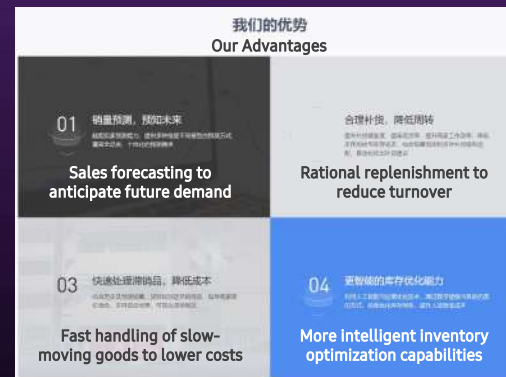


ET ~|~| (V| | (m | (f —



- b (c(q . (Cainiao integrates AI forecasting, smart warehousing, and intelligent routing to optimize delivery across China.
- U“ ” | BroFaster delivery, lower costs, higher logistics accuracy.

dW (m | (m (V |



- b (c(q | JD's AI platform predicts demand in real time, optimizes inventory, and enables just-in-time fulfillment.
- U“ ” | Reduces overstocking, shortens delivery time, increases supply chain flexibility.

美团 g“ | ((c | (l “ | (O(W“ “



- b (c(q | Meituan leverages AI for order dispatching, rider routing, and instant delivery of groceries and retail goods.
- U“ ” | Near-instant delivery, high rider efficiency, seamless consumer convenience.

h u

Section 01

.

Section 02

.Z

Section 03

.

Section 04

W

.

Section 05

Z

.

h u

1 Td | | (- + (& (“”-“ -((- “ (’ “|

AI can quickly generate concept images and renderings, allowing clients or instructors to grasp design ideas more intuitively and fostering faster consensus.

“In the second stage of a project, I used to spend significant time manually realizing ideas. AI tools now allow me to bypass much of that process, delivering a more complete concept rendering, which helps bridge the gap with instructors.”



Tongji University, Industrial Design student

“AI-generated renderings can effectively reduce workload. Not everyone can read 2D sketches, but AI visuals make it easy to quickly and intuitively show what a product will look like.”



Thermos, Product Designer

2 Td | | (| | | &“ | ’ (“(“ | ’ (~ “| ’ (” “ |

AI can quickly generate diverse design directions, inspiring new ideas and stimulating creativity.

“AI is like an inspiration assistant. It’s less about direct borrowing, and more about extracting design language from details and refining ideas.”



Tongji University, Industrial Design student

3 c ”-“ (“- ((Tc “ “ | “(’ “

AI demonstrates limited understanding of the physical world and dimensions, often resulting in outputs with logical or scale errors. Such limitations prevent direct use in engineering and mass production.

“AI tends to miss fundamentals; an SUV-with-bed rendering may end up missing wheels or a steering wheel.”



Jomec, General Manager of Styling Center

4 W“ (’ “ ((“(’ “ ((Td +

Their projects do not require production feasibility, so they are more open to experimenting with AI.

“In school, the work is more conceptual, so it doesn’t matter if the product doesn’t fully align. But during internships, precision is required, as actual industrial design deliverables demand refined details, constantly negotiated with mechanical engineers.”



