

金融行业

Agent百景图

Agent 100 in Financial Industry

New Finance New Future





作者

张翘

阿里云智能集团公共云事业部副总裁

新金融行业总经理

简介

张翘先生现任阿里云智能集团公共云事业部副总裁、新金融行业总经理，负责阿里云智能从银行、保险、证券到互联网金融等泛金融行业客户。曾在蚂蚁金服智能科技团队负责蚂蚁金服科技产品的开放合作，推动金融科技产品化从而帮助金融机构实现数字化转型，在大模型、分布式架构、移动 app、大数据、人工智能、区块链等领域有着深入的实践和理解。

自 17 年开始推动蚂蚁金服和阿里云新金融的金融科技技术输出项目，先后领导负责了工行、农发行、光大银行、南京银行、四川农信、中国人寿、中国人保、中华保险、财通证券、国银租赁等不同类型的金融行业重要的项目规划与建设。

在加入阿里云和蚂蚁金服前，张翘先生先后任职于甲骨文、Pivotal，领导参与了多个重大项目的建设，拥有丰富的企业架构设计、IT 战略规划、产品研发及团队管理等专业经验。

前言

河海不择细流，故能就其深。2025 年是国务院《新一代人工智能发展规划》发布的第 8 个年头，也是国家层面提出通用人工智能的第 2 个年头。通用人工智能是新质生产力的典型代表，当下已成为全球科技竞争的制高点。当 DeepSeek 春节期间在全球引爆开源革命，当 Qwen2.5-Max 在 Chatbot Arena 实现智能跃迁，我们恍然惊觉，中国力量已然成为这场科技革命中的引领者，并引发了产业界的系列变化：

- 杰文斯悖论的延伸：大模型能效的指数级提升非但不会降低算力的消耗，反而会催生出更复杂的智能应用场景，从而激发更大的算力需求，如同 19 世纪煤炭效率提升刺激更大能源需求那般；
- 梅特卡夫定律的演变：开源生态存在显著优势，更多的研究者和开发者通常会带来更多的使用场景、更快的模型算法创新、更丰富的工具链产品，从而全方位推动技术进步。

过去的一年里，我们看到金融机构的大模型应用已从局部试水迈向全场景深度渗透，从内部提效转向外部赋能，从知识问答、办公辅助等传统场景向投研决策、信贷审批等复杂场景延伸，以招商银行 x 阿里云为代表的合作，已实现大模型在智能投研、财富管理等高价值环节的专家级推理能力。应用场景的深入带来模型架构的持续进化，通用大尺寸模型与专业小尺寸模型并行，通过模型训练微调注入专属领域知识，提升专业能力，降低推理成本，在有限的资源下推动应用快速落地。应用场景的丰富特别是对客服场景，对算力需求以及系统开发和服务能力提出了更高要求，金融机构的算力基建进入攻坚期，这场算力竞赛的成败亦将决定企业未来的智能化水平，乃至在市场中的竞争能力。

这场始于岁末年初的智能革命，正在书写金融科技史上前所未有的范式迁移。未来企业的大模型架构是一个 MoA (Mixture-of-Agents) 混合架构，即一个强大的通用模型加一个强大的推理模型，通常是千亿或万亿参数，采用 MoE 架构，作为企业智能基座；N 个中小尺寸的稠密模型，用在不同场景 Agent 当中，可以通过模型蒸馏提升其领域能力；工具链层面，一个训练推理平台加一个 Agent 工厂，构建企业级的数据飞轮，基于在若干 Agent 在场景中沉淀的数据，使用基座模型对领域模型进行蒸馏；当这些 Agent 积累的数据不断丰富后，基座模型也能够通过强化学习进一步增强，逐步演变成企业内部自有的世界模型。这种模型架构与开发范式，不仅是对传统业务流程的一次重大升级，也为金融机构探索新的商业模式和服务模式提供了可能。

为助力金融机构依托大模型实现业务创新突破，我们精心梳理过去一年与众多金融机构携手共创的业务场景，将其整理成册。这本册子凝聚着实践经验与创新思路，期望能为金融机构在大模型应用浪潮中找准方向，开启业务拓展与升级的新篇章。

张翊

阿里云智能集团公共云事业部副总裁
新金融行业总经理

目录

01 金融场景创新的 AI 时代

开篇

时代背景	01
Agent 是大模型落地的最佳形态	03
金融行业 Agent 落地典型实践	04
金融机构大模型落地路径	08
金融行业 Agent 发展趋势	09

02 金融行业 Agent 百景图

银行

信贷 (8 Agent)	11
风控 (3 Agent)	27
AI Native 手机银行 (4 Agent)	33

证券

投研 (7 Agent)	41
投行 (2 Agent)	55
投顾 (3 Agent)	59
智能运营 (2 Agent)	65

保险

产品开发及销售 (7 Agent)	69
核保核赔 (3 Agent)	83
监管合规 (3 Agent)	89

通用

智能客服 (14 Agent)	95
智能用数 (10 Agent)	123

知识助手 (3 Agent)	143
研发助手 (6 Agent)	149
数字人 (4 Agent)	161
内容审核 (3 Agent)	169
信息检索与打标 (5 Agent)	175
培训陪练 (5 Agent)	185
办公助手 (4 Agent)	195
营销助手 (4 Agent)	203

03

结语

211

04

附录

212

金融场景创新的AI时代

金融行业政策指引与AI技术发展双轮驱动	01
Agent是大模型落地的最佳形态	03
金融行业Agent落地典型实践	04
金融机构大模型落地路径	08
金融行业Agent发展趋势	09



金融行业政策指引与 AI技术发展双轮驱动

十四五规划纲要提出，稳妥发展金融科技，加快金融机构数字化转型。2023年10月，中央金融工作会议明确，金融机构要加快数字化转型，提高金融服务的便利性和竞争力。2024年1月，习近平总书记在省部级主要领导干部推动金融高质量发展专题研讨班上的讲话指出，要加快金融数字化、智能化转型。2024年11月，人民银行联合7部委印发了《推动数字金融高质量发展行动方案》，旨在通过金融数字化、智能化转型提高金融服务的便利性和竞争力。

在数字经济浪潮汹涌的历史背景下，在**金融数字化、智能化转型的发展大势下，金融机构数字化转型是不可避免的，不存在转不转的问题，必须坚定不移推进**，积极探索与自身特点相适应的转型路径，实现自身的可持续发展。数字化转型不是纸面上的鸿篇巨制，而是一线工作的针头线脑。金融机构要写好自己的小作文，并建立与数字化转型战略相适应的协同推进、进度管控与成效评价机制，确保任务可执行、进度可跟踪、效果可评估，形成从规划设计、执行到优化的“闭环”，真正将战略认知力转变为变革推动力。

自 ChatGPT 推出以来，生成式人工智能经历了显著的技术进步和扩展。

● **认知能力方面**，模型规模持续扩大，在 Scaling Law 的驱使下，智能水平持续增强，使得大模型在自然语言理解、文本生成、逻辑推理、工具调用等方面达到了前所未有的水平，加上推理模型的出现，加速了 GenAI 在复杂金融场景的落地。

● **感知能力方面**，端到端的多模态处理成为生成式人工智能发展的另一重要方向，早期的大语言模型专注于文本输入和输出，现在端到端的多模态大模型已经能够处理和整合多种数据类型，如文本、图像、音频等，在数字人直播、营销物料生成、智能语音客服等更广泛场景得到应用。

近期 OpenAI o1 和 DeepSeek R1 这类推理大模型的出现，得到市场广泛好评，标志着生成式人工智能领域的一个重要转折点，具体体现在几个方面：

1) **Scaling Law 依然有效，从 Pre-Training 阶段扩展到了 Post-Training 和 Test-Time 阶段**，强化学习技术的创新性应用为不断迭代升级模型能力带来了新的方向；

2) **开源模型成为世界顶级模型**，在 Qwen 和 DeepSeek 为代表的开源模型推动下，模型创新和应用创新将持续涌现；

3) **推理算力需求激增**，模型的开源开放带来了大模型时代的“杰文斯悖论”，算力应用效率的提升，带来整个社会对算力需求的增长；

4) **企业大模型应用范式向“大尺寸 MoE 模型 + 小尺寸蒸馏模型 + 数据飞轮”转变。**

Agent 是大模型落地的最佳形态

构建于大模型之上的 Agent，不仅具备强大的语言理解和生成能力，还能够通过规划、记忆、工具和行动四大关键组件实现复杂的任务处理。规划模块负责拆解复杂任务为可执行的子任务，并评估执行策略；记忆模块支持短期对话上下文保持和长期用户特征存储；工具模块让 Agent 可以调用外部 API 或插件来扩展自身功能；而行动模块则是将这些规划转化为实际操作的工作流。这种多组件灵活编排的架构设计使得大模型 Agent 能够灵活适应各种业务需求，同时提供高度定制化的服务。

金融行业因其信息密集、逻辑严谨及监管严格等特点，对大模型技术落地提出了极高的要求。无论是客户服务、风险管理、还是经营决策，每一个环节都需要精准的数据分析、逻辑判断、快速响应等能力。面对这类复杂场景，单靠大语言模型本身往往很难应对，大模型 Agent 凭借其环境感知、决策规划、调用工具及执行行动的能力，有效应对上述挑战。

当把大模型应用在金融行业具体场景当中，Agent 的上述能力特点得以完美体现。一方面，基础大模型提供了强大的语义理解和内容生成能力，支持高效率 and 准确的数据处理工作，并且其推理和规划能力，能够将复杂任务进行拆解，分而治之，更重要的是，相比小模型，大模型具备很强的泛化性，能支持多类场景的快速验证与落地；另一方面，以 Agent 为核心的多智能体架构允许其根据不同的业务场景需求，调用已有的系统接口和工具，充分复用已有的 IT 资产积累，以适应不同场景下多样化需求。可以说，大模型 Agent 不仅是金融行业数字化转型的理想选择，更是推动现有 IT 体系向智能化迈进的关键力量。

金融行业 Agent 落地典型实践

金融是国之重器。2023 年 10 月，中央金融工作会议首提“金融系统要做好科技金融、绿色金融、普惠金融、养老金融、数字金融五篇大文章”，为金融如何服务实体经济高质量发展指明了方向。在这一背景下，大模型技术的应用显得尤为重要。大模型具备强大的数据处理能力、自然语言理解和生成能力，以及复杂任务执行能力。它们能够为金融机构提供精准的风险评估、个性化的客户服务、智能化的投资顾问等解决方案，从而助力“五篇大文章”的实施。

1 普惠金融 - 蚂蚁集团

蚂蚁集团自成立以来，一直致力于通过科技创新推动普惠金融的发展，旨在为全世界提供平等的金融服务机会。其使命是构建基于互联网科技的全球开放式金融服务生态系统，使每一位普通消费者和小微企业都能够随时随地获得便捷、安全且低成本的金融服务。蚂蚁集团利用大数据、云计算及人工智能等先进技术，不断探索新的服务模式和服务边界，力求覆盖更广泛的人群，特别是那些传统金融服务难以触及的群体。

在这一愿景下，蚂蚁集团推出了两款面向金融领域的大模型产品——蚂小财与支小助。蚂小财是一款智能理财助理，而支小助则是面向金融行业专家的智能业务助手，两者均采用了最新的大模型技术进行深度定制，以满足不同用户的需求并提升服务质量。

蚂小财作为一款针对个人用户的智能理财助理，它不仅能够帮助用户解读市场信息，还能结合用户的财务目标和个人偏好给出个性化的配置分析。此外，蚂小财还能够主动预测用户需求，适时地提供相关专业服务，使得用户体验更加人性化和智能化。据了解，“蚂小财”连接了基金公司、券商、财经媒体等 200 多家主流机构，以及超 1.5 万位财经创作者，提供专业的服务和内容。AI 并不改变金融本质，提升是服务的普惠度和可得性。截至 2024 年 8 月底，“蚂小财”的月度活跃用户数已经达到 7000 万人，其中有 45% 来自三线及以下城市。

支小助则专注于为金融行业的专业人士提供支持。这款智能业务助手包含多个专业版本，如“服务专家版”、“投研专家版”等，分别服务于理财顾问、保险代理等多个角色。例如，“投研支小助”每天可以帮助每位分析师高效处理超过 100 篇研究报告的信息提取工作，显著提升了工作效率。

支小助的设计理念是为了让金融机构的专业人士能够更好地利用数据和技术手段，从而提高他们的生产力和服务质量，最终惠及更多的客户群体。

随着生成式人工智能时代的到来，大模型技术在金融行业的应用前景显得尤为广阔。无论是对于普通消费者还是专业金融从业者而言，大模型都能够带来前所未有的体验变革：更自然的交互方式、更丰富的信息供给、更有效的表达形式以及更加贴身的服务定制化。未来，我们可以期待看到更多像蚂小财和支小助这样的创新产品出现，它们将共同推动整个金融行业向着更高层次的智能化迈进，同时也将进一步扩大普惠金融的服务范围和服务深度。

2 科技金融 — 网商银行

网商银行作为中国首批民营银行之一，自成立以来便专注于利用数字技术为小微企业提供金融服务。其核心使命是通过创新性的金融科技手段，解决小微企业融资难的问题，尤其是那些传统金融机构难以覆盖的长尾小微企业。凭借阿里巴巴集团和蚂蚁集团的技术能力和生态资源，网商银行致力于打造科技金融服务新模式，使得每一位小微经营者都能够享受到便捷、高效的金融服务。

大雁系统是网商银行基于数字化技术开发的一套产业链金融解决方案，旨在满足小微企业在供货回款、采购订货、铺货收款、加盟、发薪等生产经营全链路中的信贷需求及综合资金管理需求。该系统不仅帮助解决了供应链上下游小微企业的融资难题，还提高了整个供应链的资金流转效率。大雁系统的推出标志着网商银行在服务小微企业方面迈出了重要的一步，并逐渐成为行业内的标杆案例。

随着 AI 大模型技术的发展，大雁系统得到了显著增强。针对小微科创企业数据获取难、标准化程度低、噪声多等问题，大雁系统提出了依赖于大模型的产业认知框架，通过多任务统一知识抽取引擎和大规模知识校验的人机协作链路等技术手段，有效提升了服务效率和风险控制能力。例如，大模型能够自动读取大量研报、研究产业趋势，形成对小微科创企业的精准经营画像及信用评分，从而更好地服务于产业链上的“毛细血管”企业。这些技术的应用不仅让更多的小微科创企业得以被看见，更让他们能够得到更加准确的资金支持，促进了产业链的整体健康发展。

大雁系统应用大模型后，取得了令人瞩目的成绩。截至 2024 年 12 月底，网商银行已通过大模型搭建了包括汽车、医疗、建筑在内的 9 条产业链图谱，识别超 2100 万产业链上下游的小微企业。此外，大雁系统已经成功为超过 58 万家科技型小微企业提供了信贷额度，授信金额超千亿元，覆



盖全国 31 个省份，涉及国家重点鼓励的高新技术领域如电子信息技术、生物与新医疗技术、新材料技术等。不仅如此，借助大模型的力量，大雁系统大幅提升了小微科创企业的信用画像效率，使得金融服务的覆盖面和可得率得到了前所未有的提升。

进入生成式人工智能时代，大模型在“数字金融”领域的应用前景无限广阔。它不仅可以进一步优化现有的金融服务流程，提高服务效率和服务质量，还能通过深度学习和数据分析挖掘出更多潜在的价值信息，助力金融机构作出更为科学合理的决策。未来，随着大模型技术的不断进步和完善，我们有理由相信，像大雁系统这样的创新产品将继续引领行业发展潮流，为实体经济注入新的活力，推动金融服务向着更加智能化、个性化、定制化的方向发展。同时，这也将促使金融机构不断创新业务模式和服务方式，以适应快速变化的市场需求和技术环境。

3 养老金融 - 某商业银行

某商业银行长期以来致力于为老年人提供专业、全面的金融服务，确立了自己在养老金融领域的领先地位。作为市养老金发放的主要金融机构之一，某商业银行不仅承担着养老金客户的代发任务，还不断推出针对老年群体的专属理财产品和服务，充分体现了对老年客户需求的关注与尊重。

随着大模型技术的发展，某商业银行积极探索如何将这一前沿技术应用于养老金融服务中，特别是

开发了一款基于对话式用户界面(LUI)的AI Native APP。这款APP代表了从图形用户界面(GUI)向语言用户界面(LUI)的重要转变,特别适合视力下降或不熟悉智能手机操作的老年用户。通过支持语音输入,老年人可以更加自然地与APP进行交互,无需担心复杂的菜单导航。结合大模型强大的自然语言处理能力,该APP能够理解并回应用户的口语化指令,例如查询账户余额、转账汇款或是咨询最新的理财产品信息。此外,它还能根据用户的偏好和历史行为提供个性化的建议和服务提醒,如按时缴纳水电费、预约体检等实用功能。



在银发经济时代,大模型技术的应用极大地增强了金融服务的便利性和用户体验。传统的手机银行APP经过多年发展,已经变得功能繁多而复杂难用,尤其是对于老年用户来说,过多的功能堆砌反而增加了使用的难度。借助大模型技术,能够帮助老年用户快速智能筛选出最相关的信息和服务,并以直观的方式呈现给用户。例如,在养老金发放日,APP可以通过语音提示告知用户资金已到账,并询问是否需要查看详细账单或者安排定期存款计划。这种智能化的服务方式不仅简化了操作流程,也使得老年人更容易接受金融服务,帮助他们跨越“数字鸿沟”,享受科技带来的便捷生活。

总结而言,利用大模型技术开发的适老化AI Native APP标志着养老金融服务进入了一个新时代。通过对话式的交互体验,这款APP不仅解决了老年用户使用传统手机银行APP时遇到的问题,还通过智能化的服务提升了用户体验,真正做到了让每一位老年人都能轻松享受到高效、安全的金融服务。随着技术的持续进步,我们可以预见未来会有更多类似的创新产品出现,进一步推动养老金融领域的发展,满足老年人日益增长的美好生活的需要。

金融机构大模型落地路径

大模型这类新兴技术在金融机构中的落地，遵循应用价值驱动的原则，采取自上而下的建设方式，从应用场景出发，到对模型能力的要求，再到对算力基础设施的规划。这意味着企业首先需要明确应用场景和业务价值，选择真正有影响力的场景落地，再结合场景需求和模型特点，选择合适的模型架构，再考虑 GPU 算力卡等硬件资源的采购。否则会导致资源估算不准确，甚至没想清楚大模型能带来怎样的应用价值，就动手开始做，导致时间和资源的浪费，错过了技术红利期。因此，从实际问题出发，识别出关键落地场景，才能确保投资的有效性和最终的项目成功。

以客服系统为例，这一类知识密集型的场景，大模型在整个客户服务流程的不同阶段都可发挥作用。事前，大模型基于历史通话记录生成会话摘要，帮助客服人员快速了解客户的历史沟通情况，提高服务效率和服务质量。事中，通过更精准的意图识别，大模型实时分析客户的对话内容，感知客户情绪，捕捉客户需求，并提供个性化的话术推荐，增强客户体验。事后，大模型还能对通话内容进行质检，自动评估客服的表现并提出改进建议。在明确了这样一个场景后，可以再往下估算资源需求，假设某银行拥有 1000 个坐席，每个坐席平均每天处理 50 通电话，每通电话调用大模型 5 次，那么理论上该银行每天将产生约 25 万次大模型推理请求。再结合大模型的尺寸、输入输出长度、系统延迟要求、业务量波动等因素，评估出相对准确的 GPU 卡规格和数量要求。

明确了大模型的具体应用场景及其带来的价值后，接下来便是构建支持这些应用的工具链，工具链的核心作用是提升大模型开发和管理工作的效率。例如，为了更好地服务于智能客服系统，需要搭建基于检索增强生成（RAG）技术的客服知识库，与传统 FAQ 结合共同生成答案。此外，如果需要对复杂业务 SOP 做托管处理，还需要建立 Agent 平台，用于编排不同 Workflow，例如：信用卡提额 SOP、协商还款 SOP 等。参考这种应用 + 模型 + 算力 + 工具链的落地模式，确保每一个组件都是围绕着实现业务和技术目标而设计的，少走弯路，最大化有限资源的投入。

金融行业 Agent 发展趋势

回看 2024 年，大模型技术在金融机构中的应用主要仍集中在基于检索增强生成（RAG）架构的知识库和问答 Chatbot 上，包括知识助手、客服助手、合规助手等，这一类助手型应用旨在通过智能化对话的交互形式提升信息处理效率。尽管这些系统在一定程度上提高了工作效率，但其价值相对有限，尤其在面对复杂金融业务流程时显得力不从心，这甚至引起了行业内对大模型实际价值产生质疑。一方面，标准 RAG 的工程链路相对固定，与外部系统灵活交互的能力有限；另一方面，对话交互这种形式仍然只是“辅助”，并未实现作业的托管，其他工作还需要人工完成，导致对效率的提升有明显瓶颈。

岁末年初，基础大模型的能力发展迅猛，尤其是 OpenAI 发布 o1、Deepseek 发布 R1 新型推理大模型，在逻辑推理、数学计算、复杂规划等方面取得了突破性进展，展示了大模型智能水平的持续增强。这些进展给予市场极大的信心，大家坚信大模型在已有的场景中可以变得更准确，并且被应用到更加复杂的场景当中，这些场景不再限于问答交互，而是深入到客户服务、智能投顾、经营决策、合规审核等各个关键业务环节中，以 Agent 的形态提供个性化服务。这种转变标志着大模型开始真正发挥其潜力，为企业带来实质性的业务增值。

应用到更加复杂的场景当中，这些场景不再限于问答交互，而是深入到客户服务、智能投顾、经营决策、合规审核等各个关键业务环节中，以 Agent 的形态提供个性化服务。这种转变标志着大模型开始真正发挥其潜力，为企业带来实质性的业务增值。

展望 2025 年，随着基础模型能力的不断增强，金融 Agent 将全面爆发，并深刻影响整个金融行业的运营模式。在新的一年里，大模型 Agent 将不仅仅作为辅助工具存在，而是会成为重塑业务流程的关键力量，推动全系统的数字化重构。

我们认为，企业的大模型架构会是一个 MoA (Mixture Of Agents) 混合架构：

- 1) 一个强大的通用模型加一个强大的推理模型，通常是千亿或万亿参数，采用 MoE 架构，作为企业智能基座；**
- 2) N 个中小尺寸的稠密模型，用在不同场景 Agent 当中，可以通过模型蒸馏提升其领域能力；**
- 3) 工具链层面，一个训练推理平台加一个 Agent 工厂，构建企业级的数据飞轮，基于在若干 Agent 在场景中沉淀的数据，使用基座模型对领域模型进行蒸馏；**
- 4) 当这些 Agent 积累的数据不断丰富后，基座模型也能够通过强化学习进一步增强，逐步演变成企业内部自有的世界模型。**

这种模型架构与开发范式，不仅是对传统业务流程的一次重大升级，也为金融机构探索新的商业模式和服务模式提供了可能，不仅有助于提升企业的竞争力，也将为客户提供更加便捷、个性化的金融服务体验。

金融行业Agent百景图

银行

11

信贷 | 风控 | AI Native 手机银行

证券

41

投研 | 投行 | 投顾 | 智能运营

保险

69

产品开发与销售 | 核保核赔 | 监管合规

通用

95

智能客服 | 智能用数 | 知识助手 | 研发助手 | 数字人 | 内容审核
信息检索与打标 | 培训陪练 | 办公助手 | 营销助手

银行 ① 信贷 ② 风控 ③ AI Native手机银行

证券 ① 投研 ② 投行 ③ 投顾 ④ 智能运营

保险 ① 产品开发及销售 ② 核保核赔 ③ 监管合规

通用 ① 智能客服 ② 智能用数 ③ 知识助手 ④ 内容审核 ⑤ 信息检索与打标 ⑥ 培训陪练 ⑦ 研发助手
⑧ 数字人 ⑨ 办公助手 ⑩ 营销助手

银行

① 信贷

尽调报告生成助手

行业分析助手

企业股权/关联分析助手

财务报表分析助手

交易流水分析助手

信审资料查全助手

信贷授信方案助手

尽调报告审核助手

② 风控

信贷进件风控助手

生成用户可疑交易报告

风控智能助理

③ AI Native手机银行

APP零售业务助理

APP养老咨询规划

APP理财业务助理

APP信用卡服务助理

证券

① 投研

研报快读

研报观点问答

日报早评收评

深度研报写作

路演解析

行业周报观点

热门事件助手

② 投行

股权激励助手

投行法规解读

③ 投顾

投顾资讯简报

营销优先级判断

理财产品问答助手

④ 智能运营

信披报告审核

营销物料审核

保险

① 产品开发及销售

保险产品解读

条款解析助手

保险产品推荐

保险产品问答

保险营销创作

保险产品搜索

保险产品比对

② 核保核赔

智能预核保

核赔辅助

智能影像处理

③ 监管合规

条款智能核验助手

条款智转助手

对外披露审核

通用

① 智能客服

坐席通话质检

智能语音导航

机器人营销外呼

智能投诉预警

客服知识问答

人工客服助手

意图识别

智能语音分析

智能工单助手

机器人催收外呼

客服知识构建

数字人客服

App对话机器人

电销助手

② 智能用数

客户经理绩效助手

智能问数助手

数据资产问答

贷款分析助手

小微经营户潜在客户挖掘

客户经营分析助手

智能搭建助手

存款分析助手

理财潜在客户挖掘

代发户促活助手

③ 知识助手

知识问答

智能写作

复杂信息抽取

④ 内容审核

内容安全审核

制度撰写

合规问答

⑤ 信息检索与打标

企业工商信息打标

财报解析

可信搜索

企业招投标信息打标

企业知识产权信息打标

⑥ 培训陪练

交互式培训助手

对练机器人

课件生成助手

智能考试助手

智能培训问答助手

⑦ 研发助手

AI程序员

需求助手

编码助手

智能运维助手

单元测试生成助手

测试用例生成助手

⑧ 数字人

交互取证提额

视频交互陪练

直播交互

视频分享

⑨ 办公助手

工作报告助手

智能会议助手

OA小蜜

流程助手

⑩ 营销助手

营销策划助手

社群运营助手

营销素材生成

营销文案生成

银行

BANK

信贷

尽调报告生成助手、信贷授信方案助手、交易流水分析助手、企业股权 / 关联分析助手、财务报表分析助手
行业分析助手、信审资料查全助手、尽调报告审核助手

风控

信贷进件风控助手、风控智能助理、生成用户可疑交易报告

AI Native 手机银行

APP 零售业务助理、APP 信用卡服务助理、APP 理财业务助理、APP 养老咨询规划

尽调报告生成助手

Agent 概述

尽职调查报告主要是通过对借款人的人品、背景、经营情况、财务状况、现金流状况以及其他重要信息中蕴含的风险点和优劣势进行分析与识别，进而判断出借款人真实的还款信誉以及还款能力。客户经理在企业借款前需要完成企业尽职调查和手工撰写调查报告，在调查报告的撰写过程中需要从复杂的信息渠道中整合大量数据，效率较低，分析决策比较困难。

信贷尽职调查助手，采用通义千问大模型技术，实现借款客户互联网资料收集、所属行业分析、所在区域分析、银行行内信贷相关业务信息采集、尽调报告内容编排等功能，信贷尽职调查助手实现交互式自动化尽调报告生成，大大降低客户经理的工作量和专业要求。

需求分析

当前银行客户经理撰写尽职调查报告存在以下缺点：

主要采用预设模板方式，根据业务类型对客户调查数据分类、预设每一类业务指标、预设评级模型、制定审核流程，这些仅能实现固定模板下的不同客户尽调，但是分析撰写尽调报告还得需要人工重复、长耗时去处理。

部分信息通过客户经理人工在互联网相关网站查询获取，采用标签化模式，仅实现了尽调内容的采集、展示，未能结合信贷业务和信贷专家的经验进行分析，撰写尽职调查报告无法完全达到数字化风控的要求。

典型案例

模力时刻

Qwen-plus

处理 尽调报告中的撰写部分内容通过大模型生成，借助专业规则经验可以有效减少由于人为因素造成的错误，提升报告质量和数据内容的准确率。

效果 通过对尽调报告模板的配置，报告中部分内容通过配置方式直接调用 API 或大模型接口，能够快速将收集到的数据转化为结构化的报告内容，效率提升 2-3 倍。

信贷尽职调查助手中，主要功能模块包括知识库建立、尽调模板管理、尽调报告数据管理、尽调报告内容编排、尽调报告撰写、尽调报告预览、尽调报告下载等功能。具体业务流程如下：

1 信贷知识库

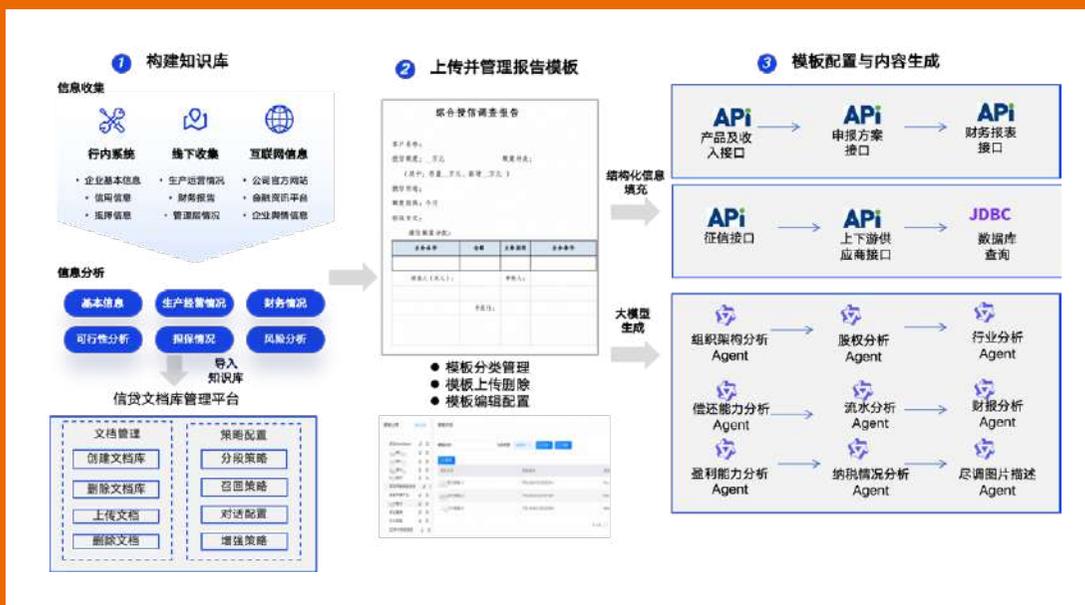
搜集行里信贷政策、产品等知识文档、行业分析报告、借款企业提供材料等构建统一的知识库。

2 尽调模板管理

支持报告 word 文件上传及模板分类管理、版本管理等功能。

3 报告模板配置与内容生成

分析报告中不同内容填充的数据源,判断银行内部业务系统接口调用、数据库访问、外部三方接口、大模型 Agent 调用,注册不同数据源接口定义,配置到模板中的不同章节,配置完模板后,就可以进行不同借款企业尽调报告自动生成功能。



价值分析

信贷尽职调查报告生成助手能够显著提升客户经理的工作效率、准确性和决策质量。以下是关于这种技术解决方案的价值分析:

提高工作效率

通过对尽调报告模板的配置,报告中部分内容通过配置方式直接调用 API 或大模型接口,能够快速将收集到的数据转化为结构化的报告内容,极大地缩短了报告编写时间,效率提升 2-3 倍。

增强准确性

对于尽调报告中的撰写部分内容通过大模型生成,借助专业规则经验可以有效减少由于人为因素造成的错误,提升报告质量和数据内容的准确率。

信贷授信方案助手

Agent 概述

客户经理在信贷授信前需要进行现场调研并形成初步的授信方案后，经常因为客户经理对信贷产品准入要求不明确、抵押资产评估有误等原因，导致提交资料多轮次打回修改，拖长了企业贷款授信周期。

信贷授信方案助手结合大模型技术，对客户经理的问题进行意图理解识别，然后分别进行制度条款、授信要点内容 RAG 检索召回，再根据条款要求进行准入条件判断，将结果组装成固定结构返回给客户经理，提高整个授信审批的通过率和效率。

需求分析

当前对公的授信环节中，普遍存在由于对公客户经理对于产品准入要求了解不全面，在设计对公授信组合方案中，无法准确的识别风险点和注入要求。需要反复与授信审批人员确认方案可行性，造成授信方案的反复沟通成本以及授信审批过程中驳回的情况出现。

客户经理经验少

客户经理对于客户能不能做、做什么产品、定什么方案往往没有头绪，无快捷的指导性工具。

文档资料类目多

行业信贷政策文件、抵质押品准入名录、小微产品种类繁多，且每年均有新增或者修订，各类文件要求较多。

资料查询效率低

客户经理及授信审批人员在日常工作中往往需要翻阅查询相关文件来核对是否符合要求，在一定程度上影响了业务申报的顺畅性以及授信审批的效率。

典型案例

模力时刻

Qwen-plus

处理 对客户经理提的问题进行意图识别，并且根据问题进行制度文件和授信方案内容生成。

效果 信贷专员能够快速找到最佳贷款产品和授信政策，提高了审批的质量和效率，提升了客户满意度和服务体验。

客户经理可以通过钉钉机器人助手结合大模型通义千问大模型能力，提供信贷授信制度问答和方案提示功能。以授信方案提示场景为例，客户经理输入的授信方案要求先通过要素拆解智能体将要素做提取，针对不同产品的制度条款，通过条款解读智能体结合知识库中的制度文件做相关产品条款的召回和筛选。再通过条款判断智能体来做量化和非量化条款的判断，最终由大模型组装成授信提示做统一展示，具体流程如下：

1 授信知识库构建

搜集相关制度文件、行内信息、信用信息、行研报告文档、行业授信政策、综合授信文档等，按照条款、自然段落对文档进行拆解与召回，支持文档订正更新能力。

2 意图识别

根据用户问题输入通过意图识别判断，识别是词条还是授信方案内容查询。

3 授信方案提示

将授信方案中的信息通过大模型进行要素拆解。要素拆解包括：企业名称、所属行业（用于做行业规划及行业前景综合分析）、贷款产品、担保方式、业务模式等。

4 要点条款召回

遍历基础要素，匹配对应的管理办法制度文件，召回要点条款，包括行研报告、行业授信方案、综合授信、敞口授信、资金池信息等；基于大模型的能力，将量化的条款结合企业相关数据做比对，非量化的条款要求直接反馈输出给出提示。

