
業 務

概覽

我們的使命

我們的使命是通過通用人工智能(AGI)創建互聯直覺的世界。

我們的願景

我們相信，隨著AI應用增加，世界變得更加美好。AI改變了我們工作、溝通和與世界交互的方式。這並非轉瞬即逝的趨勢，而是持久的轉變。本公司的名稱「雲懂你的心」(雲知聲)體現該宗旨。我們的願景是引領技術演進，探索AI領域的每一種可能性。

我們是誰

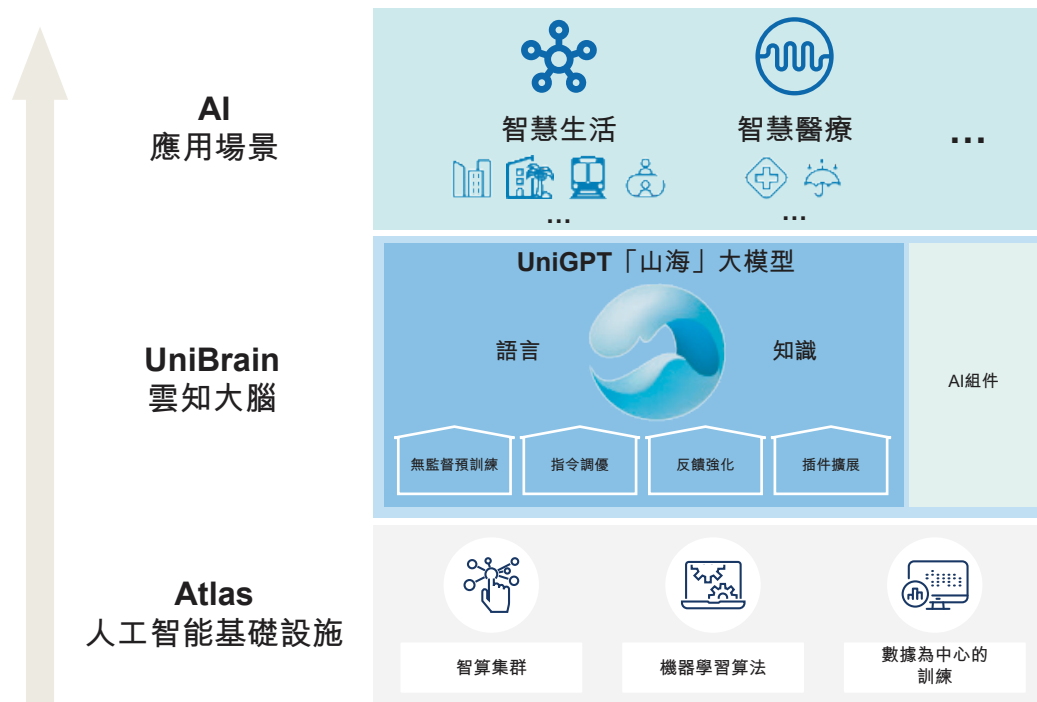
我們已在亞洲引領AGI技術，特別在被視作通向AGI關鍵路徑的大語言模型商業化方面取得成功。交互式AI本質上通過與人類直接和持續交互來運行，受益於自動和不斷提升其能力的反饋閉環，具有最大的AGI商業化潛力。在2017年Transformer算法及2018年BERT算法取得突破後不久，我們利用我們在交互式AI方面強大的研發專業知識及自成立以來獲得的市場洞察力，為廣泛的垂直行業領域客戶推出我們的首個基於BERT的大語言模型UniCore及相關AI解決方案。我們的AI技術不斷發展，具有多模態能力及大量動態、不同應用場景的用戶交互迭代。

我們戰略性地於2016年開始建立Atlas人工智能基礎設施，其基於我們的智算集群，高效地為機器學習任務動態調度強大的算力。其為前沿AI模型的開發、優化和商業化提供重要支撐。我們的智算集群目前擁有超過200 PFLOPS的計算能力，以及超過1.6PB的存儲容量。在過去幾年UniCore的演進和優化的基礎上，我們於2023年推出了一個擁有600億個參數的自有大語言模型山海大模型，作為我們技術平台雲知大腦的新核心。山海大模型具有強大的多語言及通用知識能力。通過一系列工具行業知識增強、指令調優及人類反饋強化，我們能夠將山海大模型進一步提升為一個兼具全面通用知識和行業專業能力的大語言模型，通過MaaS適配各種應用場景。

我們基於雲知大腦技術平台為智慧生活和智慧醫療中廣泛的應用場景提供了極具競爭優勢的產品和解決方案。根據弗若斯特沙利文的資料，按收入計，我們是2022年中國第四大AI解決方案提供商，但在年度收入超過人民幣500百萬元的企業中增長最快。同年，按收入計，我們在中國智慧生活及智慧醫院行業均排名第三。這些產品和解決方案的廣泛商業應用為我們提供了優質的用戶反饋，進而支持山海大模型作為雲知大腦的核心持續迭代。

業 務

下圖說明Atlas人工智能基礎設施、雲知大腦及我們的AI解決方案的應用場景。



雲知大腦

我們的關鍵技術平台雲知大腦包括我們自有的山海語言大模型及各種AI組件。山海大模型是雲知大腦的核心。經過海量文本及代碼預訓練，其擁有超過600億個參數，具有強大的通用能力。

雲知大腦的關鍵特點

- **通用性：**山海大模型以大規模生成式預訓練及參數調優的框架為基礎。其在自然語言、通用知識、推理及快速學習方面擁有強大的能力，並可處理行業複雜的應用場景，提升商業化效率。
- **適應性：**我們通過行業數據進行增量訓練和指令微調、插件和AI組件增強山海大模型的行業特定應用，而毋須進行繁重的模型重新訓練，使我們的AI解決方案適應高度專業化的場景。我們可以將山海大模型與客戶的自有知識庫、實時信息及現成的API集成，以優化其在高度專業化應用場景中的性能。
- **高效性：**根據山海大模型強大的語言、知識及學習能力，使AI解決方案具備固有的高度適應性，能夠處理多個行業的主流需求。在拓展AI解決方案的應用範圍時，我們在統一的適配框架下快速進行任務數據調優程序，毋須根據特定業務邏輯進行模型重訓及流程再開發，大大減少所需的人力投入。這種方法不同於傳統「作坊式」手動打造AI終端產品，研發效率高。

業 務

Atlas人工智能基礎設施

智算集群

強大的算力為我們開發及優化前沿AI模型提供關鍵的支持，並探索其商業應用。我們的智算集群是我們演進的基石。我們的智算集群擁有超過200 PFLOPS的計算能力，以及超過1.6PB的存儲容量，兩者均可按需無縫動態擴展。

我們的智算集群可高效動態地調度數以千計的GPU進行並行計算，並進行無縫動態擴容，以解決不斷變化的業務需求。其還針對大規模機器學習任務優化了存儲、帶寬和計算能力。特別是針對大語言模型，我們的智算集群對並行機制進行深度工程效率優化，充分利用計算和帶寬資源。我們智算集群的強大計算能力、高效且動態調度及可擴展性構成我們推進AI解決方案的競爭優勢。

機器學習算法

根據弗若斯特沙利文的資料，我們是首批將深度學習算法商業應用的公司之一。自成立以來，我們一直探索AI模型在廣泛的商業應用場景中的潛力。

根據弗若斯特沙利文的資料，於2012年，我們為率先將深度學習算法應用於語音識別產品的公司之一。自此，我們持續推動卷積神經網絡(CNN)、生成對抗網絡(GAN)、強化學習(RL)等先進AI算法的商業應用。

根據弗若斯特沙利文的資料，我們亦是亞洲最早將AI大語言模型商業化的公司之一。於2018年，我們利用新興的Transformer算法中獲得WMT2018世界機器翻譯技術評測前兩名。自此，我們就為廣泛的垂直行業客戶提供Transformer升級的AI解決方案。於2023年，我們推出我們自有的600億個參數的大語言模型山海大模型。

數據為中心的訓練方法

憑藉我們的行業知識及洞察力，我們已構建以數據為中心的AI模型訓練及優化框架，據此，AI模型選擇多樣化、優質數據反饋至AI模型的持續訓練及優化，打造數據積累及行業知識初始模型，以及反饋數據優化模型的自我強化閉環。我們認為，以有限數量的優質數據為導向的訓練較以大量未經過濾數據為導向的訓練具有更高的效率。該訓練框架消除了開發新AI解決方案時對模型重建的需求，毋須大量人為干預，從而將我們從繁重的指令調優支持中解放出來，並使我們能夠將研發工作集中在基礎模型優化上。

業 務

市場機會

AI在越來越多的應用場景中得到應用，隨著AGI技術的發展，這種應用場景將進一步加速增長。隨著大語言模型的快速演進、算力和數據的指數級增長以及應用場景的激增，AI（尤其是AGI）有望從根本上改變我們與世界的交互方式。AGI解決方案的出現改變了各垂直行業的供需兩端，而前沿科技驅動市場增長。根據弗若斯特沙利文的資料，中國的AI解決方案市場預期將由2022年的人民幣937億元增至2028年的人民幣6,332億元，複合年增長率為37.5%。我們相信AI將在以下行業發揮重要作用（其中包括）：

- **智慧物聯網**：根據弗若斯特沙利文的資料，中國物聯設備的數量於2022年為18億，為人工智能賦能增強提供重大機遇。預期智慧物聯行業的市場規模將由2022年的人民幣344億元增至2028年的人民幣3,128億元，複合年增長率為44.5%。智慧生活包括智慧住宅、商業空間、酒店及交通的人工智能解決方案，為智慧物聯網中最大的部分。
- **智慧醫療**：中國醫療行業已基本上完成初步數字化轉型，為AI創新奠定基礎。截至2021年12月31日，在中國共有36,507家醫院，智能化升級的潛在需求很大。根據弗若斯特沙利文的資料，智慧醫療市場規模預期將由2022年的人民幣36億元增至2028年的人民幣810億元，複合年增長率為67.8%。智慧醫院包括醫院及醫療機構的人工智能解決方案，為智慧醫療中最大的部分。

產品及解決方案

基於雲知大腦，我們提供易於部署的AI產品及解決方案，該等產品和解決方案基於我們不斷增強的MaaS能力，可針對廣泛的客戶進行高效調優。該等產品及解決方案有助客戶提高運營效率及生產力，為各行業的終端用戶帶來價值。具體而言：

- **智慧生活**：我們為智慧住宅、商業空間、酒店及交通等智慧生活場景提供多樣化的AI產品及解決方案。在AI解決方案的賦能下，企業可以為客戶提供沉浸式智能產品及服務，提高運營效率和服務質量，降低管理成本。於2022年度，我們的智慧生活解決方案已為373名客戶提供支持（包括中國前三大保險集團之一、深圳地鐵20號線及廈門軟件園等）。此外，我們通過MaaS向開發者和企業按需提供AI大語言模型能力。主要產品包括基於公有雲的AI能力API、定制化自有AI技術服務平台及AI模型嵌入式芯片和物聯網硬件模組。於2022年，我們AI能力的外部用戶月調用量峰值為915.8百萬。同年，我們已向開發者和客戶銷售12.8百萬顆AI芯片。
- **智慧醫療**：我們提供AI賦能醫療解決方案，如病歷語音輸入、病歷質控、單一疾病質控及醫療保險支付管理。該等AI解決方案不僅幫助醫務人員提高效率，

業 務

亦規範了醫療服務流程和決策，從而減少醫療失誤，提高醫療服務質量並保障患者權益。該等AI解決方案亦有助於防止不合理的醫療開支，支持醫改，並提高醫療保險的使用率和效率。於2022年度，我們向165名客戶提供AI解決方案。

我們的進入市場策略

我們與智慧生活及智慧醫療行業的領導者合作，開發AI解決方案，並獲得寶貴的行業洞見。我們與行業領導者的夥伴關係穩固，並取得未來項目合作。憑藉自我們的燈塔項目獲得的專業知識、經驗及市場認可度以及我們的AI解決方案的高可擴展性，我們向同一行業的其他客戶及具有類似痛點及需求的其他行業介紹我們的解決方案。此進入市場戰略使我們能夠持續通過不斷改善產品提升標桿客戶的參與度，並擴大我們的客戶群。於往績記錄期間，我們的客戶由2020年度的459家增加至2021年度的510家，並進一步增加至2022年度的538家。

財務概覽

於往績記錄期間，我們的收入主要來自銷售AI產品及解決方案。我們的收入由2020年的人民幣260.8百萬元增加至2021年的人民幣455.8百萬元，並進一步增加至2022年的人民幣600.6百萬元，複合年增長率為51.7%；我們的毛利潤由2020年的人民幣89.2百萬元增加至2021年的人民幣145.0百萬元，並進一步增加至2022年的人民幣239.9百萬元，複合年增長率為63.9%；我們於2020年、2021年及2022年的淨虧損分別為人民幣297.8百萬元、人民幣434.3百萬元及人民幣375.4百萬元；剔除股權交易中的股份支付開支、贖回負債利息的融資成本及[編纂]開支，我們於2020年、2021年及2022年的經調整淨虧損分別為人民幣209.8百萬元、人民幣171.7百萬元及人民幣183.2百萬元。

我們的優勢

AGI技術的先驅者及引領者

自成立以來，我們是人工智能技術創新的先鋒，引領中國人工智能技術的商業應用，為AGI商業化鋪平了道路。

深度學習算法的發展：我們緊跟深度學習算法（如DNN及CNN）的進步，將前沿技術融入至創新產品及解決方案中，並引領商業化應用。根據弗若斯特沙利文的資料，我們是亞洲最早(i)將深度神經網絡算法應用於語音識別產品；及(ii)發佈基於深度學習的商用雲端語音識別引擎的公司之一。於2012年，我們開始深度學習技術的商業應用，協助中國領先的互聯網搜索引擎供應商搜狗推出語音助手。2016年，我們在北京協和醫院推出語音電子病歷。我們深度學習算法的性能也獲得了著名獎項的認可，包括：

