

業務

概覽

關於我們

我們是中國領先的開放、獨立詞元供應平台，正處於AI模型在企業中快速部署及應用的時代前沿。隨著AI模型日益融入日常工作流程，對詞元的需求呈指數級增長，同時對兼具準確性、生成速度、成本效益、穩定性與安全性的高品質詞元供應的需求也變得愈發迫切。詞元供應平台整合算力資源、AI模型與系統軟件能力，提供一站式AI基礎設施服務，從而能夠大規模交付優質詞元。

根據弗若斯特沙利文的資料，按2025年的詞元年吞吐量計，我們是中國最大的獨立生態詞元供應商，在中國所有詞元供應商中躋身前五之列。憑藉我們專有的推理引擎及算力資源編排系統，我們有效聚合、優化及編排不同供應商不同硬件架構中的異構算力資源，從而將底層算力轉化為標準化、可靠及可擴展的詞元供應。藉此，我們能以行業領先的性價比，在廣泛的最先進模型中無縫獲取及部署AI能力。為滿足多元化的商業需求，我們主要通過公有雲服務與本地部署解決方案交付核心能力，為橫跨眾多行業的客戶提供服務。

下圖列示我們的平台於AI價值鏈中的位置：



業 務

截至2026年4月30日，我們的平台有超過10百萬名註冊用戶，日均詞元吞吐量約為5,785億次，於2026年4月的單日最高詞元吞吐量約為10,714億次。截至最後實際可行日期，我們已為超過13,000家企業客戶提供服務，並且我們的平台已支持累計合共170個模型。下圖列示我們迄今為止的部分主要成就：



附註：

- (1) 根據弗若斯特沙利文的資料
- (2) 按2025年的詞元年吞吐量計
- (3) 按所支持的芯片及模型數量計

我們的價值主張

AI技術棧圍繞算力資源、AI模型及系統軟件三個核心支柱展開。算力資源提供AI訓練與推理所需的處理能力；模型定義AI系統的能力與智能水準；系統軟件則將算力資源與模型連接起來，使AI能力在實際應用中落地。作為AI技術棧中的關鍵系統軟件支柱，我們連接了算力資源、模型與AI應用，為AI全產業鏈的參與者提供極具吸引力的價值。

面向客戶。我們為眾多行業的客戶提供一站式接入多元AI模型與基礎設施能力的服務，實現更高效、更具成本效益的AI應用部署。

- **簡化部署：**我們統一及標準化的API使客戶能夠通過單一接口快速獲取AI能力，大幅降低系統整合的複雜性、縮短部署週期，降低落地的技術門檻。
- **性能提升：**我們的專有推理引擎旨在提高跨不同算力環境的詞元吞吐量、降低延遲及增強運營穩定性。

業 務

- **成本優化：**憑藉我們的跨芯片及多模型適配性，我們能夠根據客戶特定的性能、延遲及成本要求靈活匹配算力資源與模型，讓客戶能在保持服務質量的同時降低單位詞元成本。
- **靈活性與可擴展性：**我們的彈性擴展架構及多元化部署選項使客戶能夠根據應用需要量身定制計算能力、AI模型選擇及部署架構，支持從研究及測試到企業級生產部署的全流程AI落地。

面向生態系統合作夥伴。我們開放獨立的定位創造了一個協作的環境，可以協調激勵措施並加強整個AI價值鏈的互通性。通過我們的全棧軟件能力（包括跨芯片及多模型適配、推理引擎優化、異構算力資源編排及標準化詞元交付能力），我們協助：

- **芯片設計商**通過在真實世界工作負載中實現高效、可擴展且可投入生產的推理能力，加速AI芯片的商業化。
- **算力資源供應商**將碎片化算力資源轉化為具備可商業交付的詞元生產能力。
- **模型開發者**在多元算力環境下實現規模化、可靠的AI推理。
- **工具鏈與應用創建者**加快產品部署與商業化落地進程。

我們極具吸引力的客戶價值主張推動了客戶群的大幅擴張。隨著服務的客戶與使用實例範圍不斷拓寬，多元的客戶需求支撐了定向技術優化，拓展的應用場景也加速了技術迭代，為獲取新客戶創造了更多機會。作為開放獨立的平台，我們持續通過擴大對更多算力資源與模型的兼容性，提升技術通用性。依托積累的技术能力與落地經驗，我們在接入新的算力與模型環境時，能夠降低適配成本、加快部署速度，進而進一步強化核心價值主張，鞏固自身在產業生態系統中的影響力。

我們的核心技術基礎

我們不斷投資開發專為AI推理工作負載而構建的專有推理引擎及算力資源編排系統。

全棧推理引擎。我們的推理引擎作為關鍵的系統軟件層直接在芯片上運行。通過專家並行(EP)、預填充與解碼(PD)分離、KV緩存優化、流水線並行、算子優化以及運行時優化與調度優化等技術，我們的推理引擎將底層算力資源轉化為高吞吐量、低延遲且具備成本效益的詞元。

算力資源編排系統。我們面向具有異構算力環境、高內存需求、低延遲推理及突發高併發需求等特徵的AI推理場景，針對性設計算力資源編排系統。我們的系統超越傳統雲能力，將計算、存儲、網絡、調度、安全及可觀測性整合至一個專注於推理實例調度、部署、擴展、監控及運行的框架中。我們的推理引擎與算力資源編排系統共同構成支持大規模詞元生產的一體化技術棧。

憑藉我們在算力資源、AI模型及系統軟件的工程能力，我們能夠在不同的模型、芯片架構、算力資源供應商及客戶環境中複製推理優化最佳實踐，同時快速適應技術迭代。

業 務

我們的產品及服務

我們已建立涵蓋公有雲及本地部署環境的全面軟件及服務產品矩陣。通過標準化及高度兼容的統一API與全棧軟件能力，為客戶提供穩定、低延遲且具備成本效益的生產級詞元及全流程技術支持。

公有雲服務。我們的公有雲服務旨在規模化交付標準化、高質量的詞元，通過單一標準接口為客戶提供對逾170個AI模型的統一及靈活獲取途徑，涵蓋包括編碼、智能體及多模態應用在內的廣泛AI應用場景。我們於2024年5月推出無服務器詞元服務，標誌著公有雲服務邁入大規模商業化階段。該平台支持兩項互補的服務－基於共享資源無服務器詞元服務及提供預留算力的專屬實例。

- **無服務器詞元服務：**我們提供基於共享資源的詞元服務，基於按詞元消耗的使用量收費。該等服務主要針對尋求以靈活及具成本效益的方式獲取AI能力的開發者、中小型企業及其他客戶，我們一般對客戶採用預付費模式（即預先充值並根據詞元使用量進行扣減），而對若干多次持續性採購的企業客戶則採用按月結算的後付費模式。
- **專屬實例：**我們於公有雲基礎設施上提供預留或專屬算力，為特定客戶預留算力資源。該等服務專為對性能或穩定性有較高要求的客戶設計，包括運行對延遲敏感、高吞吐量AI工作負載的企業用戶。定價按預留算力確定，並計及底層基礎設施成本，包括算力資源及相關運營開支。

本地部署解決方案。我們將推理引擎及算力資源編排系統在客戶自有數據中心或私有算力環境內部署。這使客戶能夠在其自身的算力資源上，跨模型使用詞元。本地部署解決方案主要針對大型企業客戶及機構客戶。我們通常透過基於訂閱或永久買斷的方式提供該等解決方案，費用包含軟件許可費（按訂閱或一次性基準）、專業實施費及（如適用）維護費，各項費用均具體視乎部署規模、系統配置及客戶要求等因素而定。我們的本地部署解決方案基於標準的模塊化平台架構，僅需要進行有限的定制即可滿足客戶的特定部署環境或運營要求。

我們公有雲服務與本地部署解決方案構成了一個能滿足多樣化客戶需求的一體化商業體系。公有雲平台使我們能夠服務廣泛的客戶群、促進跨多樣化應用場景的測試及驗證，並支持客戶的概念驗證及初始部署要求。隨著客戶使用量增加，或對性能、穩定性、數據實際存放位置、網絡邊界及供應鏈保障需要更高程度的控制，客戶可能選擇專屬實例或本地部署解決方案，從而加深參與度及擴大收益機會。

我們的商業化及財務表現

憑藉我們開放及獨立的生態定位、全棧推理優化技術及多元化產品矩陣，我們實現了快速商業化規模擴張。自2023年8月成立以來，我們的收入大幅增長，由2024年的人民幣7.3百萬元增加至2025年的人民幣55.3百萬元，增長逾七倍。

業 務

我們已與AI、互聯網、能源、金融和電信等領域的領先客戶，以及芯片設計商及模型開發者等生態系統合作夥伴建立關係。在公有雲市場，我們已成為開發者社區廣受認可的AI基礎設施平台，詞元吞吐量在全球領先平台中名列前茅。在私有雲市場，我們通過於昇騰、沐曦及摩爾線程等國內領先芯片平台進行大規模部署，與中國國內算力生態系統建立深度整合。我們於2025年2月在業內首次推出了基於國產芯片的DeepSeek R1和V3詞元服務；於2025年7月率先在MetaX C550芯片上完成Kimi-K2 Token服務的全球首次商業化部署，並基於摩爾線程芯片部署DeepSeek-V3達致最頂尖的性能。該等部署展示了我們具備將國產AI芯片理論算力轉化為可支撐大規模實景AI應用、具備商業化價值的詞元生產能力。

我們相信，我們結合技術創新、大規模工程部署及生態系統協作的的能力，使我們能夠把握不斷增長的詞元需求，維持我們的長期增長。

我們的優勢

開放獨立的平台，充當AI價值鏈中的關鍵基礎設施層

我們是一個開放及獨立的詞元供應平台，基於不同性價比的算力資源及模型提供詞元供應。我們服務廣泛的AI應用開發者及企業客戶，促進AI應用的大規模部署及商業化，為推進全球智能化轉型作出貢獻。截至最後實際可行日期，我們已支持累計合共逾170個模型，按模型覆蓋範圍計為全球領先的獨立生態詞元供應平台之一。我們開放及獨立的生態定位加強了我們的全球合作夥伴網絡、鞏固了我們的結構性競爭優勢，確立了我們在全球AI價值鏈中關鍵基礎設施層的地位。

我們相信，我們開放的定位在快速演進的AI行業中創造了長期戰略價值。在AI需求快速增長的推動下，AI價值鏈吸引了涵蓋算力資源、模型及系統軟件的廣泛參與者，導致模型迭代加速、模型架構日益多樣化、芯片持續創新及多層軟件生態系統的形成。因此，跨芯片、模型及系統軟件及基礎設施層的互操作性及高效協調，對於大規模AI部署變得日愈重要。我們通過開放的平台架構，有效整合多樣化的模型及異構算力資源，將其轉化為標準化及可擴展的詞元供應服務，從而促成整個AI價值鏈中更廣泛的協作。

我們致力於維持一個獨立的平台，連接模型層與算力層，不受限於任何單一供應商。對於開發者，我們的平台提供獲取多樣化模型能力及靈活部署選項的一站式途徑。對於企業客戶，我們的中立定位增強了運營靈活性及戰略獨立性，同時使我們能夠提供企業級的服務水平保證。我們相信，我們的獨立詞元供應服務使客戶能夠在廣泛的用例中實現更靈活、高效及可靠的詞元服務。

我們的商業化成果驗證了此項開放及獨立的平台戰略。在涵蓋能源、電信運營商、大型企業及金融機構的多元化客戶群的推動下，我們的客戶數量由2024年的逾2,400名增加至2025年的逾716,000名。我們相信，如此廣泛的採用證明了我們平台的可擴展性及商業相關性。

業 務

具備跨芯片及多模型適配能力的全棧推理引擎

我們專有的推理引擎專為高性能AI模型推理而設。它將預訓練的模型權重及計算圖加載至底層計算硬件上，通過一系列系統級優化技術處理用戶的提示詞，並高效生成響應。推理引擎作為核心執行層，將算力資源轉化為可用的AI能力。透過提供用於模型加載、執行及調度的統一接口，我們的推理引擎抽象化了GPU、NPU及ASIC之間底層架構的差異，使同一模型服務能夠在不同的計算平台上高效運行。支持廣泛的先進推理技術，並能夠高效部署DeepSeek、GLM、Kimi、MiniMax及Qwen等主流先進模型。

我們的推理引擎支持各類先進推理技術，可高效部署DeepSeek、GLM、Kimi及Qwen等主流AI模型。我們通過精細化的請求調度與資源管理，最大化提升算力利用率，降低單詞元成本，在高併發AI負載下輸出穩定的低延遲推理性能。我們的推理引擎還支持快速適配迭代中的模型架構，包括稠密模型、稀疏混合專家模型與多模態模型。憑藉我們精簡的模型適配和部署工作流程（涵蓋模型權重獲取、框架適配、性能驗證、準確性驗證和生產部署），我們能夠快速支持新發佈的主流模型，並在多元模型組合中保持廣泛的兼容性。

支持多種AI芯片的異構兼容性

我們是少數能夠在大規模的生產環境中，同時支持國際領先芯片（如NVIDIA及AMD的GPU）以及主要國產AI芯片（包括昇騰、沐曦、摩爾線程、海光及平頭哥的芯片）的詞元供應商之一。憑藉我們深厚的全棧工程能力，我們將異構算力資源，轉化為具備靈活容量分配能力的標準化、可擴展的詞元。

我們推動了國產AI芯片從理論兼容邁向生產級推理部署的進程，將國產算力資源轉化為具備商業可行性及成本效益的詞元供應。例如，我們與華為雲合作，共同為CloudMatrix超節點集群開發並優化推理解決方案，在DeepSeek-R1上實現單卡解碼吞吐量超過1,920詞元/秒。我們亦完成了Kimi-K2在基於沐曦C550芯片構建的集群上的全球首次商業服務部署。此外，我們在由摩爾線程開發的MTT S5000芯片上，於DeepSeek-V3實現了單顆芯片預填充吞吐量超過4,000詞元/秒及解碼吞吐量超過1,000詞元/秒。我們亦繼續擴大與其他國產芯片供應商的合作。特別是，在國產芯片上實現AI模型推理，需要跨AI技術棧進行大量的系統級軟件重構。我們已克服一系列複雜的技術挑戰，包括：由緊密耦合的代碼庫、缺失算子及架構差異引起的生態系統兼容性挑戰；與內存容量及帶寬瓶頸相關的內存挑戰；由缺乏高速互連基礎設施及通信庫優化不足引起的通信挑戰；以及由於低精度計算支持有限而導致的精度挑戰。透過該等優化，我們在國產芯片上實現了行業領先的大規模AI模型推理的可靠、高效及精準運行。

隨著全球對低成本、跨芯片及多模型詞元服務的需求持續增長，我們相信，我們在中國高度碎片化的算力環境中所開發的全棧優化及異構適配能力，使我們具備強大優勢，足以服務全球AI需求。

業 務

靈活的算力資源編排及最大化的詞元生產效率

我們已在業務的技術、商業及營運維度上，建立起算力資源編排及詞元生產效率優化方面的核心競爭優勢。

技術。我們已開發專為AI推理工作負載打造的生產級算力資源編排系統。在傳統雲基礎設施的虛擬化、調度及交付能力基礎上，我們針對AI推理工作負載的獨特特性進一步優化基礎設施技術棧，包括異構算力資源、大規模模型、低延遲推理及突發併發需求。我們系統性地解決了與計算、網絡、存儲、調度、安全及可觀測性相關的廣泛工程挑戰。具體而言，我們的安全容器架構以虛擬機級運行時環境取代傳統共享內核容器模式，實現近硬件級隔離，並提升租戶數據私隱及營運安全性。我們的分佈式模型分發及緩存系統匯聚集群節點之間的存儲及帶寬資源，實現大規模模型權重的高效存儲及快速分發。此外，我們的AI模型API網關將多個推理集群聚合為統一服務端點，能夠在異構算力資源之間動態路由請求、實時監測後端系統狀況並執行自動故障轉移，從而確保高度的服務連續性。

