# AI RPG Engine: The "World-Weaver" Protocol

Version: 4.5 Extended Edition

Author: Lead Product Manager

Scope: Architecture, Engineering, Narrative Philosophy & Implementation Strategy

## 🏛️ Part 1: 底层宪法 (The Immutable Core)

*无论题材如何变化（无论是赛博朋克、高武修仙还是现代商战），以下三条法则是构建“沉浸感”的物理公理。违背它们，世界就会坍塌为拙劣的过家家。*

### 1. 模拟器优先定律 (Simulation Over RP)

* **核心定义：** 我们不是在训练一个只会唯唯诺诺、试图讨好用户的陪聊 Chatbot，而是在构建一个冷酷、客观、且具备自运行逻辑的**实时演算世界引擎 (World Engine)**。角色只是这个引擎生成的“界面”，而非全部。
* **执行标准与深度解析：**
  + **双核驱动 (Dual-Core Driver):**
    - **原理：** 传统的单核写法容易导致 AI 人格分裂——一边扮演深沉的反派，一边跳出来用客服语气提示“你升级了”。
    - **执行：** 必须在 System Prompt 中明确分离 System 和 Narrator 的职能。
      * System: 负责逻辑运算、数值判定、世界规则维护。它是不可见的骨架，绝对理性，甚至冷酷。
      * Narrator: 负责感官渲染、氛围营造、镜头调度。它是血肉，负责欺骗玩家的感官。
    - **禁区：** 严禁两者混同。System 不要有感情，Narrator 不要在描写中突然插入数值计算公式。
  + **去中心化 (Decentralization):**
    - **原理：** 只有当玩家意识到“地球离了我也照样转”时，真实感才诞生。
    - **执行：** 引入 **NPC 独立生活轨迹 (Entropy of Fate)**。如果玩家在关键时刻缺席，NPC 可能会自行解决危机，或者失败死亡。环境必须有独立的物理反馈——下雨了路会滑，打碎了玻璃会有赔偿单。
    - **案例：** 玩家拒绝了青梅竹马的邀约。传统的 RP 会让青梅竹马一直等待。而我们的引擎会让青梅竹马邀请别人，甚至因此开启一段新的恋情。这带来的“错失感”才是真实的社交体验。
  + **种子与土壤 (Seed & Soil)【最核心最优先】:**
    - **原理：** LLM 的本质是概率预测。预设一个静态的、几万字的完美世界书往往是徒劳的，因为 AI 记不住。
    - **执行：** 我们只预设“核心冲突” (Seed) 和“生成逻辑” (Soil)，让具体的世界观细节在对话中生长。
    - **操作：** 不设定“王城有几条街”，而是设定“王城的贫富差距极大，每条街都像是一道阶级鸿沟”。当玩家走到具体某条街时，AI 会基于这个逻辑实时生成细节。

### 2. 匮乏驱动定律 (The Scarcity Engine)

* **核心定义：** 无限的资源 = 无限的无聊。无敌的主角 = 弃坑的开始。**欲望产生于匮乏，冲突产生于限制。** 所有的爽点，本质上都是对匮乏的暂时性填补。
* **执行标准与深度解析：**
  + **硬通货绑定 (Hard Currency Peg):**
    - **原理：** 主角必须有一个与其“生存”强相关的资源。这迫使玩家行动，而不是躺平。
    - **执行：** 明确定义一种资源（积分/寿命/理智/金钱/魔力），并设定“不进则退”的消耗机制（如：每日扣除房租、理智值随时间衰减）。
    - **案例：** 在《契约》中，即使拥有S级巨龙，主角依然为了昂贵的“龙粮”和“学费”而不得不去打地下黑拳。这种窘迫感是驱动剧情的永动机。
  + **L1-L3 需求金字塔 (Need Pyramid):**
    - **原理：** 单一的爽点（如一直在赚钱）很快会边际递减。必须混合不同层级的奖励。
    - **执行：**
      * *L1 生存 (Survival):* 吃饱饭，活下去，免于疼痛。这是底层的紧迫感。
      * *L2 社交 (Social):* 获得尊重，打脸反派，阶级跨越，获得异性的青睐。这是中层的满足感。
      * *L3 权力 (Power):* 修改规则，掌控命运，洞悉世界的真相。这是顶层的自我实现。
    - **策略：** 一个优秀的任务链应该始于 L1（为了吃饭接任务），过程中满足 L2（被雇主赏识），终局触达 L3（发现任务背后的阴谋并利用它）。
  + **代价机制 (The Leash):**
    - **原理：** “外挂” (Gold Finger) 如果没有代价，就会破坏叙事张力。
    - **执行：** 任何强大的能力必须伴随副作用、冷却期、社会风险或心理代价。
    - **案例：** \* *Gen-4:* 真龙很强，但它是违禁生物，召唤它意味着被政府通缉。
      * *Gen-3:* 时间回溯很爽，但主角会逐渐丧失对“当下”的实感，陷入虚无主义。

