China Artificial Intelligence

China AI Application Tour Takeaways: Early Evidence of Workflow-Led Monetization

Head of China Equity Research and Co-Head of Asia TMT Research

Alex Yao AC
(86 21) 6106 6505
alex.yao@jpmorgan.com
SAC Registration Number: S1730523020001

Olivia Xu
(86-21) 6106 6138
olivia.w.xu@jpmorgan.com
SAC Registration Number: S1730525060001

Jiajie Shen, CFA
(86-21) 6106 6352
jiajie.shen@jpmchase.com
SAC Registration Number: S1730520030006

J.P. Morgan Securities (China) Company Limited

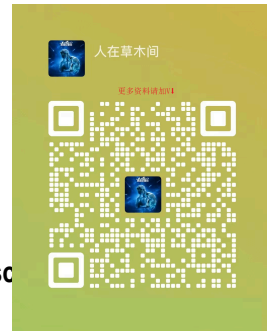
Eight meetings across autonomy, independent model development, enterprise workflow software and vertical AI applications point to early evidence of workflow-led monetization in selected use cases. The most recurring theme is that value creation appears increasingly tied to proprietary data, workflow context, customer integration and deployment know-how. Foundation models remain important, while many companies are using them as flexible inputs within broader product and workflow stacks. A useful working framework is "own the context, rent the model," although we would still treat this as a direction of travel that needs more public-market proof.

This pattern appeared across enterprise data integration, insurance and financial-services risk, cross-border marketing, independent model development and listed model-plus-infrastructure businesses. Several companies described routing tasks across different domestic and frontier models based on price, performance and use case. One listed model-plus-infrastructure company also acknowledged that raw API switching costs can be low, with customer stickiness improving after enterprise data and workflow integration. We view this as an important signal that base-model differentiation alone may face pressure over time, with the degree of pressure likely varying by use case, customer segment and workflow ownership.

The tour modestly changes our prior assumption in two directions. Previously, we viewed robotaxi as a cost line with limited near-term standalone economics, AI application monetization as still early, and infrastructure/compute as the clearest listed exposure. The meetings suggest the first two views deserve more nuance. Some vertical and enterprise AI companies described recurring revenue, value-based pricing and operating-level profitability, while autonomy companies pointed to improving unit economics in selected deployments. These datapoints are mostly management-stated, often private-company specific, and still require verification through filings, customer evidence and repeatable public-company disclosures.

As a result, our takeaway is that selected AI applications are showing early signs of commercial relevance, particularly in workflow-heavy and data-rich verticals. The key question is whether these early signals can translate into durable, scalable and software-like economics. Similarly, the tour suggests rising risk that base-model access becomes a more competitive and interchangeable input in some enterprise settings. Model vendors may still defend economics in high-value workflows such as coding, agents and enterprise software use cases where they control the user surface, workflow memory or data loop.

Our investment stance shifts incrementally. We become more constructive on AI infrastructure, compute, domestic-chip, memory and storage beneficiaries, as they can benefit from broader AI deployment even when model leadership rotates. We also become more positive on AI applications and workflow-layer companies that can translate model capability into customer-facing value through proprietary data, workflow ownership, pricing power, retention and



See page 9 for analyst certification and important disclosures, including non-US analyst disc

deep customer integration. Autonomy also moves from a pure cost-center framework toward an early-commercial hypothesis in selected ADAS and overseas robotaxi deployments, while third-party-verifiable unit economics remain necessary before assigning broader valuation credit.

Key observations

Observation 1: Early evidence suggests model access may become more interchangeable in select enterprise use cases

This was the most consistent signal across our meetings, although the conclusion still needs to be tested against public-company disclosures and customer behavior over time. Across multiple meetings, companies described base-model capability as necessary, while customer value appeared to depend more on task completion, workflow integration, proprietary data and deployment quality. One independent model developer argued that customers in the agentic era care less about using the largest model and more about reliable task completion. A listed model-plus-infrastructure incumbent acknowledged low API switching costs and located customer stickiness in enterprise-data integration. Two vertical AI application companies said that newer or cheaper models could be integrated into their products, with limited impact on customer-facing outcomes because the application value sat in vertical data, workflow design and deployment context.