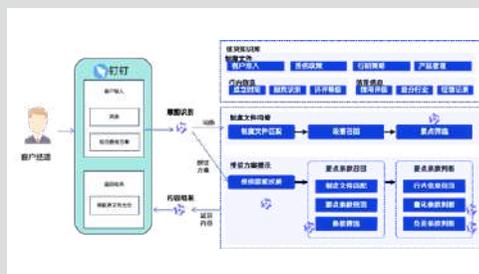
5 内容整合

集成外部接口召回企业相关信息，如 成立年限、法人信用状况等等，基于大模型的能力，对当前信息充足的条款判断是否满足。

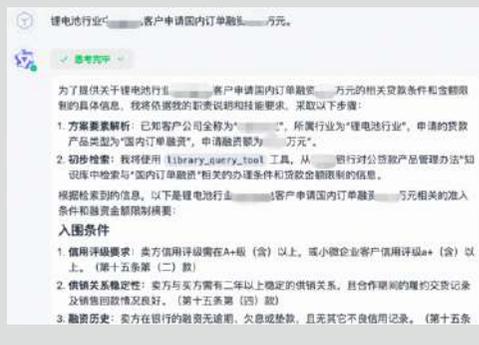
6 组装提示结果

基于大模型的能力按照给定结构组装提示结果。

总体技术链路如下图所示：



授信方案提示结果如下图所示，客户经理在查询行业准入条件和金额限时时，直接返回入围条件和融资金额限制相关条例等信息，从几个小时的资料翻阅解读政策文档到秒级返回具体条例，效率提升90%以上。



价值分析

🔍 快速检索

客户经理能够迅速获取行内内部信息、最新的准入政策、行业政策以及相关的管理办法，确保了信息的实时性和准确性，确保客户经理能够在最短时间内做出准确的风险评估和决策。

🎯 精准营销

精准定位潜在客户群体，并利用初步试评来评估不同市场活动的效果。快捷指导帮助制定个性化的营销策略，而学习搜索功能则支持基于最新市场趋势不断优化和调整营销方案，从而提升获客效率和客户转化率。

📁 高效审批

涵盖了初步评审和前期匹配环节，确保资源得到最优配置并显著缩短处理时间。通过标准化的初步评审流程，信贷专员能够快速找到最佳贷款产品。这不仅节省了大量时间，还提高了审批的质量和效率，提升了客户满意度和服务体验。最终，实现了信贷业务的质效双重提升。

交易流水分析助手

Agent 概述

在银行贷款审批过程中，客户提交的交易流水是评估其还款能力和经营状况的重要依据。信贷经理通常需要根据这些流水记录来分析客户的经营性现金流入与支出，并与企业的实际经营数据及合同履行情况进行对比，以此判断企业的真实生产经营情况及潜在风险。

基于大模型技术的银行客户交易流水分析，将信贷专家的经验固化在大模型流程中，实现了专家经验的高效复用，使得信贷人员能够在短时间内获得详尽且准确的客户交易流水分析报告，从而更有效地评估贷款风险并作出决策。

需求分析

现金流是企业生存的关键，通过流水分析可以清晰地看到企业的资金流入和流出情况，更能反映企业的真实经营成果，帮助判断该公司是否有能力偿还贷款，但实际操作来看，有以下难点：

系统化成本高

由于各银行流水单格式各异，实现自动化系统处理的成本较高，手工拼接和分析耗费巨大的人力资源和时间，目前主要依赖手工分析效率低下。

海量数据处理效率低且难以标准化

企业每月的交易流水可达上万条，客户经理难以详细准确地进行分辨和分析，这不仅耗时而且容易出错。

高度依赖经验丰富的信贷经理

有效的流水分析需要有经验的信贷经理进行专项分析，这类人才稀缺且培养周期长。此外，分析结果易受个人经验和状态影响，导致判断差异。

典型案例

模力时刻

Qwen-VL

处理 对 OCR 识别流水数据进行数据内容矫正，并且对 OCR 未识别的图片进行数据识别。

效果 将 OCR 的数据识别准确性提升 5%~10%。

QwQ-plus

处理 对流水数据进行分析统计，分析经营流水风险。

效果 能够将大量不同类型的交易流水文件分钟级就能生成流水分析结果，效率提升 90% 以上。

以下是银行信贷场景中流水经营分析的几个关键步骤，通过流水分析实现高效、标准化的交易流水分析功能：



1 智能匹配流水单格式

解决各家银行流水清单文件格式不一致的问题。通过智能匹配技术能够适应不同银行的流水单格式，确保后续分析的数据准确性和一致性。

2 智能分析经营性流水

从交易清单中识别和分类经营性交易。对清单内容进行智能分析，识别哪些关键词属于经营性关键词。通过大模型技术自动筛选出与企业经营活动相关的交易记录，为后续的风险评估提供数据支持。

3 智能生成流水分析结果

根据经营流水数据完成客户全部交易流水的分析，并返回结构化的打标结果及生成详细的报告，具体流水分析结果如下图所示：

逐条分析流水清单，判断流水类型

日期	交易类型	凭证种类	凭证号	对方户名	对方账号	摘要	借方发生额	贷方发生额	金额	是否经营性交易
2023/3/1	转账	0				转账	767.8	0	18683049	否
2023/3/1	转账	0				贷款	200	0	18616340	否
2023/3/1	转账	0				转账	1278.65	0	18511875	否
2023/3/1	转账	0				转账	137.14	0	18511481	否
2023/3/1	转账	0				400企业社	37	0	18511701	否
2023/3/1	转账	0				行存款	0	15,840.00	2110701	否
2023/3/1	转账	0				工资	0	395,799.66	6088667	否
2023/3/1	转账	0				转账	538	0	6082807	否
2023/3/1	转账	0				转账	1,996.00	0	6082807	否
2023/3/1	转账	0				转账	1,000.00	0	6037487	否
2023/3/1	转账	0				工资	395,799.66	0	20737520	否
2023/3/1	转账	0				公金扣缴	1,267.80	0	2658674	否
2023/3/1	转账	0				贷款	656	0	2051714	否

统计流水分析结果

所属路径	支出金额汇总(人民币)	收入金额汇总(人民币)	支出金额汇总(美元)	收入金额汇总(美元)
***2023.05**	3003.59	171.55	0	0
***2023.02**	270.14	0	0	0
***2023.05**	-213.12	480.02	0	0
***2023.04**	262.65	630.11	300	400
***2023.10**	78.25	383080	0	0
2023.11	-64.44	1295.6	0	0
***2023.01**	220.95	7213.6	0	0
***2023.06**	59.21	621.35	0	0
***2023.08**	113.33	652.19	0	4100
***2023.09**	7289	0.67	0	0
***2023.07**	8110	768.32	0	0
***2023.12**	112.01	3838.2	60565	60565

价值分析

提升流水分类准确率

对于交易流水中各种不同的备注信息，通过大模型方式进行交易数据的经营类型智能分类识别，客户经理不需要一条条交易内容仔细核对。

提升流水分析效率

能够将大量不同类型的交易流水文件分钟级就能生成流水分析结果，效率提升 90% 以上。

企业股权 / 关联分析助手

Agent 概述

银行对公授信时需要分析企业的股权结构、企业之间有无关联关系等信息，主要用于深入了解贷款企业的实际控制权和股东背景，帮助银行更全面地评估信贷风险。希望自动生成借款企业的股权结构信息、关联信息等信息，减少材料编写时间及信息准确率。

企业股权 / 关联分析功能通过大模型调用搜索引擎能力获取企业股权和关联信息，结合大模型能力实时绘制借款企业的股权等信息，然后大模型汇总生成股权结构 / 关联关系结构，帮助银行了解企业主要股东和关联关系等信息，深度分析不仅提高了贷款决策的准确性，也增强了银行管理信贷风险的能力。

需求分析

尽调报告中企业股权 / 关联分析时需要分析企业股权有无关联关系，有无黑名单企业，有无高风险企业。信贷时需要客户经理提供企业股权结构信息：借款人的股权结构图、主要股东或实际控制人实力介绍，传统解决方案通常是客户经理通过企查查等渠道获取，要么手工查询绘制，要么二次付费下载，属于难度高、耗时的部分，对客户经理要求极高。具体要求如下：

1 股权结构图须穿透到最终的实际控制人，需标明各出资人的出资比例（如出资人较为分散，则仅需列示主要的出资人）。

2 若有风投等投资公司进入，需补充股权取得成本、投资款到位情况、是否存在对赌协议及主要内容，判断对赌条款近期触发的可能性以及对借款人经营和财务安全性的影响。

3 关联信息综合获取企业的股东、参股公司、控股公司、联营企业，并按照表格结构显示相关信息。

典型案例

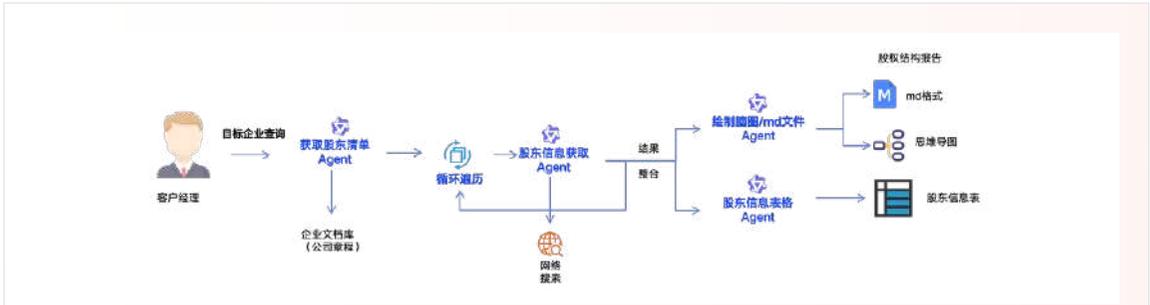
模力时刻

Qwen-plus

处理 大模型根据企业名称调取联网搜索能力轮训查询企业股权信息，整合后再绘制股权表格信息表和 Markdown 格式数据。

效果 助手能够在短时间内完成复杂的股权结构分析，从原来数小时缩短为几分钟就能进行股权分析内容生成，效率提升 90% 以上。

企业股权 / 关联分析通过调用股东信息获取 Agent 和脑图绘制文件 Agent 结合，具体流程如下：



1 股东清单获取

第一次股东信息查询，首先根据企业提供的年报、财报等知识库内容获取股东清单。

2 循环遍历获取股东信息

根据上面股东信息清单被传递到一个“循环遍历”步骤，信息再次被传递给“股东信息详情 Agent”，如果没有股东信息可以通过搜索引擎获取方式股东信息详细信息。可能用于进一步的数据验证或补充。

3 股权结构报告

处理好的信息被传递给绘制脑图/md文件 Agent，这个 Agent 负责将信息整理成脑图或 Markdown 格式的文档，同时股东信息表格 Agent 生成股东信息表。调用结果内容示例如下：



价值分析

在银行信贷授信环节中，股权分析的价值主要体现在能够帮助评估借款企业的财务健康状况、管理结构以及潜在风险。通过三方接口实现全自动绘制股权穿透图，可以显著提升这一过程的效率和准确性。以下是这种技术优化的具体价值：

提高准确性

自动化的股权穿透图生成减少了人为错误的可能性，确保了数据的一致性和准确性，自动生成的图表支持后续的人工调整和优化，使得最终的分析结果更加贴合实际情况。

节省时间

相比人工搜集编写材料，助手能够在短时间内完成复杂的股权结构分析，从原来数小时缩短为几分钟就能进行股权分析内容生成。

财务报表分析助手

Agent 概述

在银行信贷场景中财务报表分析是非常重要的，因为财务报表提供了关于企业财务状况、经营成果和现金流量的关键信息。通过深入分析这些数据，信贷机构能够评估借款企业的还款能力和信用风险，从而做出更加明智的贷款决策。

财务报表分析助手依托大模型技术，专为高效处理、深度分析及精准解读借款企业财务报表信息而设计，将资产负债表、利润表、现金流量表中通过知识库查询召回方式提取关键指标，结合审计报告内容能自动化生成详尽的智能分析报告，根据历史变化指标分析输出专业分析意见，显著提升财务报表阅读效率。

需求分析

财务报表及各项财务指标能够充分反映企业实际经营情况，因此财务解读分析能力至关重要，当前主要问题如：

分析经营风险

财务报表能评估企业价值、了解盈利能力和识别风险因素。通过分析收入、成本、利润和现金流等关键指标，可以发现高额负债、现金流问题或收益不稳定等风险因素。基于历史数据的趋势和比率分析，还可以预测未来业绩，帮助企业提前采取预防措施，降低潜在风险，确保稳健运营。

财务报表指标复杂性

财务报表包含复杂的指标，如资产负债表、利润表和现金流量表中的各项数据。这些指标不仅数量庞大，而且相互关联，需要综合分析才能得出准确结论。理解和解析这些复杂指标对全面评估企业的财务健康状况至关重要，以确保决策的科学性和准确性。

专业局限性

尽管客户经理具备一定的财务知识，但在面对繁重的数据量和复杂的指标时容易出现错误。传统分析方法往往只是简单描述数据变化，未能深入挖掘背后的真实原因和潜在风险，限制了他们对企业实际经营状态的理解和风险评估能力，影响信贷审批和风险管理精度。

典型案例

模力时刻

Qwen-VL

处理 对 OCR 识别资产负债表、现金流量表、利润表数据进行数据内容矫正，或者识别三表数据。

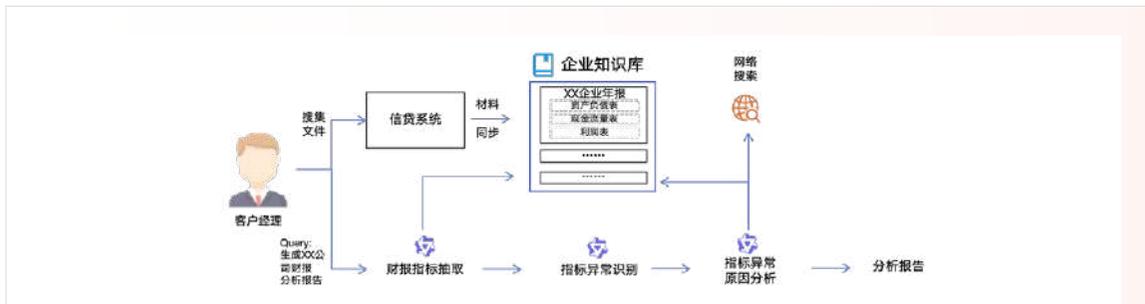
效果 将 OCR 的数据识别准确性提升 5%~10%。

QwQ-plus

处理 对财报数据进行分析统计，分析经营流水风险。

效果 自动化的财务报表分析减少了手工操作的需求，降低了人力成本，时间缩短到分钟级，财报分析效率提升 95% 以上。

客户经理把要借款企业相关材料文档上传到的企业知识库，通过财务报表分析助手生成分析报告，具体流程如下：



1 知识库构建

客户经理上传借款企业财报、公司章程、年报等文件到信贷系统，信贷系统自动同步到企业知识库，构建企业专属知识库。

2 财报指标抽取

召回借款企业资产负债表、现金流量表、利润表等信息，对比近三年历史财务数据来识别趋势、解释财务状况的变化、进行关键财务指标的比较分析，由大模型完成深入分析。

3 指标异常分析

根据指标异常将变化发生的时间和具体的变动数据项作为搜索条件的一部分。通过这些信息，在文档库中召回该时间段内的企业经营事件。例如，特定时间段内的重大合同签订、政策变动或管理层更替等事件，这些都可能对企业财务数据产生影响。

4 分析报告

结合大模型对数据变化的分析结果和从文档库中召回的企业经营事件，进行全面的综合分析报告生成。如右图是财务状况异常分析结果及异常原因。

财务状况分析
 结合上述背景信息，我们可以进一步解释财务数据的变化趋势：

- 资产规模显著扩大：**
 - 原因：公司自成立以来，通过多轮融资获得了大量资金支持，用于技术研发、设备采购和业务扩展。这导致总资产、流动资产和非流动资产均大幅增加。
- 应收账款增加：**
 - 原因：随着业务规模的扩大，客户数量和销售额增加，可能导致应收账款相应增长。此外，销售信用政策的变化也可能是一个因素。
- 存货激增：**
 - 原因：生产扩张和库存管理策略调整可能是主要原因。公司可能在为未来的订单做准备，或者由于市场需求波动而增加了库存。
- 负债增加：**
 - 原因：融资活动带来的短期借款和其他流动负债增加。虽然总负债有所上升，但大部分

价值分析

📈 提高决策效率

通过自动化处理和分析财务数据，可以快速生成报告和关键指标分析，确保数据处理的一致性和准确性，帮助客户经理或审批人员更快地做出基于授信决策。

🔍 专业洞察

利用高级分析技术，如预测分析、趋势分析等，可以从大量数据中提取有价值的信息，提供对业务绩效更深层次的理解，支持战略规划和风险管理。

⏬ 节省时间

自动化的财务报表分析减少了手工操作的需求，降低了人力成本，时间缩短到分钟级，财报分析效率提升95%以上。

行业分析助手

Agent 概述

行业分析是分析借款企业所在行业的风险特点，包括市场风险、经营风险、政策风险等，评估客户的行业风险承受能力。银行客户经理在进行行业分析时，面对种类繁多且各具特色的行业，经常会遇到行业研报材料覆盖不全、材料更新频率低等问题，导致借款企业行业分析不完善、信息不准确。

行业分析助手借助大模型能力，同时集成知识库 RAG 召回和互联网搜索引擎能力，查询企业所在行业信息生成行业分析报告，评估借款企业所在行业的整体状况，帮助银行或金融机构更好地理解企业的市场环境和潜在风险。

需求分析

客户经理在给企业借款前，对客户所处行业分析时，主要面临以下问题：

数据获取复杂性

银行在尽职调查过程中需要收集和分析大量的行业数据，包括市场趋势、竞争对手情况、政策变化等。这些数据的获取和处理过程复杂且耗时，可能导致报告的时效性和准确性受到影响。

深度和广度不足

尽职调查报告需要全面、深入地分析目标行业的现状、前景、竞争格局等。然而，银行在执行过程中可能因为资源有限或专业能力不足，导致分析不够深入和全面，无法为决策提供充分的信息支持。

风险评估的主观性

尽职调查报告中的风险评估部分需要综合考虑多种因素，包括市场风险、信用风险、操作风险等。由于评估过程涉及主观判断，可能存在评估结果不够客观、准确的问题，增加了决策的风险。

专业能力和经验差异

尽调报告的质量很大程度上取决于客户经理的专业能力和经验。不同客户经理对同一行业的理解和分析可能存在差异，影响报告的一致性和可靠性。

典型案例

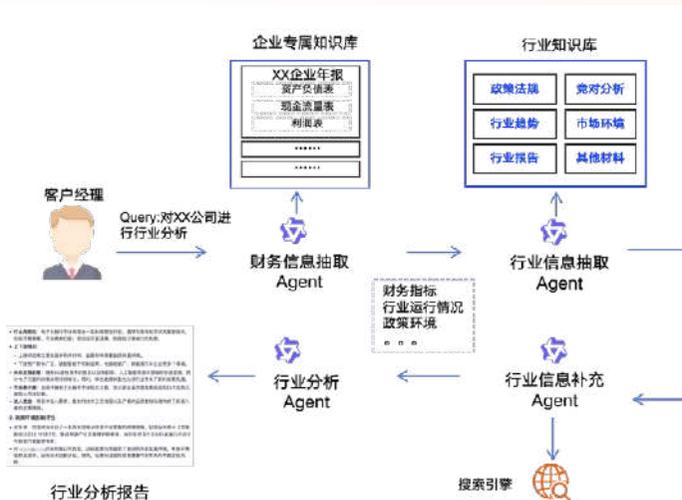
模力时刻

QwQ-plus

处理 结合行业知识库、企业专属知识库、外部互联网内容，生成行业概况、运行特点、行业周期性、上下游情况、未来发展前景、市场集中度、进入壁垒、行业政策与法规等进行调查与分析内容。

效果 减少了以前耗时耗力搜集资料、汇总材料及内容撰写的时间，从原有 1 天缩短到 2 小时，提升 75% 以上。

行业分析助手通过访问借款企业财报、行业知识库、互联网搜索引擎方式获取借款企业所处行业信息，并分析生成行业报告，具体流程如下图所示：



1 构建行业知识库

收集行业信息，如行业政策法规、行业趋势、行业报告、竞对分析、市场环境等，构建多行业知识库。

2 行业信息获取

获取借款企业的财务信息，然后根据企业所处行业搜索行业知识库，同时通过互联网搜索引擎进行资料补充，获得完整的企业相关行业信息。

3 行业分析

通过召回检索回的数据进行整合，并进行深入分析，包括分析行业政策及其影响、对比同业数据、行业发展等，将所有分析结果汇总成一份详细的行业分析报告，提供给客户经理。

价值分析

深度全面行业剖析

不仅全面分析政策因素，也对行业概况、运行特点、行业周期性、上下游情况、未来发展前景、市场集中度、进入壁垒、行业政策与法规等进行调查与分析，使分析结果更加直观、更具针对性。

同行业对比增强参照性

通过细致对比同行业上市公司的财务指标，并结合借款企业具体的生产经营状况，评估其经营、财务等方面的风险，并为贷款决策提供有力支持。

效率提升

通过行业知识库、互联网搜索方式，减少了以前耗时耗力搜集资料、汇总材料的时间，效率提升 75% 以上。

信审资料查全助手

Agent 概述

企业贷款信用审查是金融机构在向企业提供贷款服务前的一项重要风险控制措施，企业申请贷款时需要准备详实的材料，展示企业的健康运营状态、市场前景及还款能力等情况，以确定是否给予贷款以及贷款的额度和利率。

贷款信审助手利用大模型技术构建智能信审系统，辅助银行客户经理在搜集企业材料过程中，检查材料的完备性、内容时效性，检查企业是否符合信贷产品准入条件，并进行合适的信贷产品推荐。做到现场材料搜集“只跑一次”，显著提升了信审阶段的效率和准确性。

需求分析

客户经理上门对企业进行尽调材料搜集过程中，面临搜集文件多、内容不实时、不同贷款产品多次反复现场搜集材料等问题，具体如下：

资料收集多

信审阶段客户经理需要准备丰富且真实的材料，包括但不限于企业的营业执照、财务报表、税务记录、销售合同、资产证明等等。

资料实时准确

上传材料时效性验证，例如客户提供营业执照非最新版本，希望通过搜索引擎获取最新工商信息，对比资料库内容提示客户经理。

经验不足

银行金融产品众多，客户自行申请时选择产品不够精准，新客户经理对产品与客户资质了解不足，导致无效申请或者多次往复补充材料，效率降低。

典型案例

模力时刻

QwQ-plus

处理 根据信贷产品准入要求、连接联网数据，大模型自动检查贷款企业上传的材料是否完整度。

效果 不需要重复让企业提交各种材料，一步到位，初始材料审核环节提升 80% 以上效率。

客户经理在信贷申请审核前，会把客户相关信息搜集并上传到信贷系统，同时系统会根据上传内容进行完整度和时效性核验，然后根据授信产品要求核验具体内容，完成本次材料采集。

1 收集材料

申请人提供必要的材料，包括但不限于营业执照、公司章程、财务报表、税务记录、销售合同、资产证明、水电房租发票等。

2 上传材料

将收集到的材料上传至信贷系统。这一环节确保所有必要文件都被数字化并存储在系统中。

3 验证材料完整性

大模型会自动检查上传的材料是否完整。例如，它会判断是否有遗漏的文件或信息。

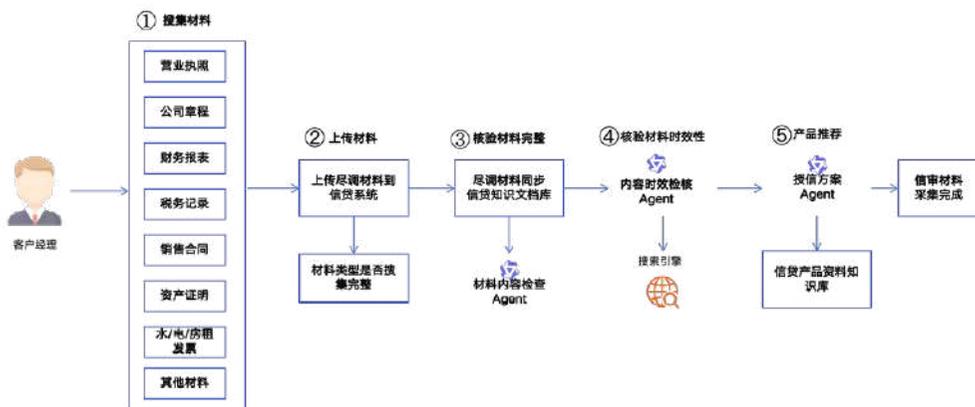
4 核验材料时效性

大模型进一步检查材料的有效性和时效性。例如，财务报表是否是最新的，税务记录是否与当前时间匹配等。

5 产品推荐

根据审核结果，系统会推荐适合申请人的信贷产品。这一步骤基于大模型对申请人资质和需求的理解。系统会查询信贷资料知识库，以获取更多关于信贷产品的信息，帮助申请人更好地理解其选择。

通过这种方式，大模型不仅提高了信贷审核的效率，还减少了人为错误的可能性，使得整个过程更加准确和高效。具体流程如下图所示：



价值分析

避免遗漏客户材料

通过引入先进的大模型技术，系统能够自动检查并提醒客户经理所需的所有材料是否齐全，有效避免因客户经理遗漏重要文件而导致的客户投诉以及多次往返补充材料的问题，缩短授信时间。

内容审核校验前置

在收集阶段进行资料内容准确性和时效性检验，一旦发现错误或过期信息，系统将立即通知客户经理进行修正，确保所有提交的信息都是最新和有效的，减少后续因为信息不准确而产生的问题。

全信贷产品规则匹配与推荐

系统内置了行内信贷产品的详细规则，并能基于客户的资料智能匹配最适合的产品。这种个性化的服务不仅能提升客户的体验感，还能提高授信的标准，使整个授信过程更加高效、精准。

尽调报告审核助手

Agent 概述

审核尽调报告时因为审核人员数量少、项目并发高等原因，也导致信贷审核周期平均在一周左右。另外审核人员对报告内容信息不对称、报告质量不高以及财务信息真实性甄别等困难，如何快速定位关键信息和风险点，确保审核过程高效准确是审核人员希望解决的。

信贷尽调报告审核助手利用大模型技术，按照预设量化及非量化指标，同步生成详细的审核意见，帮助审核人员进行快速定位关键信息和风险点，也能减少审核人员翻阅报告时间，确保审核过程高效准确。

需求分析

信贷尽调报告审核过程中，面临如下难点：

审核压力大

由于对公客户经理人数远超审核岗位，导致审核工作量巨大。每份尽调报告的审核通常需要大约一周左右，这不仅增加了审核人员的工作负担，还可能影响整体工作效率。

审核效率低

尽调报告往往超过百页，包含大量复杂的数据和信息。审核人员在这种情况下很难迅速且准确地找到所需的关键数据点，容易遗漏重要细节，从而影响审核的质量和准确性。

典型案例

模力时刻

QwQ-plus

处理 沉淀不同的审核规则汇总成专家规则，大模型对财报审查、财务比率审查、债务规则审查和现金流审查来检查企业的财务健康状况，形成审查初步报告。

效果 自动化处理大量数据的能力使得信贷审核过程大大加快，减少重复提交修改报告环节，提升审批效率，可以节约 50% 以上时间。

尽调报告内容生成后，会进行报告内容拆解和规则审核，形成审核报告。

1 报告生成

客户经理根据借款企业提供的材料进行贷款申请报告的撰写，生成最终的尽调报告。

2 报告拆分

根据不同的报告审核规则，将尽调报告自动拆分成多个不同的章节内容，方便审核规则对内容进行准确审核。

3 报告审核

沉淀不同的审核规则汇总成专家规则，形成不同规则审核智能体进行报告审核判断。通过财报审查、财务比率审查、债务规则审查和现金流审查来检查企业的财务健康状况，然后通过收入与利润审查及风险合规审查来验证数据的准确性和合规性。

4 审核报告生成

基于上述的审查智能体的反馈内容，系统提供一个综合评估，并给出具体的推荐建议。

具体流程如下图所示：



价值分析

信贷尽调报告审核助手的价值主要体现在提升信贷审批流程的效率、准确性和合规性上。该助手帮助分析和处理大量的信贷申请数据，能够快速识别潜在风险点，并提供决策支持。以下是其具体价值分析：

增强审核支持

审核助手能够在审核人员接收到尽调报告的同时，自动生成并提供详细的审核意见。这些意见基于对报告内容的深度分析，帮助审核人员快速定位关键信息，并提供针对性的建议和支持，从而显著缩短审核时间并提高审核质量。

缩短审核时效

自动化处理大量数据的能力使得信贷审核过程大大加快，可以节约 50% 以上时间，减少了人工操作的时间成本。

信贷进件风控助手

Agent 概述

信贷授信进件时需要审查客户提供企业资料，客户经理需要人工审核上传材料进行内容核对，确保其合法性和真实性，同时对进件材料进行风险评估。信贷进件风控助手小微企业资料上传内容审核对比，根据进件照片识别虚假经营场地骗贷的功能。

“信贷进件风控助手”Agent 是一种利用大模型和先进的人工智能技术，多模态大模型进行进件内容提取，并且通过照片背景识别进行相似地址的判断。当检测到异常情况时，系统能够及时发出警报，提醒信审部门进行深入调查，从而有效防范贷款欺诈行为，避免资金损失。

需求分析

进件审核原来靠人审查，信审部门审核人员较少，工作量大经常要加班，而且会漏审和误审，一旦误审会导致放款后造成资损。因此希望能够解决以下问题：

内容材料一致性审核

根据小微企业营业执照、税务登记证、法人身份证明复印件（PDF）与借款人提交的贷款申请表进行对比，查看是否一致性。

欺诈检测

进件的照片附件，识别经营场地，和照片背景，防止中介利用虚假经营场地，冒名小微企业主骗贷等。包括：一个经营场所被反复利用，同一人冒用身份进行贷款申请，经营场所与实际申请材料不一致等。

典型案例

模力时刻

Qwen-VL

处理 识别进件信息中的图片、文件等半结构化数据，提取关键信息。

效果 进件信息收集与整理环节，工作效率提升 60% 以上。

Qwen-plus

处理 基于 VL 提取的信息，做多个数据源信息的交叉比对。

效果 信审员做信息一致性对比的工作量减少 30% 以上。

借款企业提供书面材料提供个贷款机构，通过多模态方式资料审核方式实现上传材料内容审核，并且识别经营场地异常分析的情况。

1 进件材料上传

申请人员提供小微企业营业执照、税务登记证、法人身份证明复印件（PDF）、收入证明、银行流水、信用报告等材料上传到信贷系统。

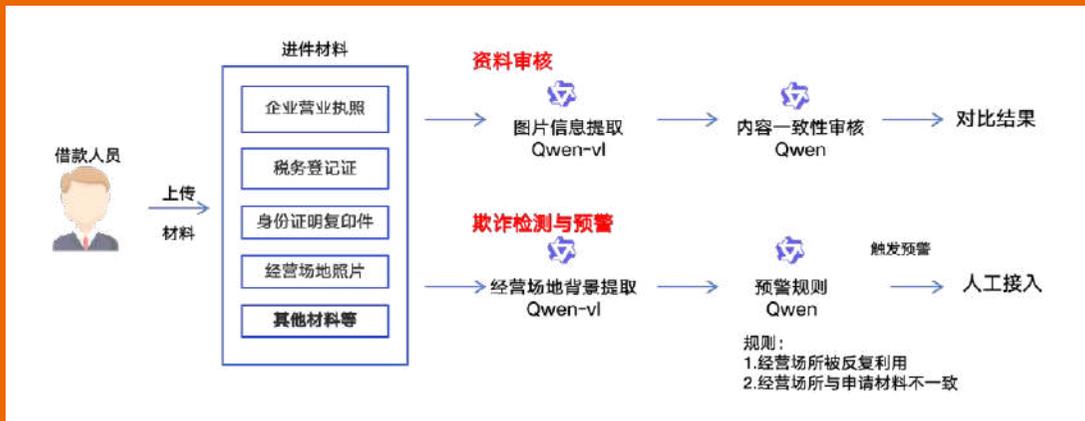
2 资料审核

利用 Qwen-VL-max 快速解析和理解客户提交的申请材料，自动提取关键信息并进行初步风险评估。这不仅加速了审批流程，还能减少人工成本。

3 欺诈检测与预警

通过模式识别和异常检测算法，Qwen-VL-max 可以辅助识别申请材料中的不一致性或潜在的欺诈行为，及时发出预警并进行人工审核，增强风控能力。

具体流程如下图所示：



价值分析

✓ 提高审批效率

通过自动读取、解析和验证提交的文件（如营业执照、税务登记证、法人身份证明等），减少了手动审核的时间和工作量，信审员工作量减少 30% 以上。

☑ 提升准确性与一致性

自动化流程减少了人工操作中的失误，确保每一份申请材料都经过严格的标准验证。

⚠ 防范贷款欺诈

系统可以验证申请人提供的证件和照片的真实性，识别同一人或同一经营场所反复用于贷款申请的情况，防止冒名顶替和虚假经营场地的使用。拦截大额贷款诈骗，每月减少百万元的资产损失。

风控智能助理

Agent 概述

在金融行业快速发展的当下，传统风控手段难以满足日益复杂多变的风险挑战。数据量暴增、风险类型多样化、客群资质变化快，导致策略模型需要高频迭代，亟需一种高效智能的风控工具，因此智能风控助理 Agent 应运而生。

智能风控助理 Agent 基于大模型开发，具备多方面强大能力。通过交互式对话的形式，它能解答风控知识疑问，进行策略的深度分析推理，并且高效自动化实现风控策略迭代过程中的函数开发、策略开发、智能建模等任务，为风控工作提供全面且智能的支持。

需求分析

主要业务需求点：

策略模型优化需求

随着金融业务创新，风控策略和模型需持续迭代。了解已有策略的数据依赖，分析策略血缘，成为改进策略精准度、降低误判风险的关键。通过工具的分析推理能力，风控人员能精准定位数据问题、完成策略诊断，针对性优化策略。

开发效率提升需求

在实际业务中，函数开发、策略生成部署、风控模型建模均耗时费力。传统开发方式依赖专业程序员，且易因沟通不畅、需求理解偏差导致项目延期。Agent 的代码开发与智能建模功能，让非专业研发人员也能快速上手，依据文档自动生成代码片段、模型框架，大幅缩短开发周期。

风控知识普及需求

金融风控领域知识体系庞大复杂，对于新手从业者而言，诸如“决策流”“规则集”等专业词汇理解困难。即使资深人员，面对不断更新的法规、技术与模型，也需要便捷的知识获取渠道。此 Agent 提供即时准确的知识问答，满足不同层次人员的学习需求。

