

雀定 — AI麻将智能算分助手

拍一张牌面，AI识别牌型、规则引擎自动算分，让麻将结算少争论、快结算、可复盘

Que Ding — an AI-powered mahjong scoring assistant: snap your winning hand, AI recognizes the tiles, and a deterministic rule engine computes an explainable score. (Formerly "Hule Ma")

项目 Project	AI 产品黑客松项目 • 2026 年 4 月 AI product hackathon project, Apr 2026	交付 Delivery	Demo 迭代 v1.0→v4.0，已上线抖音小程序，真机可扫码体验 Live as a Douyin mini-app, testable on real devices
我的角色 My Role	独立项目：一人完成产品定义（PRD、规则引擎需求文档）、用户测评与主线改版决策，并借助 AI 编程工具完成开发与抖音小程序部署 Solo project: owned all product work — PRD, rule-engine spec, user testing, pivot decisions — and shipped the build to Douyin with AI coding tools		

01

问题定义 Problem Definition

麻将是中国家庭与朋友聚会中最高频的桌游之一，但「胡牌之后」的体验长期没有被产品化解决。我们从真实牌局中提炼出三个痛点：

- **算错**：复杂番种与地区规则差异大，胡牌后心算容易出错；
- **争论**：同桌玩家计分口径不一致，结算环节频繁产生纠纷；
- **难复盘**：结果只靠口头或手写记录，整晚输赢与个人打法无法回看。

Mahjong is among the most frequent social games at Chinese family and friend gatherings, yet the post-win settlement remains unsolved: complex scoring patterns get miscalculated, inconsistent house rules trigger disputes, and results recorded verbally or on paper cannot be reviewed afterwards.

核心闭环：拍牌面 → AI识别 → 规则算分 → 生成账单 → 历史复盘

Core loop: photo → recognition → rule-based scoring → bill → history & review

目标用户：家庭/朋友聚会中的麻将爱好者；会打但不熟悉复杂计分规则的玩家；想记录战绩、了解打法倾向的高频玩家。

Target users: casual players at family/friend gatherings, players unfamiliar with complex scoring rules, and frequent players who want a personal record.

AI 产品设计：AI 负责识别，规则引擎负责确定性计算 AI Product Architecture

核心设计原则 KEY DESIGN PRINCIPLE

不让 AI 直接「猜分」。图像识别模块只负责把牌面转成标准化 tile code（含置信度与位置框），算分全部交给基于国标 MCR 81 番种表的确定性规则引擎，输出番种明细与每一番的判定证据（evidence），而非一个黑盒总分。

The AI never "guesses" a score. Vision converts the photo into standardized tile codes (with confidence scores); a deterministic rule engine built on the 81 official MCR scoring patterns computes the result, outputting an itemized fan list with evidence for each — not a black-box total.

为什么这样设计：产品要解决的核心痛点是「争论」。只有当每一分都可解释、可追溯，结算结果才有终结争议的公信力——这决定了 AI 与规则引擎的职责边界，也是整个产品最重要的一个架构决策。

Rationale: the core pain point is disputes. Only an explainable, traceable score has the authority to end an argument at the table — this single insight defined the AI / rule-engine boundary.

配套的产品层设计：

- **上下文补充流程：**81 番种中相当一部分（自摸、门前清、杠上开花、圈风刻等）无法仅凭照片判断，因此在交互上设计了用户补充牌局上下文的输入环节，而非假装「全自动」；
- **识别可修正：**识别结果逐张展示，置信度不足时引导用户手动修正后再进入算分，保证流程不中断；
- **三层演示兜底：**现场拍照不稳定 → 预置样例图；识别有误 → 手动修正入口；规则覆盖不足 → 明确声明 MVP 优先支持单一规则。

Supporting product decisions: an explicit context-input step for patterns undetectable from a photo (self-draw, concealed hand, seat wind, etc.); per-tile correction UI when confidence is low; and a three-layer demo fallback strategy (preset sample images → manual correction → clearly scoped rule coverage).

用户测评驱动的主线 Pivot User-Research-Driven Pivot

开发中期我们做了 4 人次快速测评（2 位导师 + 2 位现场同学），反馈出现了直接冲突：

用户 User	背景 Profile	关键反馈 Feedback
导师 A	麻将玩家	计分功能非常实用；建议产品名与「计分」更相关
导师 B	不打麻将	认为计分繁琐无趣，建议放大拍照功能
同学 C	麻将爱好者	很省事；建议增加新手教程
同学 D	麻将玩家	市面已有计分小程序，但认可拍照识别；建议简化为一人操作全桌

决策 DECISION

我们没有简单采纳「放大拍照」。回到真实牌桌的使用频率：**每一局都需要计分，但只有分数复杂算不清的少数时刻才需要拍照。**据此把主线改为「计分」，拍照识别降级为争议场景下的升级入口——用高频功能承载留存，用 AI 能力解决关键时刻。同时采纳了来自真实玩家（A/C/D）的可用性建议：玩家身份反馈权重高于非玩家。

We did not simply "make the camera bigger." Returning to actual table behavior: every round needs scoring, but photo recognition is only needed in the rare moments when a hand is too complex to settle verbally. So scoring became the main flow and AI recognition an escalation path for disputes — frequency drives retention, AI resolves the critical moments. Feedback from actual players (A/C/D) was weighted above the non-player's.