基於該基礎設施底座，我們專有的推理引擎可將推理延遲降低最高達70%，並將吞吐量提升三至五倍，而我們的動態量化技術可將推理計算需求降低60%至80%，從而最大化每顆芯片的詞元生產效率。我們的整體基礎設施架構提供99.95%的SLA，並支持自動故障轉移、熔斷及請求限流能力，使生產級、高併發AI應用能夠穩定運行。

商業。我們已開發多元化業務模式。

- 就開發者及企業客戶而言，我們已支持累計合共超過170個模型及行業標準協議，使我們能夠透過具競爭力的定價及廣泛的生態覆蓋快速吸引客戶。
- 就擁有自建算力集群的企業及機構而言，我們提供解決方案，以提升推理效率，同時使客戶能夠透過外部詞元服務將剩餘算力變現。
- 此外，我們採用聯合營運模式，據此，我們透過提供技術平台及品牌支持與區域智算中心合作，而合作夥伴則基於實際詞元服務量參與收入分成，從而大幅降低其技術投入及營運門檻。

不斷增長的客戶使用量吸引更多模型提供商及算力資源合作夥伴加入我們的平台，而更廣泛的模型覆蓋及算力資源可用性則進一步提升客戶採用率及平台活躍度。

營運。我們已建立涵蓋從算力資源到可擴展詞元供應全流程的營運能力。在需求端，憑藉我們廣泛的模型生態系統及面向行業的解決方案，我們服務於AI、互聯網、能源、金融和電信等領域的領先客戶，協助提升算力資源利用率，並為我們的算力資源合作夥伴創造穩定需求。在供應端，透過可視化管理工具及30多個即用型部署模板，我們將算力資源整合時間縮短至三分鐘以內，且無需手動配置，顯著降低算力資源提供商的接入複雜度。透過我們的平台將碎片化算力資源與多元化詞元需求連接起來，我們持續提升資源利用效率，並加強AI生態系統內的協同。

業 務

憑藉產品能力及多樣化應用場景實現強大的客戶獲取及留存

作為詞元供應市場的早期參與者，在全球詞元消耗需求大規模加速增長之前，我們已在AI社群中建立廣泛影響力。憑藉先發優勢及廣泛的模型適配能力，我們已吸引並留存大量成熟開發者、新興AI原生公司及企業客戶。透過持續的平台使用及集成，我們形成了強大的客戶黏性，並積累了顯著的市場認可度。

我們的平台旨在最大化不同芯片與模型組合的價值，使客戶能夠根據其對性能、延遲、質量及定價的要求，為特定應用場景選擇最具成本效益的詞元服務。透過編排異構算力資源並支持廣泛的模型，我們能夠在多樣化AI應用場景中為客戶提供經優化的詞元供應，包括編程、AI智能體、多模態生成及企業AI應用。不同工作負載需要在詞元生成速度、模型能力及成本效率之間作出不同取舍，而我們的平台使客戶能夠根據其生產需求，動態選擇最合適的解決方案。

我們為客戶提供覆蓋完整部署範圍的綜合解決方案，包括無服務器詞元服務、專屬實例及本地部署解決方案。我們的公有雲服務使客戶無需管理複雜的AI基礎設施，即可便捷接入廣泛的模型，從而推動開發者生態系統增長及平台規模效應。客戶可根據其營運需求及成本考量，在無服務器詞元服務、專屬實例或本地部署服務這三種服務模式之間作出選擇，同時亦受益於解決方案部署選項及低延遲詞元服務。與此同時，我們的本地部署解決方案為企業客戶提供更高水平的數據隔離、營運控制及定制化性能優化能力，帶來更高利潤率的收入來源及更強的企業客戶留存。該多樣化產品組合提升了AI採用不同階段的客戶轉化，並使我們能夠在客戶由概念驗證項目擴展至大規模生產部署的過程中，保持強大且具韌性的客戶留存。

此外，來自我們優質客戶群的反饋及營運洞察，持續提升我們對客戶需求及應用場景的理解，使我們能夠優化平台能力、模型選擇及資源編排策略。我們相信，這將形成強大的飛輪效應，隨著時間推移不斷提升客戶黏性、服務質量及生態系統競爭力。

植根於卓越技術的強大工程文化

我們深厚的技術專長及工程驅動文化構成了我們競爭優勢的基礎。我們的創始人袁進輝博士是一位資深AI創業者，曾在AI領域創辦多項成功事業。自成立以來，我們便將詞元供應視為AI價值鏈中的關鍵機遇，此乃基於我們堅信，隨著AI模型解鎖多樣化應用場景及詞元需求持續快速增長，可擴展、高效及高性價比的推理將成為關鍵瓶頸。

我們已組建一支高度專業化的工程團隊，其專長涵蓋分佈式系統、推理引擎及雲原生基礎設施。我們的核心能力在於，能夠高效地將快速演進的模型架構適配至異構算力資源，並在延遲、吞吐量及成本效率等方面優化系統級性能。根據弗若斯特沙利文的資料，該端到端工程能力使我們能夠在約三個月內完成國產芯片的推理引擎適配及性能優化，顯著快於傳統雲服務商。我

業 務

們專有的推理引擎、動態量化算法及分佈式緩存系統均透過工程師主導的創新開發，並迅速部署至生產環境，展現了我們強大的工程執行能力。我們維持扁平化且以工程為中心的組織架構，重大技術決策由具備一線工程經驗的團隊推動，並透過嚴格的技术評估及數據驅動分析進行驗證。

我們的管理團隊兼具來自全球領先雲計算及AI公司的深厚工程專業知識，以及強大的戰略執行能力。我們在早期即將生產級推理基礎設施確立為戰略重點。透過持續技術投入及長期工程積累，我們已在異構算力適配、推理優化及大規模資源編排方面建立差異化能力及深厚工程專有技術。我們相信，該等能力的累積及複合效應已在AI模型推理基礎設施市場形成持久的長期競爭優勢。

我們的戰略

為加強我們的競爭定位，並支持我們在AI時代的長期增長，我們擬推行以下主要戰略：

透過持續研發投入提升詞元生產效率並拓展詞元應用場景

隨著AI模型持續快速演進及AI基礎設施不斷升級，我們相信，客戶對高質量詞元日益增長的需求，將繼續成為我們業務快速增長的關鍵驅動因素。我們計劃繼續投資研發核心技術基礎，以進一步提升其在廣泛異構芯片上的效率。同時，我們尋求深化AI生態系統內的合作，並持續為全球開發者及企業客戶擴展詞元應用場景。具體而言，我們擬：

- *提升模型適配能力。*我們計劃繼續投資於適配及整合新發佈AI模型的能力，並進一步拓寬我們平台所支持的模型範圍。我們亦與模型提供商保持密切溝通及業務合作，以確保我們與新興AI技術趨勢及發展保持一致；
- *加強與異構芯片的兼容性。*我們計劃繼續擴展並優化對各類國產及國際芯片的支持，並將異構算力資源轉化為具備靈活資源調度能力的高性能、標準化詞元服務。我們亦擬進一步強化對國產芯片的優化能力，以應對AI算力資源供應受限的環境；
- *持續改善我們的算力資源編排系統。*我們計劃持續增加對算力資源編排系統的研發投入，進一步提升其可擴展性、可靠性及安全性，同時增強其在計算、存儲、網絡、調度及基礎設施管理等方面的能力。隨著AI推理工作負載的不斷演進，我們尋求進一步優化雲軟件，以適應異構算力環境及日益複雜的企業AI部署場景。
- *推進新一代推理技術套件。*我們計劃推進一套新一代推理技術，包括：(i)去中心化、基於數據流的推理架構，基於高速互聯的輕量級片上處理單元取代傳統以層為中心的協調方式，實現低延遲、高吞吐量及具容錯能力的推理；及(ii)大規模上下文記憶系統，將集群範圍內的內存及存儲資源整合為可擴展且具容錯能力的記憶池，支持近乎無限上下文的高效存儲及檢索，以支援新一代AI應用。

業 務

- *提升應用層及基礎設施層的詞元生產效率。*在應用層，隨著AI智能體正成為詞元消耗及企業智能化轉型的主要驅動因素，我們計劃構建高效的AI智能體中間件框架，以支持以更低詞元消耗開發創新型AI原生應用；及
- *持續構建開放詞元生態系統。*我們將強化作為具備全棧技術能力的開放獨立平台的定位，並尋求推動我們的模型推理及資源編排技術在AI生態系統中更廣泛採用。在公有雲領域，我們尋求更高效編排異構算力資源，並持續擴大可透過我們平台接入的算力資源規模。在私有雲領域，我們旨在提供一站式解決方案，協助更多企業釋放其內部數據及算力基礎設施的價值，並建立企業級「詞元工廠」。

擴大重點客戶覆蓋範圍並提升市場影響力

我們擬繼續擴大客戶覆蓋範圍，尤其是具備重大詞元消耗潛力的客戶。憑藉我們的核心技術能力，我們亦尋求加深對以下客戶的滲透：(i)需要靈活接入多個AI模型的客戶；(ii)對國產AI算力基礎設施有重大需求並高度重視AI及數據安全的客戶；及(iii)擁有異構算力基礎設施並尋求提升算力資源編排效率及現有算力資產利用率的客戶。

我們計劃繼續加強品牌認知度及市場影響力。具體而言，我們擬持續提升產品功能、擴展產品組合並強化服務能力，以進一步提高市場認可度及品牌影響力。我們亦擬發佈技術白皮書、案例研究、基準測試報告及行業解決方案，以展示我們的技術能力及服務優勢，並進一步提升我們的專業聲譽及市場公信力。此外，我們計劃透過主題演講、產品演示、聯合解決方案發佈及客戶案例分享等方式參與AI基礎設施行業會議、論壇及技術活動，從而提升品牌知名度並提高獲客及轉化效率。

我們亦擬繼續提升線上及線下一體化銷售及營銷能力。具體而言，我們計劃透過官方網站及開發者平台的獲客舉措、線上營銷活動、技術社群建設及內容營銷，擴展渠道銷售模式。我們亦擬通過擴充具備AI基礎設施、雲計算、企業軟件及行業解決方案經驗的銷售及解決方案團隊，強化面向重點行業及大型企業客戶的直銷能力。

強化算力資源供應能力及供應鏈韌性

隨著AI應用持續擴展，我們相信，AI算力資源供應在中短期內可能仍將維持緊張。因此，我們擬持續強化算力資源供應能力及供應鏈管理，以提升業務營運的穩定性，並降低算力資源價格波動的影響。具體而言，我們計劃擴大與更廣泛AI算力芯片型號及規格的兼容性，以及持續強化對國產AI芯片的適配能力，提升我們獲取算力資源的穩定性及靈活性。透過該等舉措，我們尋求進一步豐富並拓展自身的算力資源供應，並增強供應鏈韌性。

業 務

我們亦擬進一步深化與主要算力資源供應商的戰略合作，並探索更多合作模式，包括聯合算力營運及算力資源管理安排，以提升資源獲取效率、利用效率及整體營運協同。我們相信，該等舉措將進一步強化我們的算力資源供應能力，支持我們的持續業務增長及長期發展，並提升我們在更廣泛AI基礎設施生態系統中的地位。

憑藉海外生態系統及我們的技術優勢推進全球擴張策略

我們擬繼續拓展海外市場，並利用我們在中國所形成的全棧技術能力及產品優勢，提升我們在全球市場的競爭力。我們相信，海外市場為詞元服務帶來具吸引力的機遇，原因是海外市場具備充足的先進算力資源、成熟的AI應用生態系統，以及憑藉性價比優勢，國產AI模型在國際市場的採用率不斷提升。我們計劃以分階段方式推進國際化擴張策略，並以產品能力作為海外增長的主要驅動力。具體而言，我們擬：

- 透過我們的詞元生產效率、廣泛的模型支持能力及一體化產品組合，結合海外線上平台及數字營銷渠道，擴大海外客戶基礎；
- 透過與當地生態合作夥伴合作，加強本地化生態合作關係，以更好地服務需要本地化支持的客戶；及
- 在選定海外市場逐步建立本地化銷售及營運團隊，以提升客戶服務能力、市場響應能力及本地交付能力。

打造具全球競爭力的人才及管理平台

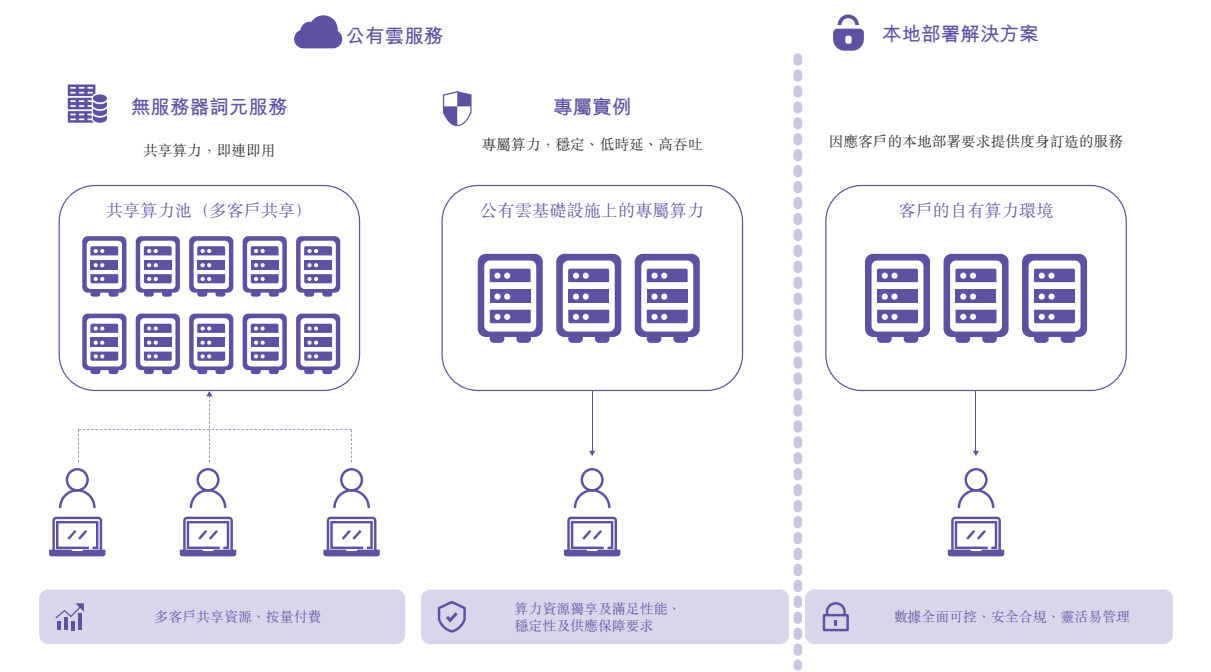
我們擬繼續強化在模型推理及AI原生雲基礎設施等關鍵領域的人才及管理能力。具體而言，我們計劃在技術、管理、營銷及其他職能方面吸引具備豐富經驗的行業專業人士、跨學科人才及具國際化視野的商業化人才，以支持我們持續的技術創新及業務擴張。我們高度重視僱員的技術專長、行業洞察、實踐經驗以及適應新興AI技術的能力，並尋求打造一支能夠支持我們在AI原生時代長期發展的高素質團隊。我們亦擬繼續培育公平、開放及以創新為導向的組織環境及企業文化，為僱員提供成長及發展機會。我們相信，該等努力將有助於我們吸引、培養及留住符合業務需求的人才，並支持我們持續進行技術創新、產品優化及業務增長。

我們的產品及服務

我們的業務主要包括兩項互補的產品及服務線，旨在滿足廣泛的客戶部署及使用需求：公有雲服務及本地部署解決方案。公有雲服務包括(i)無服務器詞元服務，客戶透過此服務按使用量獲取及使用廣泛的AI模型；及(ii)專屬實例，在該模式下，我們在公有雲基礎設施上為需要更高性能、穩定性及定制化的客戶提供預留算力。根據本地部署解決方案，我們主要提供系統軟件，使客戶能夠在其自身的基礎設施環境中部署及運行我們的平台。