典型案例

模力时刻

Qwen-plus

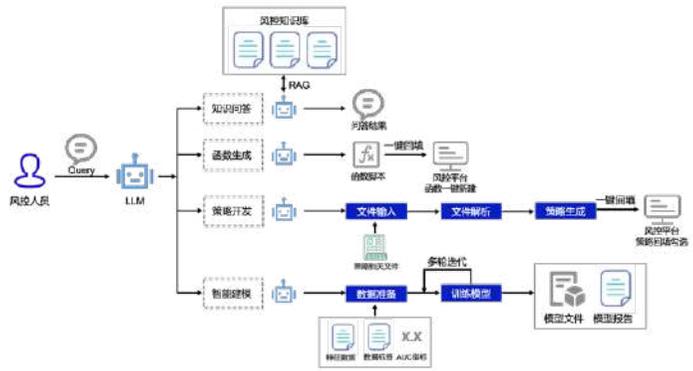
处理 大模型理解风险分析师的问题，自动化挖掘风险特征变量，并生成对应的风险策略 workflow。

效果 风险特征挖掘效率提升 100%，风险策略配置效率提升 200%。

在应用中，基于大模型的智能风控助理 Agent 具备多项核心能力，包括知识问答、函数生成、策略开发、智能建模等，旨在全面提升风控效率和质量。

它通过大模型和 RAG 等技术，实现风控知识问答。对于风控策略中频繁用到的函数，该 Agent 能根据用户 query 生

成对应函数脚本，并一键回填到风控平台的“函数管理”模块下。在整体风控策略的迭代和开发上，Agent能基于输入文件实现策略快速开发，并通过一键回填，让相关策略高效部署至风控平台。同时Agent还可进行智能建模，将特征数据、数据标签、AUC 指标值等信息输入Agent，即可开始模型训练，并自动进行多轮迭代，直到模型AUC 指标符合客户要求。这些能力共同作用，为用户提供了一个全面且高效的风控解决方案。



价值分析

智能风控助理 Agent 的价值主要体现在以下几个方面：

提升风控工作效率

自动化处理大量重复性任务，如常规建模、数据对接、策略生成等，让风控人员可以专注于更具挑战性的分析工作。

促进知识共享

作为一个持续学习的系统，Agent 不仅能回答当前的问题，还能不断积累知识，成为组织内部重要的知识库资源。

降低操作风险

通过标准化流程和自动化工具减少人为失误的可能性，确保每一个决策都有据可依。

增强决策质量

借助先进的算法和大数据分析能力，提供更加精准的风险评估结果，助力企业做出明智的商业决策。

生成用户可疑交易报告

Agent 概述

反洗钱是金融机构必须严格遵守的合规要求之一，可疑交易报告是金融机构履行反洗钱义务的重要手段，其目的是通过识别和报告可疑交易，防止洗钱和其他金融犯罪活动的发生。然而，传统的报告撰写过程存在诸多挑战，如报告质量参差不齐、撰写效率低下、信息整合困难等。

“生成用户可疑交易报告” Agent，通过自动化生成高质量的可疑交易报告，提高金融机构的合规效率和报告质量。能够准确理解用户的可疑交易特征，自动进行数据字段分析与 SQL 构建，获取相关的数据，最后生成大模型内容。

需求分析

主要业务痛点：

报告撰写效率低

由于缺乏自动化工具和标准化流程，反洗钱报告的编写往往依赖于手工操作，导致处理速度慢且容易出错，影响了整体的工作效率。

报告质量参差不齐

不同人员对法规的理解差异及数据完整性的不足，使得报告的质量不稳定。缺乏统一的标准和严格的审核机制进一步加剧了这一问题。

信息整合难度大

来自多个系统的数据格式各异，跨地域的数据共享需考虑法律合规性，加上实时性和非结构化数据处理的需求，增加了信息整合的复杂度。

专业人员依赖性强

高水平的反洗钱工作需要深厚的专业知识和经验，但此类人才稀缺且培养周期长，人员流动还可能造成关键岗位空缺，严重影响工作效率。

典型案例

模力时刻

Qwen-plus

处理 大模型理解用户问题及表结构，利用其 NL2SQL 能力获取交易数据，按提示词模板生成报告内容。

效果 数据获取准确率 90%，可以交易报告生成效率从 1 天到 1 小时。

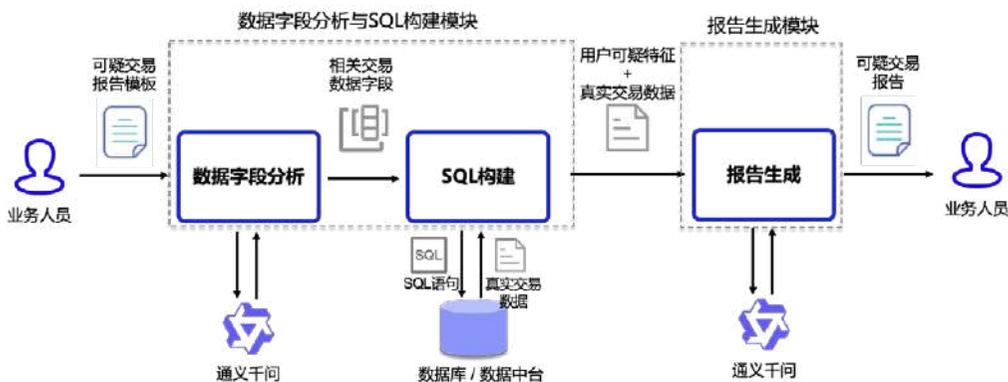
在某银行机构，“生成用户可疑交易报告” Agent 显著提升银行可疑交易监测的效率和准确性。传统的人工撰写报告方式耗时长、易出错，而 Agent 的自动化生成能力大幅缩短了报告的制作时间，减少了人工审核的工作量，降低了人力成本。

“生成用户可疑交易报告” Agent 主要包括“数据字段分析与 SQL 构建”模块和“报告生成”模块。

在“数据字段分析与 SQL 构建”模块中，大模型能够准确理解用户的可疑交易特征，并智能识别出与之相关的交易数据字段。例如，当发现某客户存在频繁大额资金转移的可疑特征时，大模型会迅速分析出需要关注的字段包括交易时间、交易金额、交易对手账户信息等，从而构建出高效的 SQL 语句，从数据库中精准提取相关数据，为后续的深入分析提供有力支撑。

在获取到用户可疑特征及真实的交易数据后，“报告生成”模块利用通义千问大模型强大的语言生成能力，结合用户可疑特征和真实的交易数据，能够自动生成内容详实、条理清晰的报告文本，不仅包含关键的交易数据和可疑特征分析，还能给出合理的结论和建议措施。

在技术上“生成用户可疑交易报告” Agent 实现了大模型与金融业务的深度融合。一方面，大模型在金融领域的应用拓展了其使用场景，证明了其在复杂业务逻辑中的强大适应性和实用性；另一方面，通过创新性地大模型应用于可疑交易报告的撰写，为金融行业提供了一种全新的智能化解决方案，推动了金融科技的发展，为其他金融机构在类似业务场景中的应用提供了参考和借鉴。



价值分析

提升效率与准确性

通过自动化工具和标准化流程，显著提升了报告的撰写速度，减少了手工操作的时间消耗和出错率，提高了整体工作效率。从庞大的数据库中准确提取与可疑特征相关的交易信息，确保数据的准确性和完整性。

标准化与质量提升

解决了因人员对法规理解差异导致的报告质量不稳定问题。通过标准化流程和自动化的报告生成机制，实现了报告质量的一致性。结合大模型的语义理解和生成能力，能够自动生成详尽且专业的可疑交易分析报告，保证了报告的深度和专业度。

简化信息整合过程

克服了来自多个系统数据格式各异、跨地域数据共享复杂等挑战，通过智能的数据字段分析和 SQL 构建，简化了信息整合的难度，使得实时性和非结构化数据处理变得更加高效。

减少对专业人才的依赖

大幅减少了对高水平反洗钱专业人才的依赖，降低了由于人员流动带来的风险，即使在缺乏经验丰富的专家的情况下，也能维持较高的工作效率和报告质量。

APP 零售业务助理

Agent 概述

基于大模型的银行“APP 零售业务助理”Agent 是专门为手机银行 APP 设计的零售业务智能客户服务解决方案。它通过 AI 技术提供智能化、个性化的服务体验。该 Agent 不仅支持账单查询、转账服务等基础零售银行业务的办理，还能通过智能对话引导用户完成复杂操作，实时解答疑问，并根据用户的历史行为推荐相关服务。其目的是改善零售端用户的交互体验，显著提升业务办理效率和服务质量。

需求分析

银行的零售业务智能客服痛点：

用户交互体验差

传统手机银行 APP 操作复杂，用户需要多次点击才能完成简单任务。

客服响应慢

人工客服响应时间长，难以满足用户即时需求。

业务办理效率低

用户需要手动输入大量信息，操作繁琐且容易出错。

个性化服务不足

难以根据用户画像和行为数据提供个性化服务。

零售业务智能助手核心功能需求：

账单查询

支持用户通过自然语言查询账单信息。

转账服务

支持用户通过语音或文本指令完成转账操作。

账户管理

支持用户查询账户余额、交易记录等信息。

智能推荐

基于用户画像和行为数据，推荐相关金融产品和服务。

典型案例

模力时刻

Qwen-72B/32B

处理 识别手机银行上客户 query 意图，调用后台对应知识库和工具生成反馈

效果 意图识别准确率 90%，转人工客服电话量下降 50%

在某银行客户，打造基于大模型智能体的手机银行智能业务助手，面向银发客群提供智能化的零售业务服务，提升手机银行 APP 的用户体验，减少人工客服压力，提高零售业务办理效率。

1 账单查询

- 用户通过语音或文本输入“查询上个月账单”等内容，自动识别用户意图，快速返回账单详情。
- 支持多轮对话，用户可进一步询问具体交易细节。

2 转账服务

- 用户通过语音或文本输入转账相关内容，如“转账给张三 1000 元”，自动识别收款人和金额，为客户提供转账相关 UI 组件卡片，支持用户完成转账操作。

3 账户管理

- 用户通过语音或文本输入“查询账户余额”等类似内容，自动返回当前账户余额和最近交易记录。

4 智能问答

- 对用户提供智能问答服务，包括问候场景、知识问答、主动对话等能力，可对企业相关产品、活动等问题进行针对性回答和详情介绍。



价值分析

☑ 增强用户体验

24/7 在线服务使客户随时随地都能得到帮助，增强了便利性和满意度。

⏏ 提高工作效率

自动化处理常见问题和任务，减少了人工客服的工作量，提升了整体服务效率。

↓ 降低运营成本

减少了对人力客服的依赖，降低了培训和管理成本。

📊 数据驱动决策

收集和分析用户交互数据，为优化产品和服务提供依据。

🛡 安全与合规

严格遵循金融行业法规，确保每一次交易的安全性和合法性。

🔄 促进交叉销售

通过智能推荐，适时向客户介绍其他可能感兴趣的产品或服务，增加收入来源。

APP 信用卡服务助理

Agent 概述

“APP 信用卡服务助理” Agent 是专门为银行手机银行 APP 设计的信用卡客户服务解决方案，它能够提提供信用卡还款、账单查询、智能问答等智能化服务，不仅支持 24/7 全天候自助服务，还能通过多轮对话理解和解决复杂问题。其目的是改善用户交互体验，简化操作流程，显著提升信用卡业务办理效率和服务质量，同时确保高准确性和安全性。

需求分析

银行的信用卡业务智能客服痛点：

用户交互体验差

传统手机银行 APP 操作复杂，用户需要多次点击才能完成信用卡相关业务。

客服响应慢

人工客服响应时间长，难以满足用户即时需求。

业务办理效率低

用户需要手动输入大量信息，操作繁琐且容易出错。

个性化服务不足

难以根据用户画像和行为数据提供个性化服务。

信用卡业务智能助手核心功能需求：

信用卡还款

支持用户通过自然语言完成信用卡还款操作。

信用卡账单查询

支持用户通过语音或文本指令查询信用卡账单信息。

智能问答

支持用户通过自然语言提问，获取信用卡相关问题的解答。

典型案例

模力时刻

Qwen-72B/32B

处理 识别手机银行上客户 query 意图，调用信用卡 Agent 智能体办理业务。

效果 意图识别准确率 90%，工具调用准确率 95%。

在某银行客户，通过基于大模型的“APP 信用卡服务助理” Agent 项目，提升手机银行 APP 的用户体验，提供智能化的信用卡业务服务，提升手机银行 APP 的用户体验，减少人工客服压力，提高信用卡业务办理效率。

1 信用卡还款

- 用户通过语音或文本输入“信用卡还款 1000 元”，自动识别用户意图，完成还款操作。
- 调用信用卡还款相关业务 UI 组件卡片，确保交易准确安全。

2 信用卡账单查询

- 用户通过语音或文本输入“查询本月信用卡账单”，自动识别用户意图，快速返回账单详情。
- 支持多轮对话，用户可进一步询问具体交易细节。

3 智能问答

- 对用户智能问答服务，包括问候场景、知识问答、主动对话等能力；
- 可对信用卡相关产品、活动等问题进行针对性回答和详情介绍。



价值分析

提升用户体验

通过自然语言交互和智能问答功能，大幅简化用户操作步骤，提升用户满意度；支持多轮对话和个性化服务，增强用户粘性。

个性化服务

基于用户画像和行为数据的智能问答，帮助银行更好地满足用户需求，提升客户满意度和忠诚度；支持多轮对话和个性化推荐，提升用户体验。

提高业务办理效率

自动化处理信用卡还款、账单查询等常见业务，显著缩短办理时间；减少人工干预，降低操作错误率。

降低运营成本

减少人工客服压力，降低客服成本；通过智能问答功能，提升用户自助服务率，减少人工客服咨询量。

APP 理财业务助理

Agent 概述

银行“APP 理财业务助理”Agent 是基于大模型的理财业务智能客服 Agent，专为手机银行 APP 设计，旨在通过 AI 技术为 C 端客户提供个性化的理财服务。该工具结合客户画像和需求，利用私域理财产品库，提供智能化的产品推荐、理财购买、理财赎回等服务，显著提升客户体验和业务办理效率。

需求分析

银行的理财业务智能客服痛点：

<p>客户画像不清晰</p> <p>难以精准识别客户的风险偏好、投资目标和财务状况。</p>	<p>产品推荐不精准</p> <p>传统推荐方式依赖人工，难以根据客户需求实时匹配合适的产品。</p>	<p>操作流程复杂</p> <p>客户在购买或赎回理财产品时，需多次操作，体验较差。</p>	<p>服务响应慢</p> <p>人工客服响应时间长，难以满足客户即时需求。</p>
---	--	---	--

理财业务智能助手核心功能需求：

<p>客户画像分析</p> <p>基于客户行为数据和财务信息，生成精准的客户画像。</p>	<p>智能产品推荐</p> <p>根据客户画像和需求，从私域理财产品库中推荐合适的产品。</p>	<p>理财购买与赎回</p> <p>支持客户通过自然语言或简单操作完成理财产品的购买和赎回。</p>	<p>智能问答</p> <p>提供理财相关的咨询服务，解答客户疑问。</p>
--	---	---	---

典型案例

模力时刻

Qwen-72B

处理 识别手机银行上客户 query 意图，调用财富管理后台 API 及资配工具。

效果 意图识别准确率 90%，会话轮数提升 50%。

在某银行客户，意图打造银行“APP 理财业务助理”Agent，提升手机银行 APP 的理财业务体验，增加客户粘性，提高理财产品销售转化率。

1 客户画像分析

- 通过分析客户的交易记录、风险测评、资产配置等数据，生成精准的客户画像，包括风险偏好、投资目标和财务状况。

- 例如，识别出某客户为“稳健型投资者”，偏好低风险、中长期理财产品。

2 智能产品推荐

- 根据客户画像，从私域理财产品库中推荐合适的产品。
- 例如，向“稳健型投资者”推荐低风险的债券基金或保本型理财产品。

3 理财购买与赎回

- 客户通过语音或文本输入“购买推荐理财产品”，自动完成购买操作。
- 客户通过语音或文本输入“赎回某理财产品”，自动完成赎回操作。
- 调用理财相关业务 UI 组件卡片，确保交易准确安全。

4 智能问答

- 对客户提供智能问答服务，包括问候场景、知识问答、主动对话等能力；
- 回答客户理财类问题，如客户通过语音或文本输入“理财产品的收益率是多少？”，自动返回准确的答案。
- 支持多轮对话，客户可进一步询问相关问题。



价值分析

📦 提升客户体验

智能客服通过 24 小时在线服务、多轮对话和个性化推荐，满足客户多样化需求，显著提升理财服务的便捷性和满意度。

⬇️ 降低运营成本

自动化处理大量咨询和交易请求，减少人工客服的工作量，降低银行运营成本。

📊 精准营销与客户留存

通过分析客户行为和偏好，智能客服能够精准推荐理财产品，提升客户留存率和产品销售效率。

⚠️ 增强风险管理

智能客服能够根据客户的风险偏好提供合适的产品建议，帮助客户合理配置资产，降低投资风险。

🌟 提升品牌形象

智能化的理财服务展示了银行的科技实力和创新能力，有助于提升银行在客户心中的品牌形象。

APP 养老咨询规划

Agent 概述

“APP 养老咨询规划” Agent 面向银行等金融机构养老金融场景，提供个人养老金的知识问答、投教问答以及引导客户进一步沟通个人养老的规划方案，为金融机构提供 APP 上的养老金融智能咨询服务能力，辅助手机银行等线上经营场景，提供能说、能懂、会引导的养老金融智能体。提供 API 支持和配置功能，支持接入金融机构标准知识库 /FAQ 库，支持配置业务目标和牵引方向。

需求分析

在零售养老金融场景下，解决当前手机银行、App 中智能机器人 /Bot 遇到的核心挑战：

如何听懂客户的诉求

识别客户的问题，针对问题进行拆解和分析，理解客户的核心意图和诉求。

如何输出专业的内容

基于客户的核心意图和诉求，组织专业的场景答案，给予客户价值内容输出。

如何有效牵引和引导客户

通过思维链等方式，针对客户的理解 / 问题进行牵引，引导客户关注相关业务重点、投教知识等。

典型案例

模力时刻

Qwen-72B

处理 识别养老场景下 query 意图，规划并执行动作，调用问答、工具、推荐等 Agent 工具。

效果 意图识别准确率 90%，工具调用准确性 96%。

在某股份制银行中，以大模型为基础，打造面向养老金融场景的手机银行 APP 智能助手。可以实现养老金融规划、智能化问答、养老资产查询、养老产品推荐等功能。通过该智能体能够理解用户的问题并提供精准、专业的养老金融服务建议。

有效提升养老金融场景下智能体的理解和牵引能力，对零售客户在个养场景的下的问题进行有效的沟通和引导，促进业务目标的达成，以上述案例为例，目标感增强提升 35%，场景深度度增强提升 29.6%。

具体包括如下场景：

1 智能问答

解答用户关于养老金账户、养老理财产品、养老保险等方面的疑问，例如“如何开通养老金账户？”、“有哪些适合我的养老理财产品？”等。

2 养老规划

根据用户的年龄、收入、风险偏好等信息，提供个性化的养老规划建议，例如“我应该如何配置养老资产？”、“我需要准备多少养老金？”等。

3 产品推荐

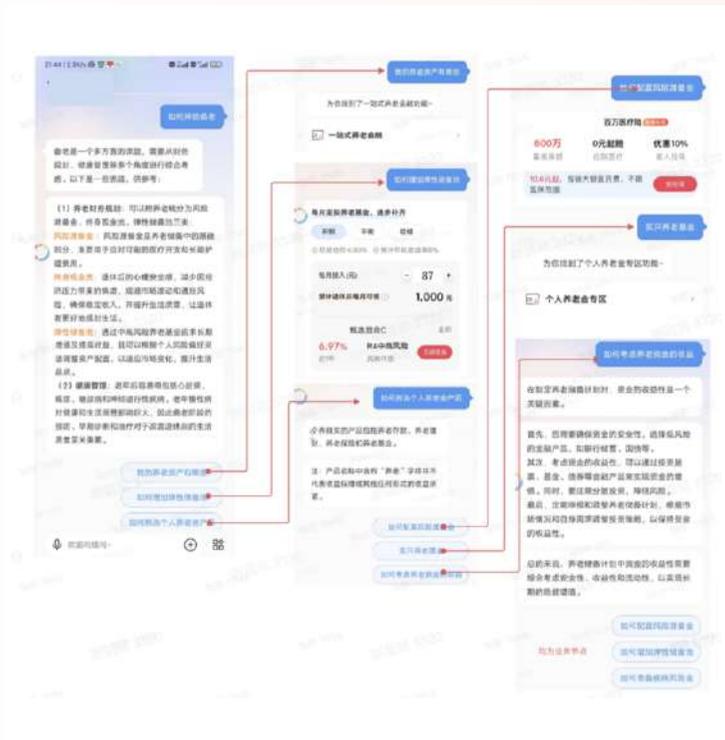
根据用户的养老规划，推荐合适的养老金融产品，例如养老目标基金、商业养老保险等。

4 风险评估

评估用户的养老风险承受能力，并提供相应的风险提示和建议。

5 资讯推送

定期推送养老金相关的政策解读、市场动态、理财知识等资讯。



价值分析

“APP 养老咨询规划” Agent 辅助手机银行、线上机器人在面客过程中的沟通，提供机器人增强 / 替代功能，牵引客户朝业务需要的方向上进行引导（例如 开户、关注产品、计算养老方案、线下约见等）。

增强用户体验

通过提供高效、准确的服务，增强了用户体验，提高了客户满意度。

推动业务增长

吸引对养老金融产品有需求的客户群体，促进相关业务的增长。

展示社会责任

体现了银行在应对人口老龄化挑战方面承担的社会责任，树立良好的企业形象。

数据驱动决策

积累大量用户行为数据，有助于深入了解市场需求，优化产品和服务设计。

证券

投研

研报快读、研报观点问答、路演解析、行业周报观点、日报早评收评、热门事件助手、深度研报写作

投行

股权激励助手、投行法规解读

投顾

投顾资讯简报、理财产品问答助手、营销优先级判断

智能运营

信披报告审核、营销物料审核

研报快读

Agent 概述

在当前信息爆炸的时代，证券公司的研究报告由于其篇幅较长及专业性强，难以有效的对大众进行教育和传播。研报快读旨在通过长图的形式展示研报的核心内容，将复杂的金融分析转化为易于理解的视觉信息，以提高传播效率并激发投资者的阅读兴趣和体验。

在生成式人工智能技术的支持下，基于阿里云通义系列大语言模型及多模态大模型的技术方案，实现复杂研报解析、核心观点提取、内容摘要生成、宣传图生成等功能，为证券公司提供智能、高效的研报快读版生成的解决方案。

需求分析

传统的长图生成办法，需要证券公司分析师 + 外部专业设计公司协作，共同实现从研报到长图的转换，存在成本高、时间长、风格难统一等问题。

成本高

现有技术依靠证券公司分析师和外部专业公司共同完成，需要付出较多的成本对研报内容先做专业的解读，以及长图 UI 的设计排版。

时间长

证券公司分析师平日工作繁忙，完成一份大几十页的研报内容摘要需要花费 2-3 周的时间；再加上和外部专业设计公司的沟通与反复修改，一次研报转长图耗时接近一个月。

风格难统一

不同分析师的背景知识和擅长赛道不一样，外部设计公司的产品在 UI 风格上也有差异，带来的结果是不同研报长图长短风格各异，而证券公司需要一种相对统一能代表公司形象的风格。

典型案例

模力时刻

Qwen-long

- 处理** 超千万长度模型能力，实现对超长研报内容的理解、结构生成、摘要及重点内容提取。
- 效果** 提取准确度达到 95%；转写文本留存率 70%，不偏离原稿含义。

Wanx2.1

- 处理** 基于研报内容与提示词指引，生成研报快读的背景头图。
- 效果** 素材制作周期上从原先 2-3 周缩短至 2-3 天。

在某证券公司，基于通义千问大模型，总结研报核心观点，提取关联图表；利用通义万相文生图能力，基于研报主题生成对应头图；最后，采用生成式 AI+ 专家编辑的模式，将大模型产出内容输出到一个可编辑应用当中，专家微调或确认后即可定稿。

1 研报的上传与管理

原始的研报文件可以统一放置在对象存储上管理，应用系统在用户登录后，按权限展示相关研报，供用户选择。

2 研报的预处理

通过预设规则，以及 OCR 等小模型的协作，对原始研报中的文字内容、图片内容、表格内容，做充分的提取，形成 markdown 结构化数据。

3 长图内容生成

将前述 markdown 中的主体内容和段落等结构化信息，送给大语言模型做长图内容的摘要和生成，此处采用 Qwen-long 大语言模型，支持超长研报的文本输入，配合系统预置的 Prompt 提示词，产出文字内容。

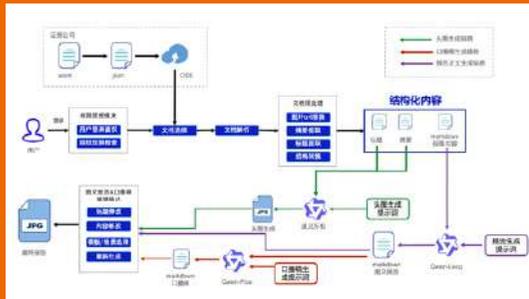
4 长图头图生成

与此同时，预处理后的标题、摘要、目录等高信息量文本，会作为提示词的一部分，送给通义文生图多模态大模型通义万相处理，生成符合特定风格和-content 相关的头图，作为长图一部分与文本做拼接。

5 口播稿生成

基于长图中的关键摘要文本，结合上下文信息做适度的内容扩展，生成适合视频场景的文字长度内容。

总体技术链路如下图所示：



实现效果如下图所示：



基于大模型自动生成内容及图表，提取准确度达到 95%；转写文本留存率 70%，不偏离原稿含义，避免合规风险；在制作周期上从原先 2-3 周缩短至 2-3 天，极大提升了效率。

价值分析

传统的长图生成办法，需要证券公司分析师 + 外部专业设计公司协作，共同实现从研报到长图的转换，存在成本高、时间长、风格难统一等问题。

增强吸引力

相比冗长的专业研报，通过生成长图，采用直观的视觉元素，提升报告吸引力，激发投资者兴趣，更容易在短视频平台上传播。

提高理解效率

使用大模型对研报内容进行提炼和摘要，结合原始研报中重点图表，加速信息吸收，提高阅读效率，节省投资者获取信息的时间成本。

提效降本

券商员工可借助大模型可自助高效完成研报长图生成，生产周期从周级将至天级，提升工作效率。

拓宽传播渠道

长图的形式更适应多平台分享，如券商的公众号和 App，扩大受众范围，增强报告影响力。

研报观点问答

Agent 概述

在证券公司的投资研究部门，有大量的研报作为分析师日常查阅和分析的信息来源。传统上大家使用研报的方式是，通过关键字检索出相关研报，然后打开逐页阅读，寻找关注的信息，形同大海捞针。

在生成式人工智能技术的加持下，可以更高效的完成此类工作。首先把研报的内容进行识别，包括文字、图片、表格、文档结构等，然后将其向量化之后存储在向量数据库中；然后当用户的查询请求过来后，精准检索、匹配和召回相关文档内容；最后把用户查询和召回的文档片段一并送给大模型进行生成式回答。

需求分析

传统的研报阅读方式有如下痛点：

检索准确性不高

传统方式主要依靠研报关键字进行匹配检索，在准确性上有所不足，另外也无法捕获到文本背后更深层次的语义关联性。

阅读耗时较长

检索到研报以后，通常要花费大量的时间去阅读，几十页的研报，可能只有一两页是用户关心的，这造成了极大的阅读时间浪费。

观点生成较慢

在查阅完研报内容后，观点需要分析师或其他用户人工编写，对总结摘要能力和文本写作能力有较高要求，且花费时间。

典型案例

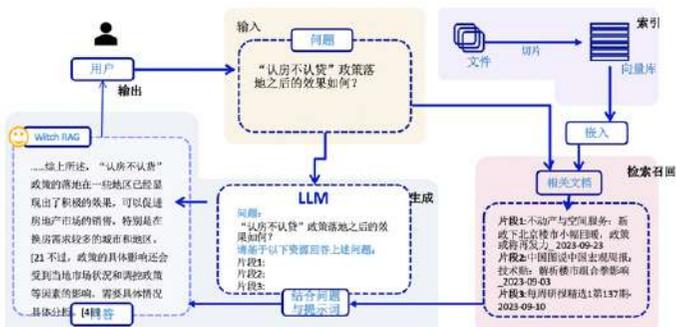
模力时刻

Qwen-plus

处理 理解用户问题，从 RAG 召回的知识片段中，筛选出语义相关内容，生成回复。

效果 研报内知识问答准确度达 90%，检索知识效率从 1 小时到 1 分钟。

右图展示了一个完整的，构建研报知识库，并基于大模型搭建 RAG (Retrieval-Augmented Generation, 检索增强生成) 的流程图：



1 文件切片与索引

大量研报被切片处理，并存储在向量库中进行索引，以便后续快速检索。

2 用户输入问题

用户提出一个问题，通常需要结合特定场景对用户问题进行改写和扩写。

3 检索相关文档

系统基于用户的问题从向量库中召回相关的文档片段。

4 结合问题与提示词

检索出的相关文档片段会被进一步处理，结合用户的问题生成提示词。

5 LLM 生成回答

结合问题与提示词，通过大模型生成最终的回答。

6 输出回答给用户

最终生成的回答通过系统返回给用户，提供详细的分析和见解。

整个流程通过结合检索技术与大模型生成能力，为用户
提供精准、详尽的答案，帮助券商的研究人员快速获取
所需的信息。

右图是主流投研工具，支小助的产品问答效果示意图：



价值分析

更准确

引入语义化知识检索，召回更多相关内容，避免遗漏关键信息。

更易用

新的 LUI 交互方式，让客户可以通过提问的方式快速获取相关知识点，不需要每一篇都去阅读，按需延展阅读即可。

更高效

不需要用户从零开始写总结摘要，大模型先生成一个基础底稿，人工编辑优化即可。

路演解析

Agent 概述

资本市场上有大量的路演会议视频，其中有很多有价值的信息，从业人员需要花大量的时间去观看视频直播或回放，获取信息的效率比较低，无法及时掌握市场动态。

在生成式人工智能时代，证券公司、上市公司、机构投资者等业内人员，通过 ASR+ 大语言模型，把路演视频内容的信息提取出来，再基于大模型完成内容总结、发言人区分、信息摘要、观点总结等任务，把结构化后的数据给到从业人员查看，极大的提升了信息获取的效率。

需求分析

在证券行业，观看路演会议视频以提取关键信息是一个耗时的过程，给从业人员的工作带来诸多不便：

时间消耗大

从业人员需要逐帧观看视频来捕捉重要信息，这一过程极为耗时。长时间的专注观看容易导致疲劳，降低信息吸收效率。

信息检索困难

在长视频中寻找特定信息点如同大海捞针，缺乏有效的索引工具。没有快速定位功能，使得查找特定话题或数据变得尤为繁琐。

内容理解依赖性强

对于非母语使用者或者专业术语不熟悉的人员来说，理解视频内容可能更加困难。缺乏即时翻译或解释功能，增加了理解障碍。

无法有效利用碎片化时间

由于视频长度较长，不适合在短暂的休息时间或通勤途中快速浏览获取要点，不能充分利用零散时间为后续工作做准备。

典型案例

模力时刻

Qwen-plus

处理 基于路演视频的文字版内容，进行关键信息提取、会议总结摘要。

效果 关键实体提取准确率 95%，会议总结摘要准确率 90%。

基于在券商研究所、上市公司、机构投资者闭环生态中构建的路演为核心的投研信息，通过大语言模型的应用与 AI 智能分析，为金融机构及投研人员提供一站式智能投研服务信息，帮助投研分析师高效便捷地开展业务，提高投研分析决策效率。



1 语音转文字

通过 ASR 或多模态音频理解大模型，把路演会议视频中的文字提取出来，过程中进行关键发言人的区分，把对话内容结构化。

2 会议内容处理

根据下游的具体任务需求，使用 Prompt，基于大模型对原始会议文字做加工。常见的任务包括会议内容摘要，让研究员快速浏览长会议内容；会议问答对抽取，让大模型从大段会议文本中，自动提取【question - answer】对，能够以问答的形式获取会议关键信息。

价值分析

使用大语言模型，把路演的非结构化信息转成文本，再加上不同提示词任务，让大模型帮助从业人员快速提取到想要的信息，提升工作效率。

📺 节约了看长时间路演视频的时间

通过大语言模型技术，加上语音小模型，能够自动转录视频中的语音为文字，并且进行语义分析以提取关键信息。这使得从业人员无需逐帧观看视频，而是可以直接阅读或搜索文本摘要，极大地节省了时间。

📖 内容更加容易理解和吸收

借助大模型强大的语言力和世界知识，对于专业术语的理解问题，大模型可以基于其庞大的知识库提供解释和上下文关联，帮助用户更好地理解复杂的专业内容。这对于提升跨文化沟通效率和专业知识普及具有重要意义。

🔍 提升了关键信息提取效率

大模型可以构建智能知识索引，通过对视频内容的深度学习，识别并标记出重要的时间节点、主题以及关键词。这样，用户可以通过简单的查询快速定位到感兴趣的片段。此外，大模型还可以根据用户的兴趣推荐相关内容，进一步优化信息获取流程。

行业周报观点

Agent 概述

传统模式下，销售、运营及投研人员需要耗费大量时间收集数据、新闻、政策信息，并进行分析和观点整理，不仅效率低下，还容易出现信息遗漏或不准确的情况。特别是以金融为代表的高度专业化行业对严谨性、专业性和可靠性都有着极为严格的要求。

面向智能投研场景下，通过整合市场上公开的各种研报和第三方数据源，结合内部基金经理的观点，利用 COM (Chain-of-Mind) 技术，将研究员的思维模式融入其中，采用 RAG 架构进行实体化的检索，实现对复杂问题的深度解析。依托大模型提供生成能力，自动生成行业周报观点，不仅提高了内容生成的速度和准确性，还显著增强了业务人员对市场动态的把握能力，提升了投研人员的工作效率和服务质量，为公司创造了更高的商业价值。

需求分析

在智能投研场景下，整合多源异构的信息，生成一份高质量研报，当前有如下挑战：

数据整合与处理效率低

在资管行业中，销售、运营和投研人员需要频繁地收集、整理和分析大量的市场信息，如行业估值、指数、新闻、政策等。传统方式依赖人工进行数据的检索、汇总和整理，不仅耗时费力，还容易出现遗漏或错误。

