

AuraMax 综合产品技术白皮书 (v2.0)

文件编号: AM-PROD-2026-001 密级: 内部公开 版本: 2.0 (详尽版) 日期: 2026-01-14

目录

- 产品愿景与核心理念
- 系统架构设计 (MCP) 架构概览 核心微服务组件 数据流与交互模型
- 核心功能模块深度解析 AuraMax Ingest: 智能数据摄取与隐私清洗 AuraMax Bio-Compute: 生物信息计算引擎 AuraMax Reasoning: 临床逻辑推理引擎 AuraMax Professional: 企业级队列分析
- 用户操作指南与应用场景 [B2C] 个人健康管理场景 [B2B] 药物研发与临床试验场景 [G2B] 政府监管与证据审查场景
- API 接口与集成指南
- 安全合规与部署

1. 产品愿景与核心理念

AuraMax

旨在解决现代医疗健康领域的核心矛盾：数据爆炸与决策能力之间的不对称。它是全球首个基于 Model Context Protocol (MCP) 构建的医疗“中间件”生态系统。

1.1 核心价值主张

- The "Un-Black Box" (去黑盒化): 在医疗领域，AI 的不可解释性是致命的。AuraMax 的所有临床建议都附带完整的“思维链 (Chain-of-Thought)”和权威文献引用 (PubMed/FDA Guidelines)，确保每一条决策都有据可查。
- Privacy-First Architecture (隐私优先架构): 采用“本地优先”策略。敏感的个人身份信息 (PII) 在离开本地网络前即被识别并剥离，使得系统天然符合 HIPAA 和 GDPR 标准。
- Scientific Grounding (科学落地): 我们不仅使用 LLM (大语言模型) 进行文本生成，更集成了 BioPython, RDKit, Pandas 等硬科学计算库，确保药物相互作用、生存分析等计算结果的数学精确性。

2. 系统架构设计 (MCP)

AuraMax 不仅仅是一个 App，它是一套分布式的微服务集群，通过 MCP (Model Context Protocol) 协议进行标准化通信。

2.1 架构概览

系统由 Client (宿主) 和多个 Server (能力服务器) 组成。

```
graph TD
    Client[AuraMax 主控端 (Web/App)] -->|MCP Protocol| Ingest[Ingest Server (摄取)]
    Client -->|MCP Protocol| Reason[Reasoning Server (推理)]
    Client -->|MCP Protocol| Bio[Bio-Compute Server (计算)]
    Client -->|MCP Protocol| Data[Data Warehouse (数据仓)]

    Ingest --> OCR[OCR 引擎] & PII[隐私清洗]
    Reason --> PubMed[PubMed API] & Guidelines[指南库]
    Bio --> Genomics[基因组工具] & Mol[分子模拟]
    Data --> FHIR[FHIR 接口] & SQL[PostgreSQL]
```

2.2 核心微服务组件

- AuraMax Ingest: 负责处理非结构化数据。能力: PDF 解析、OCR (光学字符识别)、医疗实体识别 (NER)、PII 自动编校。
- AuraMax Reasoning: 负责医学逻辑与证据合成。能力: 检索增强生成 (RAG)、证据分级 (GRADE)、多文档综合推理。
- AuraMax Bio-Compute: 负责硬科学计算。能力: VCF 基因变异解析、药物-基因相互作用检查 (CPIC)、生存曲线绘制 (Kaplan-Meier)。
- AuraMax Data: 负责数据持久化与标准兼容。能力: FHIR R4 标准化存储、队列数据管理。

3. 核心功能模块深度解析

3.1 AuraMax Ingest: 智能数据摄取与隐私清洗

- 多模态输入: 支持 PDF 实验报告、DICOM 影像元数据、手写处方照片、CSV 原始数据。
- 智能布局分析: 能够识别复杂的医疗表格结构 (如血液检查单中的“项目”、“结果”、“参考范围”、“单位”列), 并将其转换为结构化的 JSON 数据。
- Local Privacy Layer (本地隐私层): 算法: 使用 NLP 模型 (Presidio) 识别姓名、SSN、MRN、日期、地址。机制: 在数据进入云端推理模型或数据库之前, 在内存中即完成掩码处理 (Masking), 确保云端模型只看到脱敏数据。

3.2 AuraMax Bio-Compute: 生物信息计算引擎

- 药物基因组学 (PGx): 支持解析 23andMe, Ancestry, Dante Labs 等主流厂商的原始 VCF 文件。逻辑: 映射用户基因型 (Genotype) -> 对应单倍型 (Haplotype, 如 CYP2C19 2/3) -> 对应代谢表型 (Phenotype, 如 Poor Metabolizer) -> CPIC 临床建议。
- 分子结构可视化: 集成 RDKit, 可实时渲染药物分子的 2D/3D 结构, 并计算 logP, 分子量等理化性质, 辅助研发人员评估成药性。

3.3 AuraMax Reasoning: 临床逻辑推理引擎

- 实时循证 (Real-time Evidence): 通过 API 实连接 PubMed 和 ClinicalTrials.gov。
当用户提问“这款药的最新副作用研究”时，系统实时拉取最近 6 个月的 Paper Abstract 进行总结。
- 证据分级 (Evidence Grading): 系统会自动评估引用文献的质量。RCT (随机对照试验) 权重 > 队列研究 > 病例报告。输出格式明确标注：(证据等级: High - 基于 3 项三期临床试验)。