- 2017年AWE艾普蘭（中國家電及消費電子博覽會）最佳核芯獎；

業 務

- 2020年Blizzard Challenge國際語音合成大賽，我們憑藉自有的端到端語音合成技術，普通話和上海話的多項指標排名第一；及
- 2021 INTERSPEECH，在深層噪聲抑制及聲學回聲消除挑戰賽中均排名第二。

人工智能基礎設施及通向AGI的路徑：自2016年起，我們構建了面向大規模商業應用的人工智能技術體系，包括支持規模不斷擴大的人工智能模型訓練和優化的Atlas人工智能基礎平台，以及專注於人工智能模型行業增強的知識圖譜平台，並為其提供專業領域的知識。2017年Transformer及2018年BERT取得突破後不久，我們利用我們在交互式AI方面強大的研發專業知識及自2012年成立以來獲得的市場洞察力，向廣泛的垂直行業領域客戶推出我們的首個基於BERT大語言模型UniCore及相關AI解決方案。2019年初，我們升級了BERT以提升其認知能力，並在病歷質控產品的醫學知識圖譜中採用了有關能力。AI模型的演進使我們能夠擴展AI解決方案的應用，並標誌著我們從提供單一產品解決方案向全行業解決方案的轉變。我們的技術實力繼續表現卓越，獲得了以下主要獎項：

- WMT2018國際機器翻譯大賽排名前二；
- 2019年北京市科學技術進步獎一等獎；
- 2020年全國知識圖譜與語義計算大會(CCKS)，「面向中文電子病歷的醫療命名實體識別評測任務」獲得冠軍，並獲得唯一技術創新獎；
- 2022年國際先進人工智能協會(AAAI)年會，在首字母縮寫提取及首字母縮寫消歧任務中均排名第一；
- 2022年4月贏得了中國生物醫學語言理解評估(CBlue)錦標賽冠軍

於2022年，藉助Atlas人工智能基礎設施及我們堅實的Transformer算法技術基礎，我們開始探索生成式預訓練大語言模型的創新應用。於2023年5月，我們發佈我們自有的生成式預訓練山海大模型，在UniCore的基礎上大幅提升能力，並在智慧醫療及智慧生活產品中採用，標誌著AI大語言模型商業化取得了重大進展。

以堅實的基礎設施推動強大的研發及產品化能力

我們在同業中脫穎而出的關鍵因素之一是我們的Atlas人工智能基礎設施。該人工智能基礎設施執行以下各項工作：(i)根據不斷變化的業務需求動態調度算力；(ii)基於海量數據的大語言模型初始訓練；及(iii)以精選的優質用戶反饋為導向的高效優化。強大的人工智能基礎設施支持雲知大腦的不斷進步，幫助我們保持在技術進步的前沿，並提供具有強大性能的創新AI解決方案。

業 務

雲知大腦具有強大的自然語言和通用知識處理能力，這使它能夠應對各行業和產品的主流需求。雲知大腦提升了開發過程中從基礎模型到實際產品的效率，保證了我們開發的效率和產品的適配性。其方法不同於傳統的「作坊式」手動打造AI終端產品，有助我們進一步提高研發效率，縮短產品和解決方案上市時間。自2023年起，在山海大模型的支持下，我們預期我們的研發及產品化效率將進一步提升。雲知大腦通過多種軟硬件交付模式提供基於AI模型的AI能力及解決方案。其支持快速適配和定制，讓我們的智能物聯網客戶可以使用我們的AI芯片開發平台在30分鐘內為空調和抽油煙機等設備設置AI固件。自2020年以來，我們的解決方案滲透到九個新的垂直行業，包括城市軌道交通及商業醫療保險管理。

憑藉Atlas人工智能基礎設施及雲知大腦，我們在Transformer算法框架下自主訓練AI大語言模型，開發創新的AI產品及解決方案，研發效率顯著。我們經驗豐富的研發工程師有不斷提供一流研發成果的良好往績紀錄。

廣泛深入的合作支撐商業化成功

憑藉領先的AI技術及強大的研發能力，我們成功滲透各種行業，與我們稱之為「燈塔客戶」（如家電的格力、康養的平安科技、醫療服務的北京協和醫院及醫保的中國人保）的該等行業的領先公司合作。

通過與早期燈塔客戶的合作，我們獲得該等垂直行業內各種各樣的應用場景的豐富經驗。其包括高頻次、代表性場景，及零散的長尾場景。儘管各個長尾場景單獨發生的頻率較低，但合併佔有所有場景的大部分。此經驗令我們獲得了對各個行業的寶貴見解，使我們能夠快速識別常見需求，開發有針對性的解決方案及加快部署。

憑藉多年來在AI產品及解決方案商業化方面積累的市場知識，我們戰略性地將重心放在智慧生活及智慧醫療行業，兩者均具有經證實的市場潛力。於2016年，我們與格力合作進軍智能家居垂直行業。同年，我們與北京協和醫院合作，為醫院推出語音電子病歷產品。截至2022年12月31日，我們的解決方案已於廣泛的垂直行業採用，包括醫療、家居、商業空間、酒店及交通。根據弗若斯特沙利文的資料，於2022年，我們是中國第四大AI解決方案供應商（按收入計算），在年收入超過人民幣500百萬元的企業中增長最快。同年，我們在中國智慧生活及智慧醫療行業均排名第三。

於2022年度，我們的解決方案服務538名客戶。我們的財務表現穩健，證明我們商業化的成功。2022年，我們的收入達到人民幣600.6百萬元，同比增長31.8%，自2020年至2022年收入的複合年增長率達到51.7%。

業 務

反饋閉環不斷推動技術及產品優化

以真實場景的海量數據和持續的用戶反饋為指導，我們的AI模型進行高度自動化的自我強化，並不斷提高效率，推動我們的AI技術不斷迭代。高質量、海量數據讓我們充分釋放尖端模型的優勢。例如，人工智能語義理解的準確性由2018年的BERT前深度學習模型的88.04%提高至2019年的UniCore的91.02%，並進一步提升至2023年的山海大模型的96.07%。該技術進步使我們能夠將我們的人工智能解決方案的能力從感知擴展到理解、響應和生成。

我們人工智能技術的進步使我們能夠擴大我們的人工智能解決方案的應用範圍，以更強大、更完善的解決方案不斷挖掘新的行業垂直領域及應用場景。碎片化的市場需求會對AI解決方案供應商造成重大困難。通常，當面對新的應用場景需求時，由於其特定設計，重新利用現有AI解決方案變得不可行，導致大量資源投入定制甚至完全重新設計AI解決方案。憑藉雲知大腦的普遍性、適應性及效率性，以及受反饋閉環所推動，我們能夠不斷提高優化AI模型及產品的效率。

我們不斷優化的產品具有強大的性能和成本效益，精準滿足客戶需求，獲得顯著的市場知名度。截至2022年12月31日，(i)我們的智慧物聯網產品已應用於多個垂直領域(如住宅、商業空間、酒店及交通)的數千萬台設備，積累了專業知識及經驗；及(ii)我們的智慧醫療產品、醫學專業知識圖譜，包含1.6百萬個概念、3.7百萬條術語及8.4百萬種關係。

經驗豐富的管理團隊及核心人員

我們由首席執行官黃博士、董事長兼首席技術官梁博士及物聯網部門負責人康博士共同創立，研發部門負責人李霄寒博士後來加入。該等領導者具備豐富的AI研發和產業化經驗，引領我們在深度學習算法商業應用領域開拓，戰略性選擇研發和營銷重點。謝冠超先生、陳吉勝先生後續加入管理團隊，彼等均在國內外知名公司具有極為豐富的技術研發和項目管理經驗。彼等目前分別負責智慧醫療事業部及創新事業部。學術和產業背景的結合，賦予了我們管理層在AI技術和市場方面具有深刻的戰略眼光，以及廣泛的行業人脈。

在我們管理層的領導下，我們已積累一支AI專業人員團隊作為我們的核心技術人員，彼等各自在IBM、摩托羅拉、微軟及Nuance等國際知名企業獲得豐富經驗，並持有中國科學技術大學、中國科學院、北京大學、清華大學及UIUC等國內外頂尖大學的博士學位或碩士學位。

業 務

我們的戰略

加強人工智能基礎設施及模型

我們領先的人工智能基礎設施和大語言模型是我們快速增長的基礎和源動力。

我們計劃通過以下方式加強Atlas人工智能基礎設施：(i)擴大智算集群的算力，以支持現有大語言模型的演進及未來模型的訓練；(ii)緊跟尖端基礎機器學習算法的演進，以期進一步提升我們的技術系統。

我們計劃繼續升級我們的雲知大腦，以增強我們的人工智能開發能力。我們擬通過以下方式推出更強大的大語言模型並提高其效率：(i)構建參數規模更大的大語言模型，通過更高質量的數據進行訓練和優化，以更廣泛的插件及AI組件增強；(ii)訓練並優化自有的多模態大語言模型，可直接處理圖像、語音及文字等輸入，使日後更容易推出新的針對多模態場景的AI解決方案。同時，我們計劃探索事件監測、情感識別及多模態感知等技術及應用。在生成領域，我們旨在豐富大語言模型支持的語音風格，提高個性化、情感化合成，並持續提升聲像融合多模態合成技術，實現從「信息傳遞」到「輔助創造」的轉變。

吸引和培養AI人才

我們的成功深深植根於我們的前沿研發人才打造的尖端技術。

我們擬加強與一流大學和研究機構的合作。通過聯合研發、試驗課程設計、競賽及實習等方面的合作，我們旨在在合作過程中物色合適的人才並展開初步接觸。

在擴充團隊的同時，我們亦計劃維持及完善全面的人才培訓及發展計劃。我們致力於為頂尖人才提供一流的人工智能基礎設施及成長支持，以提高彼等研發工作的效率。我們專注於培養行業導向的AI人才，能夠優化AI解決方案的應用。憑藉對基礎AI技術的深入理解，預期彼等通過實踐將洞悉不同行業及客戶需求，並通過持續創新為開發新的或更好的解決方案作出貢獻。憑藉我們的系統化培訓計劃，我們致力確保我們的人才可專注於產品創新和各模塊通用能力的打造，而非以項目制的方式落入瑣碎的支持中。我們計劃培養一支研發及應用團隊，該團隊在技術及產品研發方面表現卓越，可作為我們團隊保留及傳承專業知識、經驗及文化的中流砥柱。

拓展應用場景及垂直行業

隨著AI不斷改變我們工作、溝通及與世界交互的方式，我們看到將AI應用在各行各業的眾多場景的新機遇。即使在我們的AI解決方案已滲透的行業，仍存在大量我們尚未解

業 務

決的智能升級需求。同時，廣泛的垂直行業尚未被AI解決方案滲透。我們通過與燈塔客戶合作建立的行業洞見及聲譽有助於我們吸引更多用戶及加快已滲透行業的商業化進程：

深入鑽研已滲透行業的需求：我們通過高質量的解決方案和服務，與用戶建立了穩固及密切交互的合作關係。憑藉現有的行業關係並由山海大模型賦能，我們計劃擴大AI解決方案的覆蓋範圍，以覆蓋更多應用場景，識別及解決人工智能解決方案尚未被充分滿足的需求。例如，從醫療服務擴展至醫療保險，然後再擴展至醫療監管。

將成功經驗移植至中長尾客戶：在許多已滲透的行業中，我們已與燈塔客戶建立穩固的關係。我們在與燈塔客戶的合作中獲得寶貴的專業知識。我們預期繼續把成功經驗移植到行業的中長尾客戶中，為客戶提供高效、高度標準化的解決方案，解決行業共同需求和代表性場景，並根據需求迅速部署。

擴大行業覆蓋範圍：憑藉我們對智慧家居及智慧醫療的理解，我們已成功擴展至康養領域。我們計劃加強我們的能力，並在相關領域尋求擴張機會。憑藉雲知大腦的普遍性，我們預期新垂直行業的滲透率將更加精簡及高效。

國際擴張

山海大模型適配多種語言的能力有利於我們的國際擴張。通過與不同地區用戶的交互獲得的用戶反饋為我們提供了針對性的提升山海大模型的多語言能力。

隨著山海大模型開始提供更多語言支持，我們將加強國際營運能力。一方面，具備多語言能力的山海大模型可以幫助我們解決國內客戶的國際擴張需求。另一方面，通過與海外合作夥伴合作及建立海外分銷渠道，我們期望以在中國市場已獲驗證的解決方案進軍海外市場，賦能全球範圍的智能化轉型。我們計劃專注於擁有強大數字基礎設施及對AI解決方案有高需求的海外市場。

業 務

UniBrain雲知大腦

我們已開發UniBrain「雲知大腦」，這是我們自有的中心化技術平台，為我們所有的解決方案及產品提供支持，其中包括通用大語言模型山海大模型及AI組件。雲知大腦具備感知、理解和反饋的一站式能力。它整合了(i)交互式AI為核心的多模態AI能力，(ii)基於多年實踐積累的行業洞察力和專業知識開發的行業級知識圖譜，及(iii)一站式物聯網互聯系統，為所有物聯網設備賦能。下圖概述山海大模型與我們的人工智能組件之間的交互：