Tongji University, Industrial Design student

h & A Au



Z

Research



Concept Ideation & Creation



Y

Design Generation



Z

Prototype & Presentation



Pinterest



Huaban

Inspiration Gathering



Corel Parametric



CAD



Rhino



Keyshot

3D Modeling

3D Rendering



Adobe Photoshop



Adobe Illustrator



After Effects

Visual Editing



Blender

3D Design



ChatGPT



DeepSeek



Doubao

Text to Text

Text to Image



Midjourney



Stable Diffusion



Liblib



JiMeng



Kling

Text to Image



VIZCOM



Tripo AI

Text/Image to 3D



Adobe Firefly



Stable Diffusion



Midjourney



JiMeng

Visual Editing

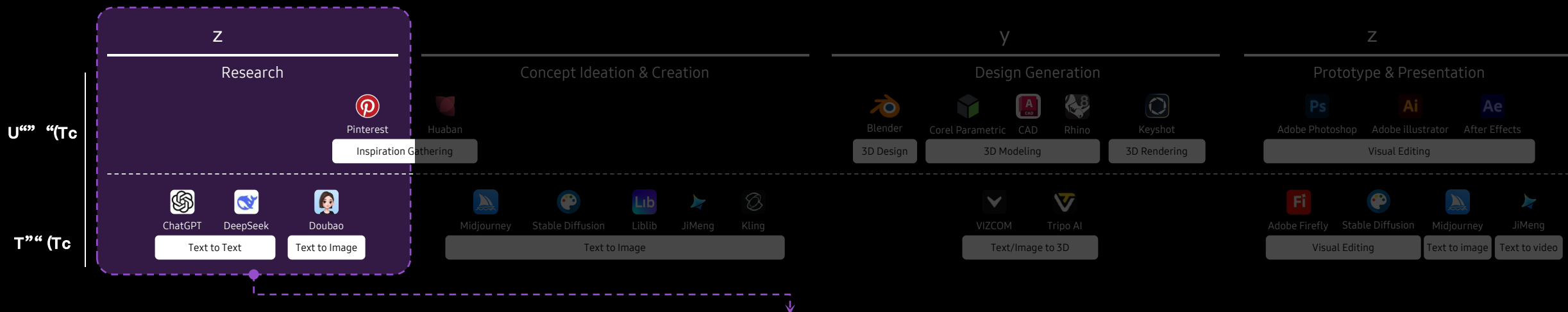
Text to image

Text to video

U“” “(Tc

T”“ (Tc

h Au h



DeepSeek

Usage Scenarios

- Rapid filtering of irrelevant information and retrieval of professional resources (e.g., technical processes) Image generation as an initial source of inspiration
- Exploration of design inspiration, reference collection, and case studies

"It helped me filter out a lot of irrelevant information. For example, when searching for aluminum alloy stamping on Baidu, I mostly get advertisements, but DeepSeek directly provided the technical process."

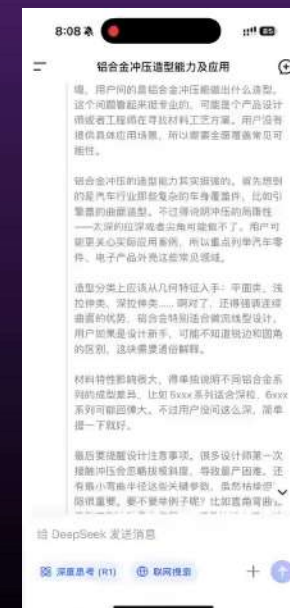


Thermos, Product Designer

"In such cases, I usually use DeepSeek to discuss directions and get recommendations for relevant artists and works."



