

=====

ITEM #78 — LLM vs DBM- 存量 AI 与增量 AI ; 全过程开源设想

Conversation Title: 存量 AI 与增量 AI

Date: 20251101

Authors: Sizhe Tan & GPT-Obot

=====

ITEM #78 — LLM vs DBM: 存量 AI 与增量 AI ; 全过程开源设想

一、LLM AI vs DBM = 存量 AI vs 增量 AI

- **存量 AI (LLM)**
 - 基于大规模既有语料训练，属于“资源开采”。
 - 特征：穷尽枚举、精炼重组，类似汉赋–唐诗–宋词的演进逻辑。
 - 优点：规模大时逼近极致表达；缺点：局限于旧有语料边界。
- **增量 AI (DBM)**
 - 基于度量空间、差分树、规则引擎，强调动态吸纳与持续演化。
 - 类似“航海发现新大陆”：每个新数据点都可生长出新分支与规则。
 - 优点：外延拓展、后浪推前浪；是开放生长而非封闭压缩。
- **二者关系**
 - 互为补充：LLM = 图书馆（存量整理），DBM = 森林（持续生长）。
 - 未来格局：存量–增量双循环，LLM 提供沉淀，DBM 提供外延。

二、全过程开源的可能性与价值

- **现状**：开源集中在成果（模型、代码），忽视了概念与过程。
- **风险**
 - 噪声过多：半成品想法充斥社区。
 - 竞争风险：设计过早泄露，易被快速商业化。
- **价值**
 - **加速演化**：减少重复试错，提升研发效率。
 - **文明积累**：透明作坊 → 社会整体艺术/科学水平提升。
 - **信任共创**：推理链与设计动机透明，更能聚合合作者。
- **历史先例**
 - 数学界通信圈（17–18世纪）：不断交流未成熟想法，学科加速。
 - Unix 社区（70–80年代）：开源的不只是代码，还有设计哲学。
 - 现代 arXiv + GitHub：已是“半过程开放”。
- **面向 DBM 的分层开源设想**
 - **理念层**：白皮书、架构图、对照表（最开放）。
 - **实验层**：伪代码、失败实验记录、部分接口（有限开放）。
 - **成品层**：核心算法与完整系统（审慎开放）。

三、总结性比喻

- **LLM vs DBM**：旧大陆的财富库 vs 新大陆的拓荒地。
- **成果开源 vs 全过程开源**：一幅完成的油画 vs 整个画室（草稿、工具、调色板、失败的画布一并开放）。