業 務

山海大模型

於2019年初，我們推出我們首個通過無監督預訓練構建、擁有數億個參數的基於BERT的大語言模型UniCore。基於過去數年UniCore的演進及進步，我們利用更大規模的文本及代碼數據訓練，開發出具有600億個參數的大語言模型山海大模型，提供強大的通用語言、知識、推理及小樣本學習能力。我們通過指令優化及從人類反饋中強化學習進一步優化山海大模型，以發展其語言生成、語義理解、知識問答、邏輯推理、數學計算和代碼生成的能力，同時符合安全規則，最終目標是實現類人AGI。針對醫療領域等特定行業，通過優質醫學教材、臨床指南、病歷等增量訓練，山海大模型顯著提高其解決專業問題的知識及能力，通過了美國醫師執照考試，MedQA數據集獲得了81分，超越GPT-4，下列截圖說明山海大模型的核心能力。

語言生成



業 務

語義理解



“明天天气贼好”和“明天天气不错”这两句话，哪句的情感更加强烈？

“明天天气贼好”的情感更加强烈。因为“贼”字表示非常、十分，而“不错”则只是表示还不错，没有那么强烈的情感色彩。

 重新生成

在此输入你想问的问题，enter发送，shift + enter 换行

请遵守《[山海大模型服务协议](#)》和《[隐私政策](#)》

所有山海的回复内容均由人工智能模型输出，其内容的准确性和完整性无法保证，不代表我们的态度和观点

知識問答



“楼台倒影入池塘”这句古诗反映了光的反射还是折射现象

这句古诗反映了光的反射现象。因为“倒影”指的是物体被水面等平行于表面的介质所反射形成的像，而不是通过介质发生了折射。

 重新生成

在此输入你想问的问题，enter发送，shift + enter 换行

请遵守《[山海大模型服务协议](#)》和《[隐私政策](#)》

所有山海的回复内容均由人工智能模型输出，其内容的准确性和完整性无法保证，不代表我们的态度和观点

業 務

邏輯推理

 | 



小明家里有两个哥哥，大哥叫王大，二哥叫王二，请问小明姓什么



小明的姓是王。因为小明家里有两个哥哥，分别叫王大和王二，而在中国文化中，兄弟之间通常会使用同一个姓氏，因此可以推断出小明的姓也是王。

 重新生成

在此输入你想问的问题，enter发送，shift + enter 换行

请遵守《山海大模型服务协议》和《隐私政策》

所有山海的回复内容均由人工智能模型输出，其内容的准确性和完整性无法保证，不代表我们的态度和观点

數學計算

 | 



今有鸡兔同笼，上有三十五头，下有九十四足，问鸡兔各几何？



设鸡的数量为x，兔子的数量为y,则题目可以转化为以下方程组： $x + y = 35$
 $2x + 4y = 94$
解方程组得到：
 $y = 12$
 $x = 23$
因此，鸡的数量为23，兔子的数量为12。

 重新生成

在此输入你想问的问题，enter发送，shift + enter 换行

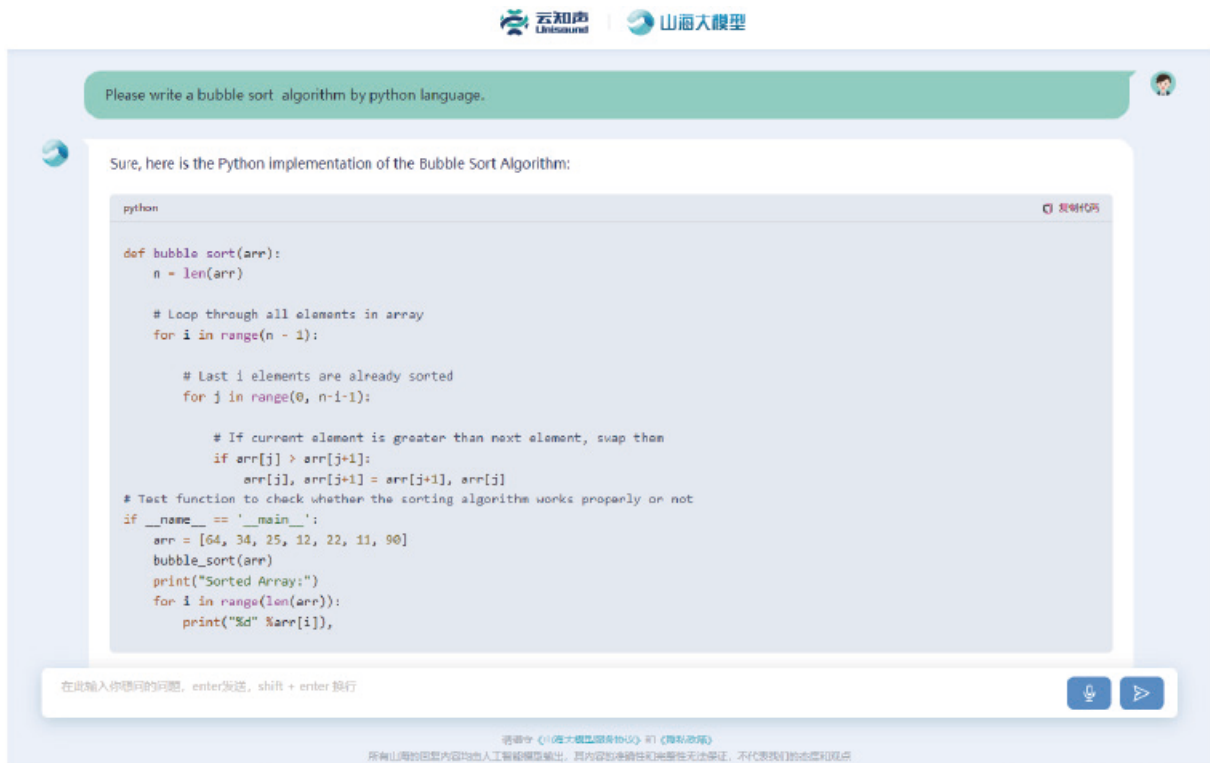
 

请遵守《山海大模型服务协议》和《隐私政策》

所有山海的回复内容均由人工智能模型输出，其内容的准确性和完整性无法保证，不代表我们的态度和观点

業務

代碼生成



在具備通用能力的基礎上，山海大模型可以通過插件擴展、行業增強和定制等方式優化，大幅提升其在智慧生活、智慧醫療等垂直行業的專業能力：

- **插件擴展：**山海大模型通過插件實現實時信息檢索和行業級的知識圖譜能力，提供專業的信息和知識整合，充分利用大語言模型的理解和生成能力，減少過時或不正確的信息結果。必要時，山海大模型可調用外部API來提升整體解決問題的能力，有效整合外部信息、知識和標準工具。
- **行業增強：**通過使用數據進行增量訓練以及指令調優，山海大模型自身具備了跨領域語言和知識，從而增強了其特定行業的知識和技能，提高了其解決專業問題的能力。
- **客戶定制：**藉助客戶信息和知識，山海大模型可以進一步增強其解決此類客戶特定問題的能力，通過輸入其信息和知識整合能力來滿足客戶的定制需求。

憑藉其認知架構的升級，山海大模型提升了基於MaaS商業模式的可能。在山海大模型之前，為了在人工智能終端產品的作坊式手動打造時代開發出新穎的人工智能解決方案，需要根據相關的業務邏輯設計和完成訓練過程，並在必要時調用相關的智能組件，如識別、理解和生成，再輔以測試和優化。這一過程通常需要大量的人力投入。此外，每當

業 務

業務調整或擴展時，訓練過程通常需要重新制定，其效率仍有待提高。利用其通用性及定制能力，山海大模型幫助我們大幅減少了在開發新的解決方案時模型微調所需的數據和人力投入。

人工智能組件

雲知大腦包括廣泛的人工智能組件：(i)支持交互式AI和其他交互的多模態感知和生成能力，作為山海大模型與應用場景交互的基礎人工智能支持；(ii)行業級的知識圖譜，涵蓋智慧生活和智慧醫療等行業，並可擴展到更多垂直領域，增強山海大模型的專業知識；及(iii)我們的一站式物聯網互聯系統，促進所有物聯網設備在一個閉環系統中的統一管理及交互。

多模態感知及生成

山海大模型的感知和生成能力可以通過人工智能組件的多模態能力得到加強，使其能夠在更廣泛的現實生活場景中提供更強大的交互、推理和決策性能。

- **多模態感知**：通過信號降噪、語音喚醒、語音識別、聲紋識別、面部識別、OCR和發音評估等技術，雲知大腦可實現多場景、複雜環境下的多模態感知能力。
- **多模態生成**：通過語音合成(TTS)和視聽合成(AVS)等技術，山海大模型產生的反饋可以自動轉換為高質自然的語音和視覺信號，實現具有類人情感的個性化反饋。

行業級知識圖譜

我們形成了行業級知識圖譜，助力高效及無縫的行業特定專業信息交互。我們的行業級知識圖譜以自然語言處理及知識圖譜技術構建，通過相關服務及應用的標準化開發套件提供可視化知識圖譜構建及管理功能。通過對行業的洞察力和跨行業的專業知識，我們可以進軍新的行業，開發相關的知識圖譜，並提高工作效率。例如，截至2022年12月31日，我們已在醫療知識圖譜平台上積累1.6百萬個概念、3.7百萬個術語及8.4百萬項關係。

我們的行業級知識圖譜具有以下主要功能：

- **智能知識管理**：有效組織管理行業知識、標準、法律法規等文檔，提供智能語義搜索、智能文檔推薦、智能知識問答等能力，提高用戶工作效率。
- **智能質量控制**：基於行業知識圖譜及標準與規範，對業務數據進行規範性審查、驗證及質量控制，提高業務質量和評估效率，反饋提高業務水平。

業 務

- **智能輔助決策：**基於行業知識圖譜和業務輸入信息，提供分析輔助與決策支持，提高人工決策質量和效率。

一站式物聯網互聯系統

隨著物聯網設備的普及，AI數字化時代見證了非結構化數據的激增。為了解決來自單獨連接的個人設備的大量碎片化以及非結構化數據所帶來的挑戰，我們建立了一站式物聯網互聯系統，所有相關的空間、設備、用戶和服務都針對特定的應用場景進行互聯，重新定義了人、事、物之間的有機交互。通過該互聯，碎片化的數據被精簡和組織，使山海大模型能夠利用其中內在的結構化信息的價值。

我們的物聯網互聯系統通過其「感知 — 處理 — 輸出」的三層結構連接物件及傳遞用戶的指令：

- **感知：**我們的物聯網互聯系統可通過雲、邊緣及終端設備訪問，並支持跨所有該等設備的信息協同計算。
- **處理：**在我們的人工智能能力的支持下，通過人工智能引擎的處理，我們的物聯網互聯系統精簡和組織非結構化數據供雲知大腦處理、分析及決策，並有效利用結構化信息的價值。
- **輸出：**在行業級知識圖譜的支持下，我們的物聯網互聯系統可提供系統的專業級信息輸出。

請參閱「— 我們的核心人工智能技術」。

我們的產品組合

我們基於雲知大腦開發產品和解決方案。我們集成芯片、邊緣SDK、公有雲和私有雲部署等多種交付和部署方式，並為開發者和客戶提供基於交互式AI的能力。根據我們基於模型的具有標準化人工智能能力的產品，我們為智慧生活及智慧醫療場景提供廣泛的AI解決方案及產品。

通過我們對人工智能能力的開發和對商業機會的長期探索，我們選擇智慧生活及智慧醫療作為我們的主要目標市場。我們運用雲知大腦開發全方位產品組合，以提供多樣化的產品和解決方案，並提供快速且具有性價比的定制。

我們通過雲知大腦將AI能力打包成面向不同應用場景和垂直行業的客戶或用戶的標準化軟硬件產品和解決方案，以雲、邊緣SDK及芯片的形式提供各種組合。例如，我們的人工智能芯片產品，將交互式AI引擎與自有或第三方芯片硬件相融合，通過雲知大腦提供定制接口。下游解決方案提供商可以通過雲知大腦提供的定制接口定制我們高度標準化的芯片產品。我們提供家電語音喚醒、語音識別及語音合成引擎、車載語音交互降噪、語音識

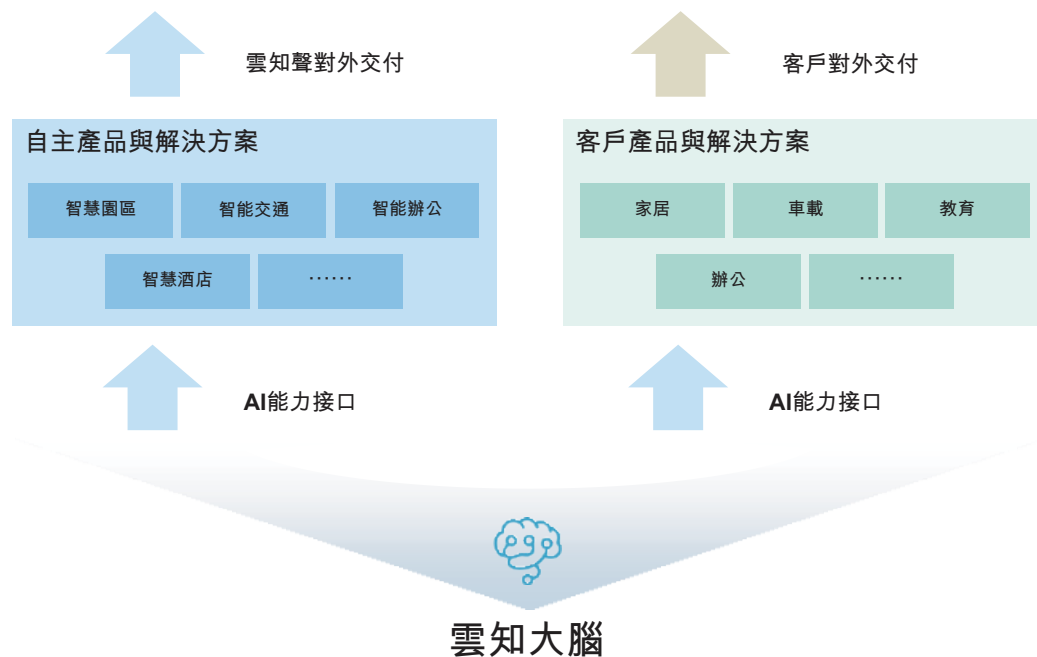
業 務

別及語音合成引擎等邊緣SDK。我們還通過雲知大腦將我們的AI能力（比如語音識別能力和發音評估能力）打包成標準化及可定制的產品，並通過雲服務為下游的解決方案提供商或品牌提供服務。

