技术分享 · 公开资料整理

AI 在银行的落地实践

算力 · 算法 · 数据 · 场景：AI 如何从单点工具，变成一套能考核的体系能力

全部内容来自公开披露信息，具体数据和案例以公开信息为准

目录

CONTENTS

体系、实践、思考。

AI 在银行的落地实践

01 体系：AI 能力的底座 P03

算力 145.69 PFLOPs、企业级知识库 5 亿条向量、近 400 个场景，以及行长对外怎么讲这件事——先看基建，再看应用。

02 实践：浙江省分行的三个阶段 P08

准备（2022–2023 技术 · 培训 · 原型）→ 寻找突破（2023–2025 场景 · 实践 · 方向 · 步骤）→ 全面规划和推广（2025 年起）。

四个案例：智能运维底座、客户经理减负、流程再造、知识产权估值。

03 思考 P20

从工具到能力的三级跳、三条与模型选型无关的经验、安全合规的边界，以及给三类角色的不同建议。

01

体系：AI 能力的底座

算力、算法、数据三层基建决定 AI 能走多远；场景层，只是水面上露出来的那一部分。

四层结构：从算力到场景

场景层	近 400 个场景应用，覆盖 6 大业务领域	**398 个**
数据层	企业级知识库：60 余类数据源、向量数据持续增长	**5 亿条+**
算法层	大模型与小模型协同，生成式与决策式 AI 混合	**千亿级**
算力层	“四地五中心”高可用、高弹性智算集群	**145.69P**

渠道服务 · 客户经营

对话式交互重塑服务体验：面向客户主推主动识别、快速响应、自然交互；面向员工，“帮得”面向零售与普惠条线全体基层客户经理，提供客户洞察、产品推荐、财务分析到客户触达的全流程支持。

运营管理 · 支持保障

网点柜面问题响应的 AI 助手覆盖率已达 99.42%，日均访问量超过 10 万人次；开发侧有 75% 的开发人员开始使用智能编码，代码采纳率在 30% 以上。这些变化直接落在可考核的指标上。

风险管理 · 产品服务

构建“AI + 风控”的全流程管理模式，在授信审批、制度大脑、智能风控等方面取得实质性进展；依托生成式大模型，2025 年审批业务受理量实现两位数增长，平均处理时间下降超过 30%。

