

定值比对AI智能体系统

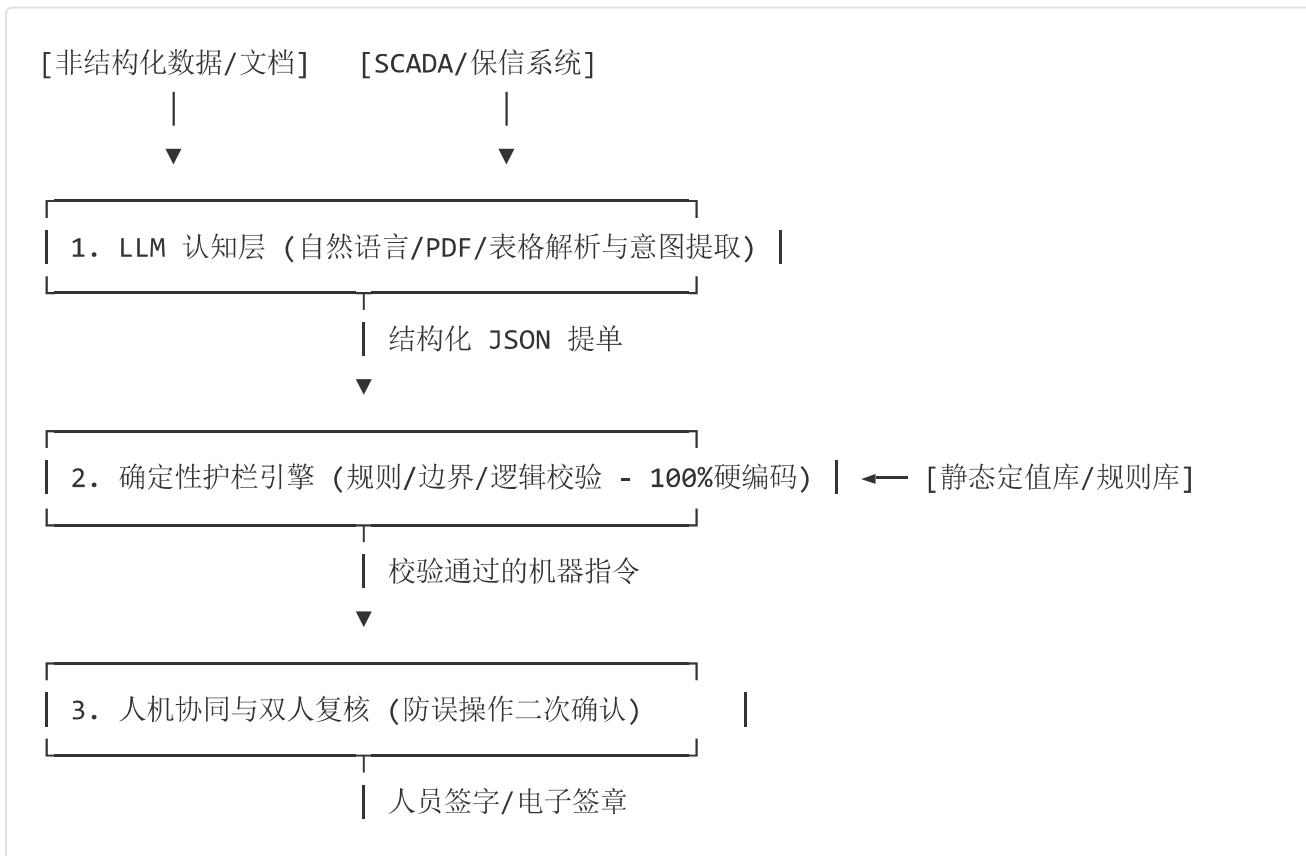
在电力系统（尤其是地市级供电局）中，继电保护定值被称为电力网安全的“第一道防线”。定值差之毫厘，可能引发越级跳闸、主变烧毁甚至电网大面积停电。**大语言模型（LLM）天然具有概率性与幻觉特征，绝对不能直接用 LLM 去充当执行器直接向控制网（I/II区）下发指令。**

但这并不意味着不能开发 AI Agent。关键在于**架构设计理念**的转变：

核心原则：“大模型做认知与解析，确定性引擎做校验与计算，人类做闭环确认。” **LLM 的角色是“超级助理（Copilot）”，绝不是“自动驾驶员（Autopilot）”。**

一、整体技术架构设计：分层防线

为了兼顾大模型的非结构化数据处理能力与电网的零差错要求，建议采用“双核驱动 + 三道防线”的技术架构：



二、三大场景的具体技术实施路径

1. 定值比对（AI 优势区）

- **痛点：** 传统纸质/PDF定值单与二次设备现场定值格式不一、版本繁多，人工核对耗时且易遗漏。
- **技术方案：**
 1. **多模态 OCR + LLM 结构化抽取：** 将 PDF、Word、扫描件形式的定值单转化为统一的 JSON 格式。
 2. **确定性算法比对：** 提取完数据后，**不再让 LLM 做比对**，而是交给 Python / C# 编写的确定性比较函数（如：数值全等比对、绝对值/相对偏差计算）。
 3. **结果输出：** 高亮标记差异字段（如整定时间、动作电流 I_{set} 、时限 T_{set} ），LLM 仅生成差异原因分析提示（如：“设备A与最新定值单比对，第 3 段过流定值偏差 0.5A，疑似未更新版本”）。