Tsinghua University Academy of Arts & Design, Glass Art student

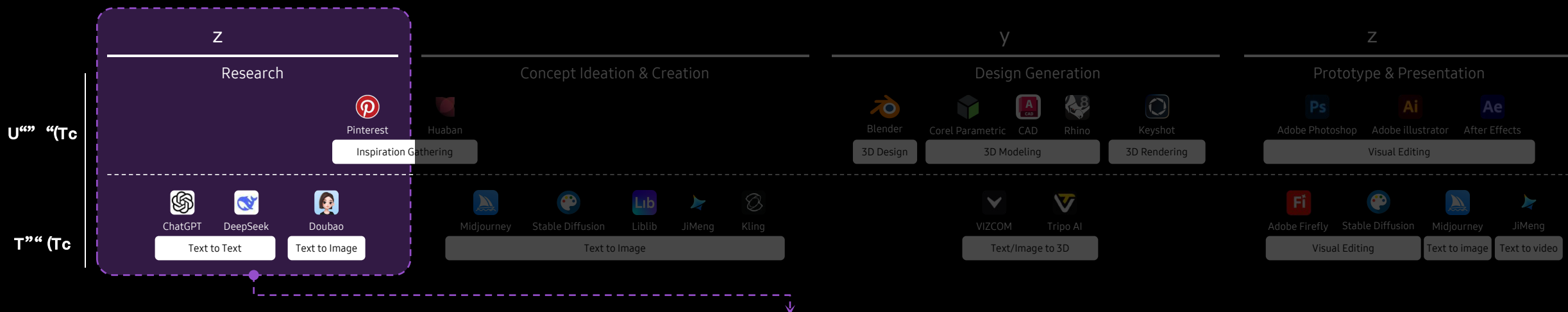


Strengths: Relatively high accuracy of information, enabling significant time savings in information screening



Weaknesses: Output quality is highly dependent on input prompts, often requiring multiple iterations to approach the desired result

h Au h



ChatGPT

Usage Scenarios

- Academic material collection and textual logic refinement
- Image generation as an initial source of inspiration



Strengths: Strong logical reasoning, suitable for structured texts and academic needs



Weaknesses: Lower efficiency in image generation, with results lacking vividness and providing limited inspiration for design

"I use ChatGPT when writing content that requires accurate references and stronger logical structure."

"I wasn't very satisfied with the images generated by GPT. They did reference the style, but it felt like an imitation without capturing the essence."



Tsinghua University Academy of Arts & Design, Glass Art student



Doubao

Usage Scenarios

- Generation of auxiliary visual sketches for rapid expression of conceptual scenarios



Strengths: Capable of transforming abstract planning into intuitive visuals, supporting product planning presentations



Weaknesses: Generated images remain at a limited level of completeness, leaning toward "conceptual illustration" rather than serving as final design outputs

"Previously, in product planning, we simply used text inside boxes. Now we can directly have Doubao generate a rough image."



Thermos, Product Designer



h

Au

Z

Research



Pinterest

Inspiration



Huaban

Gathering

Concept Ideation & Creation



ChatGPT



DeepSeek



Doubao

Text to Text

Text to Image



Midjourney



Stable Diffusion



Liblib



JiMeng



Kling

Text to Image

Y

Design Generation



Blender

3D Design



Corel Parametric

3D Modeling



CAD



Rhino



Keyshot

3D Rendering



VIZCOM



Tripo AI

Text/Image to 3D

Z

Prototype & Presentation



Adobe Photoshop



Adobe Illustrator



After Effects

Visual Editing



Adobe Firefly



Stable Diffusion



Midjourney



JiMeng

Visual Editing

Text to image

Text to video



Midjourney

Usage Scenarios

- Rapid generation of concept images to enhance efficiency in early-stage communication with clients or instructors
- Providing broader imagination and possibilities to stimulate creativity
- Fusion of different reference images in industrial design to inspire new concepts

"The rendering of lighting, materials, and interpretation of prompts is excellent. The images look very convincing, especially for imaginative and unconventional ideas."



Jomec, General Manager of Styling Center

"One advantage of Midjourney is its ability to extract the stylistic language of each form and generate multiple variations within that design language. Its fusion function can also blend two different stylistic languages into a new one, which helps spark more inspiration."