算法与数据：混合模型体系 + 数据整理

评级审查意见生成

40 秒

授信审批 AI 助手覆盖财务分析与评级审查，报告从数小时降到分钟级；审批处理时间下降 30%+。

60 余类

企业级知识库已接入的数据源——散落的非结构化数据先清理干净。

5 亿条+

知识库向量数据规模，仍在逐日增加；数据全域入湖同步推进。

真正跑通的是混合体系

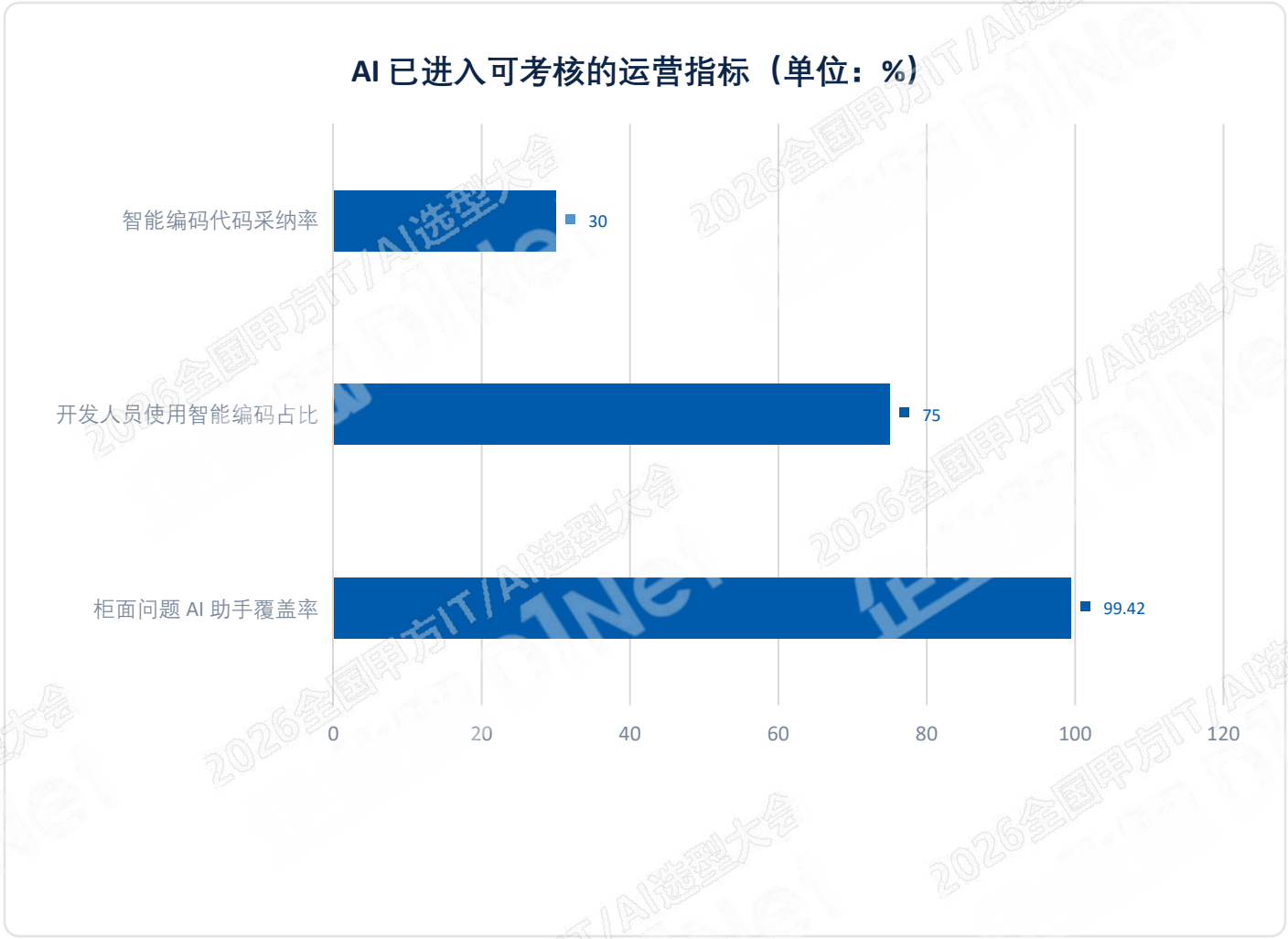
引入 DeepSeek、千问、智谱等生成式大模型，同时建设千亿级参数的建行大模型——大模型与小模型协同、生成式与决策式 AI 混合。决定效果的不是单一最大模型，而是模型分工与场景的匹配度。

JUDGEMENT

模型能力是公共品，数据整理才是差异化

官方表述很直白：AI 有没有潜力，主要看前期数据整理得细不细。

场景：近 400 个，覆盖 6 大业务领域



数据口径：截至 2025 年末；受理量增长为“两位数”表述，图中以百分比口径指标并列展示。

SCALE

99.42%

网点柜面问题响应中，由 AI 助手先给出回答方案的比例；日均访问量超过 10 万人次。

200 → 600 人

专营客户经理平均可深度服务的客户数——AI 没有替代客户经理，而是放大了他的服务半径。

来源：建行 2025 年度业绩发布会问答实录

行长怎么讲 AI：已经写进三年规划

KEYNOTE

对外表述的三个关键词

战略级

《“人工智能+”三年规划》，全行层面的战略安排

平台化

智能体平台 + “AI 众创工作室”，业务人员自己造工具

人机协同

“人 + AI”协同作业模式，人是决策主体

2025 • 成方金融科技论坛

2025-10-28

「打造数据驱动引擎，提升商业银行竞争力」

讲的是数据治理与数据资产，不是模型参数——和第一部分“数据层”那一页的判断一致。

2026 • 中期业绩发布会

2026-08-28

「600 多个应用场景，“人 + AI”协同作业模式」

助理“帮得”月活超 4 万人、累计交互近 9000 万次，带动资产规模新增提升 57 个百分点。

2026 • 中国—东盟博览会

2026-09-18

「AI 赋能金融合作，三个贯通」

能力贯通、场景贯通、生态贯通；AI 能力开始作为对外合作的接口，而不只是内部工具。

来源：2026 年中期业绩发布会、2025 成方金融科技论坛、2026 中国—东盟博览会公开报道

02

实践：浙江省分行的三个阶段

从总行的体系回到分行的一线。先用两年打地基，再用两年找突破，2025 年起全面铺开。

01 准备阶段 2022 – 2023	02 寻找突破 2023 – 2025	03 全面规划和推广 2025 年起

阶段一·准备（2022–2023）：先把地基打好

PHASE 01

准备阶段

2022 – 2023

技术 · 培训 · 原型

这两年没有“AI 项目”这个名目。
攒的是设备与算力、认知与人才，
以及一个能被业务认领的框架。

技术
先有底座

算力、模型、数据，技术条件和路线

培训
先换脑子

专业人员、管理和业务人员

原型
先挂上去

找到合适的人、合适的资源、合适的方向（写作、知识问答、数据分析、业务痛点）

阶段二·突破（2023–2025）：场景、实践、方向、步骤

01

场景

从总行已有的场景里挑能承接的、有可能的

02

实践

技术运维、开发；客户经理减负；网点赋能；业务拓展

03

方向

结合全行数字化经营业务主线

04

步骤

先内部、后外部；先减负、后赋能；先辅助、后替代

判断

分行的优势不在算法，而在网点、账户体系与地方政府数据接口——这三样总行替代不了。所以浙江的 AI 是从“接接口”开始的，不是从“训模型”开始的。

案例一·智能运维底座：先让全省 IT 资源“看得见、管得住”