2. 定值监控（AI 辅助区）

- **痛点：** 运行环境变化、设备变更后，定值越界或配合失效隐患难以实时发现。
- **技术方案：**
 1. **实时数据对接：** 代理接入保信系统，定时拉取现场二次设备的运行定值镜像。
 2. **知识图谱/规则引擎协同：** 将电网保护级配合规条文（如上下级定值配合级差规则）写入知识库与逻辑引擎。
 3. **预警诊断：** 规则引擎发现越界或不配合现象后，触发 LLM 汇总上下文，生成人性化的告警简报及处置建议供运行人员参考。

3. 定值下发（高风险管控区）

- **痛点：** 误下发、下发错误参数导致事故。
- **技术方案（严格闭环控制）：**
 - **LLM 不接触写接口：** LLM 的输出仅为“定值下发申请草稿”（JSON 格式）。

- **强规则拦截（硬护栏）：**
- **数值范围校验：** $I_{\min} \leq \text{下发值} \leq I_{\max}$ ，越界直接物理熔断。
- **版本号校验：** 下发目标版本必须大于当前现场版本且与正式批准的定值单号一致。
- **级差与互锁校验：** 判断当前方式下是否满足开关互锁要求。
- **双人复核（Human-in-the-Loop）：** 必须经过保护专责、班长双人电子签章确认。
- **工业协议选择性下发（带返校）：** 下发过程采用标准电力协议（如 IEC 60870-5-104 或 IEC 61850），遵循“选择-返校-执行”三步法。下发前先预置并读取设备返校值，比对 100% 正确后再发出执行命令；若返校不一致则自动撤销。

三、 风险评估与应对策略

风险维度	具体风险表现	技术/管理应对策略
幻觉风险	LLM 在提取或转换定值参数时产生数值错误（如把 10.5A 变成 10.5kA 或 10.50A）。	解耦隔离： 提取后立即用 Schema 校验 + 正则匹配。数值计算与比较由确定性计算模块完成，LLM 不做任何数值计算。
网络与安全 风险	AI 系统被注入攻击或越权操作生产控制区（I/II 区）。	安全分区： AI Agent 部署在 III 区（管理信息区）或 II 区的辅助分析模块， 严禁直连 I 区下发接口 。下发指令需通过正反向隔离装置及现有成熟的定值下发系统代理。
责任归属 风险	AI 生成了错误指令，人员盲目信任导致下发事故。	制度闭环： 明确 AI 仅为决策辅助工具（L2级辅助），最终确认权在保护专责。系统留存全过程日志（包含 LLM Prompt、提取结果、规则校验记录、人员确认记录），可追溯审计。
设备兼容性 风险	不同厂家二次设备（国网/南网多厂商）私有规约不同，AI 输出格式不兼容。	标准化中间层： AI 仅输出 IEC 61850 / SCL 规范或标准 JSON，底层转换由成熟的保信系统适配器完成。

四、 开发此 AI Agent 的真正优势在哪里？

既然下发不能全靠 AI，那开发 Agent 的价值何在？

1. **非结构化数据处理效率提升 90% 以上** 地市局积累了大量的 Word/PDF 历史定值单、会议纪要和现场运维记录。传统系统无法识别这些非结构化文本，而 LLM 能够秒级完成定值单提取与数字化入库。
 2. **打破系统孤岛，实现“自然语言提问”** 保护专责无需打开多个系统，只需询问 Agent：“把变电站A 10kV 1号馈线上个月的定值变更记录和现场运行定值做个比对”，Agent 即可自动跨系统调取数据并给出分析结果。
 3. **从“事后查核”变为“动态防御”** 传统靠人工每年或半年定期核对定值；Agent 可进行 7x24 小时无人值守式监控，发现异常随时提示，把隐患消灭在萌发阶段。
-

五、总结与建议

在设计该方案时，建议向局领导与业务部门明确界定边界：

1. **第一阶段（上线初期）：** 仅开放**定值比对与监控**（只读模式），不做任何下发功能，积累样本与准确率数据。
2. **第二阶段：** 开放**定值下发辅助拟单**，生成的下发任务必须送入现有的成熟定值下发系统，并保留人工审核流程。
3. **核心口诀：**“数据提取靠 AI，数值比对靠代码，下发控制靠协议，最终安全靠校验与人工”。

这样既能享受大模型带来的效率红利，又能保证电网安全的万无一失。