通過我們靈活的交付和部署方式，我們已顯著降低了基於模型的服務的實施成本，並實現了大規模的出貨量和相當大的市場份額。我們已開始提供並計劃擴展基於MaaS的產品及解決方案。

智慧生活

我們提供由雲知大腦支持、基於山海大模型的解決方案及產品。我們基於山海大模型的解決方案為特定的行業垂直領域打包合適的人工智能引擎工具箱，並通過具性價比的定制滿足大量客戶的需求，而我們基於模型的產品為開發者和中小型企业提供高度標準化的硬件和軟件，並提供直觀的安裝和部署。此外，我們亦為開發者和企業按需提供基於MaaS的AI服務。下圖說明我們基於雲知大腦模型能力支持的智慧生活產品：



基於模型的解決方案

基於雲知大腦提供的人工智能及物聯網能力，我們提供基於模型的智慧生活解決方案，連接智能音箱、集中式語音控制屏、人臉識別設備及多模態交互屏等物聯網設備。我們的智慧生活解決方案主要具備如下特點：

- **數據訓練的飛輪效應：**我們為各行業垂直領域提供智慧生活解決方案，為人工智能技術創造了廣泛的現實應用場景，尤其是交互式AI，通過語音實現了用戶

業 務

與我們平台之間最自然的交互交流方式。我們的智慧生活解決方案促進了各種空間單元(例如單個房間或整個工業園區)內具有不同用戶需求的全場景人機交互。通過雲知大腦處理的物聯網設備持續及全面的用戶反饋，我們創造了一個良性循環，我們的人工智能技術在現實生活場景中不斷得到應用和加強。因此，我們增強的人工智能引擎可以大幅改善用戶體驗，並滲透到人工智能技術不斷擴展的適用場景中。

- **統一管理的全場景覆蓋：**我們智慧生活解決方案通過對設備、人員、流程及運營進行統一管理實現全場景覆蓋，以提升運營效率，改善服務質量。通過在不同行業垂直領域及應用場景的閉環系統中，對各類設備進行統一互聯及管理，我們在同一系統內實現不同地點的多模態物聯網設備之間的交互。通過智能管理工單(覆蓋工單流轉、任務分配、進度管控全流程)統一管理各個場景下不同工作流，我們實現整合後的工作流管理，促進一站式交互，滿足所有用戶需求。基於對物聯網設備及工作流的統一管理，我們能夠集中收集、顯示及分析各項運營數據，大幅提高了用戶體驗和管理效率。
- **分佈式語音控制系統支持的多模態交互：**基於我們的AI及物聯網能力，我們提供多樣化的交互方式，包括語音命令、小程序和智能控制屏幕，促進人與空間的有機交互。就同網絡下相同空間的設備而言，我們的分佈式語音控制系統根據設備信息、激活狀態信息及環境信息，實現對多個設備的單一激活裁決。其允許同一網絡中的所有設備作為一個整體與用戶進行交互，從而改善用戶體驗。結合聲紋識別技術，我們的分佈式語音控制系統保持多個用戶與同一網絡上的設備(不論設備位置)的個別及同時交互。

業 務

通過不同層次的設備和服務的互聯，我們為不同的場景和客戶提供一站式解決方案，包括住宅、商業空間、酒店及軌交，標準化的產品及根據客戶需求的定制。以下代表性使用案例說明我們的解決方案在此類應用場景中的商業化情況：

使用案例1

深圳地鐵20號線

深圳地鐵通過其數字化工作，在其系統及車站採用先進技術。我們與深圳地鐵合作，並自2021年起成為主要技術供應商。我們為深圳地鐵20號線（於2021年推出的全數字化智能地鐵）提供票務系統智能交互解決方案。



在大部分地鐵線路原有的自動售票站的基礎上，我們增加了多種功能，包括無需喚醒的語音輸入，為乘客提供便捷的購票體驗。

業 務



我們已在地鐵20號線的自動售票站安裝語音購票功能，支持涵蓋普通話、英語和粵語的多語言語音識別。考慮到地鐵站的嘈雜背景，我們的交互式AI票務系統亦支持去噪及

業 務

抗干擾功能。一旦有乘客靠近，語音模式將被自動觸發，在我們的語音識別和語義理解技術的支持下，乘客的講話將同時在屏幕上顯示。系統會根據乘客的語音指令，自動彈出推薦的地鐵站。

使用案例2

廈門市

我們與廈門市有深入合作，為其健康步道、集美新城和軟件園三期提供智慧園區解決方案。



我們在廈門健康步道沿途的休閒場所安裝了多個智能交互屏，均能處理交互語音指令，並通過我們的視聽合成技術創建的影像提供智慧引導並推薦周邊設施。我們亦通過安裝在步道沿線景區的攝像頭提供聲控拍照服務，使市民能夠通過我們的語音喚醒及語音識別技術支持的語音指令進行拍照。

在雲知大腦和自有的多模態人工智能引擎的支持下，我們已在集美新城的各個熱點地區安裝智能交互站，為市民提供交互查詢、搜索、指導及建議。我們亦建立了集美新城的知識圖譜，整合了集美的歷史文化、地方法規政策和旅遊景點等區域性知識，為市民和遊客提供本地化的信息。

圍繞人工智能、5G、大數據、物聯網和智能硬件的發展，廈門軟件園三期已成為中國東南地區的重要技術中心。我們在園區的辦公樓、休閒場所、交通站點和主要道路上安裝了多模態的智能交互服務站，由我們與廈門市合作建立的超級計算平台和我們自有的多

業 務

模態人工智能引擎提供支持，集成了包括語音識別、語義理解、語音合成和面部識別等智能交互服務。我們亦為園區內的電梯安裝了交互式AI指令解決方案，為電梯設置定制了去噪和抗干擾算法。

使用案例3

一家金融集團

我們與一家金融集團合作，提供我們的人工智能及智慧物聯網能力，作為其私有化需求的聯合解決方案，涵蓋端到端的全鏈條支持。我們基於私有化提供人工智能解決方案，以滿足金融機構複雜的安全和隱私要求。

在基礎設施層面，我們在兩個月內為客戶建立了超算平台，大幅縮短了客戶自身可能長達一年的建設時間，並涵蓋了多個兼容系統，包括高效存儲和通信、基於Kubernetes的高效任務分配架構、用於大數據機器學習的多GPU加速，以及基於上述平台的通用機器學習結構。我們的超算平台解決方案支持業界多種主流的人工智能計算結構，滿足了人工智能任務的多樣化計算需求。

在應用層面，我們已建立一個中央人工智能平台，為雲平台和終端設備提供支持。我們為終端設備提供交互解決方案，為雲平台提供包括人工智能核心引擎、語音交互支持、業務交互支持、定制優化、大數據平台及運營平台等多維度支持。自人工智能平台建立以來，其為該集團的各個業務部門提供服務，整合了包括醫療和教育在內的常見業務技能。人工智能平台已被用於保險、銀行、行政管理和老年護理等垂直領域。

基於模型的產品

- **人工智能芯片及智能模塊：**我們提供自研自有芯片及從第三方採購的芯片，集成交互式AI算法。人工智能芯片作為標準化的芯片硬件交付，內嵌人工智能引擎，同時我們為客戶提供一系列支持工具來定制芯片的功能和性能調優。基於人工智能芯片，客戶可以在最少的技術支持下定制自己的產品。

除了我們的人工智能芯片，我們還提供嵌入人工智能芯片的智能模塊，為大型物聯網設備製造商提供定制的硬件解決方案。

- **AI能力API：**我們以公有雲AI能力API和定制化自有AI技術服務平台的形式，按需向開發者和企業提供基於MaaS的AI能力。我們以按次付費的方式收取該等服務的訂閱費。

業 務

使用案例4

智慧車載

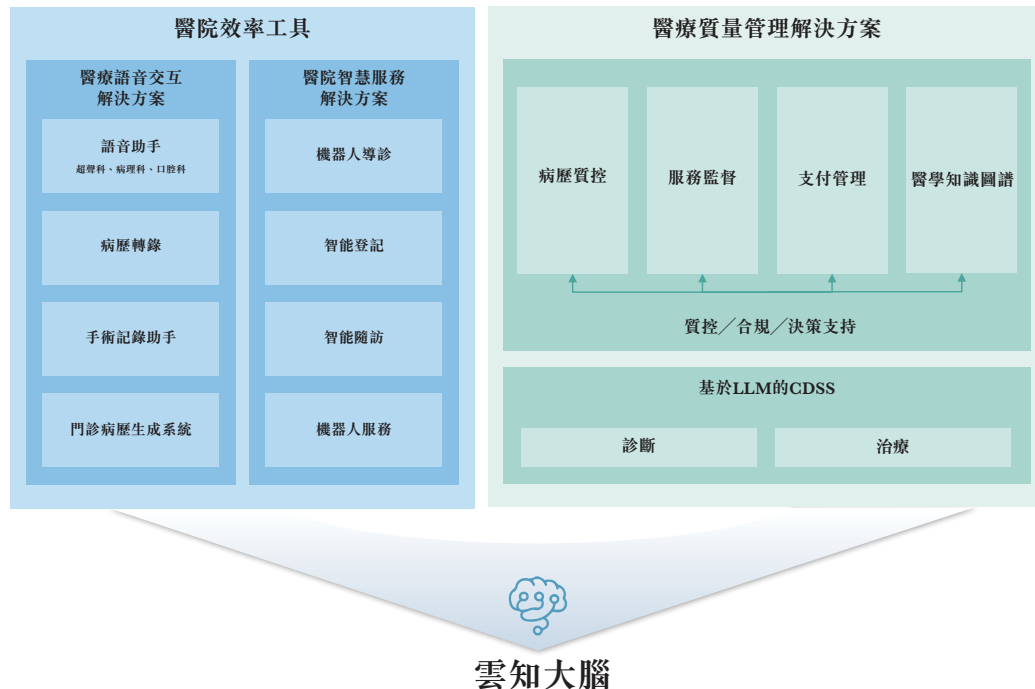
我們的智慧車載產品以智慧車載為核心打造人車生活互聯互通的出行體驗，為車企提供一站式語音交互解決方案，涵蓋語音交互SDK、車規級人工智能芯片及語音雲服務。可對接雲環境，且支持場景化大數據運營服務能力，助力傳統汽車智能升級。

我們首先提供後裝產品，與行車記錄儀、智能後視鏡和車機中控的品牌和製造商合作，整合我們的核心人工智能公有雲服務，包括語音識別、NLU和語音合成，安裝在數以千萬計的交互式AI設備上。

我們亦與ECARX在安徽成立了一家合資公司，專注於前裝產品。我們已為一線汽車製造商的過百萬車輛提供了涵蓋普通話、四川方言、粵語及英語的語音識別服務。

智慧醫療

我們的智慧醫療解決方案從提供效率工具開始，逐步擴展至為相關醫療機構、保險公司及部門提供廣泛的醫療質量管理解決方案，整合各種決策支持產品和質量管理解決方案，滿足相關機構的質量控制和合規要求，最終目標為提高醫務人員的工作效率及提升醫療機構的服務標準。下圖說明我們的智慧醫療產品：



我們構建了智慧醫療解決方案的核心能力，包括數據接入、病歷全文書理解、知識圖譜構建和應用，並將這些能力組合成不同產品投放給目標客戶，迎合其需求和市場反應。

業 務

通過持續使用我們的產品，客戶可以為我們的產品提供反饋，並形成我們解決方案的自我強化發展循環。

醫院效率工具

我們利用我們的交互式AI和大語言模型能力，為醫院提供效率工具。我們的解決方案可大幅提升醫療文書錄入和患者服務的效率。

我們提供以下解決方案以支持醫生及患者：

醫療語音交互解決方案：我們的醫療語言交互解決方案通過病歷轉錄及醫院科室語音助手等效率工具為醫生提供支持。

- **語音助手：**基於我們的AI能力和醫療知識圖譜的智能解讀，我們的語音助手為醫護人員提供便捷的語音交互功能，編製醫學檢驗報告。例如，其通過語義理解自動過濾與檢查內容無關的信息，並可直接與PACS對接，從而支持在採集和選擇圖像的同時語音輸入報告內容。我們為超聲科室提供語音識別模型，識別更準確，針對專科的專業術語和符號進行定制。
- **病歷轉錄：**我們的病歷轉錄系統通過我們的交互式能力，為醫務人員提供語音病歷錄入方法。其大幅提高了病歷錄入的效率，減輕醫護人員的工作負擔。該系統能夠為門診和住院治療提供實時病歷錄入，並能生成醫療檢查和檢測報告。其支持一套全面的功能，包括語音喚醒、智能標點和專業詞彙定制，以促進轉寫，其高準確率超過98%。此外，其亦能實現結構化醫療記錄，支持後續的數據分析和知識圖譜生成。

使用案例5

我們為北京協和醫院提供語音電子病歷系統，為全院40多個臨床和醫技科室提供分場景支持。語音識別顯著提高輸入速度，錄入準確率在98%以上，該系統尤其在神經科、免疫內科、血液科、普通內科等對病歷準確、細緻要求較高的科室，服務效果非常出色，醫護人員能夠花更多時間與病人溝通。