This updates our framework from a pure model-capability debate toward a broader capability-plus-context debate. In many enterprise use cases, customers appear willing to use whichever model is adequate, cost-effective and available, while paying for the workflow layer that embeds AI into business processes. This does not mean model quality loses relevance. It means model capability may need to be paired with proprietary workflow control, user interface ownership, data feedback loops and customer integration in order to support durable value capture.

The read-through is constructive for workflow-led AI companies and selective for model-led valuations. Base-model leadership can still matter in high-value workflows such as coding, agents, enterprise software automation and other tasks where reliability, context length, tool use and multi-step completion rates directly affect business outcomes. At the same time, generic API access may face growing pricing pressure if customers can route across multiple models with limited operational friction. For listed model companies, the key question is whether they can prove workflow ownership, retention, pricing power and rising enterprise value capture beyond benchmark leadership.

Observation 2: Select vertical and enterprise AI applications are showing early commercial signals

Several private vertical AI vendors described recurring revenue, value-based pricing and operating-level profitability. The strongest examples came from insurance and financial-services risk, enterprise data integration and cross-border marketing. We deliberately do not reproduce the specific figures because they are non-public and several companies are private or potentially in listing processes.

On our rubric, these businesses appear to sit between Early-Commercial and Scaled. This represents a potential upgrade from our prior assumptions, which treated most AI-application monetization as early-stage and difficult to underwrite. The meetings indicate that some customers are paying, and in select cases pricing is linked to measurable business value, such as risk reduction, workflow efficiency, marketing conversion or decision automation.

At the same time, the broader usefulness of LLM-driven applications remains debatable. The evidence we saw points to promising signals in workflow-heavy, data-rich verticals, with less proof for the AI applications as a whole. The figures are self-reported and still need independent verification. Some management teams also acknowledged high customer concentration, sub-scale wallet share at large clients, and lower-margin service-heavy or human-assisted cold-start segments. These caveats matter because the economics may look less software-like once customer acquisition cost, implementation work, human assistance and concentration risk are fully reflected.

As a result, our takeaway is that select AI applications are moving from narrative toward early commercial validation, supported by signs of recurring revenue, value-based pricing and operating-level profitability in some verticals. We would still treat this as an emerging thesis. The durability, scalability and software-like quality of these economics need to be tested through filings, customer evidence and repeatable listed-company disclosures.

Observation 3: Autonomy may be moving from pure cost center toward early-commercial in select use cases

Autonomy was the area where the tour most challenged our prior assumptions. We previously viewed robotaxi as a cost line with limited near-term standalone economics. The tour suggests that this view may be too conservative in select segments, although broader sector underwriting still requires stronger verification.

The distinction between ADAS and L4 matters. Mass-production ADAS appears closer to Scaled, supported by OEM adoption, real-world data loops and software-like margin potential at volume. Robotaxi is earlier. Listed L4 operators pointed to improving city-level economics, declining vehicle costs and better economics in certain overseas deployments, while company-level profitability remains a later-stage objective and regulatory gating remains material. L4 fleets therefore remain Early-Commercial in select cities and Pilot in broader deployment terms.

The investable conclusion is selective. ADAS and supplier exposure look more underwritable than L4 robotaxi pure plays because they are closer to mass production, OEM attach rates and visible revenue contribution. Overseas robotaxi deployments may offer better pricing and less competition, although current evidence remains early and management-stated. For listed L4 companies, the key test is auditable segment economics, license-city expansion, vehicle cost-down, safety record and regulatory continuity.

Observation 4: Neutral specialists are winning some OEM and enterprise sockets

The tour complicates our platform-distribution prior. We had previously assumed that platforms with distribution surfaces, cloud infrastructure and ecosystem control would capture much of the AI application value. The meetings suggest a more mixed outcome.

One independent, non-platform model developer described winning OEM and enterprise deals partly because it is perceived as neutral, customizable and less tied to a competing consumer or cloud ecosystem. Its model consumes platform cloud and map infrastructure as inputs. This points to a split-value chain where neutral specialists can own deployment context, while platforms monetize cloud, maps, compute and supporting infrastructure.