業 務

我們的公有雲服務與本地部署解決方案構成一個整合的商業系統，以滿足客戶多元且不斷演變的需求。公有雲平台使我們能夠服務廣泛的客戶群，支持概念驗證及初始部署等應用場景，並促進於不同應用場景下的驗證。隨著客戶使用量增加，或對性能、穩定性、數據物理位置、網絡邊界及供應鏈保障需要更程度的控制，彼等可選擇專屬實例或本地部署解決方案，從而加深客戶參與度，並於整個客戶生命週期內拓展收入機會。



下表載列於所示期間我們按業務線劃分的收益：

	自8月29日 至12月31日期間		截至12月31日止年度			
	2023年		2024年		2025年	
	人民幣	%	人民幣	%	人民幣	%
(千元，百分比除外)						
公有雲服務.....	—	—	1,069	14.6	29,261	52.9
本地部署解決方案.....	6	100.0	6,277	85.4	26,069	47.1
總計.....	6	100.0	7,346	100.0	55,330	100.0

公有雲服務

我們的公有雲服務旨在通過支持累計合共逾170個模型的單一集成接口，大規模提供標準化、高質量的詞元。該等服務支持廣泛的應用場景，包括編程、AI智能體及多模態應用，為客戶提供統一且靈活的AI能力接入。我們於2024年5月推出無服務器詞元服務，標誌著我們開始大規模商業化。該平台提供兩項互補的服務，即按共享資源模式提供的無服務器詞元服務，以及提供預留算力容量的專屬實例，從尋求具成本效益及按需獲取的個人開發者及初創企業以至需要專屬性能、持續穩定性或供應保障的大型企業的全方位客戶需求。

業 務

無服務器詞元服務

我們在共享資源基礎上提供詞元服務，並根據所消耗詞元計量的使用量收費。在該模式下，多名客戶同時使用同一個底層算力資源共享池。

客戶可透過以下方式使用我們的無服務器詞元服務：(i)直接API接入，即客戶直接接入我們的API以訪問我們平台上的模型；或(ii)自帶密鑰(BYOK)，即客戶在我們的平台上生成API密鑰，並在兼容的第三方工具及應用程序中使用該密鑰，使其能夠保留其現有工具及工作流程，同時選擇我們的平台作為彼等的詞元供應商。截至最後實際可行日期，用戶可透過大部份獨立應用程序及開發工具(包括Claude Code、Hermes Agent及OpenClaw)訪問我們的模型；此外，近100個廣泛使用的應用程序(如LangChain、TRAE、Dify及Cherry Studio)已預先選擇我們的平台作為其內置模型服務提供商，使用戶無需進行額外配置即可使用我們的服務。

我們無服務器詞元服務的主要特徵包括：

- **靈活且無需承諾的訪問權限。**客戶可透過單一統一接口獲取涵蓋廣泛領先模型的生態系統並在其中切換，無需前期基礎設施投資或長期承諾。用量按需擴展，使客戶能夠根據不斷變化的應用需求調整模型選擇。
- **開箱即用及快速部署。**我們的標準化接入流程(賬戶註冊、購買餘額、生成API密鑰及應用程序整合)旨在將部署及全面運行準備所需的時間縮短至分鐘內。
- **適應多樣化應用場景。**我們的平台支持廣泛的用例及整合模式，使客戶能夠將AI能力嵌入至多種工作流程、環境及業務場景，而無需為每次部署重建基礎設施。
- **具成本效益的基於詞元定價。**透過優化的編排及異構算力支持，我們的平台提高了使用率並降低了詞元成本，從而實現了具競爭力的現收現付詞元定價，而無閒置算力資源的開支。
- **端到端模型全生命周期支持。**除推理能力外，我們的平台支持模型選擇、整合及部署，允許客戶在單一環境中管理完整的模型定制及擴展流程 — 減少對各類第三方工具的依賴。
- **按設計實現的數據保護。**我們的平台以企業級數據保護為基準構建，包括在推理期間不保留客戶數據，以及提供靈活的部署選項，讓客戶清楚了解並控制其數據的處理方式。
- **自帶密鑰(BYOK)兼容性。**我們的平台支持BYOK模式，在該模式下，第三方AI工具及應用程序的終端用戶可直接在該等工具中配置我們的API密鑰，以訪問我們平台的模型，而無需修改其現有工作流程或切換應用程序。借助該模式，用戶可直接在其現有使用的工具中，自由選擇並切換我們平台所提供的各類模型。

我們的無服務器詞元服務主要針對尋求靈活及具成本效益地獲取AI能力的客戶，特別是開發者及中小型企業。我們無服務器詞元服務的客戶群高度多元化，涵蓋互聯網公司、模型開發者、各行各業的企業以及個人開發者。

業 務

我們一般(i)就客戶採用預付費模式，在該模式下，客戶提前購買額度，並根據所消耗的詞元扣除用量；或(ii)就若干多次持續性採購的企業客戶採用後付安排，據此我們提供按月結算。我們的系統生成詳細的計費記錄，指明獲取的模型、消耗的詞元數量及相應費用。輸入及輸出詞元分開定價，並可能根據所獲取的模型及輸入上下文的長度而進一步變更定價。定價通常參考模型推理服務的現行市場慣例及模型提供商的定價策略釐定，而考慮到我們的成本結構及營運效率，我們可能按該基準定價給予折扣定價。對於大用量客戶，我們可能提供協商的定價調整。共享資源池架構非常適合基於詞元的計費，因為底層基礎設施成本可分攤至多名客戶，從而實現具競爭力且與使用量成比例的定價。

專屬實例

我們向客戶提供在公有雲基礎設施上的預留算力容量，為他們提供專屬或優先使用的指定資源。

我們專屬實例服務的主要特徵包括：

- *獨佔算力資源且無資源共享。*客戶獲預留專屬算力資源，不與任何其他客戶共享，確保數據及工作負載的完全分離。這可實現可預測的吞吐量及穩定的性能，而不會出現資源爭奪。
- *基於模板的自助服務配置。*客戶可透過精選的最佳實踐模板庫配置算力能力、模型選擇及部署架構，從而實現靈活的自助服務定制，而無需深厚的基礎設施專業知識。我們可提供額外模板以滿足特定客戶要求。
- *企業級SLA及服務可靠性。*預留算力資源可提高可用性、系統可靠性及性能一致性，從而支持任務關鍵型企業應用的要求。
- *從無服務器詞元服務的順利升級路徑。*客戶可從無服務器詞元服務開始進行測試及早期部署，並隨著使用量擴大或性能需求增加而過渡至專屬實例，而不會中斷現有整合。

專屬實例通常提供予對性能、穩定性或供應保障有較高要求，以及對算力資源有大量且穩定需求的客戶，包括運行對延遲敏感、高吞吐量的AI工作負載的企業用戶。截至最後實際可行日期，我們的專屬實例客戶包括領先的科技及互聯網公司、全球製藥企業及頂級AIGC公司。

專屬實例客戶可視乎其偏好，選擇(i)於預留期繳付固定訂閱費，或(ii)根據在專屬實例內所用詞元計算按用量收費。對於潛在的專屬實例客戶，我們與彼等的技術及採購團隊緊密合作，分析彼等的使用模式(包括平均每日詞元量、峰值吞吐量需求及彼等使用情況的可變性)，以協助彼等確定最具成本效益的配置。根據該分析，我們可能推薦固定的專屬實例、具備彈性擴展能力的靈活編排或混合模式。隨著客戶用量增長及彼等對更高性能及服務穩定性的需求增加，彼等可從我們的無服務器詞元服務過渡至專屬實例。

業 務

本地部署解決方案

我們亦提供本地部署解決方案，其中，我們將推理引擎及算力資源編排系統於客戶自有的數據中心或私有算力環境內部署。我們的平台一旦在客戶環境內部署，即在客戶的數據中心內有效創建類似私有雲的AI基礎設施，使員工能透過集中、受控的接口使用多種AI模型，同時確保所有數據及推理活動完全保留在客戶的自有基礎設施內。我們的本地部署解決方案基於標準化、模塊化的平台架構構建，僅需進行有限的定制，以配合客戶的特定部署環境或運營需求。

我們本地部署解決方案的主要特徵包括：

- *客戶自有的基礎設施及完全的數據主權。* 算力資源完全在客戶自有私有基礎設施內提供及營運。客戶對其數據、工作負載及硬件保留完全控制權，使此解決方案特別適合受監管行業及涉及敏感或機密資料的用例。同時，客戶無需自行搭建或維護底層軟件棧，即可受益於我們的詞元供應平台的全部能力，從而兼具本地部署的數據控制與託管服務的便捷性。
- *支持異構運算並兼容國產芯片。* 我們的平台支持廣泛的硬件架構，特別著重於國產芯片生態系統。這使客戶能夠利用其現有或計劃的異構算力環境（包括國產芯片），同時實現高性能推理。
- *供非技術操作人員的簡易控制台。* 我們的平台提供可視化及自助式運行工具，旨在使非技術人員能夠監控、管理及運行模型部署，從而減少對專門基礎設施專業知識的依賴，並在整個組織內加速應用。

我們的本地部署解決方案通常被數據實際存放位置、網絡邊界或供應安全或系統集成具有更高要求的大型企業客戶及機構客戶所採用，並可能視乎部署規模及複雜性以軟件許可證及／或集成軟硬件解決方案的形式交付。截至最後實際可行日期，我們的本地部署客戶包括AI、互聯網、能源、金融及電訊行業的主要參與者。

我們通常透過基於訂閱或永久買斷銷售模式提供該等解決方案，費用通常包括軟件授權費、專業實施費，以及（如適用）根據部署規模、系統配置及客戶需求釐定的維護費。訂閱安排適用於以下客戶：偏好將軟件支出與年度預算周期保持一致者，或鑒於人工智能模型發展步伐加快而希望定期重新審視其平台需求者。交付後，我們提供持續的維護和支援服務。本地部署通常在幾周內完成。

業 務

主要營運指標

下表載列所示年度／期間的主要營運指標。

	截至12月31日／截至該日止年度		截至4月30日／ 截至該日止四個月
	2024年	2025年	2026年
公有雲服務			
－期內最後一個月的平均每日詞元吞吐量(十億) ⁽¹⁾	47.8	163.1	578.5
－註冊用戶人數 ⁽²⁾	127,133	9,197,439	10,282,431
－無服務器詞元服務：			
－活躍用戶數量 ⁽³⁾⁽⁴⁾	57,431	5,452,952	1,453,923
－無服務器詞元服務客戶數量 ⁽³⁾⁽⁵⁾	2,455	716,012	642,866
－專屬實例：			
－專屬實例客戶數量 ⁽³⁾	7	49	20
－平均每名專屬實例客戶收益(人民幣元).....	87,669	305,353	不適用
本地部署解決方案			
－本地部署客戶數量 ⁽³⁾	28	20	5
－平均每名本地部署客戶收益(人民幣元).....	224,184	1,303,448	不適用

附註：

* 由於本公司於2023年8月註冊成立，且於2023年並無產生重大收入，故截至2023年12月31日／截至該日止年度的經營指標並不適用。

(1) 期內最後一個月的平均每日詞元吞吐量，按該月處理的詞元總數除以該月的歷日數計算。就此而言，「詞元」包括用戶提交的輸入詞元及模型使用的輸出詞元，二者均由我們生成。

(2) 指截至期末的累計數字。

(3) 指2024年及2025年以及截至2026年4月30日止四個月的數字。截至2026年4月30日止四個月的數字並無按年化計算，因此較低。

(4) 活躍用戶指於所示年度／期間使用我們無服務器詞元服務的用戶。

(5) 無服務器詞元服務客戶指於所示年度／期間購買我們無服務器詞元服務的用戶。

我們的主要營運指標反映於往績記錄期間及2026年，我們在公有雲服務及本地部署解決方案方面的業務均實現快速擴張。

公有雲服務。我們的平台在用戶規模及詞元吞吐量方面均實現快速增長。於2024年12月至2025年12月，我們的平均每日詞元吞吐量增長241.1%，並於2025年12月至2026年4月進一步增長254.7%。於2024年12月31日至2025年12月31日，我們的累計註冊用戶數量增加約72倍，並將於2026年持續增長，反映人工智能應用的快速普及和使用詞元供應服務的開發者及企業的生態系統不斷擴大，推動我們平台的採用範圍擴大。具體而言：

- 就無服務器詞元服務而言，於2024年至2025年，活躍用戶數量增加約95倍，而付費客戶數量於同期增加約292倍。該增長主要受各行業越來越多地採用人工智能應用及對我們平台能力的日益認可所推動，這兩方面共同吸引大量開發者及企業用戶。截

業 務

至2026年4月30日止四個月，付費客戶數量接近2025年全年水平，突顯用戶變現於2026年的持續強勁勢頭。

- 就專屬實例而言，於2024年至2025年，客戶數量增加600.0%，原因是詞元消耗不斷增長且對性能穩定性及數據隔離要求更高的企業客戶過渡至預留算力資源。於2024年至2025年，平均每名專屬實例客戶收益增長248.3%，反映隨著我們的企業關係日趨成熟，商業參與度加深且合約價值增加。

本地部署解決方案。我們的本地部署解決方案本質上以項目為基礎。於2024年至2025年，本地部署客戶數量減少28.6%，而同期平均每名客戶收益則大幅增加481.4%，反映向具有較高合約價值的大型企業部署轉變。

案例研究

案例研究1：無服務器詞元服務—AI智能玩具公司

客戶背景：該客戶為一家專注於AI驅動兒童玩具的創新型公司，致力於打造智能化、具互動性的陪伴類產品。其玩具產品依賴對兒童語音指令及對話意圖的實時理解與響應。

客戶需求：兒童用戶對響應時延高度敏感，即便僅有數秒延遲，亦可能導致其互動興趣下降。因此，客戶需要能夠穩定提供毫秒級響應的推理能力。兒童語音亦存在特有挑戰，包括發音不清晰、語法表達不規範及情緒驅動型表達等，這對模型的意圖識別能力提出了較高要求。除性能需求外，客戶需要能夠隨需求靈活調整的成本結構：由於學校假期期間使用量顯著上升，而夜間等時段需求相對較低，固定基礎設施投入的成本效益有限。

我們的解決方案：客戶選用我們的無服務器詞元服務，作為其產品的AI基礎模型。我們通過單一API調用讓客戶接入我們平台上所有可用的模型，且新模型發佈後通常可於一周內完成上線，從而使客戶無需進行獨立部署，即可持續獲取前沿模型能力。

交付價值：在採用我們的無服務器詞元服務後，客戶取得顯著成效。產品的平均響應時間從秒級降至毫秒級；在高併發情況下，延遲降至500毫秒以下，直接提升了兒童的互動體驗，並推動需求翻番。按使用量付費的模式避免了閒置資源開銷，為客戶提供了輕鬆擴展的靈活性，使其能夠順利應對學校假期等高峰期，而無需過度配置資源。憑藉對新模型的即時訪問能力，團隊能夠快速測試和驗證模型能力，並將洞察轉化為產品改進，且無需承擔基礎設施運營的負擔，從而釋放工程資源，專注於產品創新與用戶體驗優化。

業 務

案例研究2: 專屬實例- AI編碼平台

客戶背景：該客戶為中國一家AI公司，專注於全棧編碼工作流程，涵蓋代碼設計、編寫、完成及測試，全流程依賴即時AI模型推理。該客戶自其產品開發的早期階段起就一直在使用我們的無服務器詞元服務。由於其AI智能體產品部署擴展，每日詞元消耗達千億級，以維持及時訪問最新模型。

客戶需求：隨著消費規模擴大，客戶現有的無服務器詞元安排變得不太適合其生產需求。在高峰期間，算力資源受到爭用，給生產工作負載帶來可靠性風險。同時，隨著使用量的增長和可預測性的增加，每個詞元的定價結構成本效益降低。該客戶需要一種方法，既能確保穩定、專用的計算能力，又能保證性能，以及更適合其規模的成本結構，同時又不承擔管理基礎設施的間接費用。

我們的解決方案：我們將客戶的核心推理工作負載轉移到我們的專屬實例服務，為其生產運營提供專屬的算力資源。由於專屬實例服務使用與我們的無服務器詞元服務相同的API接口，因此遷移不需要應用級的更改。預留的算力資源不與任何其他客戶共用，從而解決高峰期爭用問題，且服務以企業級SLA承諾為後盾。