醫院智能服務解決方案：我們的醫院智能服務解決方案實現了全面的患者服務，包括機器人導診、智能預診和智能隨訪在內的診前、診中、診後，全方位智能服務。

- **機器人導診：**導醫機器人整合了傳統的醫院導診流程，為患者提供一站式諮詢服務。其與患者交互，詢問症狀，並根據患者的具體問題引導他們到相關科室。我們亦擬利用山海大模型能力升級機器人服務，實現治療流程的全數字化覆蓋。

業 務

- **智能預診。**通過我們的交互式AI和預診系統，我們通過與患者對話，進行病史採集，為醫生提供患者病情摘要，從而提升門診效率。
- **智能隨訪：**我們的智能隨訪系統對出院後的患者安排及進行自動的治療後隨訪。其收集並記錄患者的健康狀況，發送複診提醒。該系統還為患者提供健康指導，包括出院後的醫療、護理和康復措施。

我們亦在山海大模型的基礎上開發以下產品：

- **門診病歷生成系統：**我們正在開發基於臨床醫患對話的門診病歷生成系統，該系統進行實時語音識別和角色識別以轉錄對話。基於山海大模型的分析、理解和生成能力，臨床病歷生成系統支持關鍵信息匯總、醫學術語標準化和臨床電子病歷的自動生成。
- **手術記錄助手：**我們正在開發手術記錄助手，該助手可以實時記錄醫生在手術室的對話，並通過進一步的語音識別及角色辨認來完成轉錄過程。基於山海大模型，手術記錄助手可以進行無關語音過濾、關鍵信息匯總，並自動完成手術記錄。

醫療質量管理

我們為醫院及保險公司提供廣泛的醫療質量管理解決方案，通過人工智能賦能，提高醫院的服務質量，確保遵守相關法規。利用我們在醫療領域的經驗，我們將人工智能能力與相關機構的信息系統相結合，實現對服務質量的全面管理。我們的病歷質控解決方案應用AI能力來理解病歷和其他醫療數據的內容，分析醫生的決策，並進行缺陷檢查和警報。該系統提高了病歷的質量、醫療服務質量、關鍵科室的專業表現、對醫療保險法規的遵守及成本控制。我們亦通過提供服務監管解決方案，幫助醫院提供更優質、更安全的治療服務，同時滿足相應監管要求。我們為醫院及保險公司提供支付管理解決方案，以管理彼等的醫保支付。此外，我們亦提供由醫學知識圖譜提供支持的創新應用。

病歷質控平台

利用我們的NLP技術和醫學知識圖譜，我們提供智能病歷質控平台，支持醫護人員的病歷審核、決策和病例總結，為他們提供便捷和智能的病歷審核和分析工具。我們通過應用AI能力來瞭解病歷並進行缺陷篩查，以解決病歷中存在的缺陷。

業 務



我們的智能病歷質控平台通過智能交叉檢查和驗證病歷的合規性要求，從準確性、完整性和及時性的角度提供反饋。該平台還能幫助專業人員根據有關疾病的知識圖譜識別病歷中的衝突和缺失信息，使專業人員能夠及時糾正和修正。

此外，智能病歷質控平台還能協助改善內部質控工作流程。其通過評估病歷的完整性、主要診斷的正確率、缺陷數量和其他績效數據，自動對病歷質量進行統計分析，為質量控制政策的制定和改進提供決策支持。

使用案例6

我們為北京友誼醫院提供智能病歷質控系統，幫助醫生全量的篩查病歷缺陷，規範病歷書寫風格，提高病歷質量，提高醫療質量和醫療安全。我們的系統在患者住院期間對病歷進行環節質控，並在患者出院後對病歷進行終末質控，同時自動對質量給予評分評級。

業 務

醫院出院患者數量巨大，我們的系統覆蓋醫生質控工作全過程，從全文書閱讀、檢查缺陷，到針對性地核實系統自動篩查出來的缺陷，大大提高了該項工作的效率和效果。病歷質控覆蓋度從2-5%提高至100%，檢出缺陷的查準率達到90%以上，查全率達到85%以上。

服務監管

我們為單病種質控工作提供解決方案，包括數據上報、過程管理和指標評價。其可以智能地審查數據格式，檢測出超出範圍的異常數據值。該系統幫助醫院針對不同類型的疾病進行各種數據分析，滿足管理需求，實現覆蓋治療全過程的全面質量控制。



支付管理

醫保監管平台

社會醫保基金監管平台以國家、省、市醫療衛生相關政策文件、臨床知識庫、醫保檔案專家評審結果為基礎，開發醫療規則數據庫。其應用人工智能技術對醫療機構上傳的文件進行自動審核，準確識別存在騙保及違規操作，包括但不限於過度醫療、重複檢查、濫用藥物等問題的文件。該平台支持流程優化和風險控制審核，提高基金效率，降低基金風險，提升服務質量。

醫院智慧醫保支付管理系統

我們的智慧醫保支付管理系統為醫院維持診斷、治療和收費規定的合規性提供決策支持。通過對醫療和支付記錄的智能分析，該系統自動審核治療和支付過程，檢查治療和

業 務

收費的一致性，並幫助醫院和醫生識別不合規問題。此類問題通常包括重複和不合規的用藥、過度治療、捆綁銷售以及無適應症的用藥、治療和檢查。

商保監管與保險公司稽核系統

在醫保基金監管和審計積累的最佳實踐基礎上，我們延伸到商業保險的其他業務領域，如車險人傷的監管和審計。我們利用自然語言理解、知識圖譜和其他技術，為商業保險公司提供保單理賠的審計和統計分析的智能系統。

醫學知識圖譜

利用我們的深度學習能力，我們根據國際醫學術語集，在醫學概念層次的基礎上建立了醫療行業的知識圖譜結構。我們利用從權威醫療文本（如臨床指南和藥品說明書）中提取的信息，結合臨床實踐知識，建立知識圖譜，並在概念和實體之間建立聯繫。

我們亦在山海大模型的基礎上開發以下在研產品：

- **基於大語言模型的現有產品升級：**藉助大語言模型技術在數據理解、提取、標準化、生成等方面的卓越性能，我們可以大幅提升現有產品的數據適配、智能分析、用戶交互能力。
- **基於大語言模型的CDSS產品升級：**基於山海大模型理解和分析多模態醫學數據的能力，在臨床知識圖譜的進一步支持下，我們正在開發我們的CDSS產品，該產品整合病歷文件質量控制、醫療服務質量控制和醫保費用監管，用戶界面與醫療質量管理解決方案相同，為醫生和醫院管理者提供全面統一的決策支持系統。

ATLAS人工智能基礎設施

我們的Atlas人工智能基礎設施通過大規模機器學習和大數據處理能力支持我們的研發。其由其(i)智算集群；(ii)機器學習算法及；(iii)以數據為中心的模型訓練方法支持。

- **智算集群：**我們的智算集群為大數據並行模型訓練提供計算能力的統一管理及動態調度。我們智算集群計算能力的效率及靈活的可擴展性構成我們推進人工智能算法、模型及解決方案方面的核心競爭優勢，尤其是在大語言模型方面。目前，總計算能力超過200PFLOPS，存儲容量超過1.6PB，可隨時不間斷擴展。
- **機器學習算法：**利用智算集群，我們可以探索語音、語言、知識及視覺模型的前瞻機器學習算法，並在標準基準及大規模應用數據上評估其性能。在完成新

業 務

算法的驗證後，我們將快速升級我們的人工智能引擎、訓練工具及相關模型，以保持我們的技術領先優勢。我們的機器學習算法是我們生產高精度人工智能模型及可優化人工智能引擎的核心方法。

- **以數據為中心的模型訓練法：**通過先進的機器學習算法及高效的智算集群，我們已構建以數據為中心的模型訓練平台，賦能行業級應用，進行快速模型迭代及微調。該平台對於人工智能模型的高效演進至關重要，尤其是大語言模型，其並行計算需求巨大，需要大規模高質量訓練數據以進行訓練。我們通過數據、模型和流水線並行計算來加速訓練過程，將硬件的理論效率提高到50%以上(相比之下，沒有流水線並行化的效率少於40%)。該平台亦提供重要且高效的10TB規模數據的去重、去噪及選擇，以提高訓練數據的質量及多樣性。

我們的核心人工智能技術

基於Atlas人工智能基礎設施，我們構建了以山海大模型為核心的完整人工智能核心技術體系雲知大腦，涵蓋從信號感知，到語義理解、知識圖譜、推理和決策，最後生成響應信息的人機交互閉環交互。在交互式AI核心技術的基礎上，我們拓展了人臉識別、OCR、視聽合成(AVS)等多模態感知和生成能力，持續提升人工智能核心技術。

大語言模型(LLM)

我們的大語言模型山海大模型是雲知大腦的核心，決定了該系統的認知和業務處理能力，是我們MaaS商業模型的關鍵。大語言模型通過無監督預訓練、指令調優和基於人類反饋的強化學習，構建了強大的通用基礎能力。在插件擴展、領域增強及定制支持下，我們促進了大語言模型在垂直行業的應用落地。

- **構建通用功能：**這是邁向AGI新範式的關鍵一步。通過海量文本和代碼數據的無監督預訓練，以及指令調優和基於人類反饋的強化學習技術，我們的大語言模型能夠進行語言生成、語義理解、知識問答、邏輯推理、數學計算和代碼生成，靈活應對各種任務，實現類人AGI能力。
 - **無監督模型預訓練：**基於GPT架構，我們建立了一個具有萬億級token和600億參數的GPT模型訓練框架。該框架支持併行訓練加速，包括數據、模型和流水線的並行機制。我們建立了一個看門狗系統，用於檢查模型訓練的運行狀態和健康狀況。通過海量文本和代碼數據的無監督預訓練，我們訓練出了一個強大的通用基礎模型，為後續的指令調優和基於人類反饋的強化學習奠定了基礎。

業 務

- **指令調優**：通過引入指令微調技術，大語言模型可以理解人類意圖並處理各種語言相關任務，包括語言生成、語義理解、知識問答、邏輯推理、數學計算和代碼生成，構建具有通用功能的有用且強大的模型。其亦支持LoRA、Adaptor等高效參數調優機制，可以高速進行指令調優，從而迅速擴展模型能力。
- **RLHF**：引入基於人類反饋的強化學習(RLHF)技術使大語言模型的生成結果盡可能有用、真實、無害，從而實現與人類價值觀的契合。首先，根據人類對模型生成結果的比較排序，訓練獎勵模型以適應人類偏好。然後，使用強化學習框架，將模型參數設置為策略，並使用PPO(近端策略優化)算法進行多次迭代，以優化模型參數，增強模型的安全合規性。
- **專業能力提升**：在通用大語言模型的基礎上，面對特定行業的專業問題，我們亦引入插件擴展、行業增強和定制，以有效整合實時信息、行業知識和企業數據，減少大語言模型在及時性、準確性和專業性方面的不足，並增強雲知大腦解決專業問題的能力。
 - **插件擴展**：大語言模型本身並非萬能。其缺乏實時信息捕獲能力、某些專業工具和第三方服務，例如進行複雜計算、訂購外賣和完成網上購物支付的能力。我們提供插件擴展框架，並開發了多個內置插件，如計算器、天氣查詢、鄰里查詢和網絡搜索插件。插件擴展能力允許大語言模型訪問實時信息，支持模型充當大腦中樞訪問外部工具，實現特定任務的規劃、協作和完成，亦可以訪問基於知識的插件來減少模型幻覺。
 - **行業增強**：將大語言模型應用於特定行業時，我們會進行領域增強，以優化其域內性能。我們有一套領域提升的方法，包括(i)在預訓練階段增加對領域特定高質量文本數據的增量訓練；(ii)在微調階段增加與行業特定任務相對應的微調數據；及(iii)使用插件擴展框架，基於行業知識圖譜提供多個基於知識的插件，並將其嵌入大語言模型，可以大幅減少大語言模型產生的錯誤響應，顯著提高大語言模型的信息質量和專業能力。基於該方法，我們正在為醫療、金融和法律行業訓練特定行業的大語言模型。

業 務

- **客戶定制：**我們可以為客戶提供多層次的大語言模型定制服務，包括(i)接入客戶網絡，構建企業級交互式搜索引擎；(ii)使用插件擴展將客戶的內部私有數據和服務封裝為插件，允許大語言模型集成客戶的內部數據和服務API；(iii)使用客戶積累的面向任務的註釋數據為大語言模型提供微調API，使客戶可以對大模型進行定制微調，從而提高效率；及(iv)基於大語言模型構建和管理平台，我們為客戶提供從數據選擇、模型訓練到模型部署、運營和維護的一站式服務，為客戶提供所有大語言模型功能。

行業級知識圖譜

行業知識圖譜是認知智能的核心組成部分，與山海大模型結合可以有效提高行業知識的準確性和可靠性，解決行業問題。我們建立了行業級的知識圖譜管理平台，支持知識圖譜的全生命週期管理，包括構建、融合、存儲、服務和交互。我們通過從我們處理的大量非結構化數據(如語音形式)中提取關鍵行業概念及技術詞彙，通過三個步驟轉化為行業特定的知識及見解，從而構建行業級知識圖譜：

- **標準化：**我們的知識圖譜通過反覆提取類似及高頻的行業概念及技術術語，實現自動分類及識別標準概念，並自動糾正誤讀及排版。一旦建立並記錄於知識圖譜中，標準化的行業詞彙可用於識別不同情景中的各種方言語音及手寫文字中具有相同或類似涵義的特定熱門詞彙。我們向特定行業提供的服務越多，我們處理的數據越多，我們的知識圖譜越敏感及準確，便可根據其建立的標準化概念識別專業言論。
- **分層：**通過標準化概念的行業級數據庫，我們的知識圖譜進一步提供分層和樹狀結構的概念目錄，以進一步明確不同的概念層級，並梳理同一組別內術語之間的關係。例如，肝炎和肝癌均屬於較高級別的肝病類別，而甲型肝炎和肝細胞癌分別低於肝炎和肝癌類別中的一級別。通過對分層概念和術語的清晰劃分，我們的知識圖譜為特定行業和場景建立結構化百科全書。
- **聯繫：**知識圖譜基於標準化和分層的積累成果，進一步支持聯繫的關鍵步驟，利用專業術語數據庫和百科全書。知識圖譜將不同知識庫下的不同概念連接到同一行業內不同的實踐領域，以及自動創建或總結結構化言語或文字的方式，支持對長篇及難明言語的混合概念論文的理解。