内容生成质量参差不齐

线下渠道和线上渠道需要持续输出高质量的内容来陪伴用户，但现有的内容生成工具无法保证内容的准确性和时效性，导致用户体验不佳，影响业务效果。

总之，对于销售人员来说，需要用更为高效的工具自动生成行业周报，涵盖行业估值、指数、新闻、政策等多维度信息，推动更多复杂场景和问题的解决。

典型案例

模力时刻

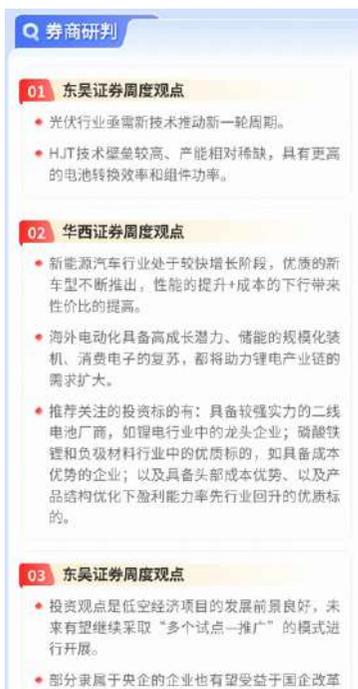
Qwen-plus

处理 基于市场资讯内容，自动生成一周要点新闻与观点解读。

效果 行业周报生成极大效率，从1天到10分钟。

行业周报观点应用针对销售人员进行行业相关信息的物料辅助生产，以提升对应运营动作的效率及效果。行业周报观点应用通过自动获取对应行业的估值、指数、新闻、政策、大事、研报观点等多维度信息，确保观点内容的时效性和事实准确性；针对其中的抽象金融问题，如光伏行业的投资时机，应用模拟研究员的思维模式，从多个角度思考和收集信息，并结合研究员标注的专业图谱，进行全面洞察和分析。

当前行业周报观点应用已覆盖全行业，辅助生产内容百余篇，渗透至各个销售陪伴等业务环节。下图为光伏行业周报应用示例：



价值分析

提升工作效率与质量

自动化生成行业周度观点报告和解读，极大减少了销售、运营和分析人员的数据收集和整理工作量，同时保证了信息的准确性和时效性。

智能化与精准化

在处理问题时更加贴近实际应用场景，提高了回答的针对性和专业性。

日报早评收评

Agent 概述

在快节奏的金融市场中，无论是个人投资者还是机构参与者，都需要及时准确地获取市场动态和专业分析来辅助决策。而传统模式下，销售、运营及投研人员需要耗费大量时间收集数据、新闻、政策信息，并进行分析和观点整理，不仅效率低下，还容易出现信息遗漏或不准确的情况。

日报早评收评应用依托大模型智能技术，具备强大的信息整合与分析能力。它如同理财师的“智能助手”，每日自动运行，精准聚焦金融市场动态，无论是开盘前的早评，还是收盘后的收评，都能快速、高效地为理财师呈现有深度、有价值的分析报告，成为理财师洞察市场、服务客户的得力工具。

需求分析

资本市场信息繁杂，从业人员要想要掌握每日全面、及时、准确的资讯，存在以下困难：

信息爆炸与时效难题

金融资讯海量繁杂，理财师需耗费大量时间从众多渠道筛选当日关键信息，难以及时把握稍纵即逝的投资机会，且容易错过开盘前的黄金准备时段。

专业解读深度与广度

解读热点事件、行情走势需深厚专业功底，理财师受限于精力，难以对各类资产、多元板块进行全方位且深入的剖析，影响为客户提供精准投资建议。

日常工作流程繁琐

手工收集整理数据、撰写早评收评报告，流程冗长，理财师无法将更多精力投入客户沟通与个性化方案制定，制约业务拓展效率。

对于理财师来说，迫切需要每日早评和收评的辅助工具，帮助他们快速获取最新市场信息并进行解读。

典型案例

模力时刻

Qwen-plus

处理 基于当日即时的市场资讯内容，自动生成分析洞察与评价。

效果 每日快报类工作，生成效率极大提升，从 1 小时到 5 分钟。

日报早评收评应用可以多源扫描权威财经媒体、研报平台、整合金融行业数据库，提供深入的行业分析、资金流向、市场行情、政策变化等信息，帮助投资者更好的理解投资环境。

1 早评精细解读

对热点事件进行深度剖析，结合历史数据、市场模型，阐释事件对股市、债市、汇市等走向的短期冲击，为理财师早会提供清晰逻辑脉络，辅助制定当日投资策略基调。

2 收评全面洞察

收盘后，迅速汇总行情数据，通过智能算法识别强势板块、异动个股，对比历史行情解析资金流向、板块联动，以可视化图表与简洁文字，展现当日市场全貌，助力理财师复盘总结。

日报早评收评不仅为理财师提供了一个高效的工作辅助工具，同时能够满足广大投资者对于即时性、权威性和实用性信息的需求，通过这一服务，用户可以更加从容地应对市场的变化，实现财富的稳健增长。



险股也呈现下跌态势，中国太保领跌，可能反映了投资者对保险行业近期业绩或政策环境的担忧。尽管部分板块回调，但整体市场仍以上涨为主，超过2700只个股实现上涨。

热点解读

白名单政策重塑房市融资，推动市场复苏

在2024年的房地产市场，中央金融工作会议催生的“白名单”政策成为行业焦点。这是一项旨在精准扶持优质项目的创新举措，由住建部建设部与金融监管总局联手推动，旨在重塑融资秩序，保障业主权益，维护市场稳定。通过设立严格的评估标准，只有符合条件的项目才能进入名单，确保资金流向风险低、交付有保障的建设项目，从而缓解开发商压力，提振金融机构信心，加速贷款拨付。

随着“白名单”规模扩张至4万亿，其效应愈发显著，直接推动项目施工，助力“保交付”目标，为市场复苏注入强大动力。10月16日审批通过的2.23万亿元贷款，彰显了政策的执行力度和成效。政策的透明度和针对性增强了投资者与消费者的信心，引导资金流向更稳健的开发商和项目，房地产资产价值得以稳定，购房者决策也更为理性。

政策不仅着眼增量，存量调整如取消限购、限售等同步进行，短期去库存，长期优化行业结构。国新办发布会的明确导向，进一步稳定了市场预期。房地产“白名单”政策的实施，就像催化剂，正在重构行业融资格局，激发市场活力，为投资者指明方向，推动市场在稳定中寻找新的增长点，助力经济健康发展。未来，我们有理由相信，房地产市场将在这一政策的引领下，煥发新生，开启新篇章。

风险提示

数据来源 Wind、iFind、澎湃新闻、新浪财经、历史业绩不代表未来表现，观点仅供参考，不构成投资建议，基金有风险，投资需谨慎。我国基金运作时间短，不能反映股市长期所有阶段，过往历史业绩不代表未来表现，投资者应详读《基金合同》、《招募说明书》等基金法律文件，了解基金的风险收益特征，并根据自身的投资目的、投资期限、投资经验、资产状况等判断基金是否和投资者的风险承受能力相适应。指数基金存在跟踪误差。

价值分析

效率飞跃

理财师每日节省数小时信息搜集整理时间，将精力聚焦客户需求挖掘、资产配置优化，客户响应速度提升，服务质量显著升级。

专业赋能

凭借 Agent 深度专业解读，理财师为客户提供投资建议时，准确率提升，复杂市场环境策略适配性更强，客户资产组合收益稳定性增强。

机构竞争力提升

多机构应用实践表明，该 Agent 助力理财团队提升整体效能，吸引高净值客户流入，创造额外业务收入增长，稳固机构在财富管理领域领先地位。

热门事件助手

Agent 概述

热门事件助手是一款依托大模型技术的智能问答机器助手，专注于热门事件的解读。无论是输入事件类问答，还是提供新闻链接，它都能迅速剖析，为用户提供深度见解。通过对海量信息的整合、分析，输出各方观点与市场分析，成为各行业人员把握时事动态的得力工具。

需求分析

资本市场上，及时获取准确的一手热门事件是一项非常重要的工作，但往往存在以下问题：

信息获取困难

信息爆炸时代，热门事件海量涌现，销售、运营、投研等业务人员难以及时筛选有价值信息，容易错失商机或投资时机。

时效性差

人工解读事件耗时长、视角局限，难以兼顾多维度分析，无法满足快节奏业务场景对时效性、全面性的要求。

分析能力不足

普通新闻获取渠道仅呈现事实，缺乏针对业务需求的深度挖掘，难以直接指导实践决策。

典型案例

模力时刻

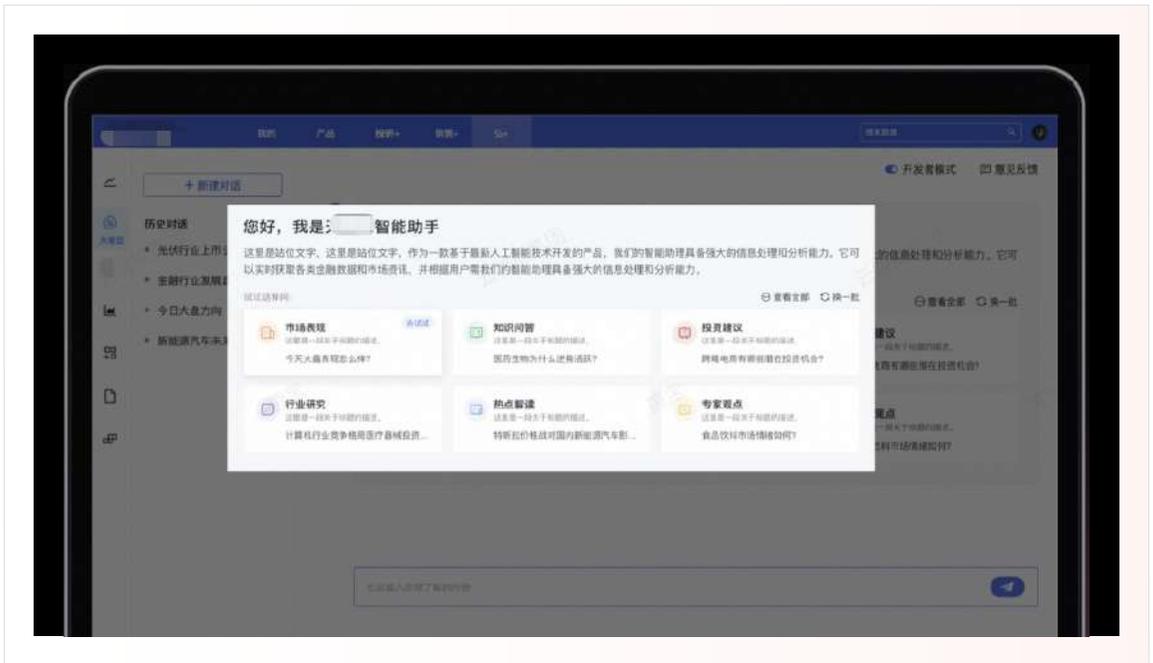
Qwen-plus

处理 基于当日即时的市场热点，自动生成分析洞察与解读报告。

效果 热点解读效率极大提升，从 1 小时到 5 分钟。

热门事件助手是针对热门事件进行解读的智能问答机器助手，其主要是针对事件类问答或者对应新闻链接进行解读分析，同时可对销售、运营、甚至投研人员进行辅助事件解读。区别于新闻链接可直接获取新闻内容，对于事件类问答首先会对该问题进行事件检索以获取对应事件内容。而后会根据大模型技术能力进行多维度拆解、分析与理解，包括但不限于市场影响、行业趋势、公众反应等，提供各方观点和专家意见，帮助用户全面了解事件的各个方面，从而得到各方观点、市场分析等多角度时间解读。

该热门事件助手已服务解读各类大小热门事件 1000 余次，极大地提升了各业务人员对于事件的时效性把控。



价值分析

提升效率

通过自动化工具代替人工操作，大幅减少寻找和处理信息所需的时间，提高工作效率。

增强竞争力

及时掌握市场脉搏有助于企业快速响应变化，抢占先机，提升市场竞争力。

优化决策

基于客观数据分析作出更明智的选择，降低因主观判断失误带来的损失，提高决策质量。

深度研报写作

Agent 概述

在投资研究领域，研究员通常需要对所覆盖的公司股票进行深入分析和评估。然而，由于每个研究员能够覆盖的公司数量有限，大量未被充分研究的个股可能存在关键信息遗漏的问题。为解决这一问题，我们引入了基于大语言模型的自动化深度研报写作工具。该工具结合了公司积累的研报写作框架与经验，实现了高效且深入的自动化研报撰写。

需求分析

资本市场上上市公司众多，每家公司的信息来源众多且分散，投研人员在分析企业数据，编写研报的过程中，存在如下困难：

覆盖范围有限

每位研究员能够覆盖的公司数量有限，导致许多个股未能得到充分研究，可能遗漏关键信息。

信息过载

对这些个股进行深入研究可能会导致信息过载，进而引发人力资源紧张的问题，增加研究人员的工作负担。

效率低下

传统的人工研报写作流程耗时较长，难以快速响应市场变化，影响决策时效性。

资源分配不均

长尾公司的研究资源相对匮乏，难以获得足够的关注和分析。

对于研究员来说，需要高效的工具自动生成深度研报，涵盖行业估值、指数、新闻、政策等多维度信息，推动更多复杂场景和问题的解决。

典型案例

模力时刻

Qwen-plus

处理 基于专业解读框架，以及企业知识库，自动生成研报内容。

效果 研报撰写效率极大提升，从1周到1小时。

公司深度研报写作工具为投资研究人员提供了一个高效且深入的途径，以探索和评估公司的详细信息及其潜在价值。在实际投研业务中各个研究员所覆盖公司股票数量有限，导致大量未被充分研究的个股存在，从而可能遗漏关键信息。同时，对这些个股进行深入研究可能会导致信息过载，进而引发人力资源紧张的问题，这无疑增加

了研究人员的工作负担。

为了解决上述问题，借助大模型技术，结合公司积累的研报写作框架与经验，实现了自动化深度研报写作。该系统利用自研 Chain-of-Mind 思维导图链路思考、规划、分析等能力进行深度研报撰写。目前，公司深度研报写作已覆盖可定制 / 自动化写作大纲、开放 / 内部信息获取、信息归纳筛选、指标图表生成、总结思考规划以及长文写作等更多模态的能力。

该公司深度研报写作应用目前已初步展现初级研究员深度研报写作能力，极大地提升了公司深度研报的覆盖，特别是长尾公司个股，提升投研人员研究行为的效率数倍。

股份深度研究报告

一、公司概况

1.1 公司由来及发展历史

并于 2012 年 5 月 7 日改制为股份有限公司。2017 年 3 月 17 日，在上海证券交易所正式挂牌上市。迈入资本市场的新阶段。公司自成立以来，一直专注于汽车内、外饰件系统零部件及模具的设计、制造和销售。作为全国百家优秀汽车零部件供应商之一，已成为众多汽车品牌的一级配套商。公司与广菲克、上海大众、一汽大众等合资品牌，以及上海汽车、吉利汽车、北汽等自主品牌建立了长期稳定的合作关系。同时，还为一汽解放、东风、福田戴姆勒、中国重汽、陕重汽等大中型商用车企业提供配套服务。

在质量管理方面，恪守“追求卓越、超越期望”的质量方针，先后通过了 ISO9000、TS16949、IATF16949 等行业体系认证，致力于为客户提供优质的产品和服务。实际控制人。

分别第三、第四大股东，分别持有公司 4.0% 和 2.6% 的股份。这种家族式的股权结构，为公司的长期稳定发展提供了有力支持。

1.2 公司主业构成

主要从事汽车内、外饰件系统零部件及模具的设计、制造和销售，公司的主营产品包括仪表板总成、门板总成、保险杠总成、预框总成、内饰件和外饰件等。

根据公司历年主营业务收入占比（如下表所示），仪表板总成和门板总成是公司的核心产品，收入占比逐年提高，2019 年，仪表板总成为 0%，2020 年占比 59.42%，到 2023 年已提升至 66.05%；门板总成从 2020 年的 13.75% 增加到 2023 年的 16.84%，此外，内饰件和外饰件业务在 2023 年开始显现，占比分别达到 3.51% 和 1.68%。

主营业务项目	2019 年	2020 年	2021 年	2022 年	2023 年
仪表板总成	0	59.42	61.82	65.64	66.05
门板总成	0	13.75	15.04	18.32	16.84
保险杠总成	0	0.41	0.28	0.62	0.87
预框总成	0	5.43	3.34	0.78	1.27
内饰件	0	0	0	0	3.51
外饰件	0	0	0	0	1.68
其他业务	16.06	13.83	12.22	9.95	10.08
其他主营业务	0	4.38	4.17	3.32	0

数据来源：中国 A 股上市公司年报

从上表可以看出，公司的主营业务重心逐渐向仪表板总成和门板总成倾斜，体现了公司在这两个领域的市场竞争力。同时，内饰件和外饰件等副业也开始贡献收入，显示出公司产品线的不断丰富和业务拓展的成效。在主营业务收入总额方面，公司连续几年保持增长态势。以下是公司历年主营业务收入总额：

1.3 公司历史股价复盘

7 近期股价表现强劲，多次创下历史新高。据资料显示，截至 11 月 8 日，本周累计上涨 8.58%，周总成交额达 17.65 亿元，收盘价为 53.75 元，随后，在 11 月 11 日的交易日内，股价再次大涨 5.21%，收于 56.35 元，创下历史新高。

与行业龙头竞相崛起。据东方财富网显示，本周上证指数上涨 5.31%，汽车零部件板块上涨 11.78%，在同行业公司中，新普集团（601699）上涨 22.77%，德业科技（002929）上涨 18.71%，奥翔佳（002259）上涨 15.83%，倍渡科技（000700）领涨涨幅位居前列。

以下是部分同行业公司股价表现：

代码	名称	最新价	涨跌幅	10 日涨幅	1 月涨幅
002085	万丰奥威	22.21 元	-1.94%	26.4%	7.82%
601689	拓普集团	52.08 元	22.77%	13.35%	21.02%
002238	德业科技	14.10 元	5.05%	4.96%	8.96%
002229	奥翔佳	3.22 元	15.83%	15.0%	11.02%
000700	倍业科技	8.96 元	18.21%	20.87%	14.18%

股价上涨的背后，是公司良好的业绩支撑和市场的积极评价。据东方财富网显示，国信证券、天风证券、国海证券等多家券商近期发布了“买入”的研报，分别给予“优于大市”“买入”等评级。AI 点评指出，新普集团在一个月内获得了多家券商的关注，其中多家给予“买入”评级。

业绩方面，公司核心客户销量增长迅速，根据国信证券研报显示，2023 年三季度归母净利润同比增长 30%，天风证券的研报也指出，公司 2023 业绩同比大增 41.54%，盈利能力改善明显。

综合来看，公司的持续上涨，反映了市场对该公司业绩增长和未来发展的信心。随着公司可研数据逐步全球化布局，核心客户销量提升，以及新产品的持续放量，股价有望保持良好表现。

价值分析

📈 提高工作效率

显著提升了公司深度研报的覆盖范围，特别是对于长尾公司个股的研究效率，使研究员能够更专注于核心任务。

🔍 增强决策支持

通过自动化工具，研究员可以更快地获取和分析所需信息，从而做出更加准确的投资决策。

🏆 提升专业水平

借助大模型的强大能力，生成高质量的研报，帮助公司在竞争激烈的投研市场中脱颖而出。

综上所述，通过引入自动化深度研报写作工具，我们不仅解决了当前研报写作中的诸多痛点，还显著提升了整体投研工作的效率和质量。这将有助于公司在复杂的市场环境中保持竞争优势，并为投资者提供更为精准和及时的投资建议，提升决策的专业程度。

股权激励助手

Agent 概述

在当前复杂多变的经济环境下，投行业务人员为企业设计和实施股权激励方案时面临着诸多挑战。首先，每个企业的业务模式、发展阶段以及人力资源状况都有所不同，这意味着每一份股权激励计划都需要量身定制，以确保能够有效地激励员工并促进企业长期发展。此外，制定股权激励方案需要综合考虑法律合规性、财务健康度、市场条件以及员工的期望值等多方面因素，这对业务人员的专业知识和技能提出了很高的要求。同时，在执行过程中，如何准确评估员工的表现，并根据业绩考核结果合理分配激励股份，也是业务人员必须面对的一个难题。

引入大语言模型技术后，可以从多个维度提升股权激励业务的效率和质量。首先，大语言模型 + 工具 Agent 可以帮助业务人员快速获取最新的法律法规信息和市场动态，确保股权激励方案的设计符合最新的监管要求。其次，通过自然语言处理技术，可以对大量历史数据进行分析，为激励对象的选择、授予额度的设定提供科学依据，提高决策的准确性。最后，利用大语言模型的推理能力，可以在短时间内模拟不同的激励情景及其可能带来的影响，帮助企业找到最优的激励策略。这样不仅提升了工作效率，也使得整个流程更加透明化和智能化。

需求分析

在企业设计和实施股权激励方案时，业务人员面临着以下主要痛点：

多方面信息需要考虑，很复杂

设计股权激励方案涉及到法律合规、财务健康、市场条件以及员工期望等多方面因素的平衡。确保所有这些元素都能得到妥善考虑，并且不会相互冲突，是业务人员面临的一个巨大挑战。任何一点处理不当都可能导致激励效果大打折扣或引发不必要的法律风险。

定制化需求高

每个企业的独特性要求股权激励计划必须高度定制化。业务人员需深入了解企业的业务模式、发展阶段及人力资源状况，才能制定出既能激励员工又能促进企业发展的方案。这需要耗费大量的时间和精力进行调研与分析。

典型案例

模力时刻

Qwen-plus

处理 基于股权激励和企业信息知识库，理解用户 query 要求，召回知识片段，进行理解、整合、分章节生成方案。

效果 问答过程，信息检索准确率 85%；生成过程，效率提升，从 3 天到 1 个小时。

基于大模型和小模型融合的股权激励助手，可以通过多轮问答（QA）和股权激励知识库、客户信息知识库，帮助客户经理快速学习和检索相关知识，基于对客户现状的理解，产出有针对性的股权激励方案。

1 业务意图理解

客户经理或业务人员以问答的形式，提交查询请求，大模型准确理解业务人员的具体需求和意图，包括请求中涉及到的专有名词，大模型分析查询中的关键词、上下文信息以及潜在的隐含需求，确保后续处理步骤能够精准对接业务人员的真实需求。例如，如果查询涉及股权激励方案的设计，模型需要识别出是关于激励对象的选择、考核标准还是行权条件等具体方面。

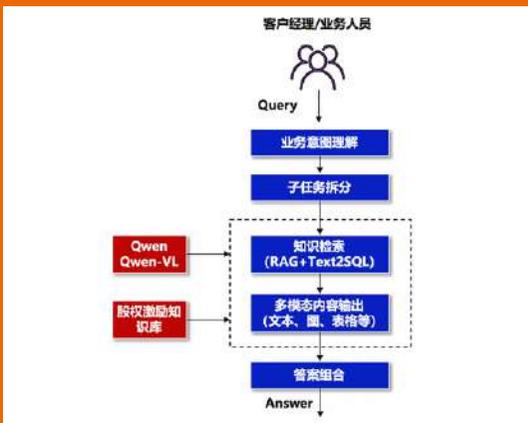
2 子任务拆分

在理解了业务意图后，系统将复杂的查询任务拆解为多个子任务。这一步骤旨在提高处理效率和准确性。每个子任务可能对应不同的数据源或处理逻辑，例如激励对象的选择、绩效考核标准的设定、行权条件的制定等。通过这种方式，可以更细致地管理和优化每个部分的工作流，确保最终输出的方案既全面又精确。

3 知识检索

针对拆分后的子任务，系统利用知识检索技术（Retrieval-Augmented Generation, RAG）结合 Text2SQL 方法从股权激励知识库中提取信息。RAG 技术能够高效地从大量文档中检索出与子任务相关的片段，以非结构化数据为主；而 Text2SQL 则将自然语言查询转化为数据库查询语句，从而快速获取所需的结构化数据。这一过程确保了信息的准确性和时效性，为后续的内容生成提供坚实的基础。

通过整合先进的技术和专业知识，股权激励助手能够显著提高客户经理的工作效率和业务转化率。它不仅提供了一个全面的知识库来支持决策制定，还通过数据分析帮助识别潜在的机会和挑战。这种综合性的解决方案有助于企业在激烈的市场竞争中脱颖而出，并实现长期可持续发展。



4 多模态内容输出

基于大语言模型 + 多模态大模型，系统产出的方案内容除了包含文本，还可以生成图表和表格等多样化形式，形象地展示激励对象、考核指标、行权进度等关键信息。多模态输出使得信息更加丰富和易于理解，有助于业务人员快速掌握核心要点。

5 方案组合生成

最后，系统将所有子任务的输出进行整合，形成一个完整的答案。这一阶段确保了最终结果的连贯性和完整性，使业务人员能够获得一个全面且结构化的解决方案。同时，系统还会对答案进行质量检查，确保其准确无误，将股权激励方案的内容以清晰易懂的方式呈现给业务人员。

价值分析

有了大模型的加持，生成股权激励方案的过程相比原有方法带来了显著的价值提升，更好的辅助从业人员在面向企业客户时开展股权激励相关业务。具体体现在以下几个方面：

📄 高效信息检索与整合

利用 RAG 技术和 Text2SQL 查询转换，大模型可以快速准确地从海量数据中检索出关键信息，并将这些信息有效地整合起来。这大大缩短了信息收集的时间，提升了工作效率。

📊 智能内容生成与优化

大模型能够自动生成多模态内容，包括文本、图表和表格等，不仅使信息呈现方式更加多样化，而且让复杂的股权激励方案变得直观易懂。此外，还能根据反馈不断优化内容，确保最终输出的质量。

🎯 增强决策支持能力

借助于大模型的数据分析和预测能力，业务人员可以获得更为科学合理的决策建议，如激励对象的选择标准、绩效考核指标设定等。这有助于提高决策的准确性和前瞻性，更好地引导企业战略实施。

投行法规解读

Agent 概述

在投资银行业务中，无论是进行首次公开募股（IPO）、再融资还是并购重组等操作，都需要严格遵守相关法律法规。然而，在实际工作中，从业人员往往面临法规查询和解读的复杂挑战。一方面，由于金融市场的快速发展以及监管政策的不断更新，相关的法律条文、规章制度也日益繁杂，这使得业务人员在寻找适用法规时常常感到无从下手。另一方面，即使找到了相应的法规文本，理解其具体含义及应用范围同样不易。

有了大语言模型技术的支持，构建一个基于检索增强生成（RAG）的知识库将极大地提升上述流程的效率与准确性。通过整合来自多个权威来源的投资银行业务法律法规，创建一个全面且动态更新的知识库。利用大语言模型强大的自然语言理解与生成能力，解析用户查询，识别客户意图，准确筛选出相关法规条款，并提供通俗易懂的解读，帮助业务人员快速掌握要点。

需求分析

投行工作人员在处理大规模专业文献或行业法规时，将面临的数据量大、格式复杂、知识专业性强等特点，需确保数据的有效利用与准确解读，面临主要难点：

数据复杂性挑战

数据格式多样

需要进行格式转换和统一处理，以确保数据能够被有效利用。

数据量庞大

可能导致检索效率低下，需要优化检索算法和数据存储结构。

数据质量参差不齐

存在噪声数据、重复数据和不准确数据，需要进行数据清洗和验证。

知识专业性难题

专业术语繁多

理解和处理难度大，可能导致召回不准确和模型理解错误。

知识体系复杂

知识点之间的关联和逻辑关系难以把握，影响推理和总结的准确性。

典型案例

模力时刻

Qwen-VL

处理 解析法规中图片、表格等半结构化内容。

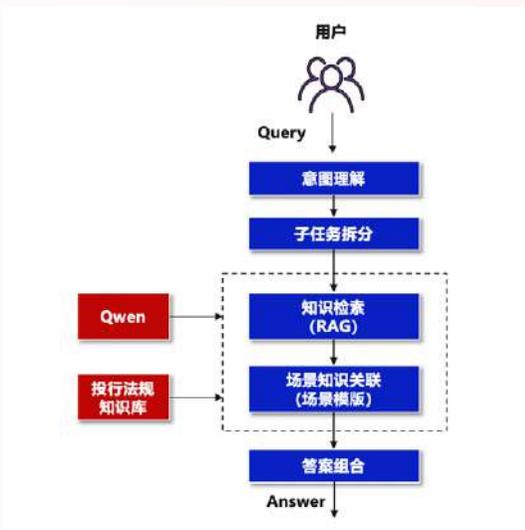
效果 语义识别准确率超过 90%。

Qwen-plus

处理 基于法规知识库，理解用户 query 会召回相关片段，进行理解、整理生成。

效果 法规知识问答准确率 85%。

在 GenAI 时代，大语言模型的出现，能对上述多源异构的半结构化数据做很好的处理，补齐原有大数据平台只能处理结构化数据的短板。LLM 应用的核心技术点包括：



1 多模态大语言模型

在数据预处理阶段，基于多模态大语言模型 Qwen-VL，对 PDF、图像类数据进行语义理解，增强原有的 OCR 只能识别字符但不理解语义的不足。

2 向量知识库

构建现行投行法规文件 2500 篇，历史全部近 8000 篇，对文本内容进行识别，针对性进行段落、版面、语义等方式切片，把切片后的 chunk 内容利用 embedding 转换成向量后存储到向量引擎当中。

3 大语言模型 RAG

接到投研业务人员的 query 问题后，进行向量化转换，通过语义检索从知识库中召回相关片段，送给大语言模型 Qwen 进行理解和答案生成，问答准确率超过 85%。

价值分析

使用大语言模型加持的投行法规解读 Agent，能解答的提升从业人员工作效率：

全面、精准的信息检索

全面获取特定法律法规相关知识，对投行业务涉及的现行全部规定进行了知识治理，以保障精细查询解读。包括法律、行政法规、证监会、证券业协会、三大交易所、股转公司等层面的全部规定。

热点信息及时感知

基于外部检索引擎等工具调用能力，大模型辅助解读新法规、发行上市审核动态、监管案例、过会排名及撤否率排名等投行法规动态。

专业性知识辅助

对于投行领域专有名词和业务知识，大模型利用其内置世界知识，以及强大的推理能力，能帮助从业人员快速了解新知识，辅助做出合理决策。

投顾资讯简报

Agent 概述

在金融市场日益复杂、客户需求愈发多样的当下，理财师的工作充满挑战，既要为客户量身打造精准的资产规划，又要高效管理众多客户账户，时刻把控风险、紧跟市场动态。基于大模型的理财师助手为此提供了创新解决方案。

投顾资讯简报 Agent，能快速整合海量金融数据，涵盖股市、债市、基金、外汇等多领域信息，精准剖析市场趋势。面对复杂的经济指标、政策变动，抽丝剥茧，提供通俗易懂的解读，辅助判断对各类投资品的影响。针对热点财经事件，即时评估投资机遇与风险，给出资产配置建议。还可深度分析客户持仓，结合市场动态，助力理财师为客户定制个性化策略，全方位提升金融服务的专业度与效率。

需求分析

金融资讯推送作为金融理财师客户陪伴服务的重要环节，存在如下痛点：

信息过载

金融市场资讯繁多，要筛选出对客户有价值的内容很困难，而且发送过多信息会让客户感到厌烦。

时效性难把握

金融市场变化快，消息发送不及时，客户可能错过投资机会或不能及时规避风险。

个性化不足

每个客户的风险承受能力、投资目标不同，很难做到咨询内容完全贴合每个客户的需求。

于此同时，如何利用好资讯推动这个触点进行进一步营销推荐，更深入的做好客户维护，也是金融理财师面临的问题。

典型案例

模力时刻

Qwen-plus

处理 基于市场实时资讯，以及客户持仓画像，生成异动提醒和个性化简报。

效果 在客服主动触达场景下，推送物料生成效率提升，从 1 小时到 5 分钟。

某大型银行在零售部门试点理财助手，为客户批量提供金融讯息分析，针对金融热点资讯，结合产品库、客户持仓、风险偏好、资产情况及当下热点，大模型生成符合客户兴趣点的内容，针对用户的个性化需求，自动筛选相关新闻进行推送。提升了理财师工作效率和客户服务的质量。

- 兴趣匹配环节通过大小模型结合的方式实现，其中小模型为主，大模型为辅（利用其泛化性突破小模型信息茧房的限制）；
- 摘要生成环节按照客户画像基于大模型生成千人多面的新闻摘要，完成兴趣匹配之后，智能体把新闻内容和客户标签作为提示按照一定的专家框架结构化生成提示词传递给大模型，由大模型生成摘要；
- 新闻解读环节在常规 RAG 的基础上打通客户信息，匹配产品支持，并参考专家经验进行生成增强；
- 最终把这几个模块的生成结果以卡片的形式推动到客户经理后台，校验后发送给管理的客群。

总体技术链路如下图所示

实现效果如下图所示：



价值分析

结合大模型和检索实时咨询的投顾咨询简报 Agent，提升理财经理获取信息的效率，每日自动生成内容，增加触达客户的营销时机，提升理财经理的管户规模，并最终带来业务目标的转化提升。具体体现在：

🕒 更及时的市场回顾

大模型按需调用外部实时资讯检索工具，获取市场、行情、热点等信息，基于大模型本身的语义理解和总结生成能力，产出对已发生信息的摘要，提升阅读效率。

🔍 更深度的事件解读

大模型 + 投研分析智能体，充分发挥人工专家的经验 and AI 的理解生成能力，对重要热点事件进行深度解读，产出可媲美专业分析师的分析内容。

👤 更加个性化的陪伴

结合客户持仓相关行业，以及当日行业资讯，自动整理相关的热门内容，形成一份简报推送给客户，提升触达时机，和投顾陪伴效果。



理财产品问答助手

Agent 概述

理财产品问答 Agent，整合企业内各类理财产品信息，利用大模型语义理解和意图识别能力，精准剖析用户的提问，检索出相关的产品说明、资金投向、收益规则、风控措施等知识，进行生成式回复。基于大模型的产品问答，不仅能实时答疑，还能依据内部策略调整同步更新，精准匹配客户需求，提升回复的准确性和实时性。

需求分析

金融产品问答是金融理财师面对客户沟通过程中最基础的工作，存在如下痛点：

查找知识耗费时间

与客户的问答交互属于一个开放式的聊天过程，客户既会问基金、股票等产品信息，也会问行情走势，或者一些买卖交易规则等。客户经理需要从不同渠道获取知识，耗费时间。

准确性要求高

金融机构对客户服务内容准确性及合规相关要求，即便是找到了相关信息，客户经理还需要进行甄别与整理，最终用自然语言把相关素材串联整理，形成可外发的内容。

客户经理管户效率瓶颈

问答会消耗客户经理大量的时间去搜集资料和整理答案，直接影响了平均每个客户经理可服务客户的数量。

典型案例

模力时刻

Qwen-plus

处理 基于用户 query 和从客服知识库中召回的相关信息，进行答案生成。

效果 答案生成准确度达到 90%。

金融机构在企业内搭建产品与服务问答知识库，完成全链路知识抽取、清洗、整理，语义化匹配客户问题及意图，快速检索答案进行响应，提升回复效率。支持客户多样性问题的理解，支持知识溯源，高效且准确的进行回复，提升了理财师的工作效率与客户满意度。