交付價值：在過渡之後，客戶在穩定性、性能和成本方面均有所改善。由於消除資源爭用，客戶實現99.95%的SLA可用性，使其每日千億級詞元工作負載得以無中斷運行。通過量身定制的實例配置結合超過90%的緩存命中率，客戶實現超過50%的成本削減。此後，客戶根據不斷增長的業務量持續擴展其預留容量，其工程團隊也能夠專注於產品開發而非基礎設施運營。

案例研究3：本地部署解決方案——國有電力企業

客戶背景：該客戶是中國最大的國有電力企業之一，業務涵蓋發電、輸電及可再生能源。隨著AI在行業內加速應用，該客戶希望建立覆蓋廣泛應用場景的綜合性AI能力，包括智能辦公運營、工業安全管理以及可再生能源運維等。

客戶需求：該客戶缺乏統一的AI平台，導致流程分散、資源利用率低且成本不斷上升。其生產場景對高吞吐量及低時延有較高要求，單靠開源解決方案無法滿足。客戶還需要一套能夠適配其異構算力資源、無需大量定制化適配工作，並能夠快速集成最新前沿模型、不受傳統部署週期限制的解決方案。

我們的解決方案：我們為客戶本地部署了AI模型服務平台，為其提供統一的模型訓練、推理、知識庫管理及智能體編排環境。

交付價值：在部署後，客戶的異構算力資源（包括國產芯片）整合至單一託管平台下，將閒置能力轉化為生產用途，並大幅簡化基礎設施。該平台可接入近100個模型，這些模型不斷更新，且推理性能大幅提升，僅DeepSeek V3就提升了50%以上。在營運方面，該平台實現智能辦公流程、企業知識庫建設及AI智能體開發，減少行政負擔並提高日常效率。值得注意的是，智能風機診斷現在只需三分之一的時間即可完成，準確率由約75%提高至95%以上。

業 務

第十八C章框架下的資格評估

我們正根據上市規則第十八C章尋求[編纂]。我們主要從事詞元供應平台的設計、開發及商業化。我們所有產品及服務均為上市規則第十八C章所界定的指定特專科技產品。我們的董事認為，經弗若斯特沙利文告知，按以下基準，我們的產品及服務屬於上市規則第十八C章所界定的特專科技行業可接納領域。

下表概述我們的產品及服務如何歸屬於人工智能（「AI」）領域，而該領域屬於上市規則第十八C章所界定的特專科技行業項下的可接納領域。

特專科技產品	特專科技行業可接納領域	主要功能分析	商業模式	商業化時間表
公有雲服務.....	人工智能（「AI」）（支持AI的技術及基礎設施）	公有雲服務旨在大規模交付標準化及高質量的詞元，透過單一標準接口為客戶提供統一及靈活的途徑，以接入累計合共超過170種模型，涵蓋編程、智能體及多模態應用等廣泛的AI應用場景。	消耗制及訂閱制相結合	商業化：自2024年5月起，並持續研發
本地部署解決方案	人工智能（「AI」）（支持AI的技術及基礎設施）	本地部署解決方案旨在將我們的推理引擎及算力資源編排系統部署於客戶自有的數據中心或私有算力環境內，使客戶能夠接入廣泛的模型並使用詞元服務。	交易制及訂閱制相結合	商業化：自2023年12月起，並持續研發

有關進一步詳情，請參閱「一 我們的產品及服務」。

業 務

商業化途徑

商業化情況

下表載列我們特專科技產品的主要商業化時間表。

日期	里程碑
2023年12月	我們開始透過軟件訂購產生收益，標誌著我們商業營運的開始。根據弗若斯特沙利文的資料，我們的商業化進程快於行業平均水平。
2024年5月	我們推出無服務器詞元服務，標誌著我們公有雲服務向大規模商業化過渡。
2024年7月	我們推出專屬實例服務，為企業客戶提供為生產級AI工作負載預留的算力資源。
2025年2月	我們率先在國產芯片上部署DeepSeek R1及V3。這項成就在行業內首次成功將國產芯片用於詞元推理業務，展示了國產AI芯片架構在生產級推理工作負載上的商業可行性及實用性，並顯著提升了我們的公眾形象。我們開始大規模商業化部署跨芯片及多模型適配。
2025年4月	我們開始向企業客戶提供我們的全棧本地部署解決方案。
2025年6月	我們正式面向國際客戶推出無服務器詞元服務，使海外開發者及企業均可使用我們的服務。
2025年8月	我們的無服務器詞元服務上線領先的AI模型路由平台。這標誌著我們國際業務的進一步拓展。
2025年12月	我們平台的註冊用戶數超過9百萬及企業客戶超過10,000家。
2026年4月	我們的單日最高詞元吞吐量首次超過萬億，平台註冊用戶數超過10百萬，顯示開發者和企業對我們的詞元服務的採用持續增長。
2026年5月	我們的海外月收入超過1百萬美元。

我們預計將於2026年底前達到上市規則第十八C章所界定的商業公司（「商業公司」）的收入要求。

業 務

商業化假設

我們預計於2026年底前成為一家商業公司受重大不確定因素影響，包括但不限於以下各項：(i)概無重大延誤或障礙影響我們的商業化計劃、產品及服務開發路線圖或未來市值執行；(ii)我們將能夠繼續及時開發、部署及交付我們的產品及服務，並具有客戶期望的品質和性能；(iii)我們的客戶及潛在客戶將繼續接受及採用我們的服務及解決方案，並按大致符合我們預期的速度完成彼等的內部評估、測試、預算審批及採購程序；(iv)我們將能夠保持對供應商及其他基礎設施合作夥伴算力資源的充分訪問權以及足夠的技術、運營、供應鏈、合規及品質控制能力，以支持我們的商業化增長；(v)監管或政策環境不會出現影響我們服務、算力資源獲取、海外業務或商業化活動的重大不利變化；(vi)詞元服務的市場需求將繼續增長，且不會因宏觀經濟、技術或行業特定情況而大幅下降；(vii)我們將能夠繼續吸引及挽留研發、基礎設施運營及商業化職能方面的主要人員；(viii)不會發生「風險因素」所述的任何重大不利事件。因此，雖然我們相信我們有合理的基礎對滿足商業公司收益要求的時間表的預期，倘證明任何上述假設不準確，則我們的商業化可能會延遲且我們的收益增長可能會受到不利影響。除上文所載假設外，我們預期截至最後實際可行日期達成收入目標並無任何重大障礙。

業務可持續性

過往錄得虧損

我們為一家根據上市規則第十八C章界定的未商業化公司。於往績記錄期間，我們成功過渡至初步商業化階段，並實現收入的快速增長。我們的收入由2023年的人民幣6千元大幅增加至2024年的人民幣7.3百萬元，並於2025年進一步增加653.2%至人民幣55.3百萬元，主要受我們公有雲服務及本地部署解決方案的客戶群不斷擴大所驅動。儘管收益增長快速，於往績記錄期間，我們於2023年、2024年及2025年錄得期內／年度虧損分別為人民幣12.2百萬元、人民幣81.9百萬元及人民幣345.5百萬元，同期經調整淨虧損（非國際財務報告準則計量）人民幣12.2百萬元、人民幣54.0百萬元及人民幣187.1百萬元，主要由於以下原因：

早期商業化階段帶來的毛利率壓力。我們是詞元供應行業中的一家新興公司，而我們平台及服務的商業化仍處於早期階段。我們的收入成本主要包括以算力租賃費形式產生的算力資源成本；於早期商業化階段，有關成本的增長速度顯著快於我們的收入，由2023年的人民幣1千元增至2024年的人民幣4.5百萬元，並於2025年進一步增加1,441.6%至人民幣68.6百萬元，其中算力資源成本佔2025年收入總成本的86.9%。因此，我們的整體毛利率由2023年的83.3%降至2024年的39.4%，並於2025年降至(24.0)%，導致我們於2025年錄得毛虧損人民幣13.3百萬元。毛利率收窄，主要是由於我們在持續改進產品及服務的同時租用算力資源而產生大量成本，而該等算力資源的使用率則逐步提升。具體而言，毛利率收窄及我們的收入成本增幅超過收入增幅歸因於以下因素：(i)我們的收入組合發生結構性變化，原因為我們的公有雲服務以較低的毛利率營運，於2025年佔我們總收入的比例大幅增加；(ii)由於為我們迅速擴張的用戶群確保充足容量並維持高服務可用性，需要龐大的算力資源成本（特別是基於公有雲服務的算力租賃費），而該等成本超逾我們於早期商業化階段初步產生的收入，導致錄得負毛利率；及(iii)由於我們吸納來自更廣泛行業的新企業客戶，致使在初始接入階段涉及較高的部署複雜程度，導致本地部署解決方案的毛利率下降。因此，於往績記錄期間，我們的成本水平相對較高。

業 務

於研發方面的大量投資。於往績記錄期間，我們大量投資於研發以支援我們的特專科技產品的開發，研發開支於2023年為人民幣10.8百萬元、2024年為人民幣64.5百萬元及2025年為人民幣209.2百萬元，為我們最大的運營成本中心成本。該等開支佔總收入的百分比由2024年的877.7%大幅下降至2025年的378.1%，顯示出初步的營運槓桿效應。我們的研發開支主要包括：(i)研發算力資源成本，指用於模型適配、底層算子庫優化及平台測試的算力租賃費、(ii)僱員福利開支，指我們研發人員的薪酬、福利及以股份為基礎的薪酬；及(iii)其他開支，主要包括專業服務費、折舊及攤銷以及差旅開支。由於詞元供應平台行業仍然快速發展且技術密集，因此於研發的持續投資對於保持性能領先、提高成本效益、擴大服務能力及支援未來商業化至關重要。值得注意的是，我們的研發能力為我們邁向盈利提供直接支撐：持續的技術突破使我們能夠提升推理能力，而此對從結構上降低單位收益成本及隨時間改善毛利率至關重要。我們預計，隨著我們繼續提升詞元供應能力，我們的絕對研發開支將增加，而隨著我們實現更大的商業規模，我們預計該等開支佔總收入的百分比將逐漸下降。

於銷售及營銷以及一般及行政開支的投資。我們的盈利能力進一步受到銷售及營銷以及一般及行政活動的大量支出影響。我們的銷售及營銷開支由2023年的人民幣110千元增至2024年的人民幣6.4百萬元，並於2025年進一步增加1,210.1%至人民幣83.7百萬元。其中，推廣型算力資源成本（即與大規模分發免費詞元代金券以獲取及引導新用戶至我們公有雲平台有關的算力成本）構成主要開支來源，於2025年為人民幣54.2百萬元，佔銷售及營銷開支總額的64.7%。分發該等代金券不僅有助提升用戶獲取，亦是一項重要的營運策略：透過產生大規模真實世界工作負載，使我們能持續優化平台可用性、調試底層基礎設施並改善整體用戶體驗，從而直接推動註冊用戶基礎快速增長，於2025年底達約9.2百萬。隨著我們平台及市場影響力逐步成熟，我們計劃結構性地降低對廣泛推廣型代金券分發的依賴，並預計銷售及營銷開支佔總收入比例將穩步下降，從而帶來持續的經營槓桿效應。我們的一般及行政開支由2023年的人民幣1.2百萬元增至2024年的人民幣9.1百萬元，並進一步增加至2025年的人民幣21.8百萬元，增幅為140.2%，主要歸因於僱員福利開支，原因為我們擴大行政團隊以支持不斷增長的業務。

實現長期盈利的計劃

我們認為，我們有能力通過以下措施從結構上改善毛利率及經營效率，從而支援我們業務的可持續性：

把握行業爆發性增長

詞元供應市場正處於拐點，需求正以前所未有的速度擴張。受以下因素推動，我們處於有利位置，可於該不斷擴張的市場中佔據日益增長的份額：

- **詞元需求的指數級增長。**根據弗若斯特沙利文的資料，中國的詞元吞吐量從2024年的約142.5萬億個詞元激增至2025年的約2,426.3萬億個詞元，按年增長1,602.6%，並預計於2030年將達到約53.2百億億個詞元，相當於2025年至2030年的複合年增長率約為638.3%。該指數級增長是由企業的AI應用迅速普及所支持，涵蓋編程、智能體、多模態生成應用。隨著AI應用持續成熟及每個用例的詞元消耗不斷加深，我們預期對高性能、具成本效益的詞元供應需求將大幅增長。

業 務

- *把握行業趨勢的領先市場地位。*作為中國最大的獨立生態詞元供應商及全球認可的詞元供應平台，隨著行業規模不斷擴大，我們處於有利位置，可佔據市場增長中極大比例的份額。我們預期強勁的行業趨勢將支持我們長期收益及毛利率改善的走勢。

加深國內客戶參與度

我們尋求透過以下方式，同時擴大我們的客戶群及加深與現有客戶的合作，以增加國內收益：

- *客戶拓展及留存。*我們服務多個細分市場的客戶，尤其注重加深對高價值客戶的滲透，此類客戶通常具有更高的使用強度、更高的單賬戶支出以及更強的長期留存特徵。於往績記錄期間，我們的高價值客戶群在使用規模方面實現大幅增長。月均詞元消耗量超過1,000億的客戶自2025年下半年起開始出現；月均詞元消耗量超過1萬億的客戶已於2026年出現，而該最高消耗量客戶群已佔我們總詞元吞吐量的一半以上，且該比例仍在不斷上升。隨著我們在更高價值客戶群中的滲透率提升且留存率保持強勁，我們預計單客戶平均收入將隨時間顯著改善。
- *向專屬實例安排遷移。*我們企業戰略的一個關鍵要素是將客戶從標準的無服務器詞元服務（我們以具有競爭力的價格提供以推動平台採用）遷移到專屬實例。專屬實例的毛利率高於標準無服務器詞元服務且毛利率更為穩定。隨著企業客戶在我們平台上擴大其AI使用規模，我們預計專屬實例會自然發展，我們認為這將是毛利率增長的重要驅動因素。
- *本地部署。*除公有雲服務外，我們的本地部署業務為我們的增長戰略提供高利潤補充。我們的本地部署解決方案產生82.5%的毛利率，作為可靠的盈利能力和現金產生來源，隨著我們繼續投資以擴大我們的平台，支援我們的整體財務狀況。

加快海外收益增長

我們相信，海外市場為我們的詞元供應服務帶來重大且滲透率相對較低的增長機遇，我們計劃透過以下舉措繼續發展我們的海外業務：

- *透過線上及線下渠道擴大海外客戶群。*我們已透過於領先的AI模型路由平台上架我們的無服務器詞元服務從而擴展海外業務版圖。在此基礎上，我們計劃參與更多線上及線下活動，以提高在主要海外市場的品牌知名度。
- *加強本地化合作夥伴關係及建立本地營運能力。*我們計劃招募具備相關市場經驗的海外銷售、售後及技術人員，以提升客戶服務能力、市場反應能力及本地交付能力。我們亦擬於選定的海外市場逐步設立及營運本地化分支機構，以支援我們的本地團隊並加強我們在該等司法管轄區的市場地位。

業 務

最大限度提高算力資源利用率

我們計劃結合技術驅動的改進及智能調度策略，最大限度地提高我們算力基礎設施的生產力：

- **異構算力適配。**我們已開發出將領先的AI模型適配並部署於國產芯片的能力，該等芯片的租賃成本大幅低於基於NVIDIA GPU的基礎設施。隨著國產AI芯片在我們算力基礎設施中的份額增加，我們預期將大大改善整體成本結構。我們的異構兼容性亦減少我們對任何單一算力資源提供商的依賴，提高供應鏈韌性以及我們管理投入成本波動的能力。
- **推理引擎優化。**我們專有的推理引擎及相關技術架構旨在最大限度地提高運算吞吐量，使每塊芯片能夠同時處理更大量的推理請求。這直接降低我們每詞元的生產成本，並為我們提供定價具競爭力的服務，同時逐步擴大毛利率的能力奠定基礎。
- **非高峰時段調度。**我們正在開發相關能力，採用分層定價模式將非高峰時段閒置的推理算力調配給高校、科研實驗室及企業客戶，用於批量推理工作負載。該方式有助於提升集群全天候利用率，分擔更大規模可計費業務的固定算力資源成本，從而直接壓縮單位成本。