業 務

多模態感知及生成

我們的多模態感知和生成涵蓋了從感知外部信號開始到回應產生信號的溝通鏈，促進了我們交互式AI核心技術所支持的人工智能溝通方法。除了我們的核心交互式人工智能技術，我們亦部署如面部識別和OCR（光學字符識別）及視聽合成(AVS)的多模態能力。

- **多模態感知**

- **語音信號處理**：基於陣列和深度學習技術，我們的信號處理技術提供了不同場景下的高性能人機語音交互和通話降噪，涵蓋各種不同數量麥克風的遠場及近場陣列處理解決方案實現回聲消除、語音測向、空域濾波及混響抑制。通過將語音信號從各種噪聲干擾中分離，提升聽感效果及語音識別準確性。
- **語音喚醒及控制**：語音喚醒是智能設備啟動語音交互的第一步。我們採用基於深度學習的模型及快速定制方案以優化模型結構，利用涵蓋不同年齡、性別、口音、語速、環境噪聲及其他複雜條件的訓練數據，提高模型的準確性及穩健性。使用置信度的決策算法，我們的語音喚醒技術具有高喚醒率及低誤警率。通過模型壓縮、量化等工程優化，顯著降低了資源及能耗，使我們的語音喚醒引擎可於各種低功耗設備上運行，大幅擴展語音交互的使用範圍。
- **語音及說話人識別**：作為人機交互系統的一個重要組成部分，語音及說話人識別技術將語音轉換為文本及說話人身份。基於端到端Transformer及混合架構語音識別技術，通過大規模語音和文本訓練，我們自有的語音識別系統準確率在普通場景下識別率可達98%。與我們強大的SSP降噪前端結合，並通過多場景和任務相關數據訓練，顯著提升識別系統在複雜場景下的穩定性。同時，通過說話人切分和聲紋認證技術，我們可以高精度地識別及分離語音流中的說話人。
- **語音分析及多模態能力擴展**：除語音處理及識別技術外，我們已開發用於口語學習場景發音評估的語音分析技術。我們亦開發了面向多模態個性化人機交互的人臉識別技術，以及面向各種打印文檔信息處理的OCR技術，其廣泛應用於智慧醫療解決方案。

- **多模態生成**

- **文本轉語音**：我們的文本轉語音技術將原輸入文本轉化為自然、高質量及個性化的語音聲波。基於深度學習和端到端的語音合成技術，我們的文本轉語音技術可實時將文本準確地轉換為清晰、自然、流暢的語音，MOS分可

業 務

達4.2。我們以快捷簡便的方式建立需要最少人工參與的文本轉語音系統。我們的個性化語音合成方法利用目標說話人的少量語音數據集對現有的源模型進行轉換和調整，從而使調整後的源模型包含目標說話人的音色。我們提供離線和在線合成(有多種音色可選)及基於由20–300個句子組成的數據庫的定制個性化音色的選項。

- 多模態視聽合成：我們的視聽合成技術在我們文本轉語音功能的基礎上開發，創造出的虛擬化身可促進智能人機交互，涵蓋自然語言面部表情、手勢、姿勢和眼神。通過提供具有不同風格和個性的定制虛擬化身，我們增強了用戶的交互體驗，並為視頻創作提供智能輔助功能。

一站式物聯網互聯繫統

我們已建立獨特的一個物聯網一站式互連系統，將所有相關空間、設備、用戶及服務互聯，用於特定的應用場景，重新定義人、事、物之間的有機互動。

- 園區：園區設置所有活動的舞台，供人們與不同的物體(包括物聯網設備)互動。我們的物聯網連接及支援系統能夠與所有適用物聯網設備在不同的空間環境(從會議室、一家醫院到整個工業園的層面)進行互聯。憑藉我們高度兼容的系統接入及連接數以百計或數以千計的物聯網設備，我們將服務無縫地結合到空間中，以我們的AI能力為基礎，以交互式AI為核心，促進自然人機交互。
- 設備：物聯網設備以大量非結構化數據的形式不斷產生有價值的信息，這需要準確及時地分解。我們的物聯網連接及支援系統與各種物聯網設備兼容，涵蓋廣泛的手提設備(如麥克風、手機及揚聲器)以至大型及複雜的設施(包括家用電器、電梯及醫療設備)。
- 用戶：我們的物聯網連接及支援系統致力為所有於我們提供服務的空間內與物聯網設備互動的人士提供無縫及直觀的用戶體驗。為盡可能取代體力勞動，我們精簡及簡化輸入說明的程序，為用戶提供一個放鬆的互動環境。我們輸出我們的交互式AI能力，使語音指令(人類最自然及直觀的指令方法)可作為相關物聯網設備的主要指令輸入，以創造難以超越的用戶體驗。
- 服務：結合我們的AI能力及行業規模知識圖譜，我們的物聯網連接及支援系統是我們多維度服務的支柱。通過整合所有不同應用場景的物聯網設備，我們的

業 務

物聯網連接及支援系統從該等設備收集非結構化數據，並通過我們的AI引擎提供該等數據，支援我們針對特定行業及應用場景定制的行業規模知識圖譜。

研發

我們開發新技術、設計新產品和解決方案以及增強現有產品和解決方案的能力，對於保持我們的市場地位至關重要。截至2022年12月31日，我們的研發團隊由281人組成，約佔我們僱員總數的70.5%。我們於2020年、2021年及2022年分別產生人民幣189.3百萬元、人民幣188.2百萬元及人民幣287.1百萬元的研發費用，佔我們同期總收入的72.6%、41.3%及47.8%。

發展

我們計劃將研究重點放在以下技術領域：

*Atlas*人工智能基礎設施

- **智算集群**：不斷優化調度效率，提高平台計算、存儲和通信能力，為彈性計算需求提供高效支持；
- **機器學習算法**：跟隨機器學習算法的最新進展，在大規模行業應用成功測試和實施後，對我們的主要技術進行升級；及
- **以數據為中心的訓練方法**：優化新算法計算框架和行業數據處理工具，提高與更多機器學習算法的兼容性，高效優化目標主要模型。

大語言模型(LLM)

跟蹤多模態大語言模型的前沿技術，以及分析、推理和決策，不斷提高大語言模型的能力，降低平台在行業應用上的成本，並提高行業知識和數據滾動的效率。

知識圖譜

潛心研究智能醫療應用，解決更多的基礎性問題，反向推理核心構造分析與知識圖譜決策技術演進，提高人機協作解決專業問題的質量及效率。

多模態感知及生成

- **多模態感知技術**：不斷提高模塊在速度、準確度、穩定性及成本效益等方面的性能，通過技術的優勢，提高數據生成及模擬的質量，提高訓練效果及效率，同時，拓展多語種、多模態的感知能力，為理解提供更為豐富有效的感知信息

業 務

- **多模態生成技術**：改進語言生成和歌聲合成質量以及視聽合成技術，實現從準確的信息傳達到豐富多模態、情感化、個性化效果等表達能力的創意水平要求

研發合作

我們與學術機構在研發、技術創新以及吸引和培養人工智能人才方面進行合作，以支持我們的快速增長。我們於2020年8月與中國科學院自動化研究所（「**CASIA**」）簽訂了框架合作協議（「**框架協議**」），據此，我們同意在以下研究方面進行合作：(i)根據醫療需求構建知識圖譜，以及構建知識圖譜所需的關鍵技術及工具；(ii)基於醫學自然語言處理及知識圖譜技術的病歷和專有醫療系統的質量控制；及(iii)低資源自然語言處理及其在垂直市場中的作用。基於模式識別國家重點實驗室的先進科研水平，**CASIA**為我們研發智能交互產品提供了堅實的技術支撐，增強了我們的競爭力和市場領導力。同時，我們旨在為合作創造良好的研究氛圍，確保研究目標切實可行，反之，研究成果亦為實際應用服務。根據**框架協議**，我們負責為合作項目提供資金，並承擔聯合實驗室的相關維護及運營費用。合作中產生的知識產權的所有權應根據每個項目來確定。

我們與上海師範大學建立了自然人機交互聯合實驗室（「**聯合實驗室**」），專注於改進我們的交互式人工智能技術。我們旨在為學生創造創新平台，支持和鼓勵創新思維和新想法的分享，此將有利於吸引人才，為我們的團隊注入具有活力和高素質的新鮮血液。合作中產生的所有專利由上海師範大學及我們共同擁有，同時該校有權免費使用該等專利。

數據安全和隱私

我們致力於遵守數據隱私法律及保護數據安全。我們主要收集和存儲與開發者及終端用戶有關的數據。

我們獲得了用戶對數據隱私政策條款及條件的承認與同意，該政策描述了我們的數據實踐，包括信息收集和使用、以及我們不會將任何數據用於與用戶的數據隱私政策規定以外的任何目的。特定項目的現場數據處理受到嚴格監控，我們已實施內部項目控制制度，以確保項目參與者嚴格遵守。此外，我們在必要情況下聘請第三方對語音數據進行分析及註釋脫敏數據。我們要求彼等根據服務合同確保數據安全，並不時對其數據處理過程進行檢查。

我們使用各種技術保護我們收集的數據，例如加密、防火牆、漏洞掃描及日誌審核。例如，我們擁有一支專業人員團隊，致力於持續審查及監控數據安全實踐。我們備存記錄所有嘗試及成功調取我們數據的數據存取日誌，並對大數據請求進行自動監控及常規手動驗證。

業 務

我們根據國家法律和法規以及我們的內部數據存儲指南將用戶數據存儲一段時間，該指南規定了我們存儲和銷毀數據的原則及程序。我們對已取消的用戶和已停止服務的用戶的數據進行匿名化和脫敏處理。

充分維護、存儲及保護數據對我們的成功至關重要。我們致力於遵守數據隱私法律，且我們已實施符合行業標準的相關內部程序及控制。我們已實施相關的內部程序及控制措施，以確保我們的IT基礎設施的安全，確保我們的數據得到保護，避免該等數據的洩漏及遺失。於往績記錄期間及直至最後實際可行日期，我們的IT基礎設施並無出現任何重大系統故障，或數據出現任何重大洩漏或遺失。

知識產權

我們認為，我們的自有域名、著作權、商標、商業秘密及其他知識產權對我們的業務運營至關重要。我們依賴專利、著作權、商標、商業秘密法律及披露限制的組合來保護我們的知識產權。截至2022年12月31日，我們於中國擁有288項註冊專利及563項專利申請、235個商標及195項軟件著作權。

請參閱「附錄四 — 法定及一般資料 — 有關我們業務的進一步資料 — 知識產權」。

除進行商標及專利註冊申請以外，我們實施一套全面的措施來保護我們的知識產權。我們已設立知識產權管理部門，並聘請專職人員監督及管理我們的知識產權。我們的員工通常需要簽訂一份標準僱傭合同，其中包括保密條款及確認彼等於受僱期間產生的所有發明、商業秘密、開發及其他過程為我們的財產，並向我們轉移其於有關工作中可能聲稱的任何所有權。我們將積極監控並針對任何未經授權使用我們的知識產權的行為提出索賠。

截至最後實際可行日期，我們並無面臨侵犯第三方於中國的商標、許可證及其他知識產權的任何重大糾紛或申索。

銷售及營銷

我們的銷售網絡

我們通過我們的直接銷售團隊銷售我們的產品及解決方案。

截至2022年12月31日，我們的銷售團隊由64名僱員組成，他們擁有豐富的行業經驗以及對我們基於人工智能的產品和解決方案的深入了解。我們的銷售團隊通常會制定年度及季度銷售計劃，以有效及系統地進行銷售。我們的銷售團隊通常於年底召開會議，確定下一年度的整體年度銷售計劃，包括銷售任務及考核指標，並根據客戶類型、市場狀況、我們涉及的行業的最新變動、競爭對手和下游行業信息以制定更詳細的季度銷售計劃。我們實施嚴格的政策及內部程序，以規範我們的銷售團隊及其在銷售我們的產品及解決方案期間的行為，並確保他們遵守相關法規。例如，我們明確禁止銷售人員擅自提供低於指導價的報價。

業 務

我們的客戶包括我們產品及服務的終端用戶、系統集成商及若干終端用戶(主要為醫院)委聘的代理。部分終端用戶於選擇供應商或服務提供商時，會聘請系統集成商。有關系統集成商通過直接與大量的供應商或服務提供商談判來幫助終端用戶，儘管在大多數情況下，終端用戶還需要批准及確認供應商的選擇，特別是就核心供應商(包括人工智能服務提供商)而言。通常，終端用戶會提出其計劃實現的目標及項目預算並聘請系統集成商，系統集成商將於項目實施過程中提供各種類型的協助，比如就融資計劃提供建議、選擇供應商、管理施工及整合不同供應商的工作成果。根據行業慣例及其採購政策，我們的部分終端用戶在選擇我們作為其供應商後，將委聘採購代理進行詳細的採購計劃及執行。然後，我們會與該等代理訂立合約，訂明我們對終端用戶的直接交付責任。