在整个客户服务的过程中，大模型在侧边栏以 Copilot 的形式，给客户经理推荐相关问题的答案和话术：

1 产品问答

产品问答包含了客户意图识别、FAQ 匹配、RAG 检索生成等核心环节。

2 意图识别

意图识别是基于通义千问的生成式意图识别代码，相比于传统基于小模型 + 规则的方式，在对话轮次、训练周期、精准度上均有提升；智能体基于意图识别的结果进行路由，对于已知问题走事先准备好的问答对进行匹配，未知问题则走 RAG 大模型生成的链路。

3 RAG 链路

RAG 链路在常规 RAG 的方案基础上，通过用户特征、产品特征结合专家经验进一步增强，在专业性上相比于常规 RAG 方案有显著提升。



价值分析

在大模型时代，构建理财知识库赋能银行理财客户经理具有显著的核心业务价值：

精准响应客户需求

利用实时语义理解和大模型技术，理财客户经理能够迅速而准确地识别客户的诉求，并基于后台海量文档快速生成答案或推荐话术。这意味着客户经理能够在第一时间提供最贴合客户需求的信息和服务，极大提升了服务的专业性和针对性。

提升客户经理服务效率

通过智能化的知识库支持，理财客户经理不再需要花费大量时间搜索信息或制定沟通策略。系统能够自动分析客户需求并推荐最优解决方案，从而让客户经理的工作更加高效，减少了准备时间和提高了服务质量。

增强客服服务体验

对于客户而言，即时获得专业且个性化的回复是提升满意度的关键。借助先进的 AI 技术和智能客服系统，银行可以确保每一位客户都能得到及时、准确的回答，这不仅加强了客户的信任感，也提升了整体客户体验。

营销优先级判断

Agent 概述

基于大模型意图识别能力，针对理财师与客户当前和历史对话内容，持仓情况等数据准确识别诉求，识别并判断营销的时点，对业务人员进行服务优先级提醒，推荐关联产品。

需求分析

金融机构对存量客户进行分层管理，有些客户经理的管户规模数巨大，1 比几百，甚至 1 比几千。这种情况导致没有充足的时间服务每个客户，甚至没有时间回答客户问题。

客户的问题也是五花八门，有的是问业务和产品相关，有的只是闲聊，或者单纯的情绪表达。客户经理很难分辨服务优先级，导致有限的时间被浪费。

典型案例

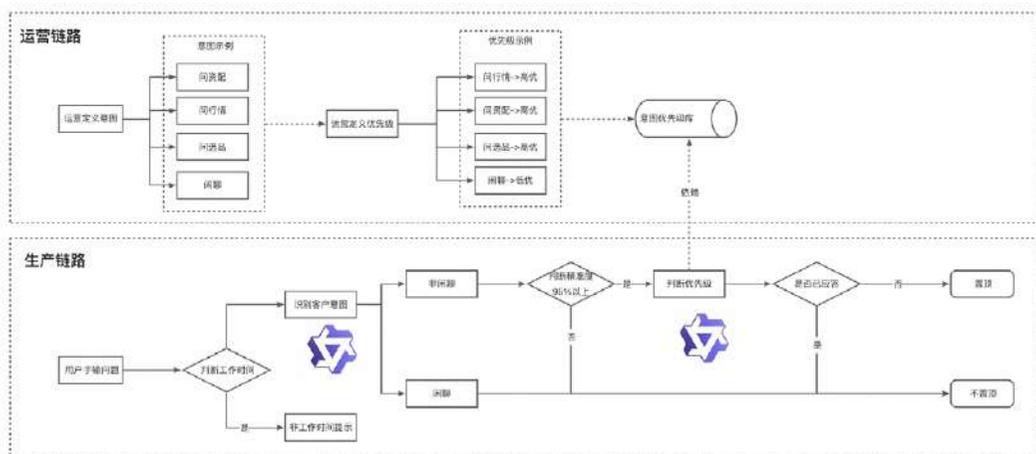
模力时刻

Qwen-plus

处理 基于客户对话历史，判断客户购买意愿和转化概率，在客服工作台上打标提醒。

效果 客服优先服务高潜客群，提升转化率。

在大模型时代，凭借 LLM 强大的自然语言理解和意图识别能力，可全面识别客户的对话历史，以及当前持仓情况，预测客户的购买意向和转化概率，进行打标，标记出的服务优先级，提示给客户经理。



判断一个客户的营销优先级，以上两条工程链路可结合使用：

1 直接识别营销时点

直接用大模型的语言力，读取客户会话历史记录，以及业务上对营销时点的判定落及，进行意图识别，和营销优先级的建议。

2 常规意图映射营销优先级

常规意图识别：通过用户 Query 识别用户问市场、问产品、闲聊等常规意图；通过配置化方式设置常规意图的沟通优先级。

识别出优先级后，智能体会根据运营设置的排序规则，将会话窗口进行排序和弹窗提醒。实现效果如下示意图：



价值分析

有了大模型加持，精准识别客户营销优先级，提升客户经理工作效率与业务转化：

提升终端客户体验

通过大模型的精准识别能力，银行能够迅速定位并响应真正有产品和服务需求的客户。这不仅确保了客户的询问得到及时且相关的答复，还让客户感受到个性化的关怀和服务质量的提升。这种即时性和针对性极大地增强了客户的满意度和忠诚度。

提升客户经理服务半径

借助大模型的强大分析功能，客户经理可以更准确地识别高潜力商机，并将其优先处理。此外，自动生成功能能够为每个商机提供定制化的话术支持，使客户经理能够在短时间内准备好与客户的沟通策略。这种方式不仅扩展了客户经理的服务范围，也显著提高了业务转化率和工作效率。

信披报告审核

Agent 概述

本 Agent 旨在服务公募基金、私募、专户理财、银行理财子和券商投行等机构客户，支持包括但不限于的单一文档信披报告审核、跨文档信披报告审核，跨文档风控指标勾稽校验、跨文档章节全部内容比对、跨文档基金合同条款审核比对等需求场景。

大模型智能信披审核系统通过先进的自然语言处理技术，自动从信披文件中提取关键指标。系统内置勾稽校验引擎，根据预设的逻辑关系对数据进行校验，确保其一致性和准确性。推理测试技术用于验证校验规则的有效性，提升审核的可靠性。最终，系统通过可视化界面展示审核结果，并支持溯源功能，帮助用户快速定位和修正问题。

需求分析

信披审核 Agent，支持常见 20 余种信披报告，包括跨文档审核（三方会计报告、招募说明书等）、财务勾稽关系审核、基础语义审核（错别字、敏感词）三个功能。

对于基金会会计、合规审核、风险控制业务人员来说，每年公司大量产品的公募基金年报、半年报、季报等信披材料都需要细致审核，但对于复杂语义、数字逻辑关系、跨期复杂数据勾稽关系校验等指标或者段落，之前的 IT 系统无法满足这些审核需求。因此，信披审核 Agent 能够充分发挥大模型在语义理解、意图识别方面的天然优势，完全能够支持相关个性化审核需求。

典型案例

模力时刻

Qwen-plus

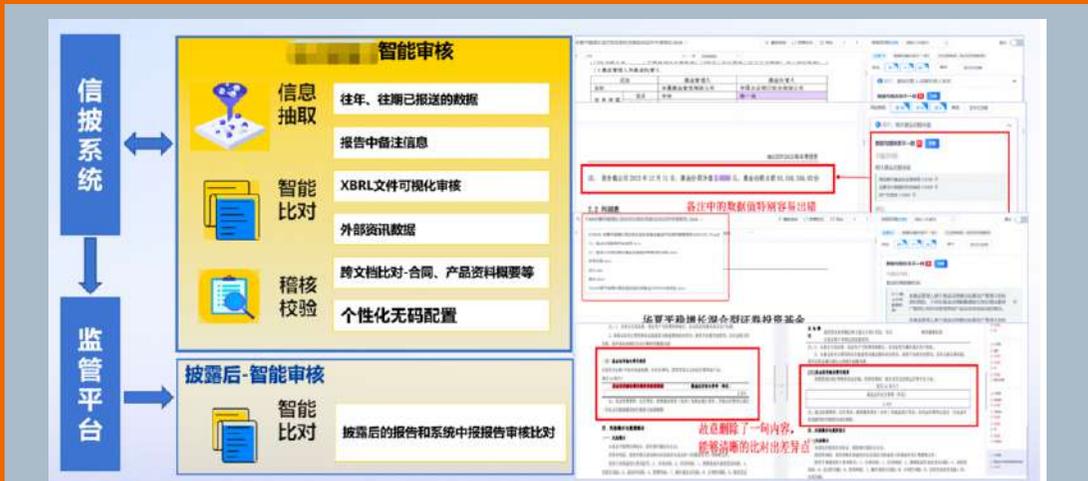
处理 基于监管合规要求，对年报、季报等信披材料内容进行内容审核，提升人工审核效率。

效果 审核准确率 99%，审核效率提升 300%。

在实际的业务流程中，大模型在信息抽取、智能比对和稽核校验等环节发挥了关键作用。例如，在券商的信披报告审核中，大模型能够快速完成年报、季报的审核，以及与产品合同的交叉审核比对。在基金的营销物料审核中，大模型能够识别宣传文档中的错别字、标点异常等精细化问题。通过这些环节的应用，大模型显著提升了审核的效率和准确性，同时也降低了运营成本。

这个方案通过结合大模型的先进技术，实现了信披审核流程的自动化和智能化，为金融机构提供了一个高效、准确、低成本的审核解决方案。信披系统：作为信息的来源，提供数据给智能审核系统；智能审核系统：核心处理单元，负责执行信息抽取、比对和稽核校验；监管平台：接收智能审核系统的结果，进行监管和记录。

某头部基金客户 1 个月内，审核了 527 篇定期报告，督察长和基金会会计的工作效率提升了约 300% 以上（4-5 个人，审核 100 份年报从 1 个月降低 5 天），审核准确率达到 99% 以上。



大模型的作用和价值:

1 信息抽取

利用大模型的高准确率(超过95%)进行信息抽取,包括往年、往期数据、报告中备注信息等。

2 稽核校验

在信息披露时,大模型能够进行智能审核和比对,发现并提示潜在的错误和风险点。

3 智能比对

通过大模型进行XBRL文件的可视化审核和外部资讯数据的比对,确保数据的一致性和准确性。

价值分析

大模型 Agent 在业务和技术两个层面带来了显著的价值。

⌚ 大幅提升审核效率与业务流转速度

大模型 Agent 通过自动化和智能化的手段,快速完成合同、定期报告等内容的审核工作,不仅缩短了单文档和跨文档审核所需的时间,也加速了整体业务流程的进展。

🔍 提高审核的一致性与准确性,增强合规性

利用标准化流程进行内容审核,大模型 Agent 有效避免了人为因素导致的漏审和误审问题,确保每次审核的质量保持在高水平,从而提高了信息披露的整体合规标准。

🔒 实现数据智能比对与稽核校验,保障信息质量

该技术能够自动提取并对比大量文本中的关键信息,如历史数据和备注信息等,并在信息披露后进行智能审核,确保所有公布的信息既准确又符合规定要求,提升了决策质量和用户服务体验。

营销物料审核

Agent 概述

营销物料审核 Agent 专为金融机构的市场部、品牌管理部、合规部以及产品开发部设计，以支持营销物料的全面审核流程。利用大模型技术，营销物料审核 Agent 能够深入理解物料内容，自动执行多项核查任务，确保营销资料的准确性、合规性以及品牌信息的一致性。Agent 的核查引擎能够识别并提示潜在的语义错误、数据引用问题、标点异常等，同时通过智能算法优化表格和格式的一致性，提升整体物料质量。

此外，营销物料审核 Agent 还具备高度的扩展性，能够适应不同金融机构的特定需求，覆盖市场推广、品牌宣传、产品介绍等多种营销场景。通过可视化的操作界面，用户可以轻松管理和监控审核进度，同时 Agent 的溯源功能使用户能够快速定位问题源头，进行有效修正。

需求分析

对于市场部、品牌管理部、合规部以及产品开发部的业务人员来说，每年需要推出大量的营销活动和宣传资料，这些资料的审核工作量巨大，且要求高度的准确性和合规性。传统的 IT 系统在处理复杂语义、品牌信息一致性、格式规范性以及跨文档数据一致性等审核需求时存在局限。因此，营销物料审核 Agent 的引入，能够利用大模型在自然语言处理、模式识别和逻辑推理方面的天然优势，有效地满足这些个性化且复杂的审核需求。

典型案例

模力时刻

Qwen-plus

处理 依据资本市场对营销物料的要求，对营销、投教等内容进行审核，提升人工审核效率。

效果 审核准确率 98%，审核效率提升 400%。

在实际的业务流程中，营销物料审核 Agent 通过大模型的深度学习能力，实现了对营销物料的全面审核。例如，在金融机构的市场部，Agent 能够快速审核即将发布的营销资料，识别并提示可能的合规问题，如敏感词使用、不规范的表述等。通过可视化的用户界面，市场部人员可以轻松地管理和监控审核进度，确保每一份物料在发布前都经过严格的合规性检查。此外，Agent 的自动化流程也大大减轻了审核人员的工作负担，使得他们能够将更多的精力投入到创意和策略的制定上。通过这种智能化的审核方式，金融机构能够确保营销物料的合规性，同时提升宣传营销物料质量。

大语言模型作为核心技术，作用和价值体现在：

1 内容审核

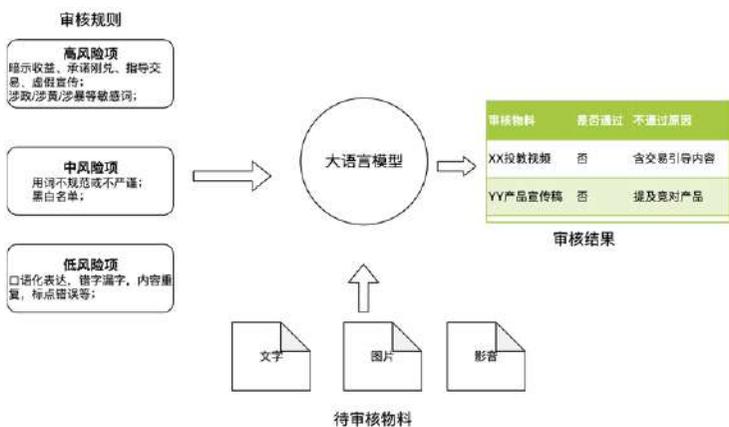
大模型理解营销物料中的文本内容，自动检测错别字、敏感词、语义错误等，确保物料内容合规。

2 格式和结构审核

分析文档的结构，大模型可以检查表格格式、空白页、内容重复等问题，保证物料的专业性。

3 风险评估

大模型根据预设的风险等级，对物料内容进行风险评估，识别高风险项如涉政、涉黄、涉暴、违反公司内外规等敏感内容。



某金融机构通过引入营销物料审核 Agent，在一个月处理了超过 600 份营销资料，市场部和合规部的工作效率提升了近 400%（原本需要 10 人团队审核的资料，现在仅需 2-3 人即可完成），审核准确率保持在 98% 以上，显著提升了营销物料的发布效率和合规性。

价值分析

营销物料审核 Agent 带来的业务价值，主要体现在提升审核效率、降低合规风险、优化物料质量上：

提升审核效率

营销物料审核 Agent 通过自动化流程，快速处理并审核营销内容，显著缩短了从制作到发布的周期。这使得营销团队能够更迅速地响应市场变化，及时抓住市场机会，极大提升了工作效率。

降低合规风险

凭借深度学习和创新的风险评估算法，该 Agent 能精准识别敏感词和不恰当表述等内容，确保所有营销物料符合行业标准 and 法规要求，有效降低了违规发布的风险，保障了信息的合规性和安全性。

优化物料质量

通过多模态处理能力和对复杂文本内容的精准分析，Agent 不仅确保了语言和内容的准确性，也增强了物料的专业性和吸引力。此外，它还促进了品牌标准的一致性维护，从而全面提升了营销物料的质量和效果。

保险

产品开发及销售

保险产品解读、条款解析助手、保险营销创作、保险产品搜索、保险产品推荐、保险产品问答、保险产品比对

核保核赔

智能预核保、核赔辅助、智能影像处理

监管合规

条款智能核验助手、条款智转助手、对外披露审核

保险产品解读

Agent 概述

在保险行业中，保险产品条款复杂，专业术语繁多，理解难度大。这不仅给用户带来了极大的困扰，也使得保险销售过程中的沟通成本大幅增加，同时还容易引发潜在的销售误导风险，严重制约了保险业务的健康发展。

保险产品解读 Agent 主要服务于保险销售人员、用户及相关利益方。通过对保险条款的深度剖析，能够精准提取关键信息，并将其转化为通俗易懂的语言，输出解读内容，有效解决了保险条款晦涩难懂、沟通成本高以及销售误导风险等问题。这不仅极大地提升了用户对保险产品的理解程度和满意度，还提高了销售效率，降低了沟通成本和销售误导风险。

需求分析

当前，保险产品解读存在诸多问题，严重影响了保险销售与服务的质量。

用户层面

保险产品条款冗长，专业术语密集，导致用户难以快速、准确地掌握保障范围、赔付条件、费用明细等关键信息。这不仅影响了用户的购买决策，还可能在后续理赔过程中引发纠纷。

销售人员层面

在向用户介绍产品时，销售人员需要花费大量时间和精力解释条款，且由于个人理解的差异，解释口径难以统一，进而增加了销售难度和合规风险。

因此，市场迫切需要一款能够精准解读保险产品的工具，将复杂条款转化为通俗易懂的内容，满足用户快速理解产品的需求。

典型案例

模力时刻

Qwen-long

处理 从长篇幅保险条款中提取保险产品要素信息。

效果 平均人工抽取成本从 40 元 / 条降低到 2 元 / 条，成本降低 20 倍。处理时间从 4 小时，降低到 10 分钟以内。

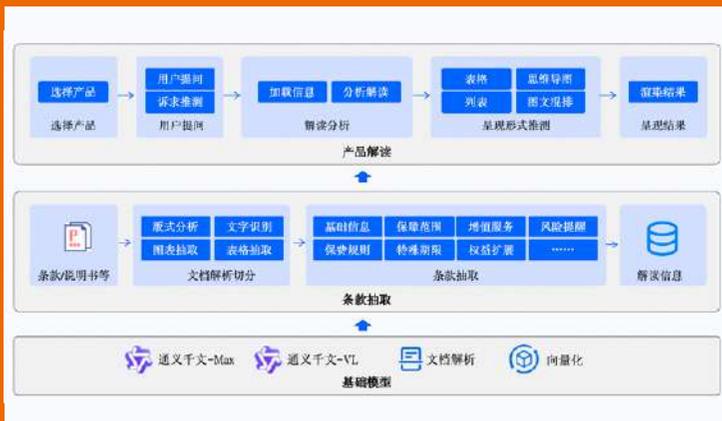
Qwen-plus/max/72B

处理 生成针对用户的个性化图文、表格、思维导图混排的产品解读。

效果

- 体验提升：从生成固定版式通用的产品解读说明书到面向单个用户的个性化。
- 制作成本：4 人天 / 产品，改进后 1 分钟 / 产品生成针对单个用户的个性化产品解读。

保险产品解读 Agent 可嵌入销售系统的用户沟通模块。用户选择产品后提出解读要求，Agent 会迅速对条款进行分析，并以简洁明了、图文并茂的方式向用户呈现产品亮点、保障细节、理赔流程等关键内容。该方案主要由以下模块组成：



1 基础模型

包括通义语言大模型、通义视觉大模型、文档解析模型以及向量化模型等。这些模型构成了整个方案的技术基础，为后续的知识处理和应用提供了强大的支持。

2 条款抽取

运用版式分析技术，精准识别文档的页面布局和结构信息；通过文档解析技术，提取文本内容与关键数据；借助 OCR 识别技术，将图片中的文字转化为可编辑文本；利用图表识别技术，解读图表中的数据信息；最终通过大模型从文档中提取解读产品所需的各种字段，并存储到数据库。

3 产品解读

根据用户选择的产品和提问，推测用户的诉求，利用大模型进行解读，并选择合适的呈现形式展示最终结果。

价值分析

☑ 用户体验提升

通过大模型将晦涩的保险条款转化为通俗易懂的内容，帮助用户快速理解保险产品，缩短理解时间，提升对保险产品的理解程度，减少因条款误解引发的潜在投诉，增强用户购买意愿，降低因沟通不畅造成的用户流失。

📈 提高销售效率

标准化的知识加工和简洁的呈现方式，使销售人员能够快速获取知识解答客户疑问，降低销售过程中的沟通成本和时间成本，提升销售效率。

条款解析助手

Agent 概述

在多元化竞争的保险行业，产品设计与管理能力是保险公司的核心竞争力。为提升产品管理能力，新一代核心系统需要实现保险条款和产品的精细化配置，其涵盖多种险种，字段数量多达数百个。这要求系统配置更规范，数据颗粒度细化到责任层级，从而实现上下游系统数据全链路自动化。这种精细化的数据处理对产品分析、精算定价、理赔分析和风险控制意义重大，是评估和提升产品盈利能力的基础。但在实际执行精细化条款解读时，面临耗时长、专业要求高、人力投入大等问题。

条款解析 Agent 凭借大模型强大的解析能力，能高效解析保险条款，并按核心系统的产品数据结构输出。这将业务专家的经验沉淀为固定的解析方案，显著提升条款解读效率，规范产品数据，有效解决保险行业在条款解读和数据处理方面的困境。

需求分析

当前条款解析存在如下问题：

时间成本高

在实际执行精细化条款解读时，整个流程耗时长，这可能会导致业务推进速度缓慢，错过市场先机，也会增加运营成本。

专业要求高

保险条款涉及补充险种设置，大量专业知识，对解读人员的专业素养要求极高。普通员工难以胜任，而专业人才相对稀缺，这在一定程度上限制了业务的开展。

人力投入大

为了完成精细化条款解读和数据处理工作，需要投入大量的人力，不仅增加了人力成本，还可能面临人员调配和管理的难题。

经验缺乏有效沉淀

原有的解读规范分散在各个专家头脑中，未形成统一、有效的沉淀机制，不利于知识传承和团队协作，也无法充分发挥集体智慧提升条款解读的效率和质量。

典型案例

模力时刻

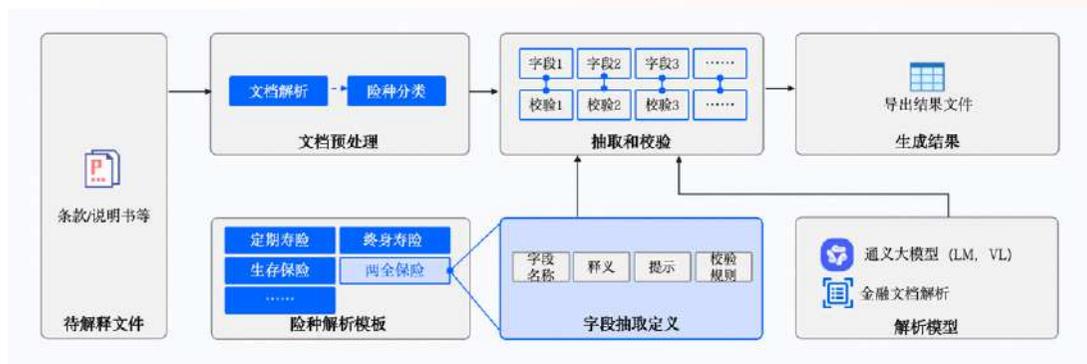
Qwen-long

处理 从长篇幅保险条款中提取保险产品要素信息。

效果 从人工抽取平均成本 40 元/条降低到 2 元/条，成本降低 20 倍。处理时间从 4 小时/条，降低到 10 分钟/条。

某财险公司借助通义千问大模型及百炼平台，成功打造了保险条款智能解析助手。该助手能够实现文本化的条款条文与费率规章的智能解析，并将其转化为符合核心系统产品数据结构要求的结构化数据，随后自动填入系统配置页面，取代了原有的人工解读和人工配置操作模式。

总体技术链路如下图所示：



1 险种解析模板

该模块配置了不同险种的的字段抽取定义，一般包括待抽取字段名称、释义、抽取提示、校验规则等。

2 文档预处理模块

主要对原始文档（如 PDF 或 Word 文档）进行文本化解析，使其转化为大模型可处理的格式，然后通过大模型对文档的险种进行分类。

3 抽取和校验模块

该模块根据文档预处理模块的险种分类结果，从险种解析模板中加载对应的字段抽取定义，分别调用通义大模型进行抽取和校验。

4 生成导出结果文件

生成可供人工进一步质检和使用的结果文件。

价值分析

效率飞跃式提升

原有每份条款文件解析耗时数天，大模型方案可缩减至数分钟，极大地加快了业务推进速度，让保险公司能及时响应市场变化，抓住市场先机。

成本明显降低

由于大模型可自动完成条款解析和数据处理工作，减少了对大量人力的需求。无需再投入众多专业人员进行人工解读，同时也避免了人员调配和管理的复杂问题。

知识沉淀与传承

将业务专家的经验沉淀为固定的解析方案，形成了统一的条款解读规范，有利于知识的传承和团队协作。

保险营销创作

Agent 概述

在保险行业的销售过程中，因保险产品的多样性和复杂性，这使得准确传达产品价值变得尤为困难。随着消费者需求的不断细分与个性化趋势的加强，销售人员不仅需要精通保险专业知识，还需具备敏锐的市场洞察力与细腻的心理分析能力，以便精准捕捉每位客户的独特需求，并据此创作出贴切的营销文案。如何既能高效又精确地执行这一任务，无疑成为了一项极具挑战性的工作。

保险营销创作 Agent 是专为保险行业营销环节打造的智能解决方案。通过大语言模型，它能够精准理解保险产品特点与用户需求，自动生成个性化营销文案、宣传海报等营销素材。该 Agent 大幅提升了营销内容创作的效率，降低了人力成本，同时显著提高了营销的精准度和成功率，为保险企业创造了更大的市场价值。

需求分析

在保险营销场景下，业务痛点主要集中在以下几个方面：

<p>1</p> <p>保险产品条款复杂，营销人员难以快速提炼核心卖点进行宣传。</p>	<p>2</p> <p>用户群体庞大且需求各异，传统营销方式难以满足个性化需求。</p>	<p>3</p> <p>营销内容创作依赖专业人员，创作周期长、成本高。</p>
---	---	--

例如，在推广一款新型健康险产品时，营销人员需要花费大量时间研究产品条款，了解保障范围、赔付条件等细节，再结合不同用户群体的健康状况、经济实力等因素创作针对性的营销文案，整个过程繁琐且低效。

典型案例

模力时刻

<p style="text-align: center;">Qwen-plus/72B</p> <p>处理 创作线索的挖掘，产品和用户推荐特征标签的提取、文本内容创作。</p> <p>效果 语义识别准确率超过 90%。</p>	<p style="text-align: center;">Wanx2.1</p> <p>处理 图片创作。</p> <p>效果 单张图平均设计师画图效率从 1 天缩短到 1 小时。</p>
---	--

为克服前述痛点，引入大语言模型的辅助内容创作，整体方案实现如下：

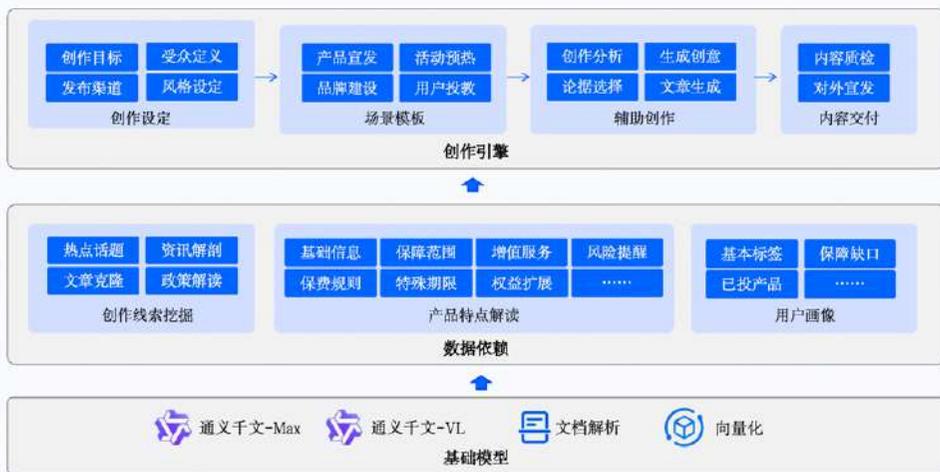
1 原始数据挖掘提供创作依赖的数据

- **创作线索挖掘**：大模型广泛抓取保险行业新闻资讯、社交媒体讨论、行业报告等海量数据，并对数据进行分析，提取热门话题、新兴趋势以及用户关注焦点等创作线索。同时，对于优质内容，解析内容创作策略架构，从中发现并学习差异化的创作角度，为保险营销内容创作提供方向指引、类质内容产出。
- **产品特点挖掘**：大模型对保险知识库文档进行深入拆解提取，识别产品文档中的关键信息，借助知识图谱，将信息进行结构化组织，清晰呈现产品特点、核心优势及标签，确保在营销内容创作中能够精准传达产品价值。

- 用户画像挖掘：整合多源数据构建用户画像，包含不限于基本信息、历史投保记录、保障缺口、操作行为等数据对用户进行分群，预测用户需求和购买倾向，为每个用户画像定制个性化的保险营销内容，提高营销的精准度和有效性。

2 创作引擎搭建满足创作需求

- 创作目标设定：定义内容的核心作用（如销售转化、品牌传播或用户教育）；同步锁定目标受众，基于用户画像及需求特征确定内容方向。根据发布渠道特性（微信朋友圈 / 小红书 / 短视频平台等）匹配内容形式与格式规范，确保适配不同终端的展示规则。最后设定风格基调，结合产品属性选择专业理性或情感共鸣的表达方式，并定义视觉元素（配色、版式、图文比例）以强化信息传递效率。
- 场景模板选择：基于目标类型调用预设框架：产品推广类通常采用“痛点触发 - 解决方案 - 行动激励”结构；活动类侧重“时限压力 - 利益点 - 参与路径”；品牌类依赖“数据 / 案例背书 - 价值主张”；用户教育类遵循“认知误区 - 知识解析 - 行为引导”。模板提供结构化内容骨架，确保核心信息不偏离业务诉求。
- 内容生成：整合目标、受众、渠道、模板四要素，通过分析历史数据与用户行为偏好，生成多组创意方向（如情感叙事、数据对比、问答互动），并自动嵌入合规依据（如条款引用、风险提示）。输出内容需包含完整逻辑链（吸引注意 - 激发需求 - 证明价值 - 促成行动），同时按渠道规则调整细节（如字数限制、跳转链接位置）。
- 交付前执行合规质检：过滤违规表述（如绝对化承诺、误导性收益），并自动替换为合规话术。针对多终端发布需求，将主内容拆解为适配不同平台的衍生版本（如长图文精简为弹幕话术、口播脚本拆解为分镜头）。利用 A/B 测试对比关键变量（标题、视觉元素），根据实时数据反馈优选方案。



价值分析

提升营销效率

通过大模型深度学习和自然语言处理，快速理解保险产品特点和目标受众需求，可持续输出大量差异化文案，自动适配社交媒体平台、邮件、短信等多个渠道，将传统人工数小时的工作缩短至分钟级。

提升内容质量

通过大模型内置知识库，快速掌握专业知识和不同文案风格，可输出风格统一、专业规范的文案内容，降低人工失误风险，并通过大模型训练保证语言表达的专业性。

提升营销效果

利用大模型用户画像和场景识别能力，针对不同用户群体生成个性化文案，提升内容相关性和用户互动性，促进对应营销目标达成。

保险产品搜索

Agent 概述

在保险行业中，传统保险产品搜索页面提供了诸如保险类型、保险公司、保障范围、保费范围及投保人年龄等多种查询条件供用户选择。用户需根据自身需求，选择或输入具体的查询条件，随后后台会根据这些条件组合成相应的查询语句，在业务数据库中检索，并返回符合条件的保险产品列表。然而，传统的搜索方法对用户需求的理解较为机械，缺乏灵活性，难以适应复杂多变的用户需求和自然语言表述。用户不仅需要具备一定的保险专业知识，还需经历繁琐的查询过程。

保险产品搜索 Agent 则采用了通义千问大模型，凭借其强大的语义分析能力，改变了传统保险产品搜索的交互模式，它能够深度理解用户意图，即便面对自然语言表述的复杂需求，也能迅速返回精准匹配的保险产品，极大地提高了搜索效率和用户体验。

需求分析

传统保险产品搜索的特点及局限性：

用户使用门槛高

用户需要具备一定的保险知识，了解系统提供的每个查询条件的含义，并知道如何合理组合这些条件以获得准确的搜索结果。例如一个上班族想要购买一份包含重疾和意外保障的保险产品。在传统搜索模式下，他需要了解并选择“保险类型”为重疾险和意外险，同时可能还需考虑“保障范围”、“保费”等多个条件。对于不熟悉保险的用户来说，这些专业术语和复杂的查询条件可能会让他们感到困惑和难以操作。

传统搜索算法局限性大

由于传统搜索算法无法从语义层面进行分析和理解，导致匹配和搜索效果不佳。例如，系统记录中有一款保险产品的描述中包含“癌症保障”，如果用户输入“恶性肿瘤保险”进行搜索，传统系统也可能无法准确匹配到这款保险产品。这种情况凸显了传统保险产品搜索在语义理解方面的局限性。

典型案例

模力时刻

Qwen-plus/72B

处理 搜索意图推理、用户 query 结构化改写、术语行业化，候选产品列表精选。

效果 用户体验效果提升，用户可基于口语化表达搜索产品，一次性搜索准确率提升。

本方案主要由以下几个核心模块构成：

1 搜索意图推理

此模块负责深度解析用户的原始搜索查询 (query)，将其意图细分为结构化需求与个性化需求两大类。

- 结构化需求处理：针对结构化需求，模块利用预设的筛选标签体系高效地从产品库中检索匹配的产品。
- 个性化需求处理：对于个性化需求，首先通过大型语言模型 (LLM) 对查询进行行业术语的精准转换，同时结合行业专家预先制定的召回模板，进一步提升召回的相关性。最终，采用 BM25 (基于词频的相似度算法) + Vector (向量相似度匹配) 的多路召回策略，综合生成产品候选列表。