提升營運效率

我們計劃隨著我們擴大業務規模，通過以下舉措優化成本結構並提升營運效率：

- **採購模式轉型。**在早期發展階段，我們優先考慮運營靈活性，採用短期、按需租賃方式獲取算力資源，但該模式的單位成本高於長期合作安排。隨著業務規模擴大，我們預計將逐步增強與算力資源供應商的議價能力，從而推動採購模式由短期按需租賃，逐步轉向單位價格顯著更低的長期年度合約，或轉向由供應商承擔部分硬件成本的聯合運營安排。
- **加強供應鏈管理：**我們持續通過推動算力資源來源多元化及優化採購及配置，增強供應鏈管理能力，從而支持運營穩定性及成本效益。我們支持多種AI芯片架構，包括NVIDIA GPU及國產芯片，並已形成在不同架構下高效利用異構算力資源的能力。這增強了我們保障算力資源供應及提升成本效率的能力。

基於上述情況，董事認為及聯席保薦人同意，我們的業務具有可持續性。

業 務

營運資金充足性

截至2025年12月31日止年度，我們的月度現金消耗率為人民幣14.8百萬元。截至2023年、2024年及2025年12月31日，我們的現金及現金等價物分別為人民幣4.9百萬元、人民幣49.7百萬元及人民幣171.5百萬元，而流動定期存款分別為人民幣零元、人民幣零元及人民幣100.0百萬元。董事認為，考慮到[編纂]的估計[編纂]淨額及我們可獲得的其他財務資源（包括現金及現金等價物、流動定期存款及按公允價值計入損益的金融資產），我們擁有充足的營運資金，可覆蓋自本文件日期起未來12個月內本集團至少125%的成本，包括研發開支、銷售及行銷開支以及一般及行政開支。

我們的核心技術

我們的核心技術基礎由我們自主研發的推理引擎及算力資源編排系統組成，兩者共同構成專為AI推理工作負載打造的一體化技術棧。我們的技術支持於不同模型、算力架構及部署環境下的可擴展且可靠的AI推理，可實現異構算力資源的高效編排，並將底層算力轉化為高吞吐量、低延遲及具成本效益的詞元生產能力。

我們已經建立了一套完整的AI基礎設施技術體系，涵蓋模型API的智能路由及可觀測性、企業級模型的高性能推理加速以及底層調優。這使我們的AI算力基礎設施能夠在大規模生產環境中以高性能、高可靠性及高可擴展性運行。我們的核心技術涵蓋多個關鍵領域，旨在解決模型商業化的算力瓶頸，為客戶提供低延遲、高併發、高性價比的算力基礎設施。

高性能模型推理引擎技術

在底層算力及計算設備層面，我們著重減低對單一硬件生態的依賴，並針對異構算力場景設計了幾項核心特性：

- **異構算力資源的深度適配與優化：**深度相容國際主流高級芯片，以及昇騰、沐曦、摩爾線程、海光及平頭哥等國內主流芯片。通過算子優化、顯存管理優化、通信計算協同優化、調度採樣優化、並行計算算法等，有效提升國產硬件的實際計算效率。這使得其在吞吐量及推理準確性方面具備可與國際先進算力相媲美的能力，從而顯著提升異構AI算力集群的整體效能。
- **多元化並行策略：**在架構層面，我們採用EP及PD分離技術來實現集群級調度。該架構支援超高併發線上請求及超長上下文處理。
- **標準化的部署管線：**我們採用高度標準化、模組化的管線設計，大幅縮短AI模型在複雜異構硬件架構上的部署週期及迭代時間。

業 務

算力資源編排技術

為了解決AI模型工作負載中常見的挑戰，如算力資源空間、碎片和多租戶隔離，我們開發了專有的編排技術：

- **動態資源供應模式**：我們正在推動算力資源由傳統「長期預留」模式向秒級動態供應模式升級。系統可以根據即時請求、負載波動、租戶需求及模型規格自動精細分配計算能力。
- **冷開機優化和安全隔離**：我們的核心技術解決了冷開機適配延遲問題，支援無縫平滑的容量擴展。同時，我們實現了為異構算力環境量身定制的底層多租戶安全隔離機制。在保證資料安全的同時，最大限度地利用全域算力資源。

模型API智能路由與可觀測技術

在系統軟件及網路運行層面，我們構建了支持全流程管理的流量閘道及精細化運營中心：

- **多維動態智能路由**：平台提供統一的模型接入和無縫的協定轉換接口，內置高級智能路由引擎。系統綜合考慮權重、優先順序、會話黏性、即時集群負載、上下文長度、具體業務維度，對海量併發請求進行動態調度，確保任務分配到最優的計算節點。
- **全鏈路可觀測精準計量**：我們在系統內嵌入統一觀測遙測探針，即時採集性能及調用數據。同時，平台具備高精度、多維度的詞元計量能力，為商業結算、成本核算、資源池規劃等提供可靠的數據支援。

研究及開發

我們開發新技術、設計新服務及提升現有服務的能力，對於鞏固我們的市場地位至關重要。與我們特專科技產品有關的所有研發活動均由內部進行。我們對該等研發活動產生的所有知識產權擁有完全所有權。我們並不依賴任何重大的獲授權第三方技術來開發我們的特專科技產品。於往績記錄期間及直至最後實際可行日期，我們概無將任何研發活動外包予第三方，亦無與第三方進行任何對我們研發特專科技產品屬重大的合作。我們已建立跨學科的研發能力，涵蓋計算機科學、應用數學、深度學習算法、分佈式系統、底層系統優化、AI框架工程及AI基礎設施開發等多個不同領域。於往績記錄期間，我們的所有研發開支均就特專科技產品而產生，於2023年、2024年及2025年分別為人民幣10.8百萬元、人民幣64.5百萬元及人民幣209.2百萬元。於最後實際可行日期，我們並不知悉有任何未決或面臨威脅的法律申索、仲裁或行政程序，可能對我們任何主要特專科技產品的研發活動造成重大影響。

業 務

我們的早期交易與研發活動

自2023年8月29日公司成立至2023年12月實現首次營收期間，我們主要致力於對推理引擎及相關技術的實質性研發。這些活動構成了該期間我們交易運營的核心基礎，並非單純的準備性活動。下表列出了該期間的關鍵里程碑和開展的實質性活動，以及它們對本公司特專科技產品開發的直接貢獻：

期間	實質性研發／交易活動	對特專科技產品的貢獻
2023年8月至 2023年9月 ...	本集團於2023年8月29日註冊成立。創始團隊隨即著手開發推理引擎及相關技術，包括項目啟動和架構設計。在2023年8月29日至2023年12月31日期間，本公司發生研發開支人民幣10.8百萬元，一般及行政開支人民幣1.2百萬元。同期，我們的研發支出總額佔同期年度總運營支出的89.1%。	建立了詞元供應平台的核心技術架構和基礎代碼庫
2023年10月 ...	完成了多個開源模型的部署和適配。開發工作側重於模型並行技術以及關鍵推理引擎組件的優化，包括採樣優化、調度優化和內存管理優化，以提高推理效率和資源利用率。	建立了並增強了我們特專科技產品的核心能力
2023年11月 ...	開發了流水線並行處理能力，擴展了對更多AI模型的支持，並進一步提升了推理引擎的性能。我們還開始初步實施FP8量化及運行優化。	增強了特專科技產品的可擴展性和模型兼容性
2023年12月 ...	啟動了跨芯片架構開發，包括開發加速器插件，以支持在異構計算架構上的部署和優化。	為跨芯片兼容性奠定了基礎
2023年12月 ...	我們開始通過軟件訂閱產生收入，標誌著商業運營的正式啟動。	首次創收驗證了我們特專科技產品的商業可行性

業 務

我們的研發過程

我們的研發過程包含六個關鍵階段－需求分析、項目立項、總體設計、迭代開發、質量驗證，以及部署與持續運營監控及反饋－載列如下：

- **需求分析**。每項研發計劃均始於對市場需求、客戶要求及技術可行性的全面分析。我們的團隊評估商業機會，評估其與我們戰略重點的契合度，並確定所涉及的技術挑戰及資源需求。此階段確立項目的核心基礎，並界定指導所有後續開發工作的功能目標。
- **項目立項**。需求一經確認，便通過內部審批流程正式進行研發項目立項。我們的研發團隊編製涵蓋範圍、技術方向、資源分配、關鍵里程碑及成功標準的項目文件。通過協作審查及同行討論，在投入重大資源之前識別潛在風險，並制定明確、商定的開發路線圖。
- **總體設計**。在獲得確認的項目授權後，我們的工程師會為擬議的產品或技術開發高層架構及系統設計。這包括界定流水線結構、API接口、調度邏輯及硬件兼容性。總體設計作為迭代開發的技術藍圖，確保相關研發子團隊之間的協調一致。
- **迭代開發**。開發工作通過一系列專注的迭代週期進行。我們的工程師實施及優化算法，在異構算力環境中進行性能基準測試，並逐步將各組件集成至功能系統中。快速原型設計及數據驅動分析是此階段的核心，使團隊能夠以敏捷的方式優化吞吐量、降低延遲，並降低單位詞元生成成本。
- **質量驗證**。在部署之前，每個產品或技術版本均須經過系統的質量驗證。這包括功能測試、針對目標SLA指標的性能基準測試、準確性驗證、安全性評估，以及跨支持的硬件架構和模型配置的兼容性測試。只有在圓滿完成所有驗證要求後，方可批准版本部署。
- **部署、運營監控及反饋**。成功通過驗證後，產品或技術即部署至生產。我們的運營團隊持續監控服務可用性、推理延遲、吞吐量及系統穩定性。我們收集及分析實時遙測數據和客戶反饋，並將運營表現的洞察反饋至下一個開發週期，以推動迭代優化。此閉環機制確保我們的平台功能與客戶及更廣泛的AI生態系統需求保持緊密契合並不斷迭代。

業 務

我們的研發團隊及核心成員

截至最後實際可行日期，我們的內部研發部門由67人組成，其中絕大部分具備學士學位或以上學歷。我們的研發部門劃分為6個核心單位，包括基礎結構領域、公有雲領域、技術運營領域、設計領域、創新產品領域及私有雲領域。我們的研發團隊由六名核心成員領導，有關詳情載列於下表：

核心研發成員	履歷
袁進輝博士.....	袁進輝博士於2003年取得西安電子科技大學計算機科學學士學位，並於2008年取得清華大學博士學位，其博士論文在中國科學院院士及中國人工智能研究領域奠基人張鉞教授指導下完成，並獲得優秀博士論文獎。自此，袁博士在將前沿研究成果轉化為可投入生產的平台及獲廣泛採用的創新技術方面建立了良好往績： - 工業AI平台工程及創新（2013年至2016年）：袁博士於2013年加入微軟（中國）有限公司，並於2014年發明LightLDA（一種高效主題模型訓練算法）。 - 分佈式深度學習平台（2017年至2023年）：袁博士於2017年創立分佈式深度學習平台項目OneFlow。於OneFlow任職期間，其主導OneFlow深度學習框架的開發，策略性聚焦提升算法研發效率及硬件利用率，推動研究創新與可擴展算力表現之間的銜接。 - AI基礎設施創業（2023年至今）：自2023年創立硅基流動以來，袁博士一直領導本公司的AI基礎設施相關工作。袁博士亦積極與生態系統合作夥伴及產業持份者保持交流，包括參與閉門技術交流，以推動協作式AI基礎設施生態系統的發展。
柳俊丞先生.....	柳俊丞先生為本公司核心技術棧的主要開發貢獻者之一，具備紮實的AI系統軟件研發能力。柳先生持有哈爾濱工業大學計算機科學與技術學士學位。
陳後江先生.....	陳後江先生於深度學習框架開發方面擁有逾10年經驗，此前曾任職於百度時代网络技术（北京）有限公司、淘寶（中國）軟件有限公司及OneFlow。陳先生持有中國地質大學（武漢）通信工程學士學位。
廖星宇先生.....	廖星宇先生專注於深度學習算法、分佈式訓練及AI基礎設施感知領域，此前曾任職於京東AI平台與研究部、OneFlow及北京輝義智慧科技有限公司。廖先生持有中國科學技術大學數學碩士學位。

業 務

核心研發成員

履歷

李一鵬博士.....	李一鵬博士專注於應用數學及分佈式深度學習系統領域。李博士目前在中國及美國擁有13項已獲授權專利，另有17項專利處於籌備申請或申請階段。李博士持有中國科學技術大學信息與計算科學學士學位，以及美國石溪大學應用數學博士學位。
張文驍先生.....	張文驍先生擁有逾十年全棧開發經驗，涵蓋系統開發、AI框架及深度學習基礎設施等領域，並為OneFlow深度學習框架創立初期的核心貢獻者之一。張先生持有四川大學軟件工程學士學位及華中科技大學軟件工程碩士學位。

上述核心研發成員為負責技術營運以及我們的特專科技產品研發的主要人員，原因為彼等對我們的特專科技產品的開發、優化及持續提升承擔主要技術領導責任。

為吸引及挽留優秀管理及技術人才，我們提供具競爭力的薪酬待遇及全面的激勵計劃。我們亦致力於員工的長期發展，提供具結構性的職業發展路徑，以深化技術專長及促進跨學科協作，並讓員工可透過參與複雜項目的實際經驗拓展技能。

為保留機構知識及支持業務持續營運，我們於研發、產品及項目管理職能中維持健全的知識保留及文件管理制度。該等持續更新的資料庫有助於團隊內部有效傳承關鍵技術知識，惟仍須遵守我們的信息安全及保密規程。我們亦以專業及有序的方式管理員工離職過渡。透過結構化交接程序，以及既定的文件管理、訪問權限控制及項目備份機制，我們致力確保關鍵技術運營的延續性，並盡量減少對團隊協作造成的影響。

我們與重要研發人員所訂立協議的主要條款載列如下：

- **知識產權歸屬：**僱員於受僱期間因履行職務，或主要利用我們的物料、技術條件及業務信息等所完成的任何發明、創作、作品或非專利技術成果（包括商業秘密），均屬於我們所有。我們有權使用或轉讓上述知識產權。僱員須積極提供一切必要的信息、材料及研究成果，並採取一切必要行動，以協助我們取得及行使相關知識產權。
- **保密：**僱員承諾不披露或使用任何保密信息。
- **競業限制及不招攬：**我們保留權利，可於僱傭關係終止後單方面施加最長兩年的競業限制期，並據此提供適當補償。於受僱期間及我們施加的競業限制期內，僱員不得從事任何競爭行為，亦不得招攬或試圖招攬我們的僱員離職。
- **爭議解決：**訂約方應透過協商解決所產生的任何問題。倘未能解決，訂約方可向我們所在地的勞動仲裁委員會申請仲裁。

業 務

知識產權

知識產權對我們的業務至關重要。我們未來的商業成功部分取決於我們是否有能力就與業務相關且具有重大商業價值的技術、發明及專有技術取得及維持專利、其他知識產權及專有保護，維護及執行我們的專利、保障商業秘密的保密性，以及在不侵犯、挪用或以其他方式違反第三方知識產權的情況下經營業務。

截至最後實際可行日期，我們於中國擁有11項已註冊專利及34項專利申請、17項軟件著作權及6項已註冊商標。截至同日，我們亦於香港擁有一項已註冊商標、於新加坡擁有兩項已註冊商標、於日本擁有一項已註冊商標及於美國擁有一項已註冊商標。截至最後實際可行日期，我們擁有全部專利及專利申請，且我們的專利及專利申請並無與第三方存在任何共同擁有或共享安排。

有關各項特專科技產品的重大知識產權，請參閱「附錄四－法定及一般資料－B. 有關我們業務的進一步資料－2. 知識產權」，當中載有與我們特專科技產品相關核心技術重大專利的詳情。於往績記錄期間及直至最後實際可行日期，我們並無牽涉任何重大知識產權侵權相關的法律、仲裁或行政程序或申索，而我們為其中的申索人或被申索人。董事確認，截至最後實際可行日期，彼等並不知悉我們涉及任何重大侵犯第三方知識產權的法律、仲裁或行政程序。