儘管我們的一部分客戶是系統集成商或代理，而不是終端用戶，但我們認為我們的業務模式並非分銷模式。如上所述，系統集成商及代理並非我們為拓寬銷售渠道而聘請的分銷商；相反，他們是我們的終端用戶為實施其項目而選擇的代理商，而有關選擇哪家服務提供商的最終決定主要由終端用戶作出。無論我們的合約是直接與我們的終端用戶還是與系統集成商或代理簽訂，在合約條款及服務範圍方面並無重大差異。當我們與系統集成商或代理簽訂合約時，我們將該系統集成商或代理(而非相關的終端用戶)確認為我們的客戶。就此而言，我們認為系統集成商或代理並非我們的分銷商，且認為其作為我們的直接客戶的參與不會引起與庫存風險、拆分或應收賬款的可回收性有關的任何擔憂。

定價

我們已為產品及解決方案開發不同定價政策。具體而言：

- **智慧生活：**我們一般會向客戶提供智慧生活解決方案的菜單式報價。每個解決方案的報價通常取決於(其中包括)供求關係、營銷定位及競爭解決方案的價格。客戶可根據自己的具體需求及預算選擇合適的解決方案。我們結合選定的解決方案，為客戶進行適當的定制修改。我們通常根據客戶選擇的解決方案的總單價及定製程度來確定定制解決方案的最終價格。
- **智慧醫療：**我們通常會考慮各種因素來確定我們的智慧醫療解決方案的價格，包括市場定位、市場情況、政策要求、客戶類型、競爭對手的定價及定制水平。對於燈塔客戶，我們通常會通過提供具有高級功能的解決方案。為進一步吸引我們已滲透的垂直領域或應用場景的客戶，我們提供具有通用功能的解決方案，以降低解決方案的整體價格。此外，我們為部分燈塔客戶提供定制服務，並收取額外的定制服務費用。

業 務

營 銷

我們相信最有效的營銷方式是不斷增強我們解決客戶未被滿足的需求的能力，並突出我們解決方案的有效性，以提高品牌知名度並推廣我們的產品及解決方案。我們已成立一個銷售及營銷部門，負責制定營銷計劃、管理銷售團隊、分析產品及解決方案的定位、培訓銷售團隊、維護現有客戶及開發新客戶以及定價。

我們的銷售團隊負責與潛在客戶進行深入溝通，通過展示優秀案例以顯示產品及解決方案的專業性，讓潛在客戶充分了解及認可我們的能力。我們亦非常重視積極維護現有客戶關係，通過現有客戶的認可，在客戶所涉及的行業進行低成本推廣，從而在場景中不斷複製，實現更廣泛的銷售。

客 戶

我們擁有龐大的客戶群，不依賴任何單一客戶。於2020年、2021年及2022年，來自我們的最大客戶的收入分別佔我們於該等期間的總收入的20.6%、21.6%及13.1%。於2020年、2021年及2022年，來自我們五個最大客戶的收入分別佔我們同期的總收入的40.1%、34.4%及30.8%。請參閱「業務 — 客戶」。

下表載列於往績記錄期間我們五大客戶的詳情：

截至2020年12月31日止年度

客戶	所售產品 及服務	客戶背景	開始業務 關係的年份	銷售金額 人民幣千元	佔總收入的 百分比
公司J	智慧生活	從事物業開發、商業物業經營、物業管理及酒店經營	2019年	53,845	20.6%
客戶A	智慧生活	從事物業開發	2020年	30,481	11.7%
客戶B	智慧生活	從事保險相關服務	2016年	8,469	3.2%
客戶C	智慧生活	從事信息技術諮詢服務、軟件開發及信息系統集成服務	2018年	6,159	2.4%
客戶D	智慧醫療	從事醫療器械銷售	2020年	5,752	2.2%
總計				104,706	40.1%

業 務

截至2021年12月31日止年度

客戶	所售產品 及服務	客戶背景	開始業務 關係的年份	銷售金額 人民幣千元	佔總收入的 百分比
公司J	智慧生活	從事物業開發、商業物業經營、物業管理及酒店經營	2019年	98,506	21.6%
客戶E	智慧生活	從事其他技術推廣服務、其他機械設備及電子產品批發、電器設備批發	2021年	25,398	5.6%
客戶F	智慧生活	從事電子產品及配件製造、電子產品銷售、電子產品及配件、計算機應用電子設備	2020年	12,027	2.7%
客戶G	智慧生活	從事智能建築系統設計服務	2021年	10,560	2.3%
客戶H	智慧生活	從事城市基礎設施投資建設、區域土地開發儲備管理及房地產經營業務	2021年	10,161	2.2%
總計				156,652	34.4%

業 務

截至2022年12月31日止年度

客戶	所售產品及服務	客戶背景	開始業務關係的年份	銷售金額	佔總收入的百分比
				人民幣千元	
公司J	智慧生活	從事物業開發、商業物業經營、物業管理及酒店經營	2019年	78,518	13.1%
客戶I	智慧生活	從事施工總承包、軟件開發、計算機系統服務、軌道交通通信信號系統集成服務和城市軌道交通裝備製造	2022年	51,701	8.6%
客戶J	智慧生活	從事電子儀器及辦公設備的銷售	2020年	20,936	3.5%
客戶K	智慧生活	從事軟件開發、信息系統集成服務、信息技術諮詢服務、集成電路設計、電氣設備批發、其他機械設備及電子產品批發	2021年	16,960	2.8%
客戶L	智慧生活	從事信息技術服務和軟件開發	2021年	16,955	2.8%
總計				185,070	30.8%

於往績記錄期間，我們與公司J有廣泛合作。我們為公司J提供智慧生活解決方案，包括智能住宅和酒店解決方案。董事確認，我們向公司J的所有銷售乃在正常業務過程中按一般商業條款及公平原則進行。

截至最後實際可行日期，概無我們的董事、其聯繫人或我們的任何股東（擁有或據董事所知擁有我們已發行股本的5%以上）於我們的任何五大客戶中擁有任何權益。

供應商

我們的供應商包括信譽良好的硬件組件、軟件、研發及營銷服務供應商。

為確保我們產品的質量，我們嚴格選擇供應商，採購並優先考慮符合行業標準的材料。我們在選擇時考慮的因素包括技術專長、產品質量、資格及證書、市場聲譽及成本競爭力。我們亦在收到使用我們設計的語音模塊組件後進行隨機抽查。我們亦對現有供應商進

業 務

行年度審查，並相應地更新彼等的詳細資料。通常，我們與供應商訂立一次性採購協議或框架採購協議。根據框架採購協議，每筆採購均單獨下達採購訂單。一般而言，我們的採購按所提供報價中規定的固定單價進行，而我們的芯片供應商要求我們全額預付協議價。我們的供應商通常負責自費安排送貨給我們。我們的供應商通常授予我們一年至三年的保修期，我們有權退回或更換有缺陷的供應品。

我們的主要供應商

我們的供應商包括信譽良好的硬件組件、軟件、研發及營銷服務供應商。於2020年、2021年及2022年，來自我們最大供應商的費用分別佔我們於該等期間的採購總額的8.5%、9.9%及20.9%。於2020年、2021年及2022年，來自我們五大供應商的費用分別佔我們同期的採購總額的25.5%、28.1%及46.2%。請參閱「業務 — 供應商」。

下表載列於往績記錄期間我們五大供應商的詳情：

截至2020年12月31日止年度

供應商	所採購的 產品及服務	供應商背景	開始業務 關係的年份	採購金額 人民幣千元	佔採購總額 的百分比
公司J	社區智能軟 硬件系統	從事物業開發、商 業物業經營、物業 管理及酒店經營	2019年	24,476	8.5%
供應商A	技術開發服 務及硬件	從事計算機軟硬件 技術開發、人工智 能技術服務、零售 交通車輛及地面產 品的研發與銷售	2020年	16,864	5.8%
供應商B	技術開發服 務、軟硬件 系統擴展與 集成	從事辦公設備軟 硬件、智能控制產 品、音視頻信息終 端產品的開發與銷 售	2019年	11,733	4.1%
供應商C	技術開發服 務、展廳	從事人工智能基礎 軟件開發和公共數 據平台建設，以及 工程技術和設計服 務	2019年	11,191	3.9%
供應商D	法律服務	服務行業公司，提 供法律服務	2019年	9,194	3.2%
總計				73,458	25.5%

業 務

截至2021年12月31日止年度

供應商	所採購的 產品及服務	供應商背景	開始業務 關係的年份	採購金額 人民幣千元	佔採購總額 的百分比
供應商A	技術開發服務及智慧屏 終端硬件	從事計算機軟硬件技術開發、人工智能技術服務、軌道交通車輛及地面產品的研發與銷售	2020年	46,990	9.9%
供應商C.....	技術開發服務、展廳	從事人工智能基礎軟件開發和公共數據平台建設，以及工程技術和設計服務	2019年	32,645	6.9%
供應商E.....	硬件	從事集成電路及相關產品的研發、設計、銷售	2019年	21,036	4.5%
供應商F.....	智能空調系統	從事空調設備專業技術領域的技術諮詢服務、人工技術服務和機電設備銷售	2021年	18,284	3.9%
供應商G	技術開發服務	從事軟件開發、數字內容製作服務、計算機軟件、硬件及輔助設備的批發	2020年	13,878	2.9%
總計				132,833	28.1%

業 務

截至2022年12月31日止年度

供應商	所採購的 產品及服務	供應商背景	開始業務 關係的年份	採購金額 人民幣千元	佔採購總額 的百分比
供應商A	技術開發服務及智慧屏 終端硬件	從事計算機軟硬件技術開發、人工智能技術服務、軌道交通車輛及地面產品的研發與銷售	2020年	106,991	20.8%
供應商C.....	技術開發服務、展廳	從事人工智能基礎軟件開發和公共數據平台建設，以及工程技術和設計服務	2019年	49,655	9.7%
供應商B.....	技術開發服務、軟硬件系統擴展與集成	從事辦公設備軟硬件、智能控制產品、音視頻信息終端產品的開發與銷售	2019年	28,553	5.6%
供應商H	技術開發服務，交互式AI開發	從事信息技術和技術服務	2021年	27,246	5.3%
供應商E.....	硬件	從事集成電路及相關產品的研發、設計、銷售	2019年	24,497	4.8%
總計				236,942	46.2%

截至最後實際可行日期，概無我們的董事、其聯繫人或我們的任何股東（擁有或據董事所知擁有我們已發行股本的5%以上）於我們的五大供應商中擁有任何權益。

客戶和供應商重疊

於2020年，公司J是我們五大客戶之一，亦是我們五大供應商之一。我們向公司J提供智慧生活解決方案，包括智能住宅和酒店解決方案，並向其採購社區智能軟硬件系統。於2020年、2021年及2022年，我們來自公司J的收入分別為人民幣53.8百萬元、人民幣98.5百萬元及人民幣101.2百萬元。

業 務

元及人民幣78.5百萬元，分別佔我們總收入的20.6%、21.6%及13.1%。於2020年，我們向公司J的採購金額為人民幣24.5百萬元，佔採購總額的8.5%。除本節所披露者外，於往績記錄期間，我們的五大供應商概無同時為我們的五大客戶。董事確認，我們向公司J的所有銷售以及向公司J的採購均在正常業務過程中按照一般商業條款及公平原則進行。

競爭

我們所處的人工智能解決方案行業競爭激烈。我們主要與其他專注於開發及商業化人工智能產品和解決方案的公司競爭。就我們所進入的各行業垂直領域而言，我們亦與傳統解決方案競爭，該等供應商在該垂直領域不以人工智能為導向。我們在未來亦可能面臨來自新進入者的競爭，此將提高競爭程度。例如，擁有大量財政資源、尖端技術能力及廣泛銷售渠道的更為成熟的技術公司可能會開發與我們直接競爭的解決方案。請參閱「行業概覽」。

物流及庫存管理

物流及倉庫

一般而言，我們的供應商負責將硬件組件和所製造產品送達我們指定的倉庫。根據與客戶簽訂的協議，部分硬件產品由供應商直接運送至客戶指定的地點，物流成本由供應商承擔。據我們所知，所有物流服務供應商均為獨立第三方。

庫存管理

我們的庫存主要包括原材料、零部件及待交付項目。我們繼續監測我們的庫存，尤其是我們的庫存周轉天數及庫齡，以使其保持在足以滿足我們客戶訂單的水平。截至2020年、2021年及2022年12月31日，我們的存貨分別為人民幣19.2百萬元、人民幣46.9百萬元及人民幣33.6百萬元。我們已制定庫存控制政策以監控我們的庫存水平並盡量減少過時庫存。通過按需從供應商處採購成品、原材料及零部件，我們能夠減少庫存並降低我們的庫存風險。然而，為避免任何供應短缺，我們可能會策略性地為若干關鍵硬件組件及原材料保持較高水平的庫存，以避免可能出現的全行業短缺。我們還積極評估市場狀況的變動，戰略性地預先儲存若干原材料，以解決潛在的供應短缺。我們的供應管理團隊及我們的業務運營團隊定期審查我們的庫存水平、庫存構成及庫存周轉率，並採取必要的行動以盡量減少陳舊的風險。

業 務

僱員

截至2022年12月31日，我們擁有399名全職僱員。下表按職能載列我們的僱員數目：

僱員職能	僱員數目	佔總數的%
研發.....	281	70.5
銷售及營銷.....	64	16.0
行政.....	54	13.5
總計.....	399	100.0

我們主要通過線上渠道招聘員工，包括社交媒體及本公司的官方網站，以及內部推薦計劃。我們致力於建立有競爭力及公平的薪酬。為有效激勵員工，我們通過市場調研不斷完善薪酬及激勵政策。我們每年對員工進行績效評估，以提供有關彼等績效的反饋。員工的薪酬通常包括基本工資及基於績效的獎金。

為提高員工的技能及知識並發掘員工的新潛力，我們投資於管理層及普通員工的繼續教育及培訓計劃，以定期更新彼等的技能及知識，以確保彼等了解並遵守我們的政策及程序以及相關法律法規。