2 产品库模块

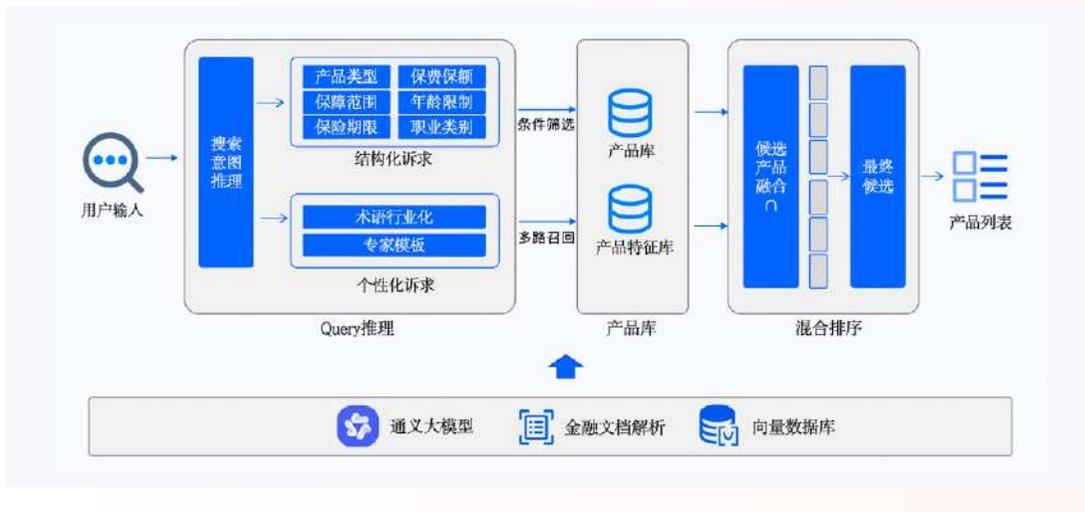
存储所有产品的基本信息用户产品的结构化和非结构化搜索。

- 筛选标签字段：设立丰富的筛选标签，便于用户根据特定条件快速筛选产品。
- 产品特征库：专门存储产品的语义化向量信息，这些信息通过先进的自然语言处理技术提取，用于支持更高级别的语义匹配和推荐。

3 混合排序模块

该模块旨在整合并优化从条件筛选和多路召回中获得的产品列表。

- 候选融合：首先，对通过条件筛选和多路召回得到的产品列表进行交集运算，确保最终候选产品同时满足用户的结构化与个性化需求。
- 最终候选：随后，利用大型语言模型对候选融合后的产品深度理解和排序，依据产品的相关性、用户偏好、历史行为等因素，生成最终的、高度个性化的产品候选列表。



价值分析

显著提升用户体验

通过采用通义千文大模型，本 Agent 能够深度理解用户的自然语言查询，无需用户具备复杂的保险知识或进行繁琐的条件选择。这种交互方式更加贴近用户的日常思维习惯，极大地简化了搜索流程，从而提升了用户的满意度和忠诚度。

增强搜索灵活性和准确性

传统的保险产品搜索系统往往受限于固定的查询条件，难以适应复杂多变的用户需求。而本 Agent 通过语义分析技术，能够准确捕捉用户的个性化需求，即使面对自然语言描述的模糊或多样化表述，也能返回精准的匹配结果。这大大增强了搜索的灵活性和准确性。

保险产品推荐

Agent 概述

在保险线上销售中，传统的产品推荐依赖既定模型和专家经验，通过用户与产品特征协同信号建模。这种方式虽能捕捉用户偏好且资源开销小，但局限明显。它主要依据历史数据和用户行为推荐，缺乏对语义和意图的深度分析，难以满足复杂多变的用户需求。例如，年轻家庭在选择人寿保险时，会综合考虑家庭成员的年龄、职业风险、经济状况及健康状况等多重因素，而这些因素在传统的推荐模型中可能无法得到充分的体现。

保险产品推荐 Agent 引入基于大语言模型的推荐系统，它能引入外部知识，更全面提取特征信息，丰富语义信号，还具备逻辑推理能力，可深入理解用户动机和偏好，提供精准、全面的产品及背景信息。该系统不仅提升推荐精准度，还特别适用于冷启动场景，有效优化了推荐系统性能和用户体验。

需求分析

保险产品的购买低频且复杂，用户需求和偏好差异大，导致保险产品推荐系统比传统电商领域更具挑战性。传统推荐系统存在诸多痛点：

可解释性差

传统推荐系统基于复杂算法模型，很难向用户解释推荐结果的依据，用户难以理解推荐逻辑，这使得用户对推荐结果的信任度和接受度较低。

冷启动和长尾问题突出

面对新用户或新产品，传统推荐系统因过度依赖历史数据，在缺乏数据时难以给出有效推荐，这限制了新用户的拓展和长尾产品的推广。

个性化不足

由于缺乏强大的自然语言理解和生成能力，传统推荐系统无法深入理解用户语境和需求，难以精准捕捉用户上下文信息，提供的推荐结果不够多样化，无法满足用户个性化需求，导致用户参与度和满意度不高，产品推荐转化率也受到影响。

典型案例

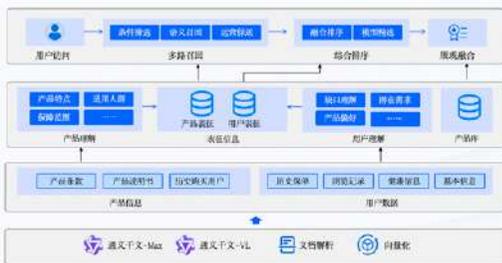
模力时刻

Qwen-plus/72B

处理 产品理解和表征抽取、用户理解和表征抽取、候选产品列表精选，形成推荐理由。

效果 解决了推荐系统冷启动问题，提升了推荐结果的可解释性，借助社会化知识提升了推荐转化率。

为了克服这些挑战，我们引入了基于大语言模型的推荐系统，该方案的整体实现如右图所示：



1 原始数据

- 产品信息：涵盖产品的详细条款、说明书以及历史购买用户的信息。这些信息是了解产品特性和构建产品画像的重要基础。
- 用户数据：包括用户的历史保单、浏览行为、基本信息（如年龄、性别等）以及健康信息。这些数据构成了用户画像的关键依据。

3 产品推荐

- 推荐请求发起：当用户访问保险 APP 或网页时，触发推荐系统请求。
- 多路召回：根据用户表征，通过条件筛选、语义匹配及运营推荐等多种手段，从产品库中精准召回符合用户需求的产品信息。
- 融合排序：对召回的产品进行排序，筛选出 Top N 候选产品，并借助大模型进行多维度推理与精选，最终确定推荐结果。

2 特征提取

- 产品理解及表征：通过对历史购买用户的画像、产品条款和说明书等材料进行深入分析，提取出产品的核心特点、保障范围等关键信息。利用模型推理，进一步推断出产品可能适用的人群。这些特征直接映射产品的功能、优势及目标用户群。
- 用户理解及表征：通过分析用户的历史保单、浏览行为、基础信息及健康数据，挖掘出用户的潜在需求、购买动机及产品偏好。这些特征信息被转化为用户的特定标签和潜在诉求，从而构建出用户表征库。

4 展现融合

在产品界面上，展示由模型精选的产品列表，供用户参考和选择。这些产品列表包含了产品的基本信息，还可以附带了清晰的推荐理由和解释，帮助用户更好地理解推荐结果，并做出明智的选择。

价值分析

📊 经济效益提升

精准的推荐提高了产品推荐转化率，更多用户选择合适的保险产品，直接增加了销售额。同时，减少无效推荐节省了运营成本，使得资源能够更高效地投入到其他关键业务环节，提升整体运营效率，从长期来看，为企业带来持续的利润增长。

👤 用户体验优化

可解释性增强让用户清楚了解推荐依据，增强信任度。个性化推荐满足用户多样化需求，冷启动和长尾问题的解决确保新用户和小众产品也能得到关注，全面提升用户在保险产品选择过程中的满意度和便捷性，有助于培养用户忠诚度，形成良好的口碑传播。

保险产品问答

Agent 概述

保险产品具有独特的复杂性，与股票、基金等参数相对结构化的金融产品不同。保险产品往往险种多，其详细信息承载于各种保险条款材料中，这些条款不仅内容专业性强、个性化高，内容晦涩难懂，普通用户极难理解。即便是保险销售人员，想要全面、准确地解答客户的多样化问题，给出精准且通俗易懂的解答，也必须提前深入研究每款产品的特性，这对销售人员的专业水平和应变能力提出了极高要求，同时销售人员需要全天候服务非常多的客户，服务效率和质量难以保证。

保险产品问答助手基于通义大模型的自然语言处理、语义理解与生成能力，可以将晦涩难懂的保险条款转化为清晰易懂的产品知识，从而构建一个全面、准确、易用的保险产品知识库。

需求分析

当前关于保险产品的问答存在如下痛点：

材料繁杂整合麻烦

保险行业的文档资料种类繁多，形式多样，包括但不限于产品条款、说明书、运营规则、监管要求、投保规则、渠道业务规则、公司品牌资料及费率表等。这些资料不仅需要依据业务逻辑进行有序分类与关联，如按保险产品编号对产品条款、说明书与投保规则进行归类，按渠道划分等。手动整理这些资料并建立有效联系，以及从不同格式文档中抽取并整合知识构建知识库，均需投入大量的人力和时间成本。

知识专业、学习成本高

保险销售人员需要掌握海量产品资料，但每款产品的知识体系独立且复杂，学习成本极高。因此，销售人员迫切需要一种高效、便捷的知识获取方式，以从而提高销售服务水平，更好地满足客户需求。

服务效率低与吞吐量有限

在保险业务中，客户数量众多，咨询问题的时间分布广泛，保险销售人员难以同时应对大量客户的全天候提问。每个销售人员需服务众多客户，导致工作压力大，服务效率难以保证。

典型案例

模力时刻

Qwen-plus/72B

处理 query 理解、query 改写、关联提问生成、答案生成。

效果 问答准确率 85%+。

整体方案包括以下几个部分：



1 基础模型

包含通义语言大模型、通义视觉大模型、文档解析模型以及向量化模型等。这些模型构成了整个方案的技术基石，为后续的知识处理和应用提供了强大的支持。

3 知识问答

通过 query 理解优化技术，深入剖析用户问题的语义、意图及关键信息；运用多路检索召回策略，从知识库中快速获取相关知识片段；采用重排序算法，对召回的知识按重要性进行排序；将筛选后的知识融入提示中，为大模型提供更丰富的上下文信息；最后借助大模型强大的生成能力，生成针对用户问题准确、清晰且易懂的回答。

2 知识加工

这一环节主要针对业务文档开展管理、解析切分与知识提炼工作，最终构建全文索引和向量索引，以满足知识问答搜索的需求。具体步骤如下：

1) 文档管理

运用大模型技术，结合保险业务领域的专业知识，对原始提交的产品条款、说明书、费率表、运营规则等文档进行细致分类与深度整合。依据不同类型知识的内在逻辑关系，构建具有鲜明保险行业特色的原始知识资料库。

2) 文档解析切分

利用版式分析技术，精准识别文档的页面布局和结构信息；通过文档解析技术，提取文本内容与关键数据；借助 OCR 识别技术，将图片中的文字转化为可编辑文本；运用图表示别技术，解读图表中的数据信息；最后通过文档分割技术，将长文档拆分为合理的片段。

3) 知识提炼

采用大模型知识增强技术，引入外部知识库和领域专家知识，对原始知识进行补充与完善。运用向量化技术，将文本转化为向量形式，构建具备高效向量检索、BM25 检索等多种能力的知识库。

价值分析

显著提升效率

通过自动化的文档管理与知识加工流程，极大缩短了从原始资料到可用知识库的转化时间。销售人员无需手动整理繁杂资料，可快速从构建好的知识库中获取所需信息，解答客户疑问，大幅提升工作效率，使保险业务咨询响应速度显著加快。

降低成本

减少了人力在资料整理和知识构建上的投入，降低了人力成本和时间成本。同时，高效的知识获取方式降低了销售人员的学习成本，缩短培训周期，使新员工能更快上手工作。

增强客户体验

客户能够得到更快速、准确、易懂的回答，解决保险产品咨询的难题，提升客户对保险服务的满意度和信任度，有助于提高客户转化率和忠诚度，促进保险业务的增长。

保险产品比对

Agent 概述

在保险销售及用户服务场景中，用户常常面临众多保险产品选择难题，难以清晰了解不同产品间的保障范围、赔付条件、保费差异等关键信息，从而导致决策困难，影响购买体验和销售效率。

保险产品对比 Agent 旨在解决这一痛点，其基于先进的大模型技术，能够快速、准确地对各类保险产品进行多维度对比分析。该 Agent 可以根据用户输入的产品名称、需求偏好等信息，实时生成清晰易懂的对比结果，以表格、图表等形式直观呈现，帮助用户快速把握产品差异，做出更明智的决策，促进保险业务的销售转化和用户满意度提升。

需求分析

在保险销售前端，用户在咨询环节常常耗费大量时间用于产品对比，用户因缺乏专业知识和便捷工具，难以自行梳理各产品的复杂条款和细节差异，容易在选择过程中产生困惑和焦虑，进而可能导致购买意愿下降。

对于保险销售人员而言，手动对比产品不仅效率低下，而且难以保证信息的准确性和完整性，难以针对用户个性化需求快速提供精准的产品推荐和对比方案。

此外，在保险市场竞争日益激烈的环境下，保险公司需要更高效地向用户展示自身产品的优势和特色，以吸引更多用户，提升市场份额，而传统的产品介绍方式已难以满足这一需求，急需一种智能化、自动化的工具来优化产品对比流程，提升用户体验和销售竞争力。

典型案例

模力时刻

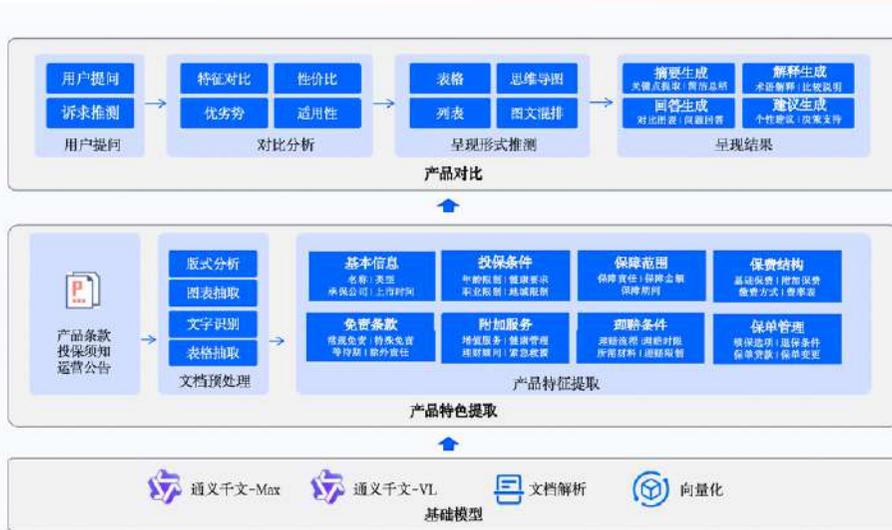
Qwen-long	Qwen-plus/max/72B
<p>处理 保险产品要素信息。</p> <p>效果 提升了用户的产品体验、平均人工抽取成本从 40 元 / 条降低到 2 元 / 条，成本降低 20 倍。处理时间从 4 小时，降低到 10 分钟以内。</p>	<p>处理 生成个性化产品对比话术、使用用户能理解的语言解释和对比保险产品。</p> <p>效果 提升了用户的产品体验、平均人工抽取成本从 40 元 / 条降低到 2 元 / 条，成本降低 20 倍。处理时间从 4 小时，降低到 10 分钟以内。</p>

在某大型保险公司的销售团队中，以往用户咨询产品对比时，销售人员需要查阅大量纸质资料或在多个系统中切换查询，然后手动整理对比信息，整个过程可能需要数十分钟甚至更长时间，用户等待耐心不足，常常中断咨询流程，导致销售机会流失。

引入保险产品对比 Agent 后，当用户提出对比需求，如“我想知道重疾险 A 和重疾险 B 的区别”，Agent 立即通

过自然语言理解模块解析用户需求，从保险产品数据库中精准提取两款产品的详细信息，包括保障疾病种类、赔付比例、等待期、保费计算方式等关键要素，运用数据对比和可视化模块，迅速生成一份清晰的对比表格，并以图文并茂的形式展示在销售终端设备上，整个过程仅需几秒钟。

大模型在信息提取和语义理解环节发挥了关键作用，确保了对比信息的准确性和全面性。其主要功能模块包括自然语言理解、数据提取与处理、对比分析与可视化等，前端通过销售系统与用户交互，后端连接保险产品数据库和大模型服务接口，形成完整的调用链路。



该方案实施后，用户咨询满意度大幅提升，销售转化率提高，同时销售人员工作效率显著提高，能够有更多时间专注于用户沟通和需求挖掘，技术创新点在于实现了保险产品对比的智能化自动化处理，突破了传统人工对比的局限性。

价值分析

提升用户体验

为用户提供快速、准确、直观的保险产品对比服务，减少用户决策时间，增强用户对保险购买过程的满意度和信心，有助于提高用户忠诚度和口碑传播。

提高销售效率

帮助保险销售人员快速响应用户需求，提供专业的产品对比建议，缩短销售周期，提升销售转化率，降低销售成本，使销售人员能够将更多精力投入到用户关系维护和业务拓展中。

增强市场竞争力

保险公司能够更有效地向用户展示自身产品的优势和差异化特点，突出产品竞争力，在激烈的市场竞争中脱颖而出，吸引更多用户选择本公司的保险产品。

智能预核保

Agent 概述

保险产品种类繁多，专业覆盖范围广泛，客户需求又各不相同，保险代理人难以全面掌握所有产品形态和承保政策。实际业务中，代理人常需借助电话、微信等方式，就政策解读、方案搭配及初步核保判断等问题，向后端核保人员咨询。然而，核保人员数量有限，其响应时效与并发处理能力难以满足业务高峰需求；且不同核保人员在经验和专业能力上存在差异，导致对前线咨询的解答缺乏统一标准，这无疑给客户体验和服务效率带来诸多挑战。

预核保是保险销售流程中的关键环节，可利用大语言模型在自然语义对话、信息收集、风险评估、辅助决策方面的能力，结合后端已有的核保引擎能力，提前评估客户的风险状况与投保资格，从而提升预核保的效率，为客户带来更为顺畅、高效的投保体验。

需求分析

当前，保险业务在售前预核保阶段存在不少亟待解决的痛点，具体表现如下：

人工经验依赖度高

保险市场上产品种类繁多，不同保险方案和定价政策之间存在显著差异。当前预核保环节在很大程度上依赖核保人员的人工经验进行判断。然而，由于核保人员的专业素养参差不齐，经验丰富程度也各有不同，这就导致对同一保险标的的风险评估结果常常出现偏差，难以保证评估的准确性和一致性。

机会成本损失明显

传统预核保流程需要投保人填写大量的表格和问卷，流程极为繁琐复杂。这不仅严重影响了保险产品的销售效率，也减缓了客户获取的速度。冗长的流程容易让投保人产生厌烦和不满情绪，部分潜在客户甚至会因等待时间过长或流程过于繁琐而放弃投保，这无疑给保险公司造成了不可忽视的机会成本损失。

人力成本居高不下

保险代理人在日常业务开展中，需要频繁向后台专业核保人员进行咨询。但人工核保评估的效率较低，面对大规模业务时，处理能力明显不足，难以满足市场的快速响应需求，进而导致反馈延迟，严重影响客户体验和业务流转速度。

典型案例

模力时刻

Qwen-plus/max/72B

处理 提取用于预核保相关参数。

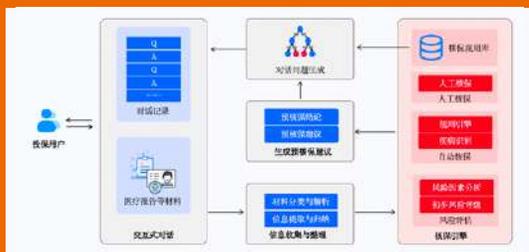
效果 转化提升：对话模式替代传统繁琐问卷按需收集信息，平均时间从 20 分钟降低到 5 分钟提升了投保转化率。

QwQ

处理 评估客户的风险状况、投保资格预判。

效果 体验提升：提前预判投保资格，给出合理引导，避免不符合要求客户进入正式核保流程再拒绝。

本方案主要由以下几个核心模块构成：



1 对话问题生成

该模块摒弃了传统逐项固定式问卷填写模式，采用大模型驱动的对话模式。系统会根据用户当前的回答以及不同保险产品的核保规则要求，自动生成后续必要问题，实现个性化提问。这种方式大幅减少了不必要的问题数量，不再局限于样板式回答，极大地降低了投保人的理解门槛，显著增强了整个交互过程的灵活性，提高了信息收集的质量和效率。

2 信息收集与整理

该模块利用大模型从客户的对话、提供的医疗报告、身份证明等资料中，完成材料分类，准确解析并提取关键信息，然后进行归纳整理，提交给核保引擎，以供核保决策。

3 核保引擎决策

基于提交的材料，结合核保规则，综合客户的风险评估结果、保险产品特点以及保险公司的核保政策，为核保人员提供可能的核保结论，包括标准承保、有条件承保还是拒保。

4 生成预核保建议

依据核保引擎的核保结论，利用大模型为代理人和客户提供针对性建议，帮助他们提前做好心理准备和应对措施。例如，建议客户进行进一步体检，或针对某些风险因素提出加费承保方案。

价值分析

提高核保效率

大模型凭借强大的自然语义对话与信息收集能力，能够快速理解客户信息，协助代理人迅速判断客户的基本投保资格。提前筛选出明显不符合投保条件或风险较高的客户，避免这些客户进入正式核保流程，从而减少核保资源的浪费，提高整体核保效率。

优化客户体验

采用更人性化的对话模式替代传统繁琐问卷，根据客户回答智能生成后续问题，降低客户理解门槛，提升信息收集效率与质量。客户在与系统交互过程中，能更轻松地表达自身情况，提前了解投保可能性。并且，大模型还能依据核保结论为客户提供个性化建议，如指导客户补充资料、选择合适保险产品等，让客户在投保过程中感受到更专业、贴心的服务，极大提升客户投保体验。

降低逆选择风险

通过整合客户多维度信息，结合保险产品特点与核保政策，大模型能够精准识别潜在的高风险客户。代理人依据大模型的评估结果，对风险明显较高且不符合承保条件的客户进行有效引导或拒绝，避免高风险客户过度集中投保，从而降低保险公司的逆选择风险，维护保险市场的公平性与稳定性，保障保险业务的可持续健康发展。

核赔辅助

Agent 概述

意外险产品的显著特点在于其条款体系的繁复性和责任界定的日益精细化，例如通常会约定医院等级、医院清单、疾病清单、药品类型 / 清单、或指定药店等内容。这些详尽的责任界定直接导致了索赔流程所需材料的复杂化。这种复杂性不仅增加了欺诈与渗漏风险，还对审核工作提出了更高要求：审核员需要记忆大量的审核规则和操作流程，还需要具备深厚的医学背景知识以应对复杂的医学情况。无疑加大了审核人员对审核标准的理解与统一执行的难度，使得历史审核中的宝贵经验难以得到高效的累积与传承。

采用大模型学习保险产品、保单及索赔信息，并结合医学知识图谱，实现理赔案件的智能审核。该过程能精确评估条款适用性、判定赔付项目并预测风险，同时清晰展现推理逻辑与依据。这一创新极大减轻了理赔审核人员的工作，使他们无需再繁琐地查阅保单、比对条款及标记风险，而是直接跃升至高效决策阶段。

需求分析

保险行业运营中存在诸多难题，主要体现在以下几方面：

条款复杂，学习成本高

保险产品数量多，综合性保险产品常超万款。条款因监管、竞争和产品迭代而变化快，条款学习需专业知识，成本较高。

专家资源不足

核赔专员中医学专业人员少。但健康险和意外险的医疗、伤残、身故赔款占比高，审核关键赔付项目缺专业医学知识支持。

结构化拆解难

保险条款内容繁杂，涉及承保、责任、免责等多方面。保险用语与产业标准有差异，各业务部门知识水平和认知不同，前后台部门对条款解读也有偏差。

典型案例

模力时刻

Qwen-plus/Max

处理 单据结构化抽取、单据分类、语义审核。

效果 提升理赔自动化比例 15%。

Qwen-VL

处理 单据结构化抽取、单据分类、语义审核。

效果 提升理赔自动化比例 15%。

某财险公司，利用通义千问大模型及通义点金 Agent 平台，打造辅助理赔审核人员的智能体，可实现大模型辅助判责、风险识别、核赔判定。功能包括试用条款 - 责任的判定、免责 / 免赔的判定、赔付范围的判定、风险识别发现等，可输出判定结论及相关依据，辅助理赔人员进行案件审核。

结合现有的理赔系统，引入大模型，在以下三个核心环节实现能力的补充与提升：

1 单据结构化处理

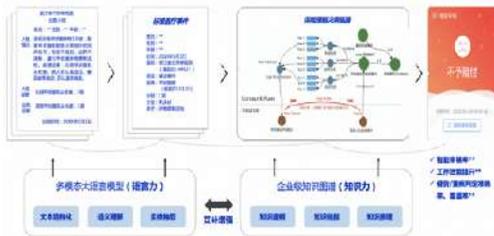
借助先进的大语言模型以及视觉大模型技术，对保险条款的具体内容以及医疗事件相关信息进行深度解析，将其转化为结构化数据。在此过程中，统一规范术语使用，并形成标准化的数据结构，从而为后续的处理和分析提供坚实基础。

2 保险条款知识库构建

通过系统整合和分析，对保险产品的多方面信息，如投保条件、保障范围、保费结构以及免责条款等进行全面梳理，形成一个统一且权威的解读体系。这一知识库将成为理赔工作中的重要参考依据，确保对保险条款的理解准确无误。

3 辅助审核决策支持

充分发挥大模型的强大推理能力，依据用户提交的理赔材料，并结合保险条款中规定的理赔要件信息，对理赔申请进行深入分析。有效弥补传统规则引擎在语义审核方面的不足，并以自然语言的形式给出精准的语义审核判断，为审核决策提供有力支持。



实现效果如图所示

价值分析

提升客户理赔体验

大模型能够快速理解和分析保险条款、保单及索赔信息，辅助理赔审核人员快速做出准确判断，风险管控前置，缩短理赔周期，提高客户满意度。

优化资源利用与降本

通过智能化审核，大模型能够减轻人工审核的负担，合理分配专家资源，降低学习成本和维护成本，同时提高整体团队的审核能力。

智能影像处理

Agent 概述

保险行业的承保与理赔等关键业务流程，高度依赖影像资料的处理。在影像信息提取与分类工作中，传统方法弊端明显。比如在处理卡证影像时，必须针对每种影像单独设计版式训练识别模型，这就需要标注海量数据集，耗费大量人力、物力和时间，成本高昂。而且，传统方法应变能力差，一旦出现新的影像版式，就难以适应，不得不重新开启复杂的训练流程，严重拖慢了后续业务的处理进度。

通义视觉大模型技术的引入，给保险行业带来了业务和技术上的重大突破。该大模型具备较好的泛化能力，可以在零样本的情况下，准确理解和高效推理图像内容，无需针对不同版式分别训练，大大简化了操作。通过这项技术，保险行业的影像处理可实现高度自动化、智能化，作业效率和识别精度大幅提高，影像数据的挖掘价值也得到进一步提升，有力确保了业务操作的合规性，为保险行业的发展提供了强大助力。

需求分析

影像处理效率低，用户体验差

在承保与理赔环节，用户提交材料时，需手动对物料诊断证明等各类材料进行分类上传，同时还要填写大量信息，操作流程繁琐复杂。传统影像处理技术，对于不同类型的影像资料，像卡证等，需要单独制定版式训练识别模型，标注大量数据集，过程漫长且成本高昂，面对新出现的版式和抽取字段更是难以适应，严重影响处理效率。

历史影像档案价值未充分挖掘

保险业务处理过程中积累了海量的历史影像数据，涵盖各类文件、照片及视频等。传统方式下，由于缺乏有效的针对不同图像挖掘技术，难以从这些沉睡的历史影像数据中提取客户个人信息、购买行为、理赔记录等关键内容。导致保险公司无法基于这些数据开展深入的客户行为分析，以提供个性化服务。

风险识别与评估能力不足

在保险业务中，虚假投保与欺诈行为时有发生。当前技术手段难以精准识别影像文件真实性、逻辑合理性等。例如在车险理赔场景中，难以精确评估车辆损坏程度与修复成本；在健康险理赔审核方面，对医疗影像中的异常识别能力不足，定责准确率低。这些问题严重影响保险业务的风险防控。

典型案例

模力时刻

Qwen-plus/Max

处理 影像件结构化抽取、影像件分类。

效果 预计提升影像件挖掘可利用率 60%。

Qwen-VL

处理 影像件结构化抽取、影像件分类。

效果 预计提升影像件挖掘可利用率 60%。

总体技术链路如右图所示，其主要功能如下：



1 单证图像归类

大模型智能影像技术针对影像资料处理的痛点，提供了直接而有效的解决方案。大模型能够自动识别单证类型，并对单证中的信息进行精确分类。这一解决方案显著提升了影像处理的效率和准确性，减少了人为错误，确保了影像分类的精准性，从而增强了对风险的管控能力。

2 图像信息提取

借助大模型的图像识别技术，对各类标准或是非标准影像、单证文件进行结构化信息提取和含义提取，如身份信息、财产证明、门诊发票、投保单、共保协议、银行征信授权书、医疗、病例、医疗检测报告、公示照片等。

3 图像语义理解

借助大模型语义理解。视觉大模型可以快速识别和分析影像数据中的关键信息，如车辆损坏程度、医疗影像异常等，从而加速理赔流程，提高理赔效率。

价值分析

效率显著提升

引入视觉大模型后，材料归类与关键信息抽取实现自动化，仅在信息缺失时需用户补充。这一变革极大地优化了承保和理赔流程中的用户体验，同时显著提高了人工操作效率，大幅减少了人工处理所需的时间和精力成本。

大幅节约成本

大模型具备出色的泛化能力，能够应对新出现的材料，从而取代传统繁琐的模型训练流程。无需进行海量数据集标注，这一优势使得单独训练识别模型和标注数据集所需的人力、物力和时间成本得以大幅降低。

深度挖掘数据价值

借助大模型的图像识别和语义理解技术，能够从海量历史影像数据中精准提取关键信息，为客户行为分析提供丰富的数据支持。基于这些数据，保险公司可以更深入地了解客户需求，进而提供个性化服务，有效增强客户粘性，提升市场竞争力。

条款智能核验助手

Agent 概述

保险产品开发中，保险条款编写制定极为关键，既受外部监管机构严格监督，又要符合公司内部规范。按规定，保险产品需通过智能检核系统完成报送流程，银保监会依据 80 多个文件里的 600 多条规则，从多维度严格检核。但当下产品条款检核工作规则繁杂，审核耗时费力，审核员容易误判、漏检，潜在风险常事后才发现，严重影响产品上市周期。

依托大模型技术，条款智能核验助手能够精准解析、提取规则，并与产品条款进行比对，从而发现问题。人工审核可根据智能体的提示，快速定位并修正条款中的瑕疵，实现事前自动化监管。这一创新模式有效提高了保险产品开发的效率与准确性，增强了时效性，降低了潜在风险，提升了报送质量。

需求分析

传统审核方式已难以满足当下业务发展需求，主要体现在以下几方面：

监管难度升级与人工审核局限

随着监管要求的不断提升，保险条款备案的复杂度显著增加，传统的人工审核模式难以高效应对大量条款文件，导致审核周期延长，严重影响新产品的上市速度。

专业知识壁垒与审核质量不稳定性

保险条款审核涉及众多专业术语和复杂法律条文，对审核人员的专业知识要求极高。然而，非专业人士难以胜任此工作，而专业人才培养周期长、成本高。此外，手工审核易出错，即便经验丰富者也可能因条款繁多而遗漏或误判，影响审核质量。

法规更新响应滞后与审核标准不统一

监管政策频繁变动，人工追踪并调整条款内容存在时间滞后，导致新产品可能无法及时符合最新要求。同时，不同审核人员对同一问题的理解可能存在差异，导致审核结果不一致，增加了审核的反复性和不确定性。

典型案例

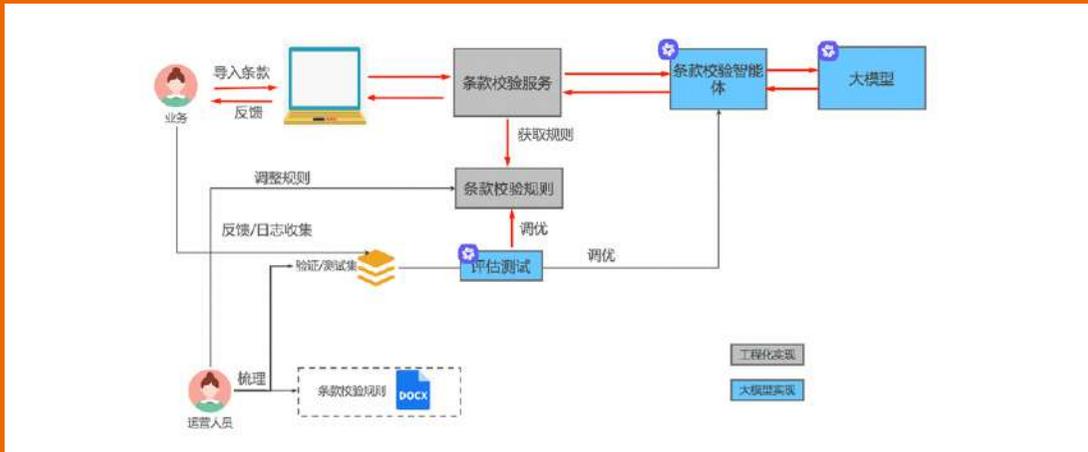
模力时刻

QwQ

处理 基于审核点规则逐项阅读条款原文，给出审核初步结论。

效果 效率提升：审核效率从 1-2 周缩短到 1 天。

为了有效应对保险条款审核中的挑战，我们提出以下基于大模型的解决方案，总体技术链路如下图所示：



1 构建全面条款规则库

- 监管规则库：全面梳理行业产品的监管规则，利用大语言模型的自然语言理解和抽取能力智能地解析和匹配监管规则，建立一个详尽的规则库。这个规则库将涵盖所有内部及外部的监管要求，确保这些规则能够快速、准确地应用到公司的所有业务中。
- 自定义规则：应对未来监管变化，增设规则编辑功能，支持新增、修改、删除等操作，确保规则库实时同步最新监管要求。

2 实施智能检核功能

基于构建的规则库，运用大模型智能检核功能，实现监管规则的自动化检核。具体包括：

- 自动化规则核验：业务上传条款文档后，大模型迅速解析文本，依据规则库逐条比对分析，确保条款合规。
- 智能化修改建议：对于不符合监管要求的条款内容，大模型将提供具体且详尽的修改建议。这些建议将涵盖用词规范、逻辑结构优化等多个方面，以帮助业务团队快速、准确地调整条款内容。

