

資策會產研聯盟
Taipei-1 算力提案構想

建立繁體中文領域型LLM

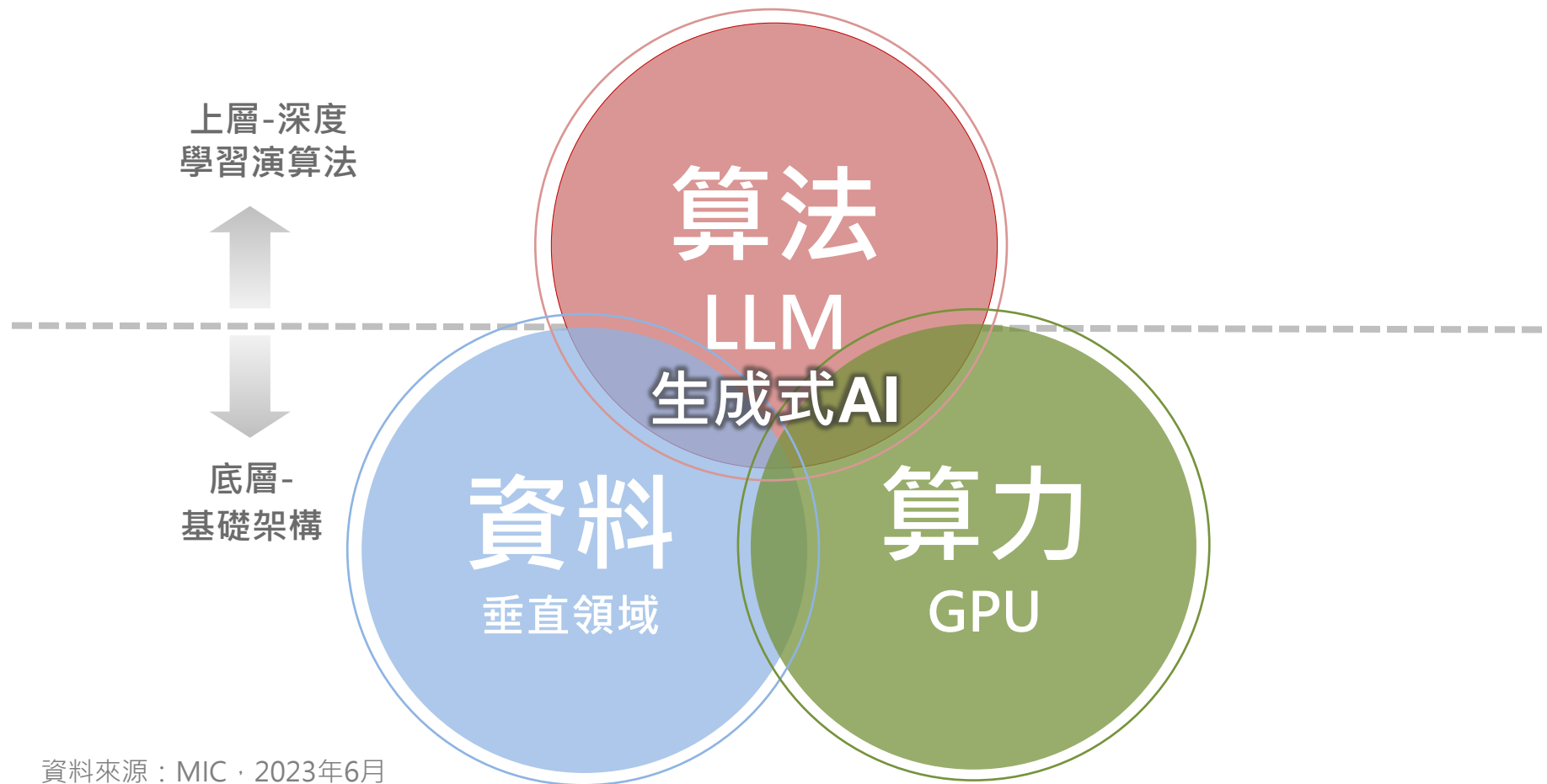
資策會 軟體院

2024.04.19





生成式AI的三大基石：資料、算力、算法