This remains relevant for platform valuation. Large platforms like Tencent and Alibaba can still benefit through cloud, maps, compute, data infrastructure and enterprise distribution. At the same time, platform-owned models may face a more selective path in OEM and enterprise procurement when customers prioritize neutrality, customization and workflow integration. Distribution remains important, yet procurement decisions may increasingly depend on trust, integration depth and vertical implementation capability.

Observation 5: Compute and inference demand appears to be broadening across use cases

Every company we met was a net consumer of compute, whether through model training, inference, ADAS development, enterprise deployment, simulation, data processing or managed infrastructure. The more important point is that demand appears increasingly deployment-driven and model-agnostic.

This supports our prior positive view on infrastructure. If enterprise and vertical AI users route across multiple models, the winning infrastructure exposure may depend less on picking the best model and more on owning capacity, memory, storage, domestic-chip adaptation, cloud orchestration and inference infrastructure. Model competition may also increase experimentation and deployment, which can support aggregate demand for inference and supporting infrastructure.

The key assumption is volume elasticity. The bull case requires AI task volume to grow faster than cost per task declines. If efficiency gains reduce compute intensity faster than use cases expand, the infrastructure read-through weakens. Based on the tour, deployment appears to broaden the demand base from frontier training toward inference, domestic-stack adaptation, memory, storage and enterprise workflow execution. That makes infrastructure the cleanest listed exposure in the current evidence set, subject to continued evidence of inference growth and utilization.

Observation 6: Consumer AI and agent commerce still look earlier-stage from an underwriting perspective

Most companies avoided direct consumer AI monetization as a near-term business model. Several cited low loyalty, intense competition, weak willingness to pay and rapid product imitation. The cross-border marketing company also framed much of its optimization work as recommendation-algorithm driven, with LLM-driven agent commerce and AI-ad revenue still at an earlier stage.

On our rubric, consumer AI and agent commerce remain Narrative-stage. The opportunity could be large, while current evidence is still insufficient for listed-company valuation credit. The key proof points should be retention, payment conversion, repeat usage, gross margin and measurable transaction uplift, beyond downloads, demos or product launches.

What we would monitor from here?

The most important question for listed model developers is whether the model layer can capture durable rent when customers can route tasks to the most suitable model and place proprietary workflow, data and context above it.

At present, scarcity and usage momentum can support listed model-developer multiples. There are few pure-play listed exposures, and usage growth is easier to measure than moat durability. The tour suggests investors may need to reframe the underwriting question. Enterprises, OEMs and model-plus-infrastructure companies repeatedly described multi-model routing, price-performance trade-offs and low raw switching costs. If that persists, the market may need to shift from valuing model usage toward valuing pricing power, retention, workflow ownership and vertical-specific value capture.

The resolution should be observable over the next few quarters. The constructive case for model developers requires evidence of defended verticals: rising net retention, stable or improving pricing, low churn, and gross margins that hold up under multi-model routing. Coding is the most obvious test case because it combines high-frequency workflow, measurable productivity value and potential developer-surface lock-in. The cautious case strengthens if usage grows while gross margin compresses, customers route increasingly by price, and model vendors fail to show vertical retention or workflow control.

The market may also be underestimating the degree to which AI applications support the infrastructure thesis. If deployers use multiple models, that pressures model exclusivity while supporting shared infrastructure. Model competition can accelerate experimentation, deployment and inference volume. The more companies rent models, the more they still need compute, memory, storage, cloud orchestration and domestic-stack adaptation.

Autonomy may also be closer to commercial relevance than the simple cost-center framing implies. The tour provided concrete management-stated examples of improving city-level economics, overseas pricing and ADAS margin potential. This remains early evidence. The gap persists because company-level losses, regulatory uncertainty and management-sourced unit economics remain material. The sector becomes more investable when city-level contribution profit, license expansion, vehicle cost-down and safety performance become independently verifiable.