3.4 AuraMax Professional: 企业级队列分析

- 生存分析: 输入: 包含“生存时间”和“结局事件”的患者 CSV 数据。算法: Python lifelines 库。
输出: 自动绘制 Kaplan-Meier 生存曲线，并计算 Log-rank test p-value，判断不同治疗组之间是否存在显著差异。
- 风险建模: 构建 Cox 比例风险回归模型，分析年龄、性别、基因突变状态等协变量对生存率的影响 (Hazard Ratio)。

4. 用户操作指南与应用场景详解

4.1 [B2C] 个人健康管理场景

典型用户: 慢性病患者、注重健康管理的极客。

核心流程：Lab Report Interpretation (化验单解读) 1. 上传: 用户通过 App 拍照上传年度体检的血液生化报告 (PDF/图片)。 2. 处理: * Ingest Server 识别其中的 45 项指标。 * OCR 提取数值，并识别出“甘油三酯”和“低密度脂蛋白”超标。 3. 推理: * Reasoning Server 结合用户年龄 (45岁，男)，生成解读：“虽然您的数值略高，但考虑到您没有吸烟史，这可能与近期饮食有关...” * 提供生活方式建议 (低碳水饮食方案)。 4. 追踪: 存入 Data Warehouse，生成历史趋势图，监控指标变化。

核心流程：Pharmacogenomics Screening (用药安全筛查) 1. 数据: 用户授权导入 23andMe 基因数据。 2. 输入: 用户输入正在服用的药物 (如：氯吡格雷/Clopidogrel)。 3. 预警: 系统检测到用户是 CYP2C19 中间代谢者，提示：“氯吡格雷在该基因型下疗效可能降低 30%，建议咨询医生更换为替格瑞洛 (Ticagrelor)。(参考：CPIC Guideline 2022)”

4.2 [B2B] 药物研发与临床试验场景

典型用户: 药企研发部门、临床研究协调员 (CRC)。

核心流程：Clinical Trial Matching (临床试验受试者匹配) 1. 定义标准: 研究员输入试验的入排标准 (Inclusion/Exclusion Criteria)，例如：“非小细胞肺癌，EGFR 突变阳性，既往未接受过 TKI 治疗，ECOG 评分 0-1”。 2. 队列筛选: * AuraMax 在院内 EHR 数据库中进行语义搜索。 * Data Server 筛选出符合基因型和病史的候选患者列表。 3. 报告生成: 自动生成符合条件的患者清单，并高亮显示可能的排除风险点。

核心流程：Real-World Evidence (RWE) Generation (真实世界证据生成) 1. 分析: 导入已上市药物的患者随访数据。 2. 计算: 运行 Cohort Analysis 模块，按不同的基因突变类型进行分层 (Stratification)。 3. 洞察: 发现药物在特定亚群 (如 KRAS G12C 突变) 中生存获益显著优于总体人群。 4. 输出: 自动生成符合学术发表标准的统计图表和描述性文本。

4.3 [G2B] 政府监管与证据审查场景

典型用户: FDA/NMPA 审评员、医保局专家。

核心流程：NDA Submission Review (新药上市申请自动化初审) 1. 摄取: 审评员上传药企提交的 CSR (临床研究报告, 通常长达数千页)。 2. 验证: * 系统提取其中的统计汇总表。 * Consistency Check: 自动比对正文描述的数据与附件原始统计表的数据是否一致。 * Literature Check: 自动检索与该药物机制相关的外部发表文献，标记是否存在未披露的不良反应信号。 3. 报告: 生成“审评辅助报告”，列出 10 个需要重点人工核查的风险点。

核心流程：Post-Market Surveillance (上市后监测) 1. 监控: 连接不良事件报告系统 (FAERS/CMD)。 2. 信号检测: 持续分析特定药物的不良反应报告趋势。 3. 警报: 当某种罕见副作用的报告率超过统计阈值时，自动触发红色警报，并汇总相关病例特征。

5. API 接口与集成指南

系统提供 RESTful API，基于 OpenAPI (Swagger) 标准。

示例：调用生存分析接口

```
POST /api/v1/cohort/analyze
Content-Type: application/json

{
  "project_id": "proj_lung_cancer_001",
  "analysis_type": "kaplan_meier",
  "groups": ["treatment_A", "control_B"],
  "time_column": "os_months",
  "event_column": "is_deceased"
}
```

响应:

```
{
  "status": "success",
  "results": {
    "median_survival": {"A": 24.5, "B": 18.2},
    "p_value": 0.034,
    "plot_url": "https://api.auramax.com/plots/km_Curve_123.png"
  }
}
```

6. 安全合规与部署

6.1 数据安全层级

- L1 (Transport): 全链路 TLS 1.3 加密。
- L2 (Storage): 数据库字段级加密 (AES-256)，密钥通过 KMS 管理。
- L3 (Compute): 支持 TEE (可信执行环境) 选项，用于极高敏感度数据计算。

6.2 部署模式

- SaaS (美国/国际): 部署于 AWS (符合 HIPAA 的专用实例)。
- On-Premise (中国/欧洲): 支持 Docker/Kubernetes
私有化部署，数据完全不出内网，符合中国《数据安全法》和欧盟 GDPR 要求。