从 v1.0 到 v4.0 的用户心智重构：v1.0 是「设置→识别→算分→复盘」的步骤式演示页；v4.0 重构为以「记一局」为中心的牌局心智——创建/加入牌局 → 连续单局结算（确认后停留在记分页，不打断牌局节奏）→ 终结整盘 → 保存完整快照（规则、局数、四家得分、单局流水）到历史。导航简化为「首页 / 记一局 / 我的」三个 tab。

v1.0 was a step-wizard demo (setup → recognize → score → review); v4.0 rebuilt the mental model around an ongoing game session: create/join a room → consecutive round settlements (staying on the scoring page so the game flow is never interrupted) → end the session → save a full snapshot to history. Navigation collapsed to three tabs.

04

MVP 取舍：少做功能，多做闭环 Scope Decisions

✓ 做进闭环 IN SCOPE (P0)

- 规则与基础点数选择
- 牌面识别（含手动修正）
- 规则引擎算分 + 番种明细
- 单局账单 → 历史记录
- 轻量复盘（牌风标签、趋势）

✗ 主动砍掉 CUT

- 完整登录注册体系
- 真实支付 / 红包 / 转账
- 多人实时对局同步
- 全地区规则一次性覆盖
- 完整复盘动画、复杂社交系统
- 拍照作为主功能（见 03 节）

取舍逻辑：「少做功能，多做闭环；少讲平台，多讲一个真实高频场景。」黑客松现场优先证明最核心的算分能力，排行榜、社交、完整复盘留作自然扩展——不是不知道有价值，而是先把最痛、最高频、最可演示的闭环跑通。

The principle: fewer features, one complete loop; not a "platform" pitch, but one real high-frequency scenario. Leaderboards, social features and full replay were consciously deferred — not unknown, but deprioritized to prove the core scoring loop live.

05

落地与验证 Delivery & Validation

- Web/H5 Demo 迭代至 v4.0，完整跑通「创建牌局 → 连续记分 → AI 识别算分 → 终盘快照 → 历史复盘」全流程；
- 上线为**抖音小程序**，真机扫码即可体验完整链路（支持抖音 App 38.2.0+）；
- 设计 3 分钟现场演示脚本，每条演示路径均有兜底方案；
- **AI 工具链实践**：开发中发现 AI 编程工具无法直接识别麻将牌、也不会算番，于是引入 GitHub 开源麻将识别包、并将国标规则文档喂入作为计算依据——识别 AI 能力边界并用工程手段补齐，是这次落地中最真实的 AI 产品经验；
- **品牌升级**：采纳测评中导师反馈（产品名应与「计分」强相关），将产品从「胡了吗」更名为「雀定」——取「雀」之麻将、「定」之一锤定音，并产出配套品牌海报。

The demo iterated to v4.0 covering the full loop and went live as a Douyin mini-app testable on real devices, with a 3-minute demo script where every path had a fallback. Responding to mentor feedback that the name should signal "scoring," the product was rebranded from "Hule Ma" to "Que Ding" — roughly, "the mahjong verdict."



抖音扫码体验完整产品

Scan in Douyin to try the live product

三条最重要的反思：

- **规则资产的选型与用户场景错位。** MVP 选择了体系最完整的国标 MCR 规则（81 番种、可解释性最强），但目标用户是家庭局和朋友局，他们打的是广东推倒胡、四川血战这类地区规则。技术上「最严谨」的选择不等于用户「最需要」的选择——规则库应该从用户场景倒推，而不是从工程完整性出发。
- **核心假设没有被数据验证。** 主线 Pivot 建立在「计分高频、拍照低频」的频率假设上，这个判断来自测评访谈和常识推演，但没有留存数据支撑。哪怕在黑客松周期内，也应该提前定义一个可衡量的代理指标（例如演示用户中主动发起第二局的比例）来检验假设，而不是只交付功能。
- **对商业化保持清醒：这个产品短期内不具备独立上线的条件。** 测评用户 D 指出市面已有计分小程序——纯计分是同质化红海，壁垒低、付费意愿弱。本产品真正的差异化在「AI 识别 + 可解释算分」终结争议的能力，但识别准确率距生产可用还有真实差距（缺少真实牌面数据集、置信度阈值与降级策略的系统设计）。如果要真正推向市场，必须先回答「用户为什么不用现有小程序」，并补齐识别侧的数据基建——在此之前，它更适合作为验证 AI 产品方法论的练手项目，而不是仓促上线。

若继续推进的路线：地区规则库（贴近真实用户）→ 真实牌面数据集与识别准确率基线 → 群组排行榜与战绩比拼（社交留存）→ 语音输入与完整 AI 复盘。

Three reflections: (1) the rigorous MCR rule set mismatched casual home players who use regional rules — rule assets should be derived from user scenarios, not engineering completeness; (2) the pivot rested on an unvalidated frequency hypothesis — even in a hackathon, a measurable proxy metric should have been defined; (3) a sober commercial judgment: standalone scoring apps are a commoditized space, the real differentiation is dispute-ending explainable AI scoring, and recognition accuracy still has a genuine gap to production (real-photo dataset, confidence thresholds, degradation design) — the honest answer to "why not existing mini-apps" must come before any launch. If continued: regional rules → recognition data baseline → group leaderboards → voice input and full AI replay.