3 输出核验结果

审核完成后，系统自动生成结论，并给出科学合理的整改建议。这些建议基于大模型智能分析，结合行业实践与监管要求，为业务团队指明清晰、可行的整改路径。

价值分析

✓ 效率与准确性提升

自动化核验与智能建议大幅缩短条款处理时间，减少错误，加速产品上市进程。

↓ 降低运营成本

减少人力需求与培训成本，优化资源配置，降低整体运营成本。

☰ 强化合规管理

快速适应监管变化，统一审核标准，确保法规遵从，有效降低法律风险。



条款智转助手

Agent 概述

在保险行业，保险条款的编写与制定不仅复杂且至关重要，还受到监管机构的严格监督。2024 年，国家金融监督管理总局办公厅相继发布了《关于启用财产保险产品智能检核系统的通知》（金办便函〔2024〕1117 号）以及《关于启用农险产品智能检核系统的通知》（金办便函〔2024〕1521 号）。根据要求，除条款原文、费率规章和精算报告原文外，还需将其转换为符合监管要求的要素表形式，并上传至系统进行审核。

条款智转助手帮助保司人员能够智能提取条款、费率规章以及精算报告内容，并按照监管要求的格式，自动生成 excel 形式的要素表。同时，依据要素规则对生成的要素表进行核验，判断是否符合要求，并给出合理建议。该助手显著提升了条款监管报备的效率和准确性，有效减少了人为错误，加快了产品上市速度。

需求分析

保险公司若想上架售卖一款保险条款，需历经手动转换、人工核验、层层审批等一系列人工操作，这使得整个流程效率低下、耗时冗长且容易出错，严重影响产品上市速度。具体业务痛点如下：

手动转换低效

条款、费率规章及精算报告需手动转换为符合监管要求的要素表格式，即每上架一个条款至少处理 6 个文档，耗费大量时间和人力。

人工核验繁琐

每个要素表内容都需要业务人员手动来编辑处理，还需要一个个检查规则，比如有字数要求和符号中英文要求的，都需要业务人员一一检查，费时费力。

层层审批耗时

条款报备需经过多轮内部审批和核查，每个人都是从 0-1 到检查所有内容，过程繁琐且容易出错。进一步延长了产品上市时间，降低了整体效率。

高错误率风险

人工操作增加了出错的可能性，可能导致条款不符合监管要求，延误审核进度或被退回重审。

所以，业务有针对性的目标建设：

要素表自动生成

将自动填写监管审查的要素表，生成检核报告，降低人工审核压力。

要素表信息检查

生成要素表时再次核验要素内容是否符合

典型案例

模力时刻

QwQ

处理 基于审核点规则逐项阅读条款原文，给出审核初步结论。

效果 效率提升：审核效率从 1-2 周缩短到 1 天。

在某保险公司，基于通义千问大模型，将条款、费率规章、精算报告按照要求提取内容，并校验每一条要素信息正确性，最后生成 Json 格式输出。
其主要功能：

1 材料预处理

主要对原始文档（如 PDF 或 Word 文档）进行文本化解析，使其转化为大模型可处理的格式，然后通过大模型对文档的险种进行分类。

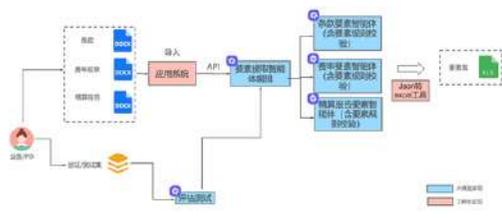
2 信息提取功能

负责从文档中提取特定的要素信息，比如保险产品的类型、覆盖范围、保费计算方法、理赔条件等。它会依据预定义的模式或规则来识别和提取这些信息。

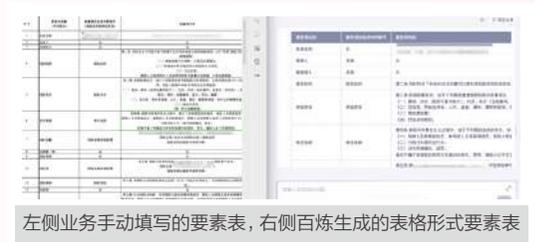
3 结构化输出

将经过校验的数据转换为结构化的格式，如 JSON。此模块负责将非结构化的文本信息转化为机器可读的结构化数据格式，以便后续处理或集成到其他系统中。

总体技术链路如下图所示：简单展示条款 / 核心系统



实现效果如下图所示：



价值分析

效率与成本优化

借助自动化生成要素表，降低人工转换和编辑的时间及人力成本，智能校验快速识别并纠正规则不符项，极大提升工作效率。

精度与合规保障

运用大模型精准提取条款、费率规章及精算报告内容，保证生成的要素表格式规范、内容准确，严格符合监管要求，降低被退回重审风险。

流程与体验升级

通过大模型规则校验结果，减少内部审批重复劳动，优化条款上报流程，缩短产品上市时间，同时减轻业务人员工作负担，助力其专注高价值任务。

对外披露审核

Agent 概述

保险公司的日常运营高度依赖互联网平台进行信息披露，涵盖产品详情、品牌推广、营销广告以及用户教育等多方面内容。这些信息不仅要符合基本语言规范，做到表述精准、清晰且无歧义，更要严格遵循金融行业的法律法规。目前，保险公司普遍采用人工与系统相结合的审核方式，力求对披露物料进行细致审查。然而，实际操作中仍暴露出诸多问题，如审核标准缺乏一致性、人工审核效率低下以及法规复杂难以全面掌握等痛点。

对外披露审核助手基于阿里云通义系列大语言、视觉多模态大模型的多模态理解与审核能力，专为保险公司的市场部、品牌管理部、合规部提供全面而高效的对外信息披露审核能力。

需求分析

当前信息披露审核存在以下痛点：

审核标准不统一

由于不同保险公司或不同审核团队可能采用不同的审核标准，导致信息披露的准确性和一致性受到影响。这种不一致性可能引发消费者的困惑和误解，进而损害保险公司的信誉。

人工审核效率低下

面对大量的信息披露需求，人工审核往往耗时较长且容易出错。这不仅增加了运营成本，还可能导致信息披露的延迟，影响保险公司的市场响应速度和客户体验。

法规复杂性

面对金融行业复杂多变的法规要求，审核人员需要不仅需要具备全面的专业知识，还时刻保持对最新规定的了解，但这一要求往往难以完全满足。

典型案例

模力时刻

Qwen-plus

处理 根据审核规则基于文本理解能力对文本、图片材料进行审核。

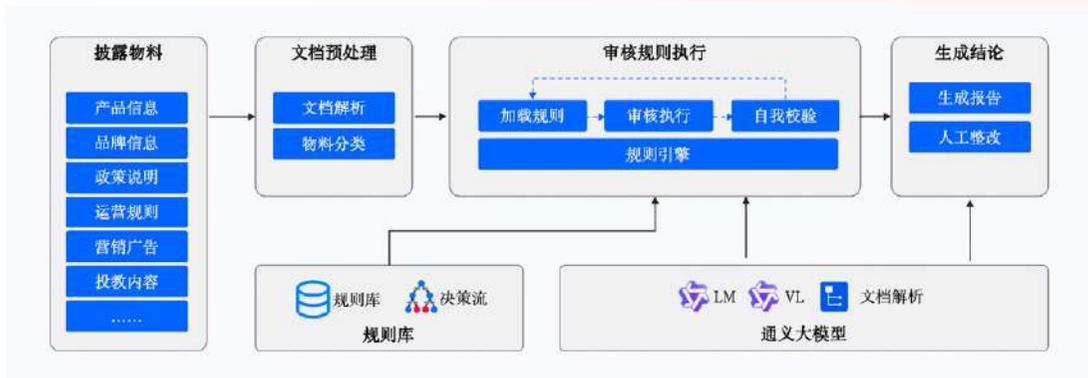
效果 提升效率，每份材料审核时间从 0.5 天提升到 10 分钟。

Qwen-VL

处理 根据审核规则基于文本理解能力对文本、图片材料进行审核。

效果 提升效率，每份材料审核时间从 0.5 天提升到 10 分钟。

整体方案涵盖以下核心组成部分：



1 规则库

依据物料使用场景进行分类构建，每个类别下的审核规则包含多个检查点，每个检查点都详细配置审核规则、校验规则、审核工具以及提示词等信息。

2 大模型

集成语言大模型、视觉大模型和文档解析模型，能够实现自然语言理解、文本与视频审核推理，并生成精准的审核结论。

3 文档解析

接收外部提交的物料，将其解析转换为大模型可识别的文档格式，并依据内容进行物料分类，为后续审核奠定基础。

4 审核执行

根据物料类别从规则库中加载相应规则，按照审核点配置定义，借助规则引擎结合大模型展开审核工作。

5 生成结论

大模型自动生成审核报告，同时提供针对性的整改建议，随后交由人工进行最后的整改完善，实现人机协同的高效作业模式。

价值分析

提升审核效率

借助大模型强大的运算和分析能力，能够在短时间内处理海量的信息披露物料，大幅缩短审核周期。以往人工审核一份复杂的产品宣传资料可能需要数小时甚至更长时间，而引入该审核 Agent 后，相同内容的审核时间可缩短至十几分钟甚至更短，极大地提高了信息披露的及时性。

降低合规风险

规则库全面整合了金融行业的各类法规要求，大模型在审核过程中能够精准比对，确保每一条对外披露的信息都严格符合法规标准。这有效避免了因违规披露而导致的监管处罚、法律诉讼等风险，降低了潜在的经济损失和声誉损害成本。

通用

智能客服

坐席通话质检、意图识别、智能语音导航、机器人营销外呼、机器人催收外呼、客服知识构建、客服知识问答
数字人客服、人工客服助手、App 对话机器人、智能工单助手、智能投诉预警、智能语音分析、电销助手

智能用数

客户经理绩效助手、智能问数助手、智能搭建助手、数据资产问答、存款分析助手、贷款分析助手
客户经营分析助手、小微经营户潜客挖掘、理财潜客挖掘、代发户促活助手

知识助手

知识问答、复杂信息抽取、智能写作

研发助手

AI 程序员、需求助手、编码助手、单元测试生成助手、测试用例生成助手、智能运维助手

数字人

交互取证提额、视频交互陪练、直播交互、视频分享

内容审核

内容安全审核、制度撰写、合规问答

信息检索与打标

企业工商信息打标、可信搜索、企业招投标信息打标、企业知识产权信息打标、财报解析

培训陪练

交互式培训助手、课件生成助手、对练机器人、智能培训问答助手、智能考试助手

办公助手

工作报告助手、OA 小蜜、智能会议助手、流程助手

营销助手

营销策划助手、营销素材生成、营销文案生成、社群运营助手

坐席通话质检

Agent 概述

在客服对话内容质检场景中，企业通常依赖人工监听和评估客服与客户的交互质量。这包括检查客服的服务态度是否友好、服务流程是否符合公司规定以及提供的信息是否准确无误等多方面。然而，传统的人工质检方法受限于时间和资源，只能覆盖到一小部分的通话记录，导致质检覆盖率低且标准难以统一，影响了整体的服务质量和客户满意度。

坐席通话质检 Agent，借助大语言模型的强大能力，可以实现对通话内容的自动化质检。通过深度语义理解和上下文感知技术，大语言模型能够自动识别并标记出潜在问题点，例如：不恰当的服务态度表达、流程执行偏差或回答错误等情况，自动生成详细的质检报告，这些功能不仅大幅提升了质检效率和准确性，降低了运营成本，也为持续优化服务质量提供了坚实的数据支持。

需求分析

客服系统原质检方案主要有两个问题：

运营成本高

质检规则配置复杂，学习和操作门槛高，导致培训和维护成本增加。传统系统要求专业人员进行精细调整，难以快速适应业务变化，进一步推高运营成本。

质检覆盖率和效率都低

基础内容质检仅能覆盖少量对话，覆盖率低且错误率高。人工抽检方式无法全面监控服务质量，容易遗漏问题，影响客户体验，降低了整体质检效率。

典型案例

模力时刻

Qwen-72B

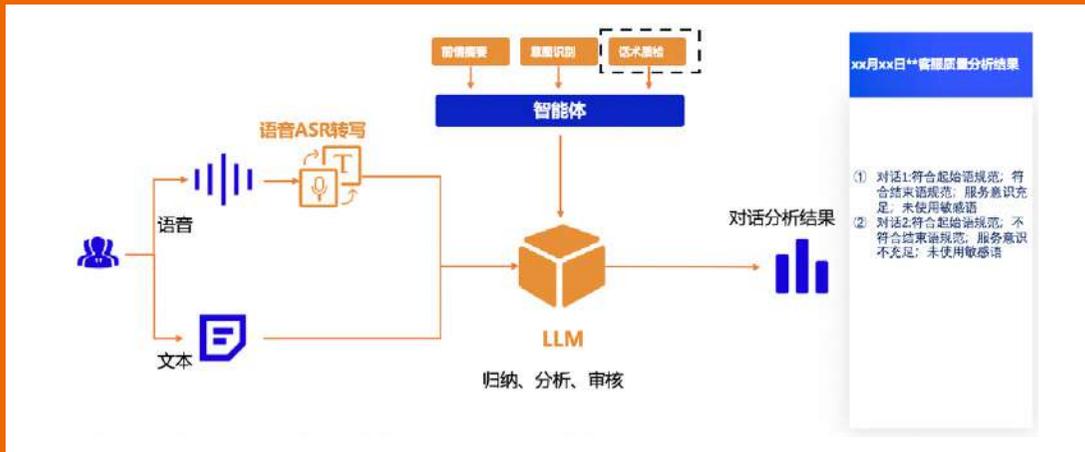
处理 基于提示词规则，对客服通话文本进行质检，产出是否扣分及其原因。

效果 质检准确率 95%。

在某银行，基于通义千问大模型 + 智能体平台，构建智能质检新方案。

大模型处理并理解客服对话文本，自动识别潜在问题和服务不规范之处；智能体根据预设规则和动态学习机制进行实时评估和反馈；对话分析、异常检测、自动评分、报告生成等功能模块，确保质检过程高效、准确且覆盖全面。

技术链路如下图所示：



质检效果示例如右图所示：

对话质量规则包括：起始语规范类、结束语规范类、未使用不文明用语、未使用敏感词、未使用态度恶劣类禁忌语、服务意识充足等。

“评分项”：“开场问候语”，“扣分原因”：“坐席缺少了‘Thanks for calling XXXX This is xxx’的标准开头，扣1分。”

“评分项”：“数据保护”，“扣分原因”：“坐席未在通话开始核对任何客户身份信息，但客户主动提供了称呼，根据规则此情况”

“评分项”：“积极聆听客户”，“扣分原因”：“坐席未出现重复询问已知信息的情况，也未若非所问，不扣分。”，“扣分”：0

“评分项”：“积极回应”，“扣分原因”：“坐席在让客户等待前后均有表达‘请您先等一两分钟时间’和‘感谢您的等待’，符合”

“评分项”：“了解客户需求”，“扣分原因”：“坐席正确理解并回应了客户关于重复投保高额医疗保险的疑问，不扣分。”，“

“评分项”：“提供准确完整的信息”，“扣分原因”：“坐席提供了无法补送的准确信息及特殊处理建议（发送邮件），未发现错”

“评分项”：“回答思路清晰/提供解决方案”，“扣分原因”：“坐席给出了清晰的解决方案即发送邮件特定邮箱，并说明后续”

“评分项”：“案件跟进”，“扣分原因”：“未在通话中提及具体的案件跟进操作如工单或邮件发送，但根据情境无需扣分。”，“

“评分项”：“正确登记小理”，“扣分原因”：“通话记录未展示小结登记过程，无法直接评估，假设未提及不扣分。”，“扣分”

“评分项”：“服务客户态度”，“扣分原因”：“坐席态度始终专业且耐心，未出现负面交流情况，不扣分。”，“扣分”：0

“评分项”：“引导客户”，“扣分原因”：“客户未提出非当前部门解决的问题，坐席无需引导，不扣分。”，“扣分”：0

“评分项”：“避免电话冷落”，“扣分原因”：“通话中未出现明显冷落，坐席处理得当，不扣分。”，“扣分”：0

“评分项”：“人文关怀”，“扣分原因”：“虽然没有紧急状况需特别关怀，但坐席语气平和，不扣分。”，“扣分”：0

“评分项”：“Hold/Mute/正确转接”，“扣分原因”：“坐席让客户等待时操作得当，未违规转接，不扣分。”，“扣分”：0

“评分项”：“确认电话号码或邮件地址”，“扣分原因”：“坐席最终同意通过短信发送邮箱地址，虽未在电话中直接确认邮箱”

“评分项”：“结束语”，“扣分原因”：“坐席结束通话时说了‘祝您生活愉快，再见’，缺少了‘还有什么帮到您’和‘感谢来”

“评分项”：“专业的对话礼仪”，“扣分原因”：“坐席沟通中未出现显著的不良礼仪，如插话抢话、未称呼对方等，不扣分。”

价值分析

降低质检门槛

基于大模型和智能体的智能质检系统显著降低质检规则构建的门槛。通过自动化的质检建模，系统能够自动生成质检规则和报告，无需复杂的配置和专业技能，减少培训和维护成本，使更多员工能够快速上手。

提升质检效果

大幅提升质检的精度和效率。大模型具备深度语义理解和上下文感知能力，可以精准识别客服对话中的问题，如服务态度、流程合规性和回答准确性。智能体实时分析大量对话数据，提供全面、客观的质检结果，显著减少人工抽检的局限性和错误率。自动评分和异常检测等功能模块确保质检过程的高效性，同时降低运营成本，提高客户满意度和服务质量。

意图识别

Agent 概述

金融机构智能客服场景中，面对海量客服需求场景，传统客服基于小模型的意图识别方案，存在意图理解能力有限，个性化服务与情感分析能力不足等问题。

本方案引入大模型进行客户意图识别，并基于语音对话记录等信息提炼形成意图树，构建完整的银行客服领域客户意图识别智能体，不仅大幅降低了运营成本，提升了客户满意度，还为金融机构建立了更高效的客户服务机制，显著增强了品牌形象和市场竞争力。

需求分析

传统小模型在智能客服领域进行客户意图识别时，存在一些局限性。主要体现在以下几个方面：

理解能力有限

传统的小模型往往基于规则或简单的机器学习算法构建而成，对于自然语言的理解能力较弱。它们可能难以准确捕捉到客户的真正意图，特别是在面对复杂、模糊或多义性的表述时。

上下文处理不足

许多传统的模型缺乏有效处理对话上下文的能力，这使得它们很难根据前后语境来正确解读用户当前的话语意义。而在实际的客户服务场景中，理解整个对话流程是非常重要的。

适应性差

随着业务的发展变化以及新词汇、新表达方式的不断出现，如果不对模型进行持续更新和训练，则其性能可能会逐渐下降。传统模型通常需要人工干预进行调整优化，成本较高且效率低下。

个性化服务欠缺

每个客户的背景信息、偏好习惯等都不同，理想状态下智能客服应该能够提供个性化的服务体验。但是，由于数据量限制及技术原因，传统小模型在这方面做得还不够好。

多轮对话支持不佳

复杂的业务咨询往往需要通过多次问答才能完成。传统的小模型可能在维持长时间对话连贯性和一致性方面表现欠佳，容易导致用户体验下降。

典型案例

模力时刻

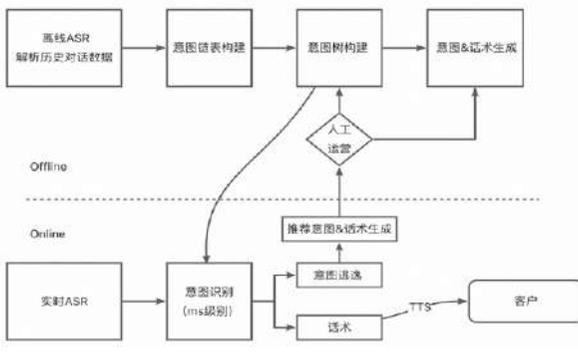
Qwen-14B/72B

处理 客服场景下，大模型理解用户 query，以及意图树，进行意图分类。

效果 意图识别准确率 95%。

在某金融机构客服中心，基于大语言模型强大的语义理解能力与规划推理能力，准确理解用户 query 意图，在多轮上下文、多主题对话中，保持对意图识别的准确性与连贯性，辅以调用外部工具与知识库，高质量完成客户的对话服务。

具体来讲，整个意图识别的工作分为离线生成和在线分类两条链路：



1 离线链路

构建指定业务场景下的完整意图树，是在线做意图识别的基础。通过挖掘历史通话记录，生成潜在意图列表，经过精炼和结构化，生成意图树，并挂载对应推荐话术。

2 在线链路

客户对话过程中，实时把通话进行文本转换，送给大模型在意图树上做识别，对于命中意图，生成话术；对于未命中的情况，大模型生成推荐意图及话术，经人工运营确认后合并到意图树当中。

价值分析

在客户意图识别场景，通过引入大模型 Agent 为智能客服系统带来显著价值：

增强智能运营能力

通过对海量对话数据的分析，基于语音、对话记录等信息提炼形成意图树，构建完整的银行客服领域客户意图体系，进一步增强系统的智能运营水平。

改善客户服务体验

提高响应速度与准确性，准确地识别出客户的真正需求或问题所在，从而提供更加迅速且针对性的服务响应。理解客户的独特需求，推送个性化的金融产品和服务建议，捕捉到顾客的情感变化，适时调整沟通方式，让互动过程更加人性化。

洞察潜在商机

通过对客户意图深度解析，发现新的市场需求点或是现有产品的改进空间，为开发新产品或服务提供依据。基于对个体用户的全面了解，银行可实施更为精确的目标市场营销活动，提高转化率的同时降低成本。

智能语音导航

Agent 概述

传统的客服电话语音导航技术通常依赖于交互式语音应答（IVR）系统，这种系统通过一系列预设的按键菜单引导用户进行操作。客户在寻求服务时往往需要按照语音提示逐步选择相应的数字键以进入不同的服务分支，这种方式不仅耗时，而且容易让用户迷失在复杂的菜单结构中，导致体验不佳。此外，传统 IVR 缺乏对自然语言的理解能力，无法处理超出预设选项之外的问题，这限制了其灵活性和效率。

智能语音导航 Agent，基于大语言模型技术，智能语音导航得到了显著增强，实现了从按键选择到语音指令的转变，使得业务办理更加直观快捷。智能语音导航可以理解用户的自然语言请求，并直接提供所需的信息或服务，无需经过繁琐的菜单选择过程。更重要的是，基于大模型的强大语义理解 and 多轮对话管理能力，智能语音导航能够准确捕捉用户意图，支持个性化服务推荐，同时还能持续学习和适应新的查询模式，极大提高了响应速度和服务质量。

需求分析

支持通过语音方式完成交互，将用户声音转换为文本并自动理解用户的自然语言中包含的准确业务需求。客户通过采用智能语音导航系统来处理日益增长的信息咨询、电话交易和服务需求，系统可有效提升用户满意度，减轻人工服务压力，降低运营成本。

智能语音导航系统可以无缝集成到企业的信息系统中，特别是与现有的 IVR 系统融合，辅助或代替人工座席完成数据查询、订单处理、业务办理等业务咨询操作，拓展及丰富企业的对外服务渠道，提供更快捷的客户服务响应能力。

典型案例

模力时刻

Qwen-72B

处理 客服场景下，大模型理解用户 query，进行路径选择与服务。

效果 意图识别准确率 95%，路径选择准确率 90%。

大模型与金融机构的呼叫平台对接，接受电话客服语音消息，通过智能语音识别 ASR 转写录音为文本，使用大模型技术自动理解用户问答内容，返回文本答案，通过语音合成 TTS 将答案合成为录音，并最终播报给用户。

大模型加持的智能语音助手，相比原有的语音导航系统带来多个方面能力的提升：

1 意图识别和分类

先用 ASR 把客户语音转文字，用大模型对用户会话内容进行分析，识别出用户的具体意图。例如，当用户说“我要办理信用卡”时，大模型能够准确识别出这是“业务办理”类的请求。大模型的强大语义理解和上下文感知能力，能

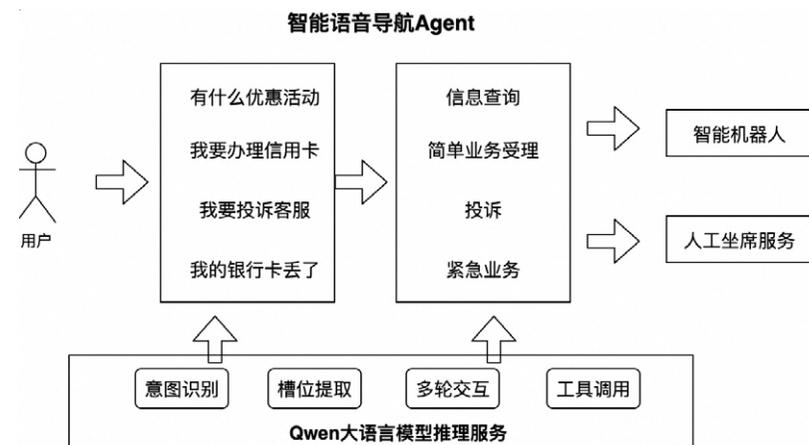
够更好的解决小模型时代意图识别准确率和意图的继承 / 跳转等问题。

2 从会话中提取关键信息

大模型从会话文本中，提取关键信息（槽位），如时间、地点、具体业务类型等。例如，从“我的银行卡丢失了要怎么处理”中提取出“银行卡”和“丢失”这两个关键信息。大模型相比 NLP 小模型，具备很强的泛化能力，能够识别出绝大多数的实体类型，不需要单独训练。

3 调用不同工具处理客户诉求

根据意图识别和槽位提取的结果，大模型智能进行工具选择，简单问题分发给智能机器人，通过 RAG 检索答案生成；复杂问题转接到人工客服按流程进行处理。



价值分析

大模型在智能客服语音导航场景下的业务价值主要体现在以下几个方面：

提升客户满意度和体验

通过更精准的意图识别和分类，大模型能够迅速理解客户的真正需求，并提供直接且相关的解决方案。例如，当客户表达“我要办理信用卡”时，系统不仅能立即识别出这是一个业务办理请求，还能基于上下文感知能力，为客户提供个性化的服务选项或引导流程，减少客户的等待时间和操作步骤。

提高运营效率和服务质量

大模型从会话中高效提取关键信息（槽位），使得客户服务团队可以快速获取必要的背景资料，无需客户重复叙述。这种能力极大地提高了处理速度和准确性，尤其对于复杂问题的转接处理更加流畅，减少了因信息不全导致的服务延误，从而提升了整体服务质量。

优化资源配置和成本控制

根据意图识别和关键信息提取的结果，大模型能够智能地分配任务给最适合解决问题的资源，无论是自动化工具还是人工客服。这种方式不仅保证了简单查询由机器人高效解决，还确保了复杂问题得到专业人员的关注，实现了人力资源的最优配置，降低了长期运营成本。

机器人营销外呼

Agent 概述

在传统的电话营销外呼场景中，企业通常依赖大量的人力资源进行客户联系和销售活动。这种模式不仅耗费时间，而且效率低下，因为人工坐席需要逐一拨打潜在客户电话，并试图在有限的时间内完成产品的介绍和推销。此外，由于缺乏有效的筛选机制，许多通话可能并不会带来有价值的商业机会，导致资源浪费。同时，人工质检的局限性也使得服务质量难以得到有效监控和提升，从而影响了整体的客户体验和转化率。

大语言模型在营销外呼中应用带来多方面的提升，通过个性化沟通、流程自动化、数据驱动的优化以及合规性控制等多方面的支持，可以显著提升外呼效率和效果，降低运营成本，同时提升客户满意度和转化率；企业应根据自身需求，灵活整合和应用这些技术，打造高效、智能的外呼体系。

需求分析

传统的电话外呼，使用 NLP 小模型和规则，面临以下挑战：

对话理解不足和混淆意图

传统外呼系统在对话理解和自然语言处理（NLU）方面存在局限性，难以准确识别和处理多意图和混淆意图，导致对话流程复杂且容易出错。

剧本流程配置复杂

对话管理（DM）的剧本流程配置成本高、维护难度大。严格的流程处理使得系统无法灵活跳转，限制了对话的灵活性和适应性。

机械感强，缺乏人性化

对话生成（NLG）过程中的回复内容固定，缺乏个性化和情感表达，使得对话显得机械，无法满足用户的情感需求和互动体验。

典型案例

模力时刻

Qwen-72B

处理 外呼场景下，基于用户会话，识别客户意向与情绪，提升业务效果和用户体验。

效果 意图识别准确率 90%，情绪感知准确率 95%。

某消金公司，通过大模型针对历史营销电话生成意图树，同时接通大模型的 ASR+TTS，实现大模型智能外呼，同时针对每通电话的交流情况，进行商机挖掘和执行计划 SOP 流程打通，加速业务的推动效率和降低风险。

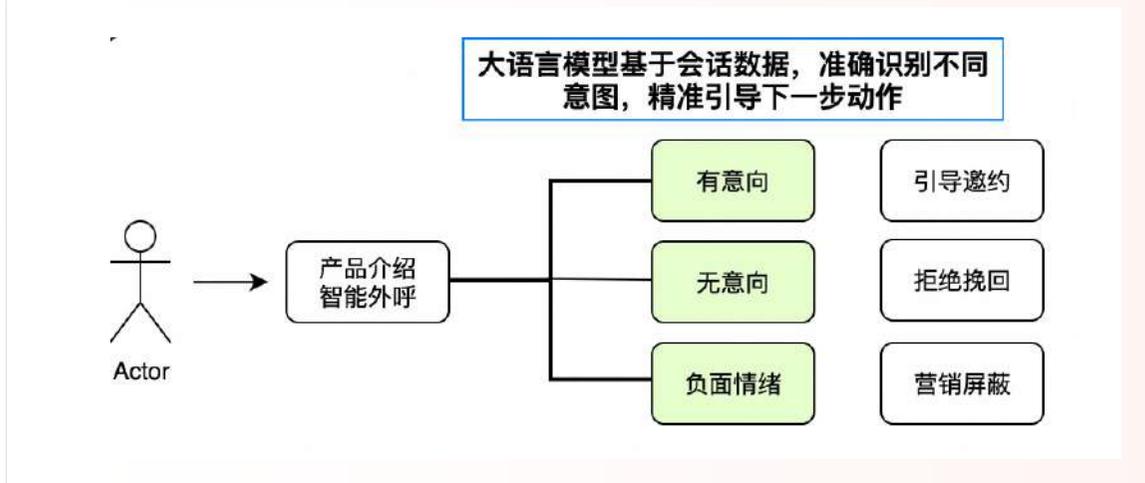
在采用了大模型技术后，外呼系统关键业务指标有所提升：

1 识别率大幅提升

大模型的识别率超过 8%，相较于传统 NLP 和人工平均水平有显著提高。这表明大模型能够更准确地理解和处理客户对话，减少了错误识别的情况。

2 接通转化率显著提高

大模型的接通转化率超过 3%，远高于传统方法。这意味着通过大模型进行的外呼不仅更容易被客户接听，而且更有可能转化为有效的沟通或行动，提升了整体业务效率和客户满意度。



价值分析

大模型加持的智能营销外呼不仅让企业能够更专注于挖掘高净值客户的潜力，同时也极大地提升了人效比。通过精准识别客户需求和偏好，利用根据不同用户特征定制的个性话术，可以显著提升客户体验，增加用户对品牌的喜爱度，并有效提高收益同时降低客诉率。此外，相比传统的小模型，大模型在意图理解和业务路径选择上展示了更高的准确性，使得客户服务更加个性化、响应更加精准。

提升高净值客户开发效率

借助大模型的高效分析能力，企业能更精确地识别并锁定高净值客户群体，实施针对性的营销策略，从而大幅提升转化率和客户生命周期价值。

优化用户体验与满意度

基于用户特征的个性化话术定制，使每一次互动都更加贴合客户需求，增强用户的好感和忠诚度，有助于建立长期稳定的客户关系，促进二次销售和服务升级。

增强意图理解与服务精准度

大模型提高了对外呼叫过程中意图识别的准确性，确保了业务路径选择的合理性，为企业提供了更为精准的服务解决方案，减少了误操作和服务偏差，进一步降低了客户投诉的风险。

机器人催收外呼

Agent 概述

传统的电话催收外呼场景往往依赖大量的人力资源，由专门的催收人员逐一拨打欠款客户的电话，试图通过对话沟通促使对方履行还款义务。这种方式不仅耗时费力，而且由于人工操作的局限性，难以保证每次通话的质量和效果。此外，面对海量的客户数据，传统催收方式很难做到精准分类和个性化处理，导致效率低下，容易错过最佳催收时机。同时，人工催收过程中还可能存在语言表达不规范、情绪管理不当等问题，影响客户体验并可能引发风险。

通过大模型技术优化催收外呼流程，不仅可以提高催收效率和成功率，还能提升客户体验，确保合规性，降低运营成本。关键在于合理规划和实施，通过技术与业务的深度融合，实现催收业务的智能化、自动化和个性化。企业应根据自身实际情况，选择合适的技术方案，并进行持续优化，以充分发挥大模型的优势，解决催收外呼中的痛点和需求。

需求分析

传统金融机构，在做电话催收时面临着各种痛点：

沟通效率低

人工外呼需要大量人力，通话时间管理难，催收效率有限。

客户满意度不高

重复性高、缺乏个性化的沟通方式，可能导致客户抵触情绪。

合规风险大

催收过程需要遵守严格的法律法规，违规操作可能带来法律风险。

人员培训和管理成本高

新员工需要时间培训，且质量难以统一。

情感识别和应对不足

无法实时识别客户情绪，缺乏有效的应对策略。

典型案例

模力时刻

Qwen-72B

处理 外呼场景下，基于用户会话，识别客户还款意愿，推荐对应话术。

效果 意图识别准确率 95%，通话轮次增加 50%。

某金融机构通过大模型针对历史催收电话生成意图树，在线接通大模型和 ASR+TTS，实现大模型智能催收外呼，同时针对每通电话的交流情况，进行商机挖掘和执行计划 SOP 流程打通，加速业务的推动效率和降低风险。



1 基于对话历史构建催收意图树

让大模型基于历史催收通话记录，自动挖掘意图标签，经过聚类、筛选，构建出一棵完整的意图树，在线上的通话中依据此进行意图识别。

2 在线的催收外呼 Agent

在每一轮交互中，结合上下文，判断客户意图，精准调用 FAQ 检索、RAG 问答、SOP 执行等不同业务链路，提升服务准确率。

3 离线外呼语料处理

每天的通话记录，事后可以用大模型对其进行挖掘，包括新增意图的挖掘，合并回意图树；基于通话内容挖掘二次营销的商机等；对通话内容进行合规性监察，确保安全合规。

价值分析

智能催收外呼让企业更聚焦高净值客户的深入挖掘，通过大模型实现自动化拨号和智能对话。利用情感分析技术实时监测客户情绪，动态调整催收策略。建立合规监控模块，确保催收过程符合法律法规。通过数据分析优化催收目标和策略，提高催收成功率。