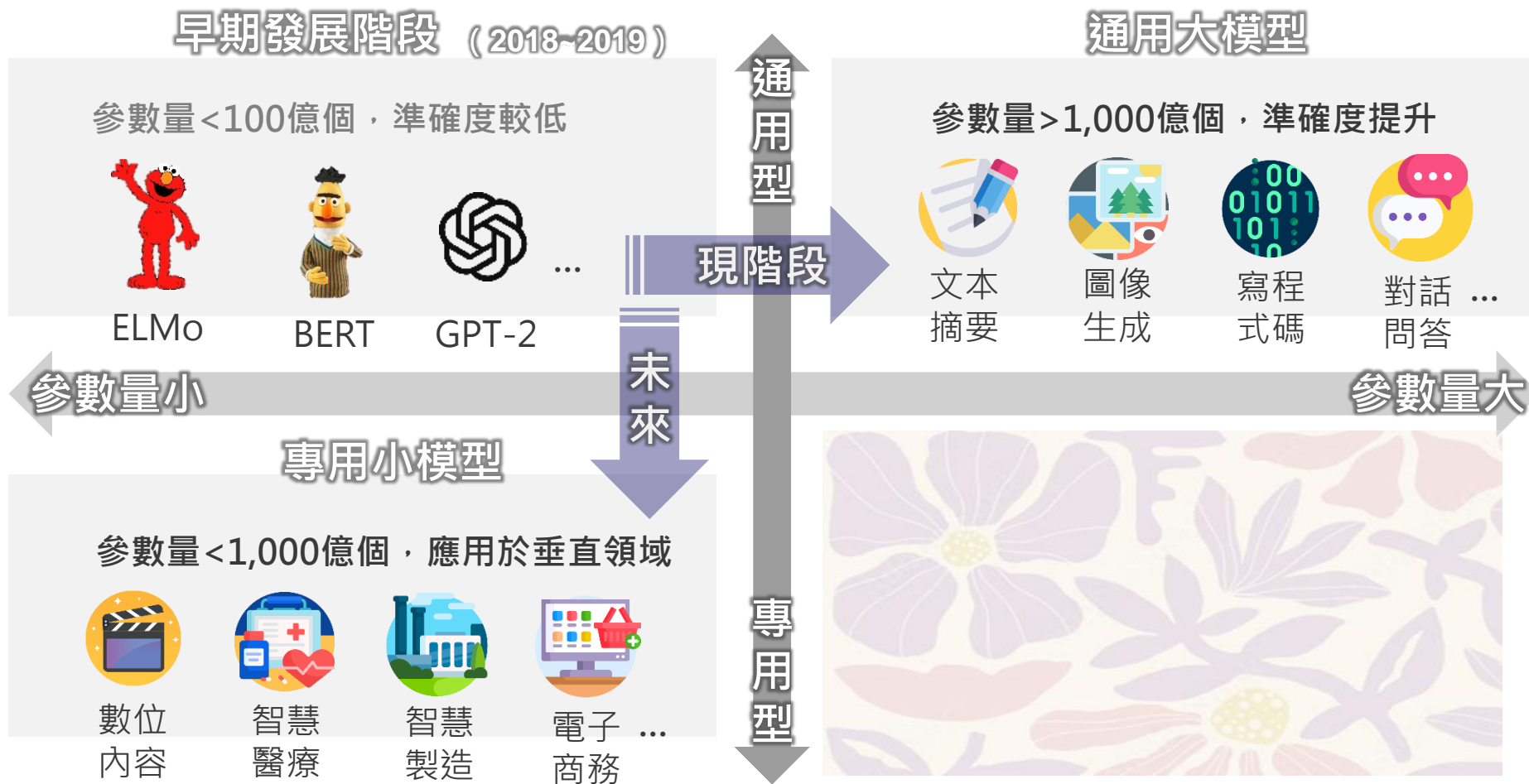


資料來源：MIC，2023年6月

- 生成式AI仰賴三大基石共同協作：資料（Data）、算力（Computing）與算法（Algorithm），由資料和算力構成底層基礎架構後，疊加深度學習演算法，建構出生成式AI模型