按照中國法規的規定，我們參加由有關的地方市政府及省政府組織的各種僱員社會保障計劃，包括住房、養老金、醫療、工傷、生育及失業福利計劃。我們與行政人員、經理及僱員訂立有關保密、知識產權及不競爭的僱傭合約及協議。

概無我們的員工目前由工會代表。我們認為，於往績記錄期間及直至最後實際可行日期，我們與僱員保持良好的工作關係，且我們在為營運招募員工方面並無遭遇任何重大勞資糾紛或任何困難。

保險

我們根據中國法律及監管要求以及我們對營運需求及行業管理的整體評估購買保險。我們受中國社會保險制度約束，並為員工繳納基本養老保險、基本醫療保險、失業保險、工傷保險及生育保險五種保險。我們還為董事購買董事責任保險，為針對董事的任何法律訴訟提供適當的保障。我們認為我們的保險範圍足夠，因為我們持有中國法律法規要求的所有強制性保險政策，並符合我們行業的慣例。按照一般的市場慣例，我們並無投購任何業務中斷保險或產品責任保險，根據中國法律，該等保險並非強制性的。我們並無為我們的網絡基礎設施或信息技術系統投保關鍵人物保險，且並無為我們的財產投保任何保單。於往績記錄期間，我們並無提出與我們的業務有關的任何重大保險申索。請參閱「風險因

業 務

素 — 我們的投保範圍可能不足以涵蓋客戶提出的所有損失或潛在申索，從而影響我們的業務、經營業績及財務狀況」。

企業社會責任與環境保護

我們致力推動企業社會責任及環境保護。我們須遵守中國的環境保護及職業健康及安全法律及法規，違反該等法律及法規可能導致主管部門的各種行政處罰，如警告、罰款、責令整改、責令披露相關資料或作出公告或責令暫停業務。請參閱「監管概覽」。

環境保護

我們委聘第三方減碳顧問，協助我們為僱員制定整體減碳培訓計劃。根據該安排，上海環境能源交易所、中國質量認證中心上海分中心及上海經濟信息中心的專家為我們的僱員提供有關碳市場及碳排放的全面培訓。我們的若干僱員已獲得上海環境能源交易所股份有限公司頒發的培訓證書。

在業務運營中，特別是碳峰值、碳中和政策和舉措的背景下，我們識別出了工業園區對低碳和零碳解決方案日益增長的需求。我們為企業及公營機構客戶開發碳管理平台。該平台整合了減碳諮詢管理、碳排放數據追蹤及用戶碳足跡管理等模塊。該平台為我們的雲知大腦模組之一。其目前正在內部應用，作為基於MaaS服務的服務提供予客戶。

由於我們並無從事產品製造，我們目前並無任何有關健康、工作安全及環境的重大責任，且預期我們將不會就此產生任何重大負債，從而可能對我們的業務及經營業績造成任何重大不利影響。於往績記錄期間及直至最後實際可行日期，我們的營運在所有重大方面均符合相關中國環境保護及職業健康及安全法律及法規，且我們並無因違反環境保護及職業健康及安全法律及法規而遭受任何罰款或其他處罰。

董事會已根據上市規則通過全面的環境、社會及管治政策，當中載列我們的企業社會責任目標，並就於日常營運中實踐企業社會責任提供指引。

社會責任

為抗擊COVID-19疫情，我們為北京、廈門、廣西、山東及武漢的醫院提供智慧醫療服務。通過我們先進的語音交互功能，醫務人員能夠與虛擬通信有效合作，從而降低感染風險。向企業及市政當局捐贈並安裝非接觸式電梯。安裝聲控電梯系統，讓乘客在無身體接觸的情況下控制電梯，降低感染風險。我們亦向北京、上海、廈門、泉州及三明的社區

業 務

捐贈為防疫而開發的智能機器人。這些機器人為個人和醫療機構提供電話協助，簡化了與COVID-19疫情相關通信的巨大工作量。

員工關懷

我們致力建立公平、透明及鼓勵的工作環境。我們不斷檢討內部工作場所政策，以確保遵守相關法律法規以及我們團隊的文化及價值觀。

我們旨在為僱員提供具競爭力的薪酬。除薪金及現金獎勵外，我們亦為僱員提供全面的福利待遇，包括節日及節日禮品、超出法定最低假期的額外年假津貼、補充商業醫療保險及年度旅遊。

業務可持續性

於往績記錄期間，我們實現了強勁的收益增長。我們的收益由2020年的人民幣260.8百萬元增加至2021年的人民幣455.8百萬元，並進一步增加至2022年的人民幣600.6百萬元。我們於2020年、2021年及2022年的毛利率分別為34.2%、31.8%及39.9%。於往績記錄期間，我們的客戶由2020年度的459名增加至2021年度的510名，並進一步增加至2022年度的538名。於往績記錄期間，我們的毛利率保持穩健並有所增長。

撇除(i)股份支付開支，(ii)贖回負債利息的融資成本；及(iii)[編纂]開支等項目的影響後，我們於2020年、2021年及2022年分別產生經調整淨虧損人民幣209.8百萬元、人民幣171.7百萬元及人民幣183.2百萬元，分別佔同年我們收益的81.8%、38.0%及30.8%。經調整淨虧損為非國際財務報告準則計量。請參閱「財務資料 — 非國際財務報告準則計量」。

我們的淨虧損主要是由於往績記錄期間產生龐大的研發開支。於往績記錄期間，隨著我們的業務迅速增長，我們研發開支的絕對金額有所增加。我們已開始實施審慎的措施來管理成本及經營開支。

根據我們的業務運營及未來擴張情況，我們密切監控現金狀況是否充足。於往績記錄期間，我們主要通過股東出資及[編纂]前投資融資來滿足現金需求。請參閱「歷史、重組及公司架構 — [編纂]前投資」。截至2020年、2021年及2022年，我們的現金及現金等價物和以公允價值計量且其變動計入當期損益的金融資產分別為人民幣146.2百萬元、人民幣152.7百萬元及人民幣95.8百萬元。我們於2023年完成D3輪融資，籌集超過人民幣700百萬元。我們的總現金餘額足以支付我們用於經營活動的淨現金流量，為我們擴大業務經營提供了充足的流動資金。因此，在考慮我們可用的財務資源後，我們相信我們擁有充足的營運資金，包括充足的現金和流動資產。

我們於往績記錄期間錄得淨虧損及淨經營現金流出，且預計該情況可能持續，直至我們實現更大規模。具體而言，我們的目標是通過以下方式實現盈利：(i)豐富和擴大我們

業 務

的產品和解決方案；(ii)擴大客戶群；及(iii)利用我們的Atlas人工智能基礎設施和雲知大腦技術平台提高運營效率和規模經濟。

物業

我們的公司總部位於中國北京。截至2022年12月31日，我們並無任何自有物業，且於中國租賃11處物業。我們租賃的物業主要用於企業管理、研發及辦公用途。

截至2022年12月31日，我們租賃的物業的賬面價值均未達到或超過我們合併總資產的15%。根據香港上市規則第5章及公司條例(豁免公司及文件遵從條文)公告第6(2)條，本文件獲豁免遵守公司(清盤及雜項條文)條例第342(1)(b)條關於將土地或樓宇的所有權益列入估值報告的規定。

牌照、批文及許可證

截至最後實際可行日期，如我們的中國法律顧問所告知，我們已獲得我們在中國經營業務所需的所有重大牌照及許可證，且有關牌照保持有效。我們的中國法律顧問已告知我們，對有關牌照進行續期並無重大法律障礙，惟前提條件是我們須遵守相關法律規定並按照適用的中國法律法規規定的要求及時間表提交相關申請。

法律訴訟及合規性

法律訴訟

我們不時受到並將受到日常業務過程中產生的各種法律或行政申索及訴訟。訴訟或任何其他法律或行政程序，無論結果如何，均可能導致大量成本及我們資源的轉移，包括我們管理層的時間及注意力。請參閱「風險因素 — 與我們行業及業務有關的風險 — 我們可能不時面臨申索、糾紛、訴訟及其他法律及行政程序」。

於往績記錄期間及直至最後實際可行日期，不存在可能個別或合共對我們的業務、財務狀況及經營業績造成重大不利影響的針對我們或我們的董事的未決或威脅將提起的法律訴訟。

合規性

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與我們業務相關的適用法律及法規，且並無涉及董事認為可能會發生個別或總體上對我們的整體業務產生重大不利影響的任何不合規事件。於往績記錄期間，我們未受到任何與業務運營相關的重大處罰或調查。

業 務

風險管理及內部控制

我們致力建立及維持穩健的風險管理及內部監控系統，包括符合我們營運動態的定制政策及程序。我們定期檢討風險管理策略及其執行情況，以核實其有效性及充分性。我們已在各業務範疇採納及實施全面風險管理框架，當中包括以下主要方面。

信息系統

信息系統的完整性對我們的營運至關重要。我們已建立完善的內部協議及控制措施，以保護我們的IT基礎設施、保護我們的數據及防止潛在洩漏或損失。於往績記錄期間及直至最後實際可行日期，我們並無經歷任何重大系統故障或數據外洩。

我們的數據受到已建立的信息系統安全管理框架的保護。該綜合系統包含網絡及數據安全措施、防病毒協議、系統變更審批程序以及嚴格的數據備份及存檔方法。我們的中央數據中心定期對多個安全系統進行備份，以降低數據丟失的風險。我們每三個月更新備份存儲介質，以確保性能穩定。為進一步確保數據完整性，我們定期進行備份恢復測試，並使用例行系統檢查、高安全標準密碼政策、訪問控制系統及定期數據恢復測試等措施。我們每年審閱備份政策及程序。

我們的IT運維部門監督IT系統及基礎設施的保存，確保數據使用、存儲及保護符合我們的內部規則及所有相關法律法規。我們持續為員工提供信息安全培訓，並實施應急機制以評估重大風險、制定災難應對計劃及定期進行應急演練。我們對持續風險管理的承諾確保了我們營運的安全性及用戶數據的保護。

合規措施

合規風險表明我們未能遵守相關法律、法規、規則及指引而可能遭受法律、監管制裁、重大財務損失及聲譽損害。我們已在全面風險管理體系的範圍內設計並實施穩健的合規風險管理框架。目的是有效識別及管理合規風險，並確保我們的營運遵守適用法律及法規。

我們已將詳細的反賄賂合規風險管理措施納入我們的監管合規管理體系。該等措施包括識別、評估、監控及報告賄賂。我們亦定期向僱員提供反賄賂合規培訓。

我們的業務行為守則明確禁止在業務活動中進行任何形式的商業賄賂，包括現金、實物、服務提供、價格優惠或折扣。其亦限制接受或提供禮品、餐飲、娛樂、旅遊、購物及在業務往來中超越標準商務禮節的其他形式的使客戶滿意的行為。

業 務

此外，我們已設立反賄賂舉報熱線，接獲的任何舉報必須轉交獨立部門進行後續調查。

財務報告

為管理財務報告風險，我們已實施一套會計政策，包括財務報告管理、預算管理、財務報表編製以及財務部門及員工管理。我們已制定特定程序以制定該等會計政策，而我們的財務部門會根據該等程序定期審閱管理賬目。我們進一步確保財務部門僱員接受定期培訓，以加強彼等對我們財務管理及會計政策的了解，鼓勵彼等於日常營運中有效實施。

審計委員會經驗及資格以及董事會監督

我們的審計委員會持續監督我們風險管理政策的實施，以確保我們的內部監控系統能夠處理業務風險。我們的內部審計部門為此程序的組成部分，負責審查監控成效、向審計委員會匯報事宜及協助實施解決策略。重大事宜會及時轉交委員會，並在有需要時向董事會匯報。請參閱「董事、監事及高級管理人員 — 董事委員會 — 審計委員會」。

獎項及認可

自成立以來及直至最後實際可行日期，我們已在產品、技術及創新方面獲得獎項及認可，其中重要的獎項及認可載列於下文：

- 2018年、2020年及2021年《科技創新2030 —「新一代人工智能」重大項目》的承辦單位，該項目為中華人民共和國科學技術部組織的國家重點研發計劃；
- 在2018年WMT國際機器翻譯大賽（全球領先的機器翻譯比賽）的英譯中排名第二；
- Atlas超級計算平台於中國榮獲2018年人工智能與實體經濟深度融合創新項目獎；
- 2019年吳文俊人工智能科學技術進步獎；
- 2019年北京市科學技術進步獎一等獎；
- 2020年全國知識圖譜與語義計算大會(CCKS)「面向中文電子病歷的醫療命名實體識別評測任務」冠軍，並獲得唯一技術創新獎；
- 2019年及2020年度中國健康信息處理獎(CHIP(健康信息處理大會))；

業 務

- 2020年Blizzard Challenge國際語音合成大賽普通話和上海話多項指標第一；
- 語言與知識智慧競賽2020閱讀理解任務第一名；
- 2021年度INTERSPEECH深層噪聲抑制(DNS)挑戰賽及聲學回聲消除(AEC)挑戰賽第二；
- 2021年獲認定為北京市、上海市、國家級專精特新「小巨人」企業榮譽稱號的企業；
- 社交媒體健康信息挖掘(#SMM4H)研討會暨2021年共享任務，與國際計算機語言協會北美分會2021(NAACI)-SMM4H評測同地舉行。國際計算機語言協會北美分會2021(NAACI)-SMM4H評測為利用社交媒體數據進行健康信息學的國際競賽；
- 2021年廈門市科學技術進步一等獎；
- 2022年國際先進人工智能協會(AAAI)年會「首字母縮寫詞提取」及「首字母縮寫詞消歧」冠軍；及
- 2022年4月贏得了中國生物醫學語言理解評估(CBlue)錦標賽。