案例二·银行 AI 的第一批成功场景，往往是员工减负

引入之前：客户经理的重复劳动

- 收取纸质报表
- 手工逐项录入
- 发票逐张核验
- 汇总撰写材料

这些环节技术含量不高，但占用了客户经理大量时间，也最容易出错。

引入之后：助手把可结构化的活接走

- 财务报表智能处理
- 混合发票批量核验
- 客户经理只做判断与确认

引进大数据与智能 AI，把客户经理从简单重复劳动中解放出来，从业务端与管理端双管齐下。

NEXT

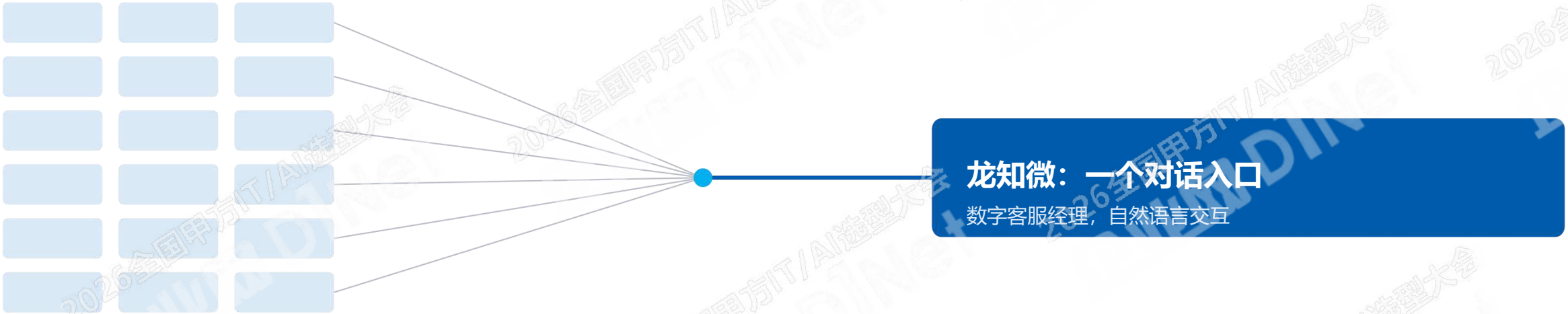
后续聚焦的五个方向

01 深耕客群
02 便利操作
03 精细管理
04 智能风控
05 效能提升

> 先让员工少做无意义的活，AI 才有机会进入真正需要判断的环节。

来源：浙江日报 2024-09-24

案例三·把 1 万个服务入口，收敛成一个对话入口



怎么做的

“龙知微”是建行打造的虚拟数字人 IP，以真实员工为原型，融合动作捕捉、三维建模与实时渲染技术，具备“准确答、秒回复”“零接触、远程办”“大字体温情服务”三大特点，可提供 7×24 小时不间断服务。

覆盖到什么程度

支持 200 余项业务的自助查询与办理，覆盖个人、对公、信用卡等 1 万余个服务场景；已应用于部分分行的智慧柜员机，为客户提供从咨询、办理到陪伴的一站式服务，问题解决率达 98%。

浙江省内落地

温州分行与建行远程智能银行南宁分中心合作打造，在瓯海支行营业厅大屏在线值守；杭州分行把它用于金融消保宣教，走进“两司两员”观影活动与浙江科技大学开学第一课。

来源：央广网 2026-03-18、金融时报 2026-09-10

案例四·让“说不清”的专利，变成可授信的资产

2025 年预授信金额（评估模型与人工智能深度融合后）

1600 亿元

浙江是唯一获批专利密集型产品培育和推广国家试点的省份，认定的专利密集型产品数量居全国第一。

600 亿元+

“建智贷”融资服务累计授信规模

415 项

首批完成价值评估的专利密集型产品，对应意向授信 300 亿元

AI 解决的是“估值”，不是“审批流程”

一台手机上可能集成数百上千个专利，单个专利难以估值、也难以处置，因此长期无法作为主担保。浙江省专利密集型产品培育和推广中心与银行联合设计开发的全国首个专利密集型产品价值评估模型，改为“组合专利 + 产品”的评估路径，用产品销售现金流反推专利组合估值，再用人工智能批量放大评估能力——这才让 415 项产品在三周内变成可授信的额度。