Investment implications

We are most constructive on AI infrastructure, compute, domestic-chip, memory and storage beneficiaries. This remains the cleanest read-through because the demand signal does not require us to pick the winning foundation model. Model-agnostic deployment, inference growth, domestic-stack adaptation and memory/storage bottlenecks all point in the same direction. The call strengthens if listed cloud and hardware companies begin to attribute growth more clearly to inference, deployment and domestic-chip mix. It weakens if architecture efficiency cuts cost per task faster than task volume scales, or if capacity loosens enough to pressure pricing.

We are also more positive on AI applications and workflow-layer companies that can translate model capability into customer-facing value through proprietary data, workflow ownership, pricing power, retention and deep customer integration. The tour suggests that application-layer monetization is starting to show early proof points in select verticals, especially where AI is embedded into measurable workflows such as risk assessment, enterprise data integration, marketing conversion and operational automation. The key underwriting variables are revenue quality, customer concentration, implementation intensity, net retention, pricing structure and gross margin after customer support and human-assisted workflows.

For model-led names, the read-through is more nuanced. Base-model capability remains important, especially in coding, agents and enterprise software workflows where task completion quality directly affects customer willingness to pay. The incremental question is whether model companies can convert capability into workflow ownership and customer stickiness. In Zhipu's case, the Overweight thesis needs to rest on coding as a defended workflow vertical, supported by developer retention, pricing power and measurable enterprise value creation. For MiniMax and other model-led names, the same framework applies: vertical context, workflow

control and repeatable usage quality matter more than benchmark performance alone.

We are selectively constructive on autonomy, with position sizing tied to evidence quality. ADAS is more underwritable than L4 robotaxi because it is closer to mass production, OEM attach rates and supplier economics. L4 robotaxi has moved from "no path" to "early-commercial hypothesis" in selected cities and overseas deployments, while public-market underwriting still requires auditable city economics, license expansion, vehicle cost-down and regulatory stability. The call breaks if an accident triggers a multi-quarter licensing pause, if cost-down stalls, or if city-level contribution economics fail to generalize.

We are mildly positive on platforms as cloud, maps, distribution and infrastructure suppliers, while taking a more selective view on platform-owned models in the application layer. Platforms can benefit if neutral specialists consume their cloud, maps and compute. Platform model units may still lose OEM and enterprise sockets where customers prefer neutral, customizable vendors. We would monitor OEM design wins, enterprise procurement share, cloud inference growth and the split between infrastructure revenue and model/application revenue.

We would continue to discount consumer AI and agent-commerce narratives until monetization becomes more visible. The burden of proof should be retention, payment conversion, repeat usage, gross margin and measurable transaction uplift, beyond downloads, demos or product launches.

Risks and counterarguments

The key constructive numbers are not independently verified. The profitability, margin and break-even claims behind the vertical-app and autonomy observations are mostly management-stated, and in some cases tied to private companies. These claims may not survive audited filings, prospectuses or public-company segment disclosure. Until then, they should inform hypotheses more than valuation.

The context moat may prove less durable than expected. Enterprises and large platforms may insource workflow orchestration, benchmark vendors more aggressively as inference costs fall, or renegotiate pricing once the underlying model input becomes cheaper. If customers capture more of the value created by AI workflow tools, application-layer margins may compress.

Infrastructure demand may be less elastic than the tour implies. The bull case assumes task volume grows faster than cost per task declines. If model efficiency, sparsity, edge inference or architecture improvements reduce compute intensity faster than deployment scales, cloud, chip, memory and storage beneficiaries may disappoint.

Robotaxi remains regulatorily gated. City-level contribution economics, even if accurate, may not generalize across geographies, weather conditions, safety standards and regulatory regimes. Accident-driven licensing pauses remain a material risk and can delay commercialization even if the technology improves.

Vertical-app economics may be more services-heavy than software-like. High top-customer concentration, sub-scale wallet share, manual implementation work, human-assisted workflows or influencer-assisted segments can all make reported profitability less scalable than it appears.

Platforms may reassert themselves. Neutral-vendor advantage may be temporary if platforms subsidize model access, bundle cloud and application tools, or use distribution to regain OEM and enterprise sockets. If that happens, independent application vendors could face pricing pressure even if their products remain useful.