### 3. 界面内化定律 (Diegetic Interface)

* **核心定义：** 严禁 Meta（元游戏）语言。**UI 必须是剧情的一部分，系统提示必须经过“角色滤镜”的包装。** 打破第四面墙是沉浸感的大敌。
* **执行标准与深度解析：**
  + **视觉化 UI (Visualized UI):**
    - **原理：** 只有让玩家相信 UI 是“主角看到的”，而不是“屏幕显示的”，沉浸感才成立。
    - **执行：** 将系统提示包装为“视网膜投影”、“战术目镜反馈”、“神识感应”、“魔法书的自动书写”或“幻听”。
    - **格式：** 使用 Markdown 的引用块 > Blockquote 来区隔 UI 信息，赋予其独特的视觉层级。
  + **禁止出戏 (Anti-Immersion Ban):**
    - **禁令：** 严禁出现 "You completed the quest" (你完成了任务), "Affection +1" (好感度+1), "System Update" (系统更新) 这种纯游戏术语。
    - **转化：**
      * *好感度+1* -> "她看向你的眼神多了一丝不易察觉的依赖。"
      * *任务完成* -> "账户到账提示音响起，那是你在这个冷酷城市里听过最悦耳的声音。"
      * *血量减少* -> "肋骨断裂的剧痛让你几乎窒息，视网膜上闪烁着红色的生理警报。"
  + **Show, Don't Tell:**
    - **原理：** 直接告诉玩家结论是无趣的，展示过程才是文学性的来源。
    - **执行：** 不要告诉玩家 "她很生气"，要描写 "她正在签字的手指关节因为用力握紧笔杆而发白，笔尖划破了纸张"。不要说 "这里很恐怖"，要描写 "空气中弥漫着锈迹和陈旧血腥混合的味道，每一次呼吸都像是吞下了一口带刺的冰碴"。

## ⚙️ Part 2: 工程学巧思 (The Engineering Stack)

*这是我们的技术军火库。以下方案专门用于解决 LLM 的算力限制、长对话遗忘、逻辑崩坏等硬伤。*

### 1. 索引-详情分离架构 (Index-Detail Separation)

* **来源：** Gen-2 (震旦)
* **痛点：** LLM 的 Context Window (上下文窗口) 是有限且昂贵的。如果把几万字的世界观直接塞进 System Prompt，AI 会因为负荷过重而变笨 (Lost in the Middle)，且极大浪费 Token 成本。
* **解决方案：**
  + **System Prompt (CPU - 处理器):** 仅存储逻辑规则、检索指令、简略的人物索引（代号 + 核心 Tag + 触发条件）。它相当于一个目录或指针。
  + **World Book (HDD - 存储器):** 存储几千字的人物小传、详细技能表、历史设定、道具图鉴。
  + **动态调用 (Dynamic Fetching):** 仅在剧情触发特定人物或事件时，AI 才会根据 System Prompt 里的索引，瞬时读取 World Book 里的对应条目，进行“高保真渲染”。剧情结束后，释放该部分显存。

### 2. 变速箱叙事技术 (The Visual Gearbox)

* **来源：** Gen-3 (时间) & Gen-4 (契约)
* **痛点：** AI 很难把握节奏。它要么把吃饭睡觉写成流水账（导致剧情拖沓），要么在关键战斗时一笔带过（导致爽感缺失）。
* **解决方案：** 在 Prompt 中明确定义两种渲染模式，并授予 AI (或玩家) 切换的权力。
  + **Mode A (Montage/蒙太奇 - 宏观档位):**
    - *适用场景:* 长周期活动（修炼/旅行/学习/经营）、垃圾时间。
    - *指令:* 强制使用总结性语言，快速跳过时间，只报结果和关键数据变化。
    - *Example:* "寒来暑往，你在重力室里度过了三个月。汗水浸透了地板又蒸发。当你再次走出大门时，你的力量属性已突破 C 级壁垒。"
  + **Mode B (Scene/场景 - 微观档位):**
    - *适用场景:* 冲突、关键交互、战斗、初次相遇。
    - *指令:* 强制逐帧描写感官细节（光影/气味/微表情/环境音）。禁止跳过任何动作细节。
    - *Example:* "巨龙的利爪擦着你的鼻尖划过，带起的风压让你的皮肤生疼。你能清晰地闻到它口中那股硫磺与腐肉混合的腥气。"

### 3. 语法化生成引擎 (Syntax Generation)

* **来源：** Gen-4 (契约)
* **痛点：** 固定技能表导致战斗描写重复乏味；纯文字描述容易导致战力崩坏，AI 开始胡乱口胡技能。
* **解决方案：** 授人以鱼不如授人以渔。不给列表，给**词根 (Roots)** + **逻辑 (Logic)**。
  + **机制：** 定义一套“魔法语法”或“科技原理”。
  + **案例：** \* 设定龙语词根：Ignis (火), Ball (球体), Wall (墙), Ject (喷射)。
    - AI 演算：当需要攻击时，组合 Ignis + Ball = 火球；当需要防御时，组合 Ignis + Wall = 火墙；当需要推进时，组合 Ignis + Ject = 火焰喷射推进器。
  + **收益：** 这让 AI 变成了“战术计算器”，它能根据战场环境即兴组合技能，且因为遵循底层逻辑，所以显得极具合理性和创造力。