Tongji University, Industrial Design student



Generating diverse wheel-hub designs



Pinterest reference (left) & Fused design (right)



Strengths:

- Strong sense of atmosphere in images, with convincing representation of lighting and materials
- Fast image generation speed, enabling the rapid production of multiple reference options



Weaknesses:

- Lack of physical logic, often leading to design flaws (e.g., missing wheels or disproportionate structures in vehicle generation)
- No support for precise local modifications, requiring manual adjustments and post-processing

h

Au

Z

Research



Pinterest



Huaban

Inspiration

Gathering



ChatGPT



DeepSeek



Doubao

Text to Text

Text to Image

Concept Ideation & Creation



Midjourney



Stable Diffusion



Liblib



JiMeng



Kling

Text to Image

Y

Design Generation



Blender

3D Design



Corel Parametric

3D Modeling



CAD



Rhino



Keyshot

3D Rendering



VIZCOM

Text/Image to 3D



Tripo AI

Z

Prototype & Presentation



Adobe Photoshop



Adobe Illustrator



After Effects

Visual Editing



Adobe Firefly



Stable Diffusion

Visual Editing



Midjourney

Text to image



JiMeng

Text to video



Stable Diffusion

Usage Scenarios

- Applied when there is a clear design requirement, generating more concrete and precise images to improve design efficiency
- From modeling render to hand-drawn sketch, or rendering based on modeling image
- Local refinement: allows locking specific areas of the image for replacement or iterative detail adjustments

"By inputting a modeling screenshot as reference and specifying a hand-drawn style prompt, Laura can apply a hand-drawn model to reverse-generate a sketch-like line drawing."



Tongji University, Industrial Design student

"Stable Diffusion allows partial modification of an image while keeping the rest intact."



Jomec, General Manager of Styling Center



Generate multiple detail views of handles and car body



Modeling screenshot (left) & Rendering generated from it (right)



Strengths:

- High controllability, with support for local deployment and customized model training
- Enables localized modifications on existing designs

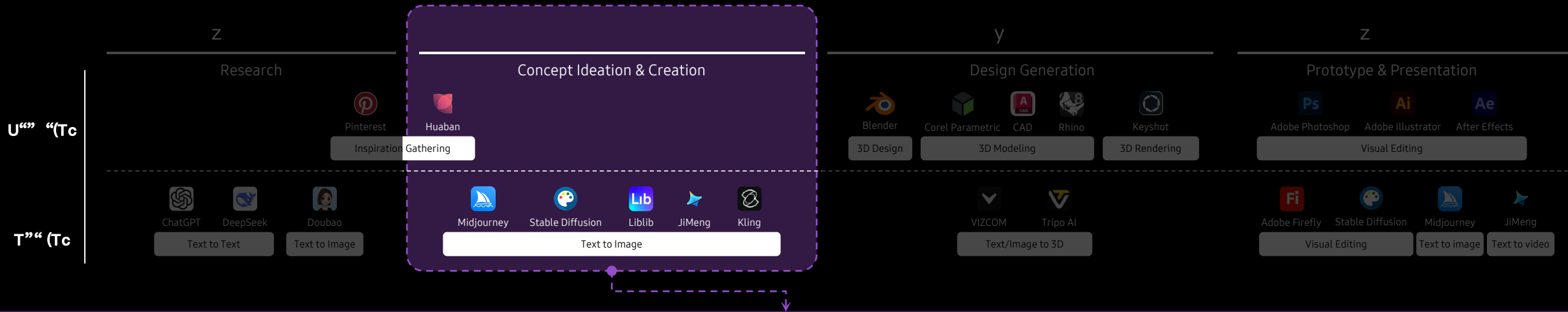


Weaknesses:

- Outputs tend to be highly concrete, making them less effective for inspiration
- Strong dependency on prompts, requiring repeated adjustments and leading to lower efficiency

h

Au



Liblib

Usage Scenarios

- Used for quickly generating concept sketches and expanding creative thinking



Strengths:

- Easy to use, capable of rapidly providing a large number of visual references
- Offers diverse model types with the option for custom model training



Weaknesses:

- Output quality is inconsistent, with many images unsuitable for direct use

"I use Liblib to supplement references. It's very efficient, but many images are not directly usable.."