銷售及營銷

銷售及營銷團隊

截至2025年12月31日，我們已建立一支由21人組成的銷售及市場推廣團隊，業務覆蓋中國及海外市場。銷售團隊按照業務分部及客戶層級進行組織，以提升客戶觸達能力及響應效率，並為公有雲客戶及本地企業客戶提供專責覆蓋。我們的業務發展及銷售團隊與研發團隊緊密合作，以確保技術能力能夠有效傳達予客戶，並轉化為商業合作機會。

我們主要透過直銷及渠道合作夥伴銷售我們的產品及服務。鑒於我們業務的性質——即透過雲端或軟件授權方式提供詞元服務——我們的商業模式並不涉及實體產品分銷或傳統分銷商渠道。我們亦已透過領先的AI模型路由平台擴展我們的海外業務版圖，使我們能夠向海外開發者及企業提供我們的服務。

售後服務

我們認為，售後服務是客戶體驗及平台可靠性管理的重要組成部分。我們已設立專門的售後及運營團隊，負責監控平台穩定性、回應客戶服務請求，並管理公有雲及本地部署中的事故處理工作。

對於公有雲客戶，我們持續對平台可用性、延遲及吞吐表現進行實時監控。當發現異常情況時，我們的運營團隊會即時收到警報，並致力於識別根本原因及恢復正常服務水平。對於本地部署客戶，我們透過結構化的服務流程管理售後互動。客戶在部署後遇到問題時，可透過指定渠道聯絡我們的服務團隊，而每宗個案均會透過我們的內部服務管理系統進行跟蹤。我們承諾於一至四小時內回應事故，並於12至48小時內解決，視乎事故的嚴重性而定。我們設有專責的交付及支援工程師團隊，按需要提供遠程及現場支援。

於往績記錄期間，我們並無發生任何客戶合約終止或重大客戶投訴。

業 務

營銷

我們的營銷策略旨在持續加強品牌知名度、擴大於開發者及企業客戶中的覆蓋範圍、推動產品採用及客戶轉化，並進一步鞏固我們於詞元供應行業的領先地位。以我們的獨立生態系統策略為基礎，我們已建立一個涵蓋開發者、企業客戶及行業合作夥伴的多層次市場發展框架，並透過品牌建設、生態系統運營、產品推廣及商業發展舉措，推動業務持續增長。

在開發者生態系統層面，我們透過開發者活動、內容營銷及媒體推廣，持續擴大用戶覆蓋範圍，並提升產品知名度及開發者參與度。藉由提供廣泛的模型選擇、高性價比的推理服務以及流暢便捷的開發體驗，我們吸引開發者及創新團隊採用我們的產品及服務，帶動用戶基礎及詞元量的增長。

在企業客戶層面，我們透過行業活動、技術峰會、客戶互動活動及解決方案推廣，展示我們的技術能力、產品優勢及行業應用案例，從而促進企業客戶對我們產品及服務的了解與採用。對於具有商業潛力的客戶，我們的營銷、業務發展及銷售團隊會緊密合作，開展需求評估並匹配合適的解決方案。

我們亦積極與模型供應商、芯片設計商、雲服務供應商及其他行業合作夥伴開展生態系統協作。透過聯合推廣、技術合作及共同推進生態系統項目開發，我們致力於拓展市場機遇、提升我們的生態系統影響力，並增強我們平台的整體價值。

客戶

於2023年、2024年及2025年，我們於往績記錄期間各年度來自五大客戶的總收入分別為人民幣6千元、人民幣6.2百萬元及人民幣24.9百萬元，分別佔我們有關年度總收入的100.0%、85.0%及45.0%。於同年，我們於往績記錄期間各年度來自最大客戶的收入分別為人民幣5千元、人民幣4.5百萬元及人民幣7.5百萬元，分別佔我們有關年度總收入的83.3%、61.1%及13.6%。

下表載列我們於往績記錄期間五大客戶的詳情：

客戶	產品／服務	開始建立 業務關係 年份	信貸期	付款方式	收入 (人民幣千元)	佔總收入 百分比
截至2023年12月31日止年度						
客戶A ⁽¹⁾	本地部署解決方案	2023年	下單付款	銀行轉賬	5	83.3%
客戶B ⁽²⁾	本地部署解決方案	2023年	下單付款	銀行轉賬	1	16.7%
總計					6	100.0%

業 務

客戶	產品／服務	開始建立 業務關係 年份	信貸期	付款方式	收入 (人民幣千元)	佔總收入 百分比
截至2024年12月31日止年度						
客戶C ⁽³⁾	本地部署解決方案	2024年	5或10個工作日	銀行轉賬	4,487	61.1%
客戶D ⁽⁴⁾	本地部署解決方案	2024年	預付款	銀行轉賬	1,040	14.2%
客戶E ⁽⁵⁾	專屬實例	2024年	預付款	銀行轉賬	377	5.1%
客戶F ⁽⁶⁾	本地部署解決方案	2024年	預付款	銀行轉賬	175	2.4%
客戶G ⁽⁷⁾	本地部署解決方案	2024年	5個工作日	銀行轉賬	164	2.2%
總計					6,243	85.0%

客戶	產品／服務	開始建立 業務關係 年份	信貸期	付款方式	收入 (人民幣千元)	佔總收入 百分比
截至2025年12月31日止年度						
客戶H ⁽⁸⁾	本地部署解決方案	2025年	5個工作日	銀行轉賬	7,547	13.6%
客戶I ⁽⁹⁾	本地部署解決方案	2025年	30日	銀行轉賬	7,100	12.8%
客戶J ⁽¹⁰⁾	專屬實例	2025年	30日或交付前付款	銀行轉賬	4,138	7.5%
客戶K ⁽¹¹⁾	本地部署解決方案	2024年	14個工作日	銀行轉賬	3,189	5.8%
客戶L ⁽¹²⁾	本地部署解決方案	2025年	15個工作日	銀行轉賬	2,903	5.2%
總計					24,877	45.0%

附註：

- (1) 客戶A為一名個人，其電郵地址與一家私營生成式AI軟件公司相關聯。
- (2) 客戶B為一名個人，其電郵地址與一家教育平台相關聯。
- (3) 客戶C為一家總部位於中國的公司，主要從事信息系統整合服務、人工智能基礎資源與技術平台、人工智能行業應用系統整合服務、集成電路設計及相關技術服務。
- (4) 客戶D為一家總部位於中國的公司，主要從事技術開發、軟件開發、計算機系統服務、信息系統整合及相關技術服務。
- (5) 客戶E為一家總部位於中國的公司，主要從事信息系統整合及AI行業應用系統整合。
- (6) 客戶F為一家總部位於中國的公司，主要從事軟件及信息技術服務。
- (7) 客戶G為一家總部位於中國的公司，主要從事信息系統整合服務及技術推廣。

業 務

- (8) 客戶H為一家總部位於中國的公司，主要從事用於AI訓練及推理、通用計算和圖形渲染的全棧芯片產品及相關軟件棧和計算平台的研發、設計及銷售。
- (9) 客戶I為一家總部位於中國的公司及其聯屬人士，主要從事信息與通信技術基礎設施、智能設備、雲計算及相關ICT解決方案、產品及服務。
- (10) 客戶J為一家總部位於中國的公司，主要從事技術服務、軟件開發、計算機系統服務及相關計算機軟硬件業務。
- (11) 客戶K為一家位於中國的研究機構，主要從事原創科技創新研究。
- (12) 客戶L為一家總部位於中國的公司，主要從事信息系統整合、運營及維護(O&M)、軟件開發及數據處理。

據我們所深知，於往績記錄期間各年度，我們的五大客戶均為獨立第三方。截至本文件日期，概無董事、其聯繫人或據董事所知擁有我們已發行股本5%以上的任何股東於往績記錄期間各年於我們的任何五大客戶中擁有任何權益。

無服務器詞元服務合約的主要條款

我們與客戶訂立的無服務器詞元服務合約的主要條款包括：

- **服務。**我們一般向客戶提供無服務器詞元服務，並提供技術支援，包括應用程序部署及程序調試。
- **付款方式。**客戶一般透過以下方式購買我們的無服務器服務：(i)發送電子郵件並向我們轉賬支付費用，或(ii)在平台上為其賬戶充值。在某些情況下，我們可能允許客戶在付款前使用服務。
- **服務限制。**客戶享有專屬及不可轉讓的權利，將服務用於個人用途或他們代表之實體的用途，惟須受(其中包括)以下禁止事項所限：(i)逆向工程、解碼或反編譯；(ii)未經我們事先書面同意轉讓或分授API密鑰；(iii)侵犯我們的知識產權；(iv)使用我們的服務處理、傳輸或儲存非法內容或資料；及(v)破壞或損害我們的服務器或基礎設施。倘我們釐定來自該服務與客戶的互動數據違反法律、法規或本合約，我們保留依法處置該等數據的權利。
- **知識產權。**客戶理解並確認我們繼續享有該服務項下的所有知識產權，且除非合約另行許可，否則禁止將該等知識產權用於任何用途。客戶可使用所生成的結果，前提是其使用不會轉讓或侵犯我們或第三方的任何知識產權。
- **保密。**客戶不得披露我們的任何機密資料或透過該服務可獲取的其他用戶的任何機密資料。
- **終止。**通常情況下任何一方均有權按相關合約訂明的通知期提前發出通知，以終止合約。

業 務

專屬實例合約的主要條款

我們與客戶訂立的專屬實例合約的主要條款包括：

- **服務。**我們一般向客戶提供專屬實例，並提供技術支援，包括應用程序部署及程序調試。就若干客戶而言，我們亦提供SLA，以確保延遲、吞吐量及穩定性達到最低標準。
- **付款方式。**客戶可支付(i)預留期的固定訂閱費，或(ii)根據在專屬實例內消耗的詞元計量按使用量收費。
- **服務限制。**客戶享有專屬及不可轉讓的權利，將服務用於個人用途或他們代表之實體的用途，惟須受(其中包括)以下禁止事項所限：(i)逆向工程、解碼或反編譯；(ii)未經我們事先書面同意轉讓或分授API密鑰；(iii)侵犯我們的知識產權；(iv)使用我們的服務處理、傳輸或儲存非法內容或資料；及(v)破壞或損害我們的服務器或基礎設施。倘我們釐定來自該服務與客戶的互動數據違反法律、法規或本合約，我們保留依法處置該等數據的權利。
- **知識產權。**客戶理解並確認我們繼續享有該服務項下的所有知識產權，且除非合約另行許可，否則禁止將該等知識產權用於任何用途。客戶可使用所生成的結果，前提是其使用不會轉讓或侵犯我們或第三方的任何知識產權。
- **保密。**客戶不得披露我們的任何機密資料或透過該服務可獲取的其他用戶的任何機密資料。
- **續期。**合約一般將自動續期，除非任何一方在屆滿日期前按相關合約訂明的通知期發出通知以終止合約。

本地部署解決方案合約的主要條款

我們與客戶訂立的本地部署解決方案合約的主要條款包括：

- **服務。**我們通常在客戶本身的數據中心或私有算力環境中部署我們的推理引擎及算力資源編排系統，並在部署後提供持續維護及支持。
- **付款。**費用一般包含軟件許可費(按訂閱或一次性基準提供)、專業實施費及(如適用)維護費。
- **產品及服務限制。**客戶在使用產品及服務時應遵守適用法律。如客戶違反合約的相關條款，我們有權取消授權而毋須承擔任何責任。
- **知識產權。**我們擁有有關軟件的知識產權。禁止客戶進行逆向工程、解碼或反編譯，且除非法律允許或我們另行准許，否則禁止將該等知識產權用於任何用途。

業 務

- **保密。**雙方同意嚴格按照合約使用對方的機密資料，且除非事先獲得另一方的書面同意，否則禁止向第三方披露該等資料。
- **終止。**雙方可互相協定終止合約。合約一般將自動續期，除非任何一方在屆滿日期前按相關合約訂明的通知期發出通知以終止合約。我們亦有權按相關合約訂明的通知期提前發出書面通知，單方面終止合約。

供應商

於2023年、2024年及2025年，我們於往績記錄期間各年向五大供應商的採購總額分別為人民幣2.2百萬元、人民幣11.4百萬元及人民幣117.3百萬元，分別佔我們有關年度採購總額的100.0%、87.6%及70.8%。於同年，我們於往績記錄期間各年向單一最大供應商的採購額分別為人民幣2.2百萬元、人民幣6.8百萬元及人民幣33.8百萬元，分別佔我們有關年度採購總額的100.0%、52.1%及20.4%。

下表載列我們於往績記錄期間五大供應商的詳情：

供應商	所提供 產品／服務	開始業務 關係的年份	信貸期	付款方式	採購額 (人民幣千元)	佔總採購額 百分比
截至2023年12月31日止年度						
供應商A ⁽¹⁾	算力資源	2023年	不適用	銀行轉賬	2,169	100.0%
總計					2,169	100.0%
供應商	所提供 產品／服務	開始業務 關係的年份	信貸期	付款方式	採購額 (人民幣千元)	佔總採購額 百分比
截至2024年12月31日止年度						
供應商A ⁽¹⁾	算力資源	2023年	10個工作日	銀行轉賬	6,774	52.1%
供應商B ⁽²⁾	算力資源	2024年	預付款及10日	銀行轉賬	2,206	17.0%
供應商C ⁽³⁾	算力資源	2024年	預付款	銀行轉賬	1,091	8.4%
供應商D ⁽⁴⁾	算力資源	2024年	預付款及10日	銀行轉賬	700	5.4%
供應商E ⁽⁵⁾	雲服務	2024年	預付款	銀行轉賬	604	4.6%
總計					11,375	87.6%

業 務

供應商	所提供 產品／服務	開始業務 關係的年份	信貸期	付款方式	採購額 (人民幣千元)	佔總採購額 百分比
截至2025年12月31日止年度						
供應商F ⁽⁶⁾	算力資源	2025年	預付款	銀行轉賬	33,830	20.4%
供應商G ⁽⁷⁾	算力資源	2025年	預付款	銀行轉賬	25,328	15.3%
供應商H ⁽⁸⁾	算力資源	2025年	每月結算	銀行轉賬	24,491	14.8%
供應商I ⁽⁹⁾	算力資源	2025年	預付款	銀行轉賬	21,005	12.7%
供應商J ⁽¹⁰⁾	算力資源	2025年	預付款	銀行轉賬	12,616	7.6%
總計					117,270	70.8%

附註：

- (1) 供應商A為一家總部位於中國的公司，專注於深度學習框架、AI操作系統、軟件開發及相關技術服務。
- (2) 供應商B為一家總部位於中國的公司，專注於算力服務、雲服務、互聯網數據服務、軟件開發、信息系統集成及相關技術服務。
- (3) 供應商C為一家總部位於中國的公司，為跨行業客戶提供雲計算、IDC託管、IT增值、IT外包及通信集成服務。
- (4) 供應商D為一家總部位於中國的公司，專注於網絡技術及相關雲/IT服務。
- (5) 供應商E為一家總部位於中國的公司，專注於雲計算服務。
- (6) 供應商F為一家總部位於中國的公司，專注於全棧雲服務，包括公有雲、私有雲、專屬雲、混合雲及邊緣雲服務。
- (7) 供應商G為一家總部位於中國的公司，專注於技術服務、軟件開發、計算機系統服務、人工智能應用軟件開發、人工智能基礎軟件開發及信息系統集成。
- (8) 供應商集團H指供應商H及其聯屬公司。供應商H為一家總部位於中國的電信公司，提供電信服務。
- (9) 供應商I為一家總部位於中國的公司，專注於雲計算及相關技術服務。
- (10) 供應商J為一家總部位於中國的公司及其聯屬人士，專注於信息與通信技術基礎設施、智能設備、雲計算及相關ICT解決方案、產品及服務。

據我們所知，除供應商A及供應商I外，於往績記錄期間各年度我們的所有五大供應商均為獨立第三方。截至最後實際可行日期，除供應商A及供應商I外，概無董事、其聯繫人或據董事所知擁有我們已發行股本5%以上的任何股東於往績記錄期間各年度於我們的任何五大供應商中擁有任何權益。