来源：央广网 2024-09-27、浙江日报 2025-11-13

阶段三·全面规划和推广（2025 年起）

PHASE 03

全面规划和推广

2025 年起，AI 从“各条线自己试”收拢为统一的战略安排：明确《“人工智能+”三年规划》，建设智能体平台，组建“AI 众创工作室”，把零散的场景做成可复用的资产。

600+ 个场景

平台层：让业务人员自己造工具

智能体平台

把场景沉淀为标准化的能力单元，一次建设、多处复用

“AI 众创工作室”

业务人员以对话和拖拽搭建应用，需求方与建设者的边界消失

三种开发模式

零代码、低代码、高代码并行，按场景复杂度选路径

来源：建设银行 2026 年中期业绩发布会公开报道

成效层：指标已经在动

4万	助理“帮得”月活跃用户超 4 万人，累计交互近 9000 万次
57	个百分点——“帮得”带来的客户资产规模新增提升幅度
3h	普惠金融“一镜到底”：采集 20 分钟，流程从 2–3 天压到约 3 小时
600万	普惠客户受惠于流程缩短，服务覆盖面进一步扩大

来源：建设银行 2026 年中期业绩发布会公开报道

来源：建设银行 2026 年中期业绩发布会公开报道

阶段三·分行「人工智能+」建设思路



承接

《建行“人工智能+”三年规划》→ 四大体系总分行协同 → 分行行动方案

路径

五层递进架构；三类能力分类施策：总行引入 / 分行自建 / 总分共建

目标

2026 年起分步提升场景覆盖率，最终全场景 AI 化（具体指标已脱敏）

「AI 大脑」：AI 底座总体设计思路



总体框架

一个中心、两条主线、三大目标、四维路径、五项原则

五层递进

总行 AI 基础层 → 分行 AI 能力层 → 分行 AI 公共服务层 → 核心服务层 → 业务场景层

三大属地目标

智能场景规模化覆盖 · 本地 AI 安全管控闭环 · 区域特色知识资产沉淀

浙江分行的收益，是多做了原本做不了的业务

15 / 12

智能运营中心累计“揽活”项目 / 已全部或部分落地

153 人次

约为全行每年节约的工作量

全流程线上化・智能化办理

系统内首家引入服务水平协议（SLA），企微客群集约化运营两个月内完成四期建设。

2025 年全年，为浙江省内科技企业投放贷款

2057 亿元

同年，分行与浙江省市场监督管理局（知识产权局）共同设计全国首个专利密集型产品价值评估模型；嘉兴分行落地首单以专利权作为主担保模式的授信业务。科技企业的评价方式，从看“三张财务报表”扩展到第四张“科技评价表”。

衡量标准

不是“省了几个人”，而是“原先因为估值难、尽调贵而做不了的业务，现在能做多少”。

来源：央广网 2025-06-24

03

思考

把前面看到的做法抽象成几条能带走的判断：怎么衡量落地进度、哪些经验可迁移、边界在哪里，以及给不同角色的具体建议。

三条可迁移经验：与模型选型无关

01

场景优先

先挑输入输出明确的活。财务分析、发票核验、开户预约这类任务，判断标准清晰，模型输出的对错肉眼可验，适合作为第一批场景。反过来，先说“我们要用大模型改造全行”，通常推不动。

02

知识库做底座

60 余类数据源、5 亿条向量，是长期工程而不是一个项目。它决定了下一次做新场景时，是复用还是从头再来。公开数据里最容易被忽略、却最难追赶的，就是这一层。

03

人机协同，人做决策主体

官方表述是“确保使用人作为决策主体的责任人”。AI 给建议、人签字，责任链条不模糊。这条界线划清楚之后，业务部门才敢真正把 AI 的产出用进流程里——否则它永远停留在“参考一下”。

这三条都不依赖模型选型，但都要过同一道门：安全合规——分级授权、输出可审计，人始终是决策主体。它们依赖的是组织愿意先花时间，把数据和流程整理清楚。

启示：给三类角色的三条不同建议

给技术团队

难点在数据与集成，不在模型。取数链路、系统对接、权限与审计、历史脏数据清理，占掉的工作量远超模型选型与提示词调优。如果把 AI 项目当成“选个好模型然后接进去”，预算和排期大概率会失控。

给业务团队

先解决员工减负，再谈客户惊喜。从公开案例看，效果最确定、内部阻力最小的场景，几乎都是先帮客户经理和柜员少干重复的活——先把内部信任建立起来，再往外推。

给组织

让业务人员成为建设者。智能体众创平台支持零代码、低代码、高代码三种模式，不具备编程基础的业务人员也能通过对话交互、积木式拖拽搭建智能应用——需求方与建设者的边界正在消失。

银行 AI 的进展，最终体现在有多少原本做不了的业务，变得可做了

而支撑这一点的，是算力、是数据、是流程，更是愿意先把基础整理清楚的那几年。