Tsinghua University Academy of Arts & Design, Glass Art student



JiMeng



Kling

Usage Scenarios

- Used for generating preliminary imagery



Strengths:

- Produces four images at once, offering higher quantity and efficiency
- Greater diversity in visual content and forms



Weaknesses:

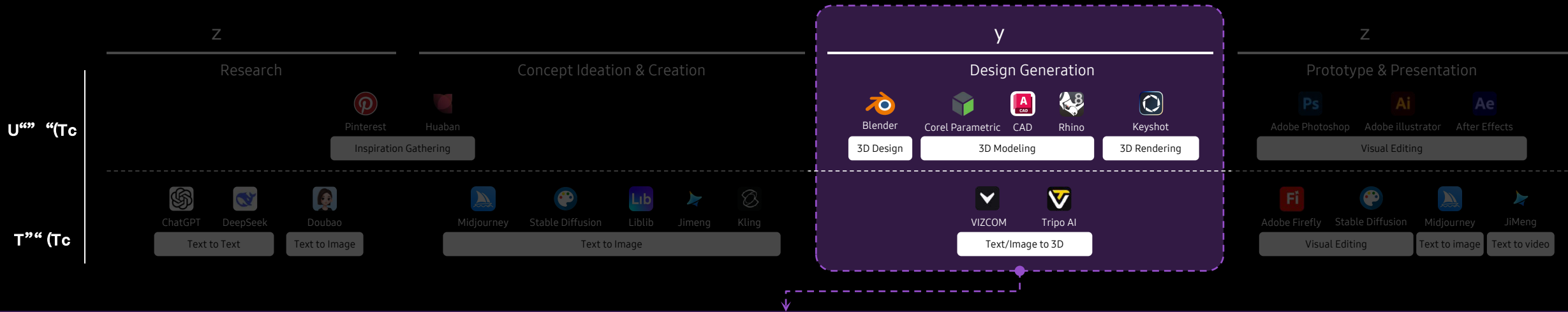
- Inaccurate interpretation of prompts, often resulting in disorganized outputs that do not meet design requirements
- Limited functionality, lacking image fusion features (in contrast to Midjourney)

"I have used JiMeng before. First, it does not support image fusion, so its functionality is limited. Second, the generated results are relatively poor, with inaccurate interpretation."



Tongji University, Industrial Design student

h Au b



Vizcom

Usage Scenarios

- Transforming hand-drawn sketches into rapid 3D renderings, assisting designers in improving efficiency during the modeling stage



Strengths:

- Significantly shortens the time required to move from 2D sketches to 3D renderings



Weaknesses:

- Limited output precision, requiring subsequent adjustments with professional modeling software
- Insufficient understanding of complex processes and structures



Hand-drawn sketch to 3D rendering



Tripo AI

Usage Scenarios

- Rapid conversion of text/images into 3D models



Strengths:

- Ability to directly generate 3D models



Weaknesses:

- Insufficient model quality, with poor textures and mappings, unable to meet the requirements of detailed industrial design

"In the past, when preparing proposals for clients, I had to go through sketching, modeling, and rendering without skipping any step. With Vizcom, as long as I draw a detailed and accurate sketch, it can directly generate the rendering."



Thermos, Product Designer



Tsinghua University Academy of Arts & Design, Glass Art student

"I used Tripo AI to generate a wine bottle. The resulting model was irregular, with messy topology. It didn't feel like a form that could be applied to industrial production."

h Au h

Z

Research



Pinterest



Huaban

Inspiration Gathering

Concept Ideation & Creation



ChatGPT



DeepSeek



Doubao

Text to Text

Text to Image



Midjourney



Stable Diffusion



Liblib



Jimeng



Kling

Text to Image

Y

Design Generation



Blender

3D Design



Corel Parametric



CAD



Rhino



Keyshot



Keyshot

3D Rendering



VIZCOM



Tripo AI

Text/Image to 3D

Z

Prototype & Presentation



Adobe Photoshop



Adobe Illustrator



After Effects

Visual Editing



Adobe Firefly



Stable Diffusion



Midjourney



Jimeng

Visual Editing

Text to image

Text to video

U“” “(Tc

T”“ (Tc



Adobe Firefly

Usage Scenarios

- Applied in final assignments or presentations to quickly generate scene images, making product visuals more complete