我們與供應商的協議主要條款包括：

- **服務。**我們的供應商通常提供算力資源及相關技術支持服務，其服務範圍、配置、期限及費用載於相關訂單、服務清單或平台記錄中。

業 務

- **付款。**我們通常通過預付款項或月度付款安排結算供應商費用。費用通常根據固定資源包、實際使用量、帶寬使用量或供應商平台記錄計算。
- **服務支持。**供應商通常提供運維服務、技術支持及服務可用性承諾。
- **合規及數據保護。**我們通常需要遵守適用法律、監管要求、平台規則及網絡安全政策以使用供應商的服務。供應商通常需要對我們的業務數據及機密資料保密。
- **期限。**供應商協議通常具有固定服務期限或框架條款。提前終止可能需要事先書面通知。

重疊客戶及供應商

於往績記錄期間，據董事所深知及確信，我們有兩名主要業務夥伴同時為我們的客戶及供應商。

客戶I(亦為供應商J)為我們於2025年的第二大客戶，我們於2025年的第五大供應商，並與客戶I同屬一集團。我們主要向客戶I提供AI模型芯片適配服務，而供應商J主要透過集團內不同實體向我們提供算力資源。於2023年、2024年及2025年，我們向客戶I的銷售額分別為零、零及人民幣7.1百萬元，分別佔我們總收入的零、零及12.8%。同期，我們向供應商J的採購額分別為零、零及人民幣12.6百萬元，分別佔我們總採購額的零、零及7.6%。

供應商I為我們於2025年的第四大供應商，亦為我們的客戶。我們主要就各項事件於2025年向供應商I提供無服務器詞元服務，而供應商I主要向我們提供算力資源。於2023年、2024年及2025年，我們向供應商I的銷售額分別為零、零及人民幣8元，分別佔我們總收入的零、零及少於0.1%。同期，我們向供應商I的採購額分別為零、人民幣419千元及人民幣21.0百萬元，分別佔我們總採購額的零、3.2%及12.7%。

根據弗若斯特沙利文的資料，在詞元供應行業中，客戶與供應商重疊屬常見情況。董事確認為，與上述重疊客戶及供應商的所有交易均按個別基準進行，相關銷售與採購之間並無關聯，亦不互為條件。所有該等交易均於日常業務過程中按正常商業條款及公平交易原則訂立。

數據安全及隱私

我們平台設計的一項核心原則是我們不會存儲或使用客戶交互數據 — 包括提交至我們平台上可用模型的提示以及由該等模型生成的輸出。

在我們的日常業務過程中，我們收集並維護用戶的有限個人資料，其中包括：(i)創建賬戶時提供的註冊資料，如姓名及聯繫方式；及(ii)因應實際業務營運要求收集的身份驗證及付款資料。該等個人資料與交互數據分開存儲，並按照中國適用的數據隱私及安全法律及法規處理，而就我們的國際平台而言，則按照我們經營所在司法管轄區的法律及法規處理。儘管我們同時運營國內及國際平台，但來自各平台的客戶個人數據乃分開存儲。交互數據根據算力資源的可用性進行隨機調度，並在完成後立即刪除。除適用法律要求或經用戶事先同意外，我們不會向第三方提供用戶的個人數據。

業 務

對於使用我們的本地部署解決方案的客戶，所有通信及推理請求均在客戶自己的基礎設施內存儲及處理，而不會暴露於我們的共享雲環境。

我們已實施一系列技術及組織措施來保護在我們平台上處理的數據安全，包括：

- 訪問控制。我們在內部系統中維護個人資料的訪問控制系統，僅限制授權人員訪問。僱員對數據的訪問根據角色及職能受到嚴格限制。
- 網絡及系統安全。我們維護防火牆及網絡安全協議，以防止因網絡攻擊導致的資料丟失或洩漏，且我們定期進行數據安全風險評估。
- 計算隔離。我們的雲基礎設施實施跨租戶的計算隔離及網絡隔離，以防止在多租戶環境中未經授權訪問或洩露客戶數據。我們持續評估及加強我們的隔離架構，包括探索增強容器級別的安全性。
- 僱員保密義務。所有僱員在加入本公司時均須簽署保密及知識產權轉讓協議，涵蓋保密、競業禁止義務以及僱傭期間所創造知識產權的處理。參與重要客戶項目的僱員另須簽署項目特定的保密協議。
- 數據合規管治。我們已建立數據合規系統，涵蓋各部門的數據處理方法、信息安全事件應急響應及應變計劃，並建立內部工作小組，對數據合規事宜進行定期監控。

於往績記錄期間及直至最後實際可行日期，我們並無遇到信息技術及軟件系統的任何重大故障、意外系統崩潰、中斷或安全漏洞，亦無經歷任何重大數據洩露、數據丟失或未經授權使用客戶個人資料。此外，於同期，我們概無因數據保護或網絡安全而遭受任何訴訟、仲裁或重大行政處罰而導致重大不利影響。我們相信，由於我們嚴格的內部控制、合規管治以及我們不會存儲或使用客戶交互數據的基本設計原則，我們的業務營運在所有重大方面均符合我們經營所在所有市場現行的數據安全法律及法規。

環境、社會及管治事宜

環境、社會及管治(ESG)管治

我們嚴格遵守有關人工智能、生態環境保護、勞工及就業以及商業道德的國家及地方性法律、法規及監管文件，並穩步提升我們的ESG管理表現。董事會為我們ESG管治的最高監督及決策機構。其主要負責審查重大ESG相關事宜、監督ESG實踐以及監控我們ESG目標的進展。在我們的運營及重大決策中，我們評估ESG風險與機遇，並通過自上而下的方法推動可持續發展理念

業 務

及負責任商業實踐融入整個價值鏈，同時恪守最高標準的商業道德。董事會通過定期會議及圍繞專題事項的溝通，持續監督ESG相關工作。此外，我們每年至少開展一次ESG培訓，以提升董事會對ESG監管框架及行業最佳實踐的認識，並持續增強董事會在可持續治理方面的專業能力。

為更好地推進我們的ESG倡議，我們已成立由各部門管理人員組成的ESG工作小組。其主要職責包括：(i)識別ESG風險及機遇並制定風險緩解措施；(ii)制定ESG相關政策並跟蹤其落實情況；(iii)評估氣候變化的影響並制定應對策略；(iv)協調ESG數據管理及日常ESG行政工作流程；及(v)為所有部門的僱員組織ESG相關培訓。

我們目前正參照香港聯交所上市規則（《上市規則》）的相關要求制定我們的ESG政策。**[編纂]**後，我們將建立常態化的ESG信息披露制度，並定期發佈ESG報告，以切實保障我們利益相關者的權益。

ESG風險管理

董事作出一切合理查詢後確認，截至最後實際可行日期，我們在所有經營司法轄區內均全面遵守所有適用的ESG法律法規，且不存在對我們的運營、財務、表現或聲譽造成重大影響的重大不合規情況，亦未受到任何重大監管處罰。

為持續提升風險管控的有效性，ESG工作小組負責具體的ESG風險評估、根本原因分析及制定相應的應急響應機制，相關成果至少每年向董事會提交一次。展望未來，我們將不斷完善ESG風險控制體系，為企業的可持續發展築牢堅實基礎。

重要性評估

我們根據香港聯交所的ESG監管指導文件及可持續發展會計準則委員會(SASB)標準，同時結合我們的實際情況，構建我們的ESG重大議題庫。我們向我們的董事、高級管理層、供應商及其他相關方等利益相關方發放問卷，並結合專家判斷及董事會集體審議，確定ESG議題的重要性排序。根據評估結果，我們認為產品責任、研發與創新、知識產權保護、供應鏈管理、發展與培訓、勞工準則、反腐敗及氣候變化等議題是我們需要重點關注的關鍵ESG議題，並在實施ESG管理、控制及實踐時特別強調上述議題。

環境

基於我們的業務性質，我們不從事製造活動，我們的環境風險相對較低。我們已制定內部環境保護政策並倡導綠色低碳理念。我們推行綠色辦公，規範電氣設備的使用及管理，並鼓勵僱員採用綠色出行方式。截至最後實際可行日期，我們已獲得ISO 14001:2015環境管理體系認證。於往績記錄期間，我們概無遭受任何與環境相關的投訴或處罰，且我們於往績記錄期間在環境合規管理方面的累計投入達人民幣40千元。

業 務

環境管理目標

為配合中國國家雙碳戰略，我們以2025年為基準年，承諾到2028年將能源消耗強度降低5%、到2030年降低10%，並承諾到2028年將溫室氣體排放強度降低5%、到2030年降低10%。為確保該等目標的落實，我們將通過提高能效、清潔能源替代及提升僱員環保意識等措施，分階段推進目標的實現。

廢棄物管理

我們的日常業務營運不涉及有害廢棄物的排放，僅產生無害廢棄物，主要包括僱員日常辦公活動產生的生活辦公廢棄物。無害廢棄物由我們辦公室所在大廈的物業管理部門集中收集及處置。同時，我們鼓勵無紙化辦公等方式，以減少廢棄物的產生。

能源管理及水資源利用

我們嚴格遵守有關能源管理的法律法規，持續推動資源有效利用。我們日常營運中消耗的能源包括外購電力及辦公用水。鑒於我們的業務性質，我們幾乎所有算力均通過市場化租賃獲取。隨著我們的業務持續擴展，自2024年以來，員工人數及辦公場所規模均不斷擴大，直接導致能源消耗規模相應增加。為進一步踐行綠色經營理念，我們在全體僱員中常態化實施節能管控措施，引導僱員妥善落實水電節約要求，從而持續降低營運端的資源消耗。於往績記錄期間，我們的能源消耗如下：

指標	單位	2023年	2024年	2025年
總耗電量.....	千瓦時	1,162.50	9,432.49	30,352.00
用電強度.....	千瓦時／人民幣千元收入	193.75	1.28	0.55

(1) 自我們開始運營以來，我們辦公場所的所有用水相關費用一直由所在樓宇的物業管理服務提供商承擔。鑒於辦公大樓並無安裝獨立水表計量我們的實際耗水量，目前尚無法獲取用水量相關數據。

除日常營運外，我們將綠色理念融入產品的全流程管理，涵蓋從技術研發到商業應用的各個環節。透過結合包括高性能推理引擎、無損精度動態量化、多級緩存及智能路由，以及異構算力資源統一調度等自主研發的技術，我們大幅減少詞元生產過程中的算力資源消耗，從而直接降低能耗。對於已部署的模型，包括MiniMax-M2.5及DeepSeek-V3.2，我們的技術棧相較未經優化的基線，在預留實例（專用算力）上的推理計算量可降低60%至80%。在實際系統級基準測試中，MiniMax-M2.5的單位詞元能耗為每百萬詞元49.2瓦時（按數據中心層面、電源使用效率為1.3計算），較MLPerf Inference v4.1審計的Dense模型參考值（Llama2-70B為每百萬詞元80.7瓦時）低約39%。按參數歸一化口徑計算，這相當於實現約10.7倍的能效提升。在緩存命中場景下，計算量及能耗節省均超過約90%。此外，我們已將此前受限於軟件生態系統兼容性的芯片整合至統一管理的算力調度池中，大幅提升整體算力資源利用率，並減少因閒置產能造成的能源損耗。

業 務

針對商業應用場景，我們利用雲原生軟件的彈性調度能力，為AI模型推理任務自動匹配最優資源，並對無需運行的模型及時將資源釋放至零，從而避免衍生無效算力並消除數據中心內相應的能源損耗。此外，依托算力資源的跨區域佈局，我們根據當地能源稟賦及客戶的業務需求，對算力實施動態分配。算力資源日間會優先部署於光伏電力豐富的地區；夜間則轉移至風力發電充裕的站點，藉此提升可再生電力的消耗比例。展望未來，我們將持續動態優化我們的算力調度策略，優先利用太陽能及風能等綠電資源豐富地區的算力資源。

應對氣候變化

我們密切關注氣候變化所引致的各類潛在影響，並審慎評估及判斷氣候因素為我們的自身營運、業務發展及財務表現帶來的各項風險與不確定性。為系統管理氣候相關風險，我們由綜合管理部牽頭，每年開展氣候風險與機遇的識別及評估，涵蓋：(i)實體風險——主要指氣候變化帶來的物理影響，包括極端天氣事件引發的急性風險以及全球變暖帶來的慢性風險；及(ii)轉型風險——主要指向低碳經濟轉型帶來的影響，包括法律及政策風險、市場風險、技術風險及聲譽風險。評估亦會分析不同氣候風險對我們日常運營以及中長期發展的傳導影響。針對已識別的高影響等級氣候風險，我們持續提升風險管理能力，並穩步增強氣候韌性及經營穩健性。

截至最後實際可行日期，我們的營運或財務狀況概無因氣候變化相關風險而遭受任何重大不利損失。下表載列了我們識別的氣候相關風險、其相關影響及已採取或擬採取的應對措施：

風險描述	影響程度	影響期間	潛在影響	緩解措施
極端天氣可能影響員工 通勤並導致辦公室暫停 營運，對我們的穩定 營運構成潛在風險	低	短期	營運成本增加	配備充足的應急物資，落實常態化的工作場所安全隱患排查，積極開展應急演練，並完善異地遠程應急辦公預案。
全球氣候變暖可能導致 我們能源消耗高	低	中期	營運成本增加	採用節能設備並建立能耗檢查機制，實施常態化能耗管控，以進一步挖掘降耗潛力。

業 務

風險描述	影響程度	影響期間	潛在影響	緩解措施
隨著雙碳目標的不斷演進以及監管機構對計算行業的碳排放和能耗實施上限管控，我們的算力合作夥伴若在碳管理方面不合規，可能會對相關合作造成限制	中	中期	算力租賃 成本上升	常態化追蹤相關法律法規，監測上游算力合作夥伴的溫室氣體排放，並加強全鏈條氣候合規管理。
極端天氣及能源波動可能推高電價，上游算力供應商會將增加的電力成本轉嫁至租賃報價。受ESG管理制約，下游客戶對以可再生能源驅動的計算需求亦日益增加.....	中	中期	算力租賃 成本上升	拓展以可再生能源驅動的合資格低碳算力資源；與合作夥伴合作推動算力負載的綠色電力替代。
若我們的推理能效比未能持續提升，可能會削弱我們的核心競爭力	高	長期	研發投資增加， 經營收入減少	設立專項研發預算，以確保每詞元的能耗優化基準；並透過利用區域能源差異實施跨區算力錯峰調度，從而降低成本及能耗。
若發生任何環境不合規事件，將損害我們的聲譽，導致優質客戶流失並阻礙投融資	中	中期	經營收入減少	規範氣候相關的環境信息披露，以提高我們低碳營運的透明度。建立針對氣候問題的常態化輿論監測機制。

溫室氣體排放

我們營運期間的溫室氣體排放源主要分為兩類：(i)外購電力；及(ii)僱員商務差旅產生的間接溫室氣體排放。於日常營運中，為減少溫室氣體排放，我們實施精細化的能源消耗管理及控制措施：充分利用自然採光優化室內照明能耗，並嚴格控制空調溫度調節引起的能源損耗。於往績記錄期間，我們的溫室氣體排放量如下：

業 務

指標	單位	2023年	2024年	2025年
範圍一溫室氣體排放	千克二氧化碳當量	0.00	0.00	0.00
範圍二溫室氣體排放 — 外購電力	千克二氧化碳當量	691.69	5,612.33	18,059.44
範圍三溫室氣體排放 — 僱員商務差旅	千克二氧化碳當量	—	—	7,508.01
溫室氣體排放總量	千克二氧化碳當量	691.69	5,612.33	25,567.45
溫室氣體排放密度	千克二氧化碳當量／ 千元人民幣收益	115.28	0.76	0.46

- (1) 上述指標的計算乃基於《溫室氣體核算體系》及《IPCC國家溫室氣體清單指南》等相關規定。鑒於本公司為非製造業企業，結合我們的實際營運概不涉及直接能源消耗，故範圍一直接溫室氣體排放並不適用。
- (2) 範圍二間接排放主要包括我們外購電力所產生的溫室氣體排放。相關指標乃參考中國省級電網的平均排放因子計算。
- (3) 範圍三溫室氣體排放 — 類別6「商務差旅」涵蓋僱員搭乘航班及高鐵的商務出行。由於目前的數據收集限制，現階段僅收集到2025年的全部僱員差旅數據。展望未來，我們將完善差旅記錄管理，以持續鞏固碳排放核算的數據基礎。