垂直領域專用模型將蓬勃發展



資料來源：MIC · 2023年6月

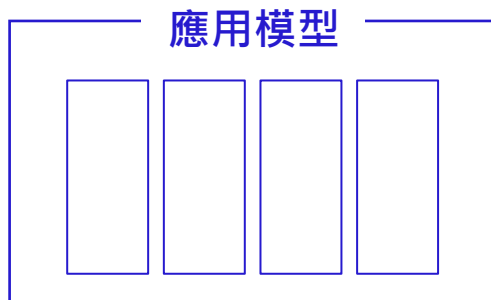
- 現階段技術發展以通用大模型為主，然而垂直領域需客製化模型解決特定任務，可在大模型基礎上微調並適當降低參數後開發專用模型，未來將呈現垂直領域專用模型蓬勃發展趨勢



台灣產業契機 – 垂直領域LLM

AI is everywhere

- AI邁向平民化與終端化後，勢必帶動各行各業**AI應用**的創新發展
- 隨著AI商業化的潛力，啟動新一波市場競逐，帶動**AI新創**的蓬勃發展

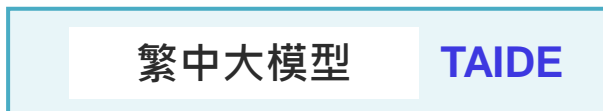


Domain LLM

- 提昇**特定領域**的準確度
- 降低**新創企業**進入門檻
- 減少**重複投資**的成本



BIG Tech.