Strengths:

- More time- and labor-efficient compared to the traditional use of Photoshop



Weaknesses:

- Some details may be inaccurate, such as inconsistent lighting or spatial relationships

“Previously, when placing models and lighting into a scene, I usually relied on Photoshop—finding backgrounds, adding lighting, and cutting in figures—which was very time-consuming. Now AI can complete this process much more quickly.”



Tongji University, Industrial Design student



Generate background environment with product integration



AI-generated CMF color schemes for aircraft seats



Stable Diffusion

Usage Scenarios

- Tested during the delivery stage for localized CMF (Color, Material, Finish) replacements, enabling rapid exploration of different color combinations



Strengths:

- Eliminates the need to manually select materials or color maps within modeling software
- Generates diverse color schemes, providing useful references

“In most cases, I need to manually select different materials in modeling software, or even create material maps myself. With AI, I can input a detailed model via ControlNet to generate various CMF options, saving this process. The AI-generated schemes are diverse, and sometimes they even provide unexpected yet valuable color combinations.”

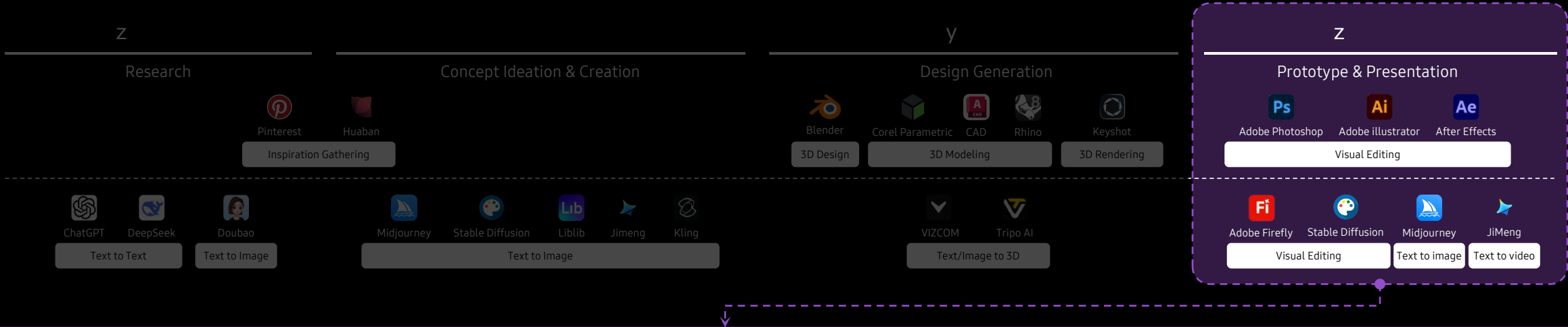


Tongji University, Industrial Design student

h Au h

U“” “(Tc

T”“ (Tc



Midjourney



JiMeng

Usage Scenarios

- Midjourney is used to generate keyframes of product videos, while JiMeng creates video segments based on these keyframes



Strengths:

- Midjourney provides high-quality visuals with a strong cinematic feel
- JiMeng performs well in maintaining character and scene consistency, with smooth transitions between shots



Weaknesses:

- Midjourney does not support ControlNet, resulting in poor product consistency across frames
- JiMeng restricts video duration (5–10 seconds), requiring manual stitching, which often leads to a lack of coherence between the first and last frames of different clips.



Keyframe images generated by Midjourney



Video clips generated by JiMeng

“I use AI to generate key product scene images and then animate them. For video generation I use JiMeng, while for keyframes I rely on Midjourney.”



Tongji University, Industrial Design student