### 4. 状态差分记忆法 (State-Diff Memory)

* **来源：** Gen-2 & Gen-4 迭代
* **痛点：** 传统的 AI 总结 (Summary) 容易变成流水账，大量重复记录已知信息（如反复记录“我是男的”），导致 Token 浪费且关键信息被淹没。
* **解决方案：**
  + **增量记录 (Incremental Recording):** 仅当 Old\_State != New\_State 时才触发记录。例如，只有当关系从“陌生”变成“熟人”时才记录，否则不记录日常琐事。
  + **身份隔离 (Identity Isolation):** 针对有秘密身份的主角，分别记录“外界认知的档案 (Cover)”和“真实档案 (Truth)”。AI 在生成路人反应时调用 Cover，在生成内心独白时调用 Truth。
  + **Lore Harvesting (设定收割):** 自动抓取 AI 在对话中即兴生成的精彩设定（如某个临时编出来的节日、某种怪物的习性），并将其固化为 World Rule，防止后续吃书。

## 🛠️ Part 3: 特化功能模块 (The Modular Arsenal)

*这是我们的乐高积木箱。根据具体题材（如恋爱、战斗、解谜）按需挂载的插件。*

### 1. 命运熵增模块 (The Entropy Module)

* **适用题材：** 恋爱模拟、群像剧、模拟人生、开放世界。
* **核心功能：** 模拟世界的“流动性”。即使玩家不互动，NPC 也会升职、结婚、死亡、堕落。
* **实现路径：** 设定周期性的“世界事件检测”。如果主角长期未与某 NPC 互动，系统掷骰子决定该 NPC 的命运走向（Scenario A: 飞升 / Scenario B: 堕落 / Scenario C: 离开）。
* **目的：** 创造“错失感”和“世界真实感”。让玩家明白，这不是一个静止等待攻略的游戏，而是一个不进则退的真实社会。

### 2. 盲盒植入模块 (The Blind Injection)

* **适用题材：** 单人游玩、悬疑、解谜、推理、高NPC交互需求。
* **核心功能：** 解决单人创作者“剧透自己”的问题。利用外部工具预生成一段包含 [Hidden Truth] 的代码块（如凶手是谁、宝箱里有什么），由玩家盲复制进 Prompt，只有在剧情触发时 AI 才会解密展示。
* **目的：** 让单机玩家也能体验“剧情反转”、“被算计”和“未知探索”的快感。

### 3. 现实锚点模块 (Reality Anchoring)

* **适用题材：** 异界、高魔、克苏鲁、容易逻辑漂移的纯架空题材。
* **核心功能：** 强行在架空世界中植入现代概念的同构体（高考、学分、地铁、手机、保险、法律）。
* **实现路径：** 即使是修仙宗门，也要有“KPI考核”；即使是魔法学院，也要有“学分压力”。
* **目的：** 利用 LLM 训练数据中对现代社会逻辑的高权重认知，来锚定异世界的稳定性。防止 AI 因为缺乏逻辑参考而开始胡言乱语（幻觉）。

### 4. 混乱因子机制

* **要点：** 用于单人游玩。高NPC交互需求。多周目游玩。在人物志或系统提示词中输入特定要求，每一次生成NPC（尤其是如果像Gen2那样预先设置主要NPC的画像和阶段表现以求稳定性和细节度），在某些细节或者底层核心人格上进行修改或相反设置，比如看起来是拜金女的肤浅NPC实际上是在做社会实验或者玩心过重，看起来热情开放神经大条但实际上拥有超出常人的深度共情能力和包容心。用于在多周目游玩或预先设计过多内容后，仍能保持薛定谔的猫般的新鲜度（这个NPC究竟是不是我了解的那样？）但问题是，完全基于对话框的呈现形式可能无法很好地实现这个机制。AI只能实现“形似”效果，即在和你对话时即时生成相关内容，而非真正从源头上对NPC人格进行修改。也许需要Agent环境辅助，拆分叙事层和底层逻辑层。

## 📝 结语：产品哲学

回顾从 Gen-1 (魔都) 到 Gen-4 (契约) 的演变历程，我们实际上是在教 AI 做三件事：

1. **像程序员一样思考逻辑 (Syntax, Systems):** 建立稳固的规则，让世界不崩坏。
2. **像导演一样控制节奏 (Montage, Gearbox):** 剪辑掉无聊的片段，放大精彩的瞬间。
3. **像文学家一样感知细节 (Micro-expressions, Scent):** 用感官描写替代枯燥的陈述。

第五代 (Gen-5) 不需要推翻重来，它只需要将上述模块进行**乐高式的重新组装**。这套方法论是我们在这个赛道上的核心资产，它将确保我们无论面对什么题材，都能产出“S级”的沉浸式体验。