資策會產研聯盟執行策略

建立領域類訓練資料集，架構於Taide模型，發展領域型可信任中文LLM



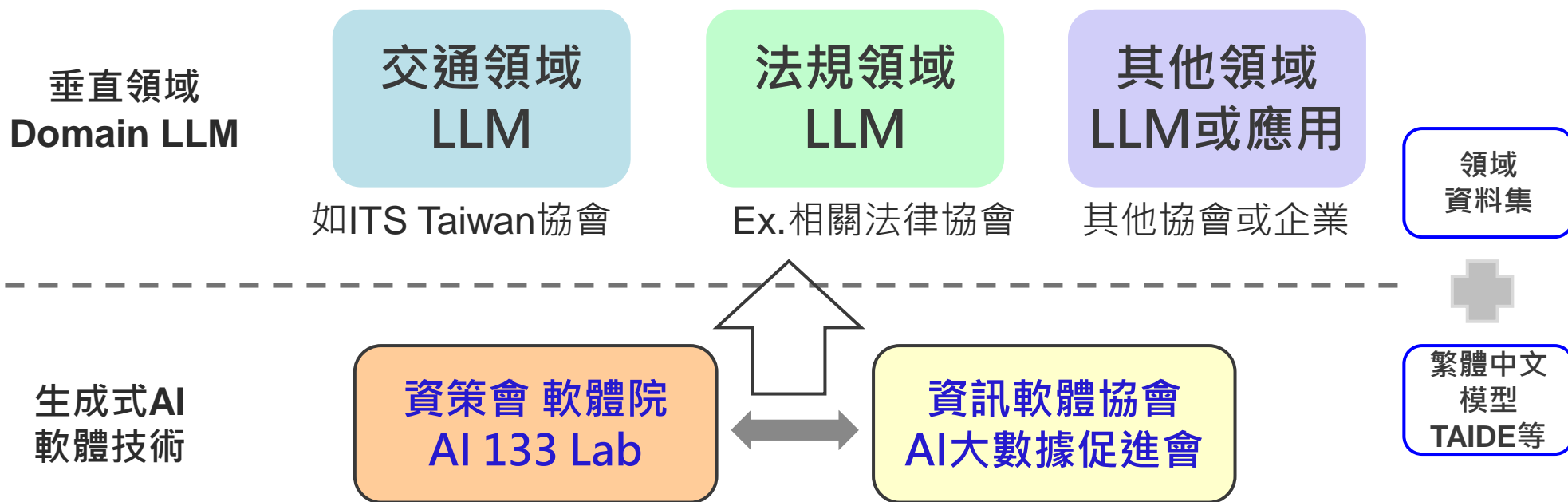


資策會產研聯盟組隊構想

專案主題：建立繁體中文領域型LLM

專案團隊：資策會 + 軟體協會

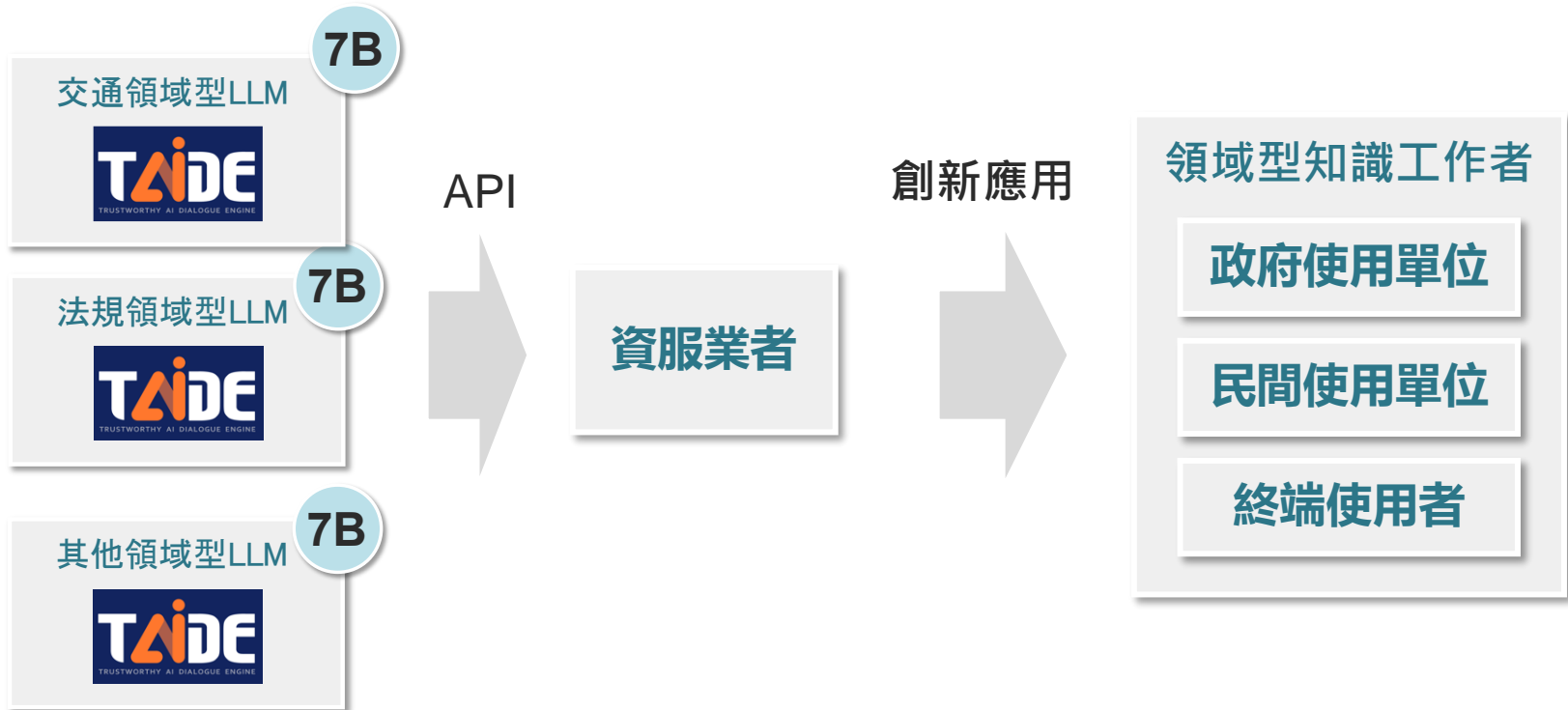
領域夥伴：各垂直領域夥伴





資策會產研聯盟預期效益

1. 建立垂直領域資料集
2. 建立兩項領域之繁體中文語言LLM模型
3. 掌握發展生成式AI(LLM)之多項關鍵技術
4. 賦能資服產業技術能量，擴散帶動領域產業升級轉型



專案諮詢窗口：財團法人資訊工業策進會、軟體技術研究、賴易聖 先生 E-Mail : davidlai@iii.org.tw