社會

勞動關係

我們嚴格遵守《中華人民共和國勞動法》等法律法規，並已建立內部政策，以保障員工的合法權益，促進和諧、穩定及互利的勞動關係。我們嚴禁僱用童工及強制勞工，並設立了全流程僱傭合規審查機制。此外，我們杜絕基於性別、年齡或種族等因素對求職者及現職僱員給予任何差別待遇，秉持平等機會及擇優錄取的原則。截至2025年12月31日，我們的勞動合同簽署率達到100%，社會保險覆蓋率達到100%。我們聘用了1名殘疾僱員，針對其能力及需求提供相匹配的工作機會。我們員工架構的明細載列如下：

類別	細分組別	截至12月31日		
		2023年	2024年	2025年
性別	男性僱員	16	32	74
	女性僱員	1	11	26
年齡	≤ 30歲	11	17	32
	31-49歲	6	23	66
	≥ 50歲	0	3	2

業 務

我們每年不定期召開全體員工大會，以收集及處理僱員的申訴與意見，從而持續提升僱員的歸屬感及職場福祉。此外，經綜合考慮市場基準、僱員表現及職位貢獻等因素後，我們制定具競爭力的公平薪酬及福利待遇，當中涵蓋基本薪金、浮動獎金、社會保險及住房公積金、年度獎金、長期獎勵及附帶福利。於2023年、2024年及2025年，員工流失率分別為0%、15.69%及7.34%，主要因個人職業發展所致。

發展與培訓

我們秉持以人為本及以實踐為導向的人才培育理念，高度重視僱員發展。我們緊扣戰略發展推行分層及分類培育，並依托導師制度及實務項目經驗鍛煉及發展員工隊伍。我們持續完善學習賦能機制，培訓架構涵蓋管理、綜合及專業能力。我們的培訓方式以內部培訓為主，輔以外部培訓，而培訓內容包括專業業務及技術培訓以及安全培訓。於往績記錄期間，僱員培訓覆蓋率每年均達100%。

職業健康與安全

我們高度重視僱員的職業健康及安全，致力為其提供安全、健康及舒適的工作環境。我們已設立僱員心理健康關懷機制，指派專人負責溝通及情緒輔導。為進一步強化對突發健康事件的應變能力，我們已購置AED（自動體外心臟去顫器）急救設備，放置在核心辦公區域，同時邀請外部培訓人員為全體僱員提供關於CPR（心肺復甦法）及AED操作的專項培訓。截至最後實際可行日期，我們已取得ISO 45001:2018職業健康及安全管理體系認證。於往績記錄期間，我們概無發生重大安全事故，亦無任何因工亡故或因工受傷的事件。

供應鏈管理

我們致力構建穩定、合規及高效的供應鏈體系。我們制定內部管理政策，並實施嚴格的准入及動態管理機制。於引入新供應商時，彼等必須通過多重內部控制程序，當中包括技術審查。我們在評估IDC營運商、雲端服務供應商及算力中心等核心供應商時，重點關注其營運能力及服務可持續性，並優先考慮具備良好ESG管治的合作夥伴。評估範圍涵蓋資質、資訊安全、質量及環境管理。我們亦優先選擇採購可再生能源的供應商。在日常供應商管理方面，我們已設立常規的供應商監督及評估機制。同時，採購團隊與技術團隊定期合作，按月對供應商的服務能力進行評估及確認。對於不合規的供應商，我們將根據合約條款追究其責任。此外，我們要求供應商簽署《廉潔協議書》，明文禁止商業賄賂及不當利益輸送，並設立了專門的舉報渠道，以確保供應鏈體系規範、透明及良性運行。

產品責任

我們已制定《產品質量應急處置預案》及《業務連續性管理辦法》等制度，建立多層級的應變工作流程，當中明確界定不同級別事件的匯報途徑、處理時限及責任分工，以確保突發質量事件能得到迅速回應及妥善解決。鑒於AI模型推理基礎設施行業的技術特徵及合規要求，我們已建立質量管理體系及信息安全控制體系，以有效保障客戶的合法權益以及安全穩健的業務營運。截

業 務

至最後實際可行日期，我們已取得ISO 9001:2015質量管理體系認證、能力成熟度模型集成第三級認證、ISO 42001:2023人工智能管理體系認證，並獲授包括「重質量守信用企業」等其他八項AAA級證書。我們每年進行定期體系監察審核及內部質量審查，以及時優化及完善質量管理措施，從而構建全流程的閉環質量管理體系。

此外，我們已構建分層及多維度的客戶服務體系，並制定《客戶權益保護管理制度》，以提升客戶滿意度。截至最後實際可行日期，我們概無發生任何因產品責任而引致的重大事件，亦無任何導致重大人身傷亡或財產損失的相關索償事件。

信息安全與數據隱私保護

我們嚴格遵守《中華人民共和國網絡安全法》等法律法規，已制定信息安全管理及隱私保護框架。平台工程部作為我們網絡安全職能的集中管理機構，統籌協調信息安全管理工作。此外，我們配備專職網絡管理人員執行日常網絡運維及安全實施等實際工作。截至最後實際可行日期，我們已取得ISO 27001:2022信息安全管理體系認證、ISO 27701:2019隱私信息管理體系認證，並完成信息系統安全等級保護備案，建立起符合國家法律法規及行業標準的信息安全保護體系。

我們圍繞數據全生命週期搭建了閉環安全管控體系，制定數據分類分級、許可權管理等專項標準。我們遵循最小許可權原則分配存取權限，並於規定情況下取得用戶的知情同意。我們在技術層面上依托計算與網絡隔離界定資源安全邊界，嚴格執行最小化存儲要求，在項目結束後對客戶數據進行刪除或匿名化處理。針對公共雲服務，我們提供高等級安全防護。在本地部署模式下，軟硬件一體化產品部署在客戶端，相關軟件、硬件、模型及數據絕不離開客戶的內網，從根源上杜絕數據外洩的風險。

知識產權保護

我們高度重視知識產權佈局和自主創新保護。我們制定《專利管理制度》，以保護知識產權。我們通過現金紅利激勵員工積極申請知識產權，引導技術創新。對於關鍵崗位，我們通過協定競業禁止條款、保密及知識產權轉讓等方式，有效防範核心技術及商業機密流失的風險。

社區投資

我們依托自有技術推動生態系統發展並創造社會價值，重點聚焦普惠AI、開發者生態系統協作、公益科研支持及開源社區賦能。我們已向平台所有用戶免費開放部分參數規模低於9B的開源模型，使開發者、學生、獨立創作者及中小團隊能夠以低成本使用大語言模型，從而降低AI创新的前期投入和門檻。此外，受限於預先批准的分配限額，我們亦向公益及科研合作夥伴免費提供詞元資源支持，有效降低AI在科研及公益項目中的使用成本。截至最後實際可行日期，我們已進行超過10個重點支持項目，覆蓋創業營、校園分享會、黑客松及海外路演等多元場景。我們以詞元資助為核心支持方式，累計發放詞元資助總額約人民幣1.4百萬元，顯著降低了創新實踐及創業啟動門檻。

業 務

商業道德

我們倡導誠實守信的企業文化，嚴格遵守所有相關法律法規，並營造反欺詐企業文化及環境。我們在《員工手冊》和勞動合同中明確界定員工職業道德標準，並將相關內容納入新員工入職培訓，同時提供便捷且保密的不當行為舉報渠道。於往績記錄期間，接受職業道德及反貪腐培訓的員工及董事覆蓋率，以及我們已向其傳達職業道德及反貪腐政策的供應商覆蓋率，均維持在100%。截至最後實際可行日期，我們未收到任何與商業道德相關事件的舉報。

競爭

中國新興詞元供應市場正處於大規模AI推理需求快速增長階段，在AI應用逐步演進至持續執行任務及日益複雜的多步驟工作流程的推動下，呈現爆發式增長。按詞元吞吐量計，中國詞元供應市場規模由2024年至2025年增長1,602.6%，並預期於2030年達約53.2百億億個詞元，2025年至2030年的複合年增長率為638.3%。

我們憑藉作為獨立生態詞元供應平台而脫穎而出。有別於封閉生態限制用戶僅可使用其專有算力資源及模型，我們始終保持中立及開放的定位。我們連接多元化的異構算力資源及多模型生態，而不會直接與下游應用供應商競爭。憑藉我們於跨芯片適配、推理引擎優化及異構算力資源編排方面的深厚技術能力，我們最大限度提升詞元生產效率，並確保於多種部署模式下實現穩定且可擴展的交付。按年度詞元吞吐量計，我們於2025年為中國第四大詞元供應平台，市場份額為1.5%，並於中國所有獨立生態詞元供應平台中排名第一。

獎項及認可

於往績記錄期間及直至最後實際可行日期，我們因服務、技術及創新獲得獎項及認可，其中重要的獎項及認可載列如下：

獎項/認可	授予機構	獲獎年份
2026卓越人工智能引領者獎Top 30...	世界人工智能大會	2026年
國家高新技術企業	全國高新技術企業認定 管理工作領導小組辦公室	2025年
北京市數字基礎技術標杆企業	北京軟件和信息服務業協會	2025年

業 務

獎項/認可	授予機構	獲獎年份
北京市專精特新中小企業.....	北京市經濟和信息化局	2025年
北京市創新型中小企業	北京市經濟和信息化局	2025年
北京市科技型中小企業	北京市科學技術委員會及中 關村科技園區管理委員會	2025年
2025年度模型服務MaaS標杆案例...	中國人工智能產業發展聯盟	2025年
50家聰明公司.....	《麻省理工科技評論》	2025年
21家高成長性創新公司	《中國企業家》雜誌	2025年
中國最具價值AGI創新機構TOP 50..	極客公園	2024年及2025年
新芽50 & 人工智能50	清科控股及投資界	2025年
2024年中國最值得關注AIGC企業& AIGC產品雙料獎項	量子位	2024年

僱員

截至2025年12月31日，我們擁有100名全職僱員，均位於中國內地。此外，41名全職僱員擁有碩士及以上學歷。下表載列我們按職能劃分的僱員人數：

僱員職能	僱員人數	佔總數的百分比(%)
研發	67	67.0
銷售及營銷.....	21	21.0
運營	5	5.0
一般及行政.....	7	7.0
總計	100	100.0

業 務

我們的成功在很大程度上取決於我們吸引、挽留及激勵合資格人才的能力，而我們相信，我們的高素質人才庫是我們的核心優勢及競爭優勢之一。我們根據高標準、嚴流程的招聘流程，採用線上招聘、內推計劃、第三方招聘等多種方式，針對相關職位的人才需求，選拔合資格候選人。

我們投資於繼續教育及培訓計劃，包括定期及因地制宜的內部及外部培訓，以提升其專業知識及管理技能，增強其技能水平，並使其與各自崗位的行業標準保持同步。我們亦組織活動以幫助僱員更好地了解我們的文化。

我們根據中國法律參與政府規定的僱員福利計劃，包括為養老金、醫療、失業、工傷、生育及住房公積金提供的社會保險。

我們相信，我們與僱員擁有良好的工作關係。我們的僱員均無工會代表。於整個往績記錄期間及直至最後實際可行日期，我們並無發生任何對我們的業務運營產生重大影響的罷工、停工或勞資糾紛。

保險

截至最後實際可行日期，我們為所有完成試用期的全職僱員投購商業保險。商業保險主要包括補充醫療保險，其他承保事項包括意外事故、交通事故等。我們並無投購任何業務中斷或產品責任保險，根據中國法律此等保險不屬於強制投保險種，且該安排符合一般市場慣例。我們並無投購主要人員人壽保險（涵蓋我們網絡基礎設施或信息技術系統損壞的保單或我們物業的任何保險單），亦未投購任何財產保險。

於往績記錄期間，我們並無就我們的業務提出任何重大保險申索。倘發生任何未投保的業務中斷、訴訟或自然災害，或對我們未投保的設備或設施造成任何重大損壞，均可能對我們的經營業績造成重大不利影響。有關詳情，請參閱「風險因素－與我們的業務及行業有關的風險－我們有限的保險範圍可能使我們面臨重大成本及業務中斷」。

執照、批文及許可證

我們持有所需的執照、批文及許可證，並監控持續的合規情況，以確保獲得所有必要的授權。於往績記錄期間及直至最後實際可行日期，自我們業務運營的相關監管批文簽發之日起，並無發生任何可能會對我們重大執照、許可證、批文及證書的維持與續期造成不利影響的重大意外或不利變動。下表載列我們業務營運所需的重大執照及許可證的詳情。

執照/許可證	持有執照/ 許可證的實體	授予日期	屆滿日期
增值電信業務經營許可證.....	本公司	2024年8月	2029年7月
境內深度合成服務算法備案.....	本公司	2025年9月	不適用

業 務

執照/許可證	持有執照/ 許可證的實體	授予日期	屆滿日期
境內深度合成服務算法備案.....	本公司	2026年3月	不適用
生成式人工智能服務登記.....	本公司	2025年12月	不適用
信息系統安全等級保護備案證明...	本公司	2024年10月	不適用

物業

截至最後實際可行日期，我們於中國並無擁有任何物業。截至最後實際可行日期，我們主要在中國租賃三項實際使用中的物業，總建築面積為2,956.74平方米，用作我們的辦公場所。我們相信，目前融資足以滿足當前的需要。

法律訴訟及合規

於往績記錄期間及直至最後實際可行日期，我們並無牽涉我們認為會對我們的業務、經營業績(包括我們的特專科技產品的研發)、財務狀況或聲譽及合規情況造成重大不利影響的任何實際或待決法律、仲裁或行政訴訟(包括任何破產或破產清算訴訟)。

於往績記錄期間及直至最後實際可行日期，我們並無任發生何重大違反法律及法規的行為，亦無發生任何系統性不合規事件，而董事認為，整體而言可能對我們的業務、經營業績及財務狀況產生重大不利影響。於往績記錄期間及直至最後實際可行日期，我們已在所有重大方面遵守相關中國法律及法規。

風險管理及內部控制

我們已建立且目前維持風險管理及內部控制制度，該等系統由我們認為適合我們業務運營的政策及程序組成。我們致力於不斷改善該等制度。我們在業務運營的多個方面採取並實施風險管理政策。董事會負責建立及更新我們的內部控制制度，而我們的高級管理層則監察各附屬公司及職能部門內部控制程序及措施的日常實施。我們的董事認為，我們已設立充分有效的內部控制程序，以確保日後遵守相關法律法規。

人力資源風險管理

我們已制定涵蓋人力資源管理各個方面(包括招聘、培訓、職業道德及法律合規)的內部控制政策。我們的行業急需經驗豐富的僱員，尤其是研發人員。主要人員的離職可能會對我們造成不利影響。

我們所有的僱員均與我們訂立僱傭協議，當中載有保密及競業禁止條款。我們亦要求僱員遵守更嚴格的職業道德標準。我們向所有僱員提供員工手冊，當中載有每名僱員必須遵守的行為守則。

業 務

財務報告風險管理

我們已實施一系列財務報告風險管理會計政策，並建立了嚴格的內部報銷、財務報告管理及審批政策。具體而言，財務部門對發票、匯票、票據及其他財務文件實施特定審核及核實程序，以確保我們收到及使用的原始文件的真實性。財務部門亦會檢查文件上顯示的金額及時間是否與相關合同一致。我們有嚴格的內部審批流程，幾乎所有審批事項均通過公司內部平台在線完成，在保障高效運營的同時，實現了有效的端到端監控。

法律合規及知識產權風險管理

我們的運營風險管理涉及遵守中國法律法規，特別是有關信息服務業以及保護知識產權和避免侵犯第三方知識產權的責任的法律法規。我們擁有一支由經驗豐富的法律專業人士組成的團隊，以確保我們的合規性並控制知識產權相關風險。我們的法律團隊負責審批合同、監控中國法律法規的最新情況，並確保業務運營持續遵守相關法律法規。我們的法律部門亦協助確保我們及時向相關部門提交所有必要的商標、版權及專利申請。