提升催收效率与成功率

通过自动化的拨号和智能对话，以及基于数据分析的策略优化，大模型使得催收过程更加高效，显著提高了对欠款的回收速度和成功率。

实时情绪响应与个性化催收策略

借助情感分析技术，系统能实时感知客户情绪并据此调整催收策略，实现更加人性化的互动，有助于建立信任关系，提高客户还款意愿。

强化合规性与风险管理

内置的合规监控模块确保所有催收行为符合相关法律法规，帮助企业在追求催收效果的同时，有效规避潜在的法律风险，保障企业的长期稳定运营。

客服知识构建

Agent 概述

在智能客服场景中，面对海量用户咨询的需求，传统客服系统存在知识库更新成本高、功能有限的问题。本方案通过引入先进的自然语言处理技术和金融级智能体能力，实现了文档内容的自动抽取及问答对的快速生成，有效解决了信息更新慢、准确率低和响应不及时等痛点。

核心技术依托通义千问大模型与智能体平台，提供了开箱即用的非结构化信息处理、知识库构建与管理、检索问答以及持续运营等功能。这一转变使得从传统到智能的知识库，在知识沉淀、管理和应用的全链路上完成了智能化升级。这不仅大幅降低了运营成本，提升了客户满意度，还为银行建立了更高效的客户服务机制，显著增强了品牌形象和市场竞争能力。

需求分析

客服系统基于某供应商的传统知识库，存在如下问题：

语义理解不足

依赖于关键词匹配或简单的模板规则，难以真正理解用户查询的语义，导致对复杂或非标准表达的理解能力差。

更新维护成本高

随着业务变化和数据结构更新，传统的 FAQ 和 N2SQL 映射需要频繁的手动调整，这增加了维护成本和时间开销。

准确性较低

由于缺乏深度学习和自然语言处理技术的支持，传统系统的查询解析和 SQL 生成准确性不高，容易产生错误的结果。

扩展性差

当业务增长或需求变化时，难以快速扩展以支持新的功能或更复杂的查询需求。

灵活性有限

固定的 FAQ 条目和预定义的查询模板无法灵活适应用户的多样化提问方式，尤其是当用户提出的问题不在预设范围内时，系统往往无法给出准确的回答或有效的 SQL 转换。

典型案例

模力时刻

Qwen-72B

处理 客服坐席辅助场景下，把知识向量化后通过 RAG 召回给大模型回答参考。

效果 知识召回率 95%。

在某银行，基于通义千问大模型 + 智能体平台，构建以远程银行需求为背景打造全行知识中心，目前以远程银行各基础产品、管理知识为基础进行知识中心的建设，功能包括知识库智能体、RAG 知识库、提示词工程等。技术链路如下图所示：



1 构建知识库内容

利用通义千问大模型的强大能力，系统能够自动挖掘上传文档中的 FAQ 知识点。通过算法分析，系统可以智能识别并提取用户高频咨询的问题与答案。提供两种挖掘方式：一是从文档中抽取关键信息，二是基于客服对话日志进行分析。挖掘出的知识点在经过业务督导的人工审核后，将被正式录入知识库，确保内容的准确性和适用性。

2 Agent 配置与优化

知识库 Agent 通过注册模型服务，实现与现有模型的无缝对接。管理员可以设定调用大模型的具体时机、超时时间和模型响应阈值等参数，以优化性能和资源使用。配置允许系统根据实际需求灵活调整，确保服务的稳定性和响应速度。

3 Prompt 设计与集成

智能体平台提供了专门的工具用于开发和测试 Prompt，帮助设计有效的提示语句。通过 API 动态获取 Prompt，确保其与业务逻辑解耦，灵活性更高，进一步提升大模型的应用效果。

价值分析

大模型 Agent 的引入为智能客服系统带来显著价值：

提升服务效率

深入理解用户提出的问题，解析其中的语义和意图，从而更准确地定位用户需求提升客服准确性与效率，减轻人工负担，增强用户满意度。

确保信息实时

可以从大量的非结构化文档中自动提取关键信息，通过文档抽取，快速更新知识，确保信息准确与时效，增强智能客服系统的响应能力和知识储备深度。

优化知识更新

基于现有的数据和内容自动生成高质量的问答对，极大地丰富了智能客服的知识库，保持知识库动态更新，优化服务状态，随时准备应对各种用户咨询。

意图识别

Agent 概述

在日常的客服问答场景中，客户通常通过电话、在线聊天或邮件等方式联系企业寻求帮助。面对各种咨询和问题，人工坐席需要快速准确地理解客户需求，并提供有效的解决方案。然而，由于业务知识的复杂性和多样性，这对人工坐席的专业知识和应变能力提出了较高的要求。同时，高流动性的人工坐席队伍也给企业的培训和管理带来了挑战，导致服务质量参差不齐，影响了客户满意度。

客服知识问答 Agent，引入大语言模型技术，通过智能信息辅助、对话管理和操作建议等功能点来增强客服系统的智能化水平。大语言模型能够实时分析客户提问并提供精准的答案推荐，减少人工坐席查找信息的时间；在对话过程中，它还能根据上下文动态调整响应策略，确保沟通的连贯性和个性化。此外，针对复杂的业务流程，系统可以自动生成操作指南，指导坐席完成相应任务，极大地降低了对员工经验的依赖。

需求分析

客服系统依赖纯人工处理导致效率低下、管理成本高，且因从业人员素质不一，服务质量难以标准化，客户体验不稳定。通过大模型与智能体的结合，降低对员工技能的要求，拓宽人才选择范围，使服务结果更加一致可控，提升处理效率和服务质量。

总体需求如下：

提升客服系统效率与服务质量

实时分析客户意图，客服服务过程中客户提出的问题，自动推送相关应答话术或知识，提高响应速度和服务准确性。利用大模型的语义理解和概要总结能力，进行会话小结和工单总结，确保服务流程的高效性和一致性。

数据驱动的智能运营优化

提供数据看板功能，管理者可以通过查看各场景的大模型调用次数、每笔交易的服务时长、热点知识排行以及坐席反馈质量排行等关键指标，进行统计分析和优化决策。

典型案例

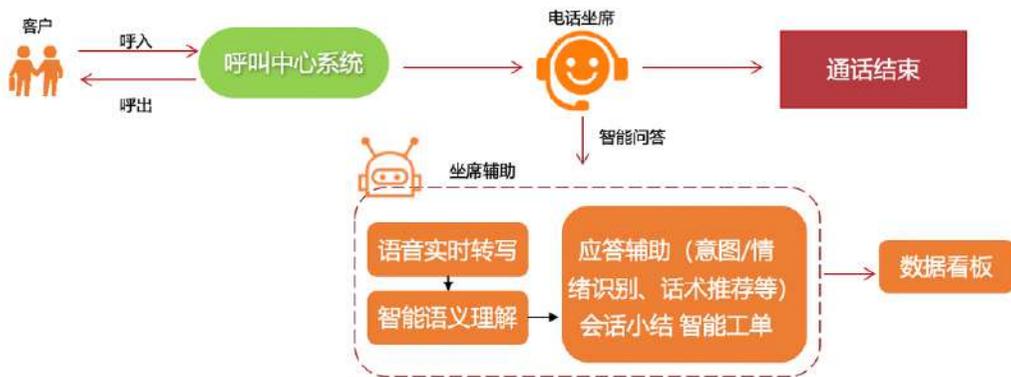
模力时刻

Qwen-72B

处理 客服坐席辅助场景下，把知识向量化后通过 RAG 召回给大模型回答参考。

效果 知识召回率 95%。

在某银行，基于通义千问大模型 + 智能体平台，构建坐席的精细化管理和辅助。
技术链路如下图所示：



1 实时文本转换

对话内容通过 AI 引擎实时将通话 ASR 处理成文本。

2 智能分析处理

语音识别结果输入通义千问大模型，基于预定义的 Agent 任务对信息进行加工分析，并结合增强型知识库完成文本处理。

3 自动总结生成

大模型对通话内容进行总结分析，自动生成小结摘要和工单。

4 数据统计展示

智能体支持按应用和日期维度的数据统计分析，展示调用次数、服务时长、热点知识及坐席反馈质量等关键指标。

价值分析

降低客服的门槛与成本，让服务质量和客户满意度更加可控：

系统能够精准识别客户需求并即时推荐相关知识，极大提升业务办理效率。

减少对员工高技能的要求，拓宽人才选择范围，同时降低培训成本和人力支出。

服务结果更加一致和标准化，每次交互符合最佳实践，提高客户满意度和运营效率。

数字人客服

Agent 概述

近年来数字人因其互动性、人格化和不间断服务等特性，在金融领域迅速得到应用并产生显著价值，获得了市场的广泛认可。数字人已深入应用于 C 端客户服务、品牌合作等多个方面，涵盖智能服务、虚拟直播和视频生产等维度，与业务深度融合，创造了实际效益。

大语言模型技术取得突破后，无疑是给数字人配备上了先进的“大脑”，集成自然语言处理技术，能够实时理解客户意图，提供个性化营销建议，并自动处理复杂业务请求，此工作模式极大地提升金融服务的效率和用户体验。

需求分析

在某金融机构，计划采用大模型和智能知识库作为数字人“大脑”，通过集成先进的自然语言处理技术，使数字人能够实时理解用户的意图和情感，从而提供更加自然、流程的交互体验。同时，支持 H5、WEB 和 API 等多种接入方式，与数字大屏系统无缝集成，实现智能音视频的多模态互动。从而提升客户体验和运营效率，增强服务的多样性和灵活性。

典型案例

模力时刻

Qwen-72B

处理 理解对话上下文，基于用户 query 检索相关知识，完成答案生成。

效果 答案生成准确率 90%。

在某金融机构，基于通义千问大模型 + 智能体平台，构建数字人“大脑”。

1 知识内容对接与集成

为了实现知识内容的无缝对接，Agent 通过 H5、WEB 和 API 等接入方式与数字人系统进行集成。新输入的内容将通过通义千问大模型处理，基于预定义的 Agent 任务对输入信息进行加工分析，并结合知识库完成文本内容的处理。

2 信息回传与处理

经 Agent 处理后的内容信息，通过数字人系统的接口回传给数字模块，确保处理结果能够实时反馈并应用于后续交互中，维持对话的连贯性和准确性。

3 多轮对话场景支持

采用 Agent 智能体的实现方式，数字人系统能够支持复杂的多轮对话场景。增强对话的灵活性，为未来更多场景的扩展提供坚实的基础，不断适应新的业务需求。

技术架构图如下图所示：



价值分析

数字人结合大模型和智能体，显著提升金融服务的效率与质量。

业务价值

数字人可以 7*24 小时不间断服务，覆盖 C 端用户服务、品牌合作等多个领域，涵盖智能客服、手机银行等多维度应用，与业务深度融合，创造实际效益。其互动性和人格化特征提高客户满意度和品牌忠诚度。

技术价值

大模型和智能体提升对话的精准性和连贯性，支持复杂多轮对话场景。同时灵活接入数字大屏系统，实现智能音视频的多模态互动。技术组合不仅降低运营成本，还提升服务的智能化水平和响应速度。

人工客服助手

Agent 概述

面向金融机构的运营中心、客户服务部门、远程银行部，辅助坐席服务客户，解答客户关于的各类问题；随着金融业务的持续创新与拓展，客户群体不断壮大，业务复杂度呈指数级上升，坐席人员所面临的知识挑战愈发艰巨，不仅要熟知大量金融业务知识，还需在分秒必争的客户交互中，迅速精准地组织答案，提供专业且准确的回答，对人工客服的知识储备与响应能力提出了较高的要求。

依托大语言模型技术，人工客服问答助手通过检索企业知识库，生成答案 / 话术润色，提升人工客服知识查找、组织答案效率，提升人工坐席客户服务效率，有效释放坐席；此外，缩短客户等待时间，提升用户体验。

需求分析

知识查询效率低

知识库庞大且复杂

金融机构知识库包含海量信息，涉及众多业务领域与产品，传统检索方式依赖关键词匹配，往往导致大量无关结果，坐席人员难以快速找到准确答案，浪费大量时间与精力。

知识更新不及时

金融业务法规、产品条款等信息更新频繁，传统知识库更新机制滞后，坐席人员可能依据过时信息提供服务，引发客户不满与业务风险。

答案组织效率低

手动整理答案耗时

即使找到相关知识，坐席人员仍需花费大量时间手动整理、编辑答案，尤其是在业务高峰期，严重影响服务效率与客户等待时间。

使用和理解成本高

缺乏有效的检索工具

目前大部分金融机构的 SOP（标准操作程序）知识库维护依靠 Word、Excel 工具，检索效率低。

理解和使用成本高

SOP 知识库中含有大量术语，表述过于专业，坐席理解和理解成本高。

典型案例

模力时刻

Qwen-72B

处理 坐席辅助场景，基于用户 query 检索相关知识，完成答案生成。

效果 答案生成准确率 90%。

1 客户输入的问题

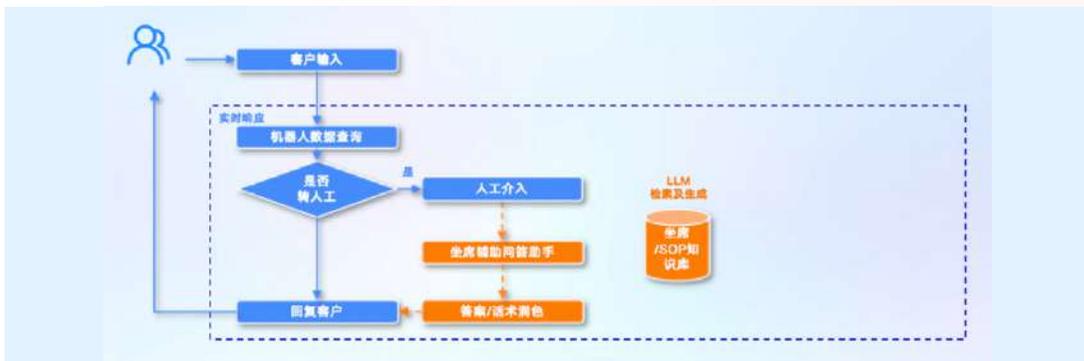
先走大模型 RAG 做知识召回，然后问答助手会生成最终回复话术；若召回结果为空，可流转人工服务。

2 在人工服务过程中

问答助手会支持知识问答、原文溯源展示等功能，便于人工坐席快速校验回复内容的准确性。

3 为了提升问答效果

采用大语言模型和传统小模型结合的架构，事前对用户 query 做改写，事中对召回的知识做排序，事后对话术进行润色等，提升回答质量。



价值分析

在人工客服坐席辅助场景下，使用大模型技术能够显著提升业务效率和服务质量：

大幅提高了知识查询与获取的效率

大语言模型能够智能检索企业庞大的知识库，通过理解上下文提供精确答案，而非依赖简单的关键词匹配。这种方式不仅减少了无关结果的数量，还让坐席人员可以迅速找到所需信息，极大缩短了响应时间，提升业务吞吐量。

增强了答案组织与提供的速度和准确性

利用大模型自动生成或润色答案的功能，坐席人员可以从繁琐的手动编辑工作中解放出来，尤其是在处理复杂金融产品咨询时，能更快速地提供专业且准确的回答，有效提升了服务效率并减少了客户等待时间。

降低了理解和使用知识的成本

大模型可以通过简化术语解释和优化知识呈现方式，使得复杂的金融概念更容易被理解和应用。这降低了新员工培训难度以及所有坐席人员理解 SOP（标准操作程序）的门槛，同时提高了整体服务质量，确保每一位客户都能获得清晰、专业的解答。

App 对话机器人

Agent 概述

针对金融机构的 App 应用、远程银行服务及互联网金融服务等业务部门，提供对话机器人助手服务已成为提升客户体验的关键。这类问答机器人不仅需要具备强大的在线问答能力，还需深刻理解客户意图，以实现精准的业务目标牵引和场景化线上营销。通过智能化的服务，这些机器人能够在满足用户需求的同时，促进金融产品的推广和服务的优化，为用户提供无缝的数字银行体验。

采用大语言模型结合检索增强生成（RAG）技术，可以显著提升问答机器人的性能与服务质量。大模型的自然语言处理能力可直接提高意图识别的准确性，召回相关知识，提升问题答复的准确率，促进业务目标达成，从而为金融机构带来更高效的客户服务和更高的用户满意度。

需求分析

在当前在线问答机器人的应用中，存在几个主要痛点影响了服务效率和客户体验：

意图理解不准确导致输出答案无效化，这不仅浪费了客户的时间，还降低了客户对自助服务的信任度，增加了转人工服务的需求。

由于处理速度慢和答案质量不高，客户的等待时间显著增加，进一步加剧了负面体验，同时也阻碍了业务目标的有效牵引与转化率的提升。

这些问题共同作用，使得在线问答机器人难以实现其预期的服务优化和业务增长目标。

典型案例

模力时刻

Qwen-72B

处理 对话机器人场景，基于上下文和当前 query 理解意图，检索相关知识完成答案生成，以及调用工具办理业务。

效果 意图识别准确率 90%，工具调用准确率 95%。

使用大模型技术，对原有的问答机器人对话系统进行升级，加强其语义理解能力、多轮对话能力、上下文意图识别、话术推荐等能力，提升问答效果。

1 问答机器人增强

提升原有小模型技术架构的准确性，满足业务场景效果要求。

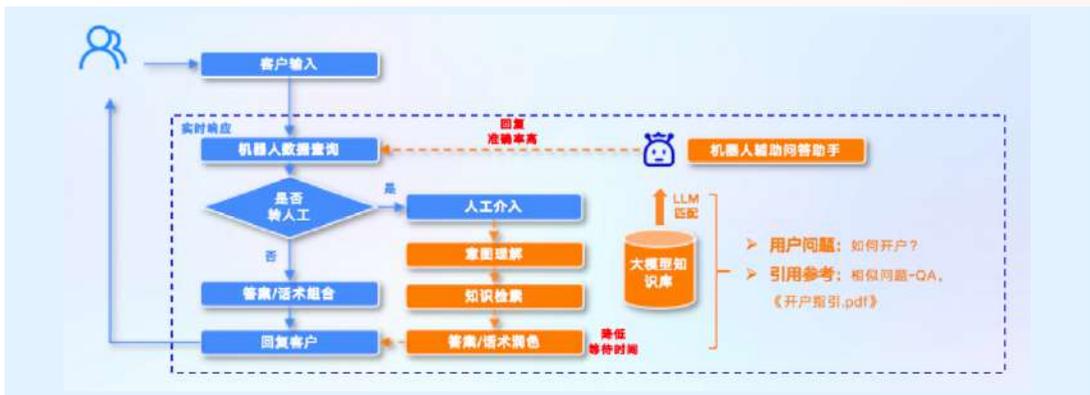
- 智能回答：采用严格匹配 FAQ 知识库的答案，严格保证回答内容的准确性与合规性；
- 智能推荐：支持配置推荐问，根据上一个问题，系统推荐下一个用户可能要问的问题，减少用户输入，进一步提升用户体验；

- 智能引导：多个渠道、多次咨询不同问题，适时引导，提供业务场景营销能力延展。

2 大模型时代

知识运营变得重要，大模型强大的语言理解和处理能力，可减少知识运营中的繁琐工作。

- 知识智能运营：知识生产 - 知识审核 - 知识入库的一整套业务流程操作，提升知识维护的效率；
- 知识重复校验：支持标准问、相似问 / 变体问的维护，并支持知识的重复检验，自动化提示用户，降低知识的维护难度；
- 知识即刻更新：支持知识的即刻更新、上线，简化知识库上线的流程，降低了时间等待，进一步提升客户咨询体验。



某金融机构在使用大模型升级其对话机器人系统后，机器人问答准确率可做到 90% 以上，运营后准确率 95% 以上，标准答案采纳率提升 20%，提升人力效率 60% 以上。

价值分析

大模型在金融 APP 对话机器人上带来的业务价值主要体现在以下几个方面：

提升问答交互的质量与效率

大模型增强了原有小模型技术架构的准确性，确保了智能回答内容的准确性和合规性。此外，通过智能推荐和引导功能，不仅可以减少用户的输入操作，还能根据用户的行为适时提供相关问题或营销信息，极大提升了用户体验和服务效率。

优化知识管理流程

利用大模型强大的语言理解和处理能力，简化了从知识生产到审核再到入库的知识运营流程，提高了知识维护的整体效率。同时，支持标准问、相似问及变体问的维护，并能自动进行重复检验，降低了知识维护的难度和成本，确保知识库的高质量与及时更新。

即时响应客户需求

知识即刻更新的功能使得新信息能够迅速上线，减少了因等待知识库更新而造成的延迟，确保客户咨询时获得最新、最准确的信息。这不仅提升了客户的满意度，还加强了金融机构对市场变化的快速反应能力，有助于实现更高的业务转化率和客户忠诚度。

智能工单助手

Agent 概述

面向坐席，包括金融机构的运营中心、客服部门、远程银行部，智能工单助手智能整合客服沟通记录，一键生成智能工单。目前，工单多为坐席人工翻阅聊天记录、手动摘录关键信息，再按照固定格式拼凑成工单内容，操作繁琐且极易遗漏重要细节，耗费大量时间精力，填写率不高，沟通记录中的有效信息未能深度挖掘。

智能工单助手采用大语言模型技术，支持接入客服沟通数据，一键生成工单，坐席人员工单生成时间答复缩短，提升坐席工作效率，降低运营成本。

需求分析

面对日益增长的客户服务需求与复杂多样的业务场景。坐席人员在处理客户咨询与业务办理过程中，需要人工填写工单，填写率不高，客服沟通历史沟通记录中的数据未充分挖掘；智能工单助手支持一键填写工单，提升坐席的工作效率，辅助坐席人员更好的跟进事项。

多渠道数据接入

能够无缝对接客服系统、电话录音系统、在线聊天平台等多种渠道的客服沟通记录，确保全面收集客户与坐席之间的交互信息，包括文字、语音、图片等多种形式的的数据。

工单智能填写

将整合后的客服沟通记录信息自动填充到工单模板的相应字段中，无需坐席人员手动录入，大大提高工单生成的速度与准确性。对于一些需要进一步确认或补充的信息，系统会智能提示坐席人员进行核实与完善。

业务系统连接

智能推送到下游业务系统，便于业务人员及时跟进。

典型案例

模力时刻

Qwen-72B

处理 基于客服通话记录，判断工单类型，提取关键槽位信息。

效果 工单模板选择准确率 95%，槽位提取准确率 90%。

大语言模型在客服工单生成中的关键技术流程可以总结如下：

2 工单模板分类识别

- 客服电话记录输入到 Qwen 大模型后，模型首先进行工单模板的分类识别。
- 根据通话内容，模型选择最合适的工单模板（如工单模板 -1、工单模板 -2 等）。

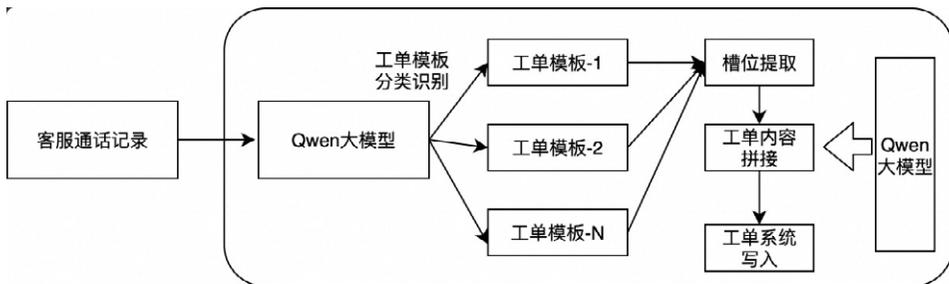
2 槽位提取与内容拼接

- 选定模板后，模型从通话记录中提取关键信息（槽位），填充到模板中。
- 这些信息被自动填充到模板的相应位置，生成完整的工单内容。

3 工单系统写入

- 生成的工单内容通过模型进一步优化和校验后，写入工单系统。
- 工单系统接收并存储这些工单，以便后续处理和跟踪。

下图为在一通客服电话结束后，大模型基于通话记录自动生成工单的流程：



价值分析

在客服工单智能生成中，大模型的业务价值主要体现在以下几个方面：

大幅提升工单生成效率

通过一键自动生成工单，显著减少了坐席人员手动创建工单的时间，从原来的 10-15 分钟缩短至 1-2 分钟；这种效率的提升不仅加快了客户服务速度，也极大地降低了人力成本和运营开支。

优化客户体验与业务流程

大模型能够迅速将客户需求转化为工单，并自动推送至相应的下游业务团队，确保问题得到及时处理；快速响应客户需求，缩短了业务周转时间，提升了客户的满意度和企业的整体运营效率。

智能投诉预警

Agent 概述

在当前企业运营中，客户投诉的及时预警对于维护品牌声誉、提升客户满意度和优化产品服务至关重要。然而，传统的投诉处理系统常常面临诸多痛点，如数据量大且结构复杂、实时性差、分类不准确等。借助大模型，可以有效解决这些痛点，满足投诉预警的需求。

大语言模型在投诉预警场景下发挥着重要作用，其强大的自然语言处理能力使其能够理解和处理各种语言表达，适应多样化的投诉内容。同时，该模型具备高效的数据处理能力，可以快速处理海量数据，实现实时分析和预警，确保企业能够在第一时间采取行动。此外，通过深度学习技术，大模型还能进行精准的情感分析与分类，准确识别情感倾向并分类不同类型的投诉，帮助企业更好地理解客户情绪，及时有效地解决问题，从而提升客户满意度和忠诚度。这些功能共同作用，不仅提高了企业的响应速度和处理效率，也为维护企业形象和预防潜在风险提供了有力支持。

需求分析

投诉预警的主要痛点：

数据量庞大且多样化

客户投诉数据来源广泛，包括社交媒体、客服热线、电子邮件、在线评价等，数据量巨大且格式多样，处理难度大。

实时性要求高

及时发现潜在的危机性投诉，快速响应，以防止负面情绪蔓延。

分类与识别困难

投诉内容复杂多样，涉及不同的产品、服务和问题，人工分类耗时且易出错。

典型案例

模力时刻

Qwen-72B

处理 基于用户通话内容，感知客户情绪，投诉提前预警。

效果 情绪识别准确率 85%，预警正确率 90%。

某金融机构使用大模型在投诉预警场景，进行客户的通话内容进行风险预警和策略应对，降低潜在的监管投诉和舆情风险。使用大模型针对每通电话的交流情况进行分析，打标、风险评估，并与风控的 SOP 流程打通，加速业务的推

动效率和降低风险。

通常采用的技术实现方案是 LLM+Prompt+RAG。在 Prompt 提示词当中，明确列出需要大模型识别的情绪类别、语言风格、行动倾向等，LLM 会基于其强大的语言理解能力，从海量会话记录中进行识别与打标；另外在企业知识库中可以存储一些关键词库、历史风险话术等，使用 RAG 检索召回，输入给 LLM 进行识别与判断。

```

{
  "description": "高压线-客户服务对话中的风险防控问题检测",
  "response": {
    "analysis_process": "客户表达了强烈的不满并声称要'去报警'，这通常意味着客户对当前的解决方案或沟通方式不满意，并可能采取进一步的行动。客服[ ]虽然重申了公司的还款政策，但未针对客户的报警声明做出适当回应，没有表现出对客户情绪的理解或提供进一步的帮助来缓解客户的情绪，也没有提及会向上级或特殊部门反馈此情况，这可能遗漏了风险防控的步骤。",
    "original_text": [
      {
        "speaker": "客户",
        "start_time": "2024-05-23 14:58:46",
        "text": "我要去报警"
      },
      {
        "speaker": "[ ]",
        "start_time": "2024-05-23 14:59:10",
        "text": "您好，[ ]借款是不支持协商/延期还款，建议您按照系统展示的还款时间和金额及时还款。如果暂时资金不足，可以找家人朋友周转一下。"
      }
    ],
    "violation": "1",
    "violation_explanation": "客户表达的报警意向是一种风险信号，表明客户满意度极
    
```

如上图所示，是大模型产出的一段分析识别结果，明确指出了风险信号，并提示客服人员要及时介入并处理。如果是在线业务，可以直接在客服工作台上进行提示；如果是离线挖掘，可以生成对应工单给到客服人员跟进处理。

价值分析

智能投诉辅助通过借助大模型的强大数据处理和分析能力，企业可以有效提升投诉预警系统的精准性和实时性，及时应对客户反馈，优化产品和服务，维护品牌形象。实施过程中需注重数据质量、模型选择与优化，并结合实际业务需求，构建适应性的预警体系，以实现持续改进和业务增长。

☑ 提升响应速度与决策质量

通过即时分析大量客户反馈，企业能够在问题初期就做出快速反应，减少负面信息扩散的风险。大模型提供的深入洞察帮助管理层制定基于数据驱动的决策，增强解决问题的有效性。

☑ 优化客户服务体验

准确的情感分析和分类功能使得企业能够更好地理解客户需求和情绪，提供更加个性化的服务解决方案。这种对客户需求的高度敏感促进了客户满意度的提升，进而增强了客户的忠诚度。

🔊 促进产品和服务创新

持续收集并分析来自客户投诉的数据，有助于发现现有产品和服务中的不足之处，为企业的产品改进和服务优化提供依据。这种基于用户反馈的迭代过程推动了企业的创新发展，提升了市场竞争力。

智能语音分析

Agent 概述

在客服领域，通话记录分析是一项关键活动，它通过对客服与客户之间的对话进行录音转写和数据处理，帮助企业评估服务质量、监控业务流程效率以及洞察客户需求。这项工作不仅涉及到对单个通话的详细审查，还包括跨多个通话的数据挖掘，以发现趋势、识别常见问题及衡量客户满意度。

智能语音分析 Agent，使用大语言模型技术，极大地增强了通话记录分析的能力。首先，通过深度学习算法，大模型能够高效地将语音转化为文本，并在此基础上实现精准的语义理解和上下文感知。这使得系统可以自动检测出服务态度、流程合规性和回答准确性等方面的问题。其次，利用大数据分析手段，大模型支持对海量通话记录进行全文检索和复杂规则组合分析，从而生成详尽的质量报告。这些增强功能不仅提高了质检工作的效率和覆盖率，还确保了评估标准的一致性，为企业提供了更加科学、客观的决策依据。

需求分析

大型企业机构的客服热线每天约处理数万个通话，而质检人员对这些通话录音只能通过人工测听的方式进行抽检，测听抽检比率低，质检工作量大、效率低且覆盖率低，难以有效评价整体服务质量。与此同时，目前的经营分析系统缺乏对客服热线录音数据的处理和分析，而录音数据能直接反映客服热线的效率和质量以及客户的满意度和忠诚度。所以当前的经营分析系统中，客服质量分析难以实现有效监控客户服务的效率和质量，无法为经营决策提供最直接的、最有效的参考依据。

随着大模型技术的出现，国内语音分析产品的发展，结合业务实际需求，引入智能语音分析 Agent，根据质检规则实现坐席通话录音的质量检测，成为企业提高坐席综合素质必不可少的手段之一。语音的全文检索、质检规则的多样化组合分析、全面的质检规则和准确的质检结果反馈，能够针对坐席的通话录音提供详细的分析和准确的结果反馈，以便实现对坐席的通话质量实行宏观的把控，并对企业的业务经营分析提供数据依据。

典型案例

模力时刻

Qwen-72B

处理 基于客服通话记录，基于提示词规则集检测客服人员工作质量，识别客户满意度。

效果 客服话术质检准确率 90%，客户满意度识别准确率 95%。

智能语音分析 Agent，可应用于银行、证券、保险、互金等多个行业的呼叫中心，对不同业务进行语音质检分析，分析服务的规范性，了解客户服务情况，挖掘服务数据的价值。

简要流程描述如下

1 分析系统

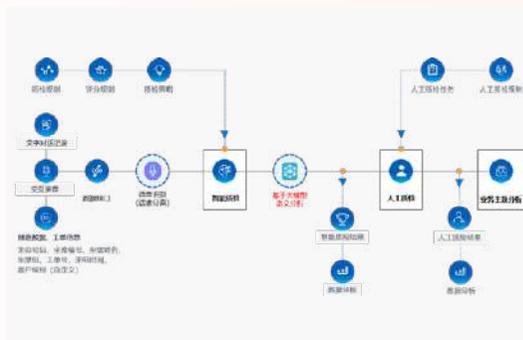
业务系统通过系统对接，将随路和录音文件传进语音分析系统。

2 属性信息

语音分析系统的语音识别引擎对录音进行转写，将录音转写为文本，同时得到录音属性信息。

3 质检结果

语音分析系统的大数据分析引擎结合大模型根据随路数据和质检规则进行质检，得到质检结果。



4 分析结果

语音分析系统的大数据分析引擎结合大模型对转写文本和质检结果，进行大数据分析，得到数据分析的结果。

价值分析

通过引入大模型技术，智能语音分析 Agent 显著提升了客服语音分析的能力，并为企业带来了以下业务价值：

数据结构化与深度洞察

利用大模型的自然语言处理能力，智能语音分析 Agent 能够将海量的客服电话录音高效、精准地转化为文本格式。这不仅使得非结构化的音频信息变得易于检索和分析，而且还能挖掘出深层次的信息，如情感倾向、意图识别等，从而为企业的数据分析提供了坚实的基础。

智能化质检与效率飞跃

大模型赋能的语音分析系统实现了对全量通话记录的自动化质检，极大地提高了质检的覆盖率和准确性。该系统基于预设的质检规则，可以快速筛选出潜在的问题通话，并提供详细的文本分析报告。质检人员可以依赖这些智能分析结果进行二次审核，或直接查看转写文本以加快质检流程，从而显著提升工作效率。

服务品质优化与客户满意度提升

通过智能语音分析 Agent 的应用，企业能够实时监控和评估客服团队的表现，确保服务质量达到高标准。大模型支持的情感分析功能可以帮助识别出客户的情绪变化，使企业能够及时采取措施解决客户问题，增强客户体验。此外，通过对优秀服务案例的学习，企业还可以不断改进培训计划，促进员工技能的发展，进而提高整体服务水平。

商业智能与战略决策支持

智能语音分析 Agent 利用大模型技术深入解析大量的客服对话数据，为企业提供有价值的商业见解。无论是市场趋势预测、客户需求洞察还是产品改进建议，都能从详尽的数据分析中获得。这种基于大数据的洞察力有助于企业制定更加精准的市场营销策略和产品发展规划，最终推动业务增长和竞争力的提升。

电销助手

Agent 概述

在金融领域，电话机器人被广泛应用在多个场景中，如营销推广、客户服务和信贷催收。在营销方面，电话机器人能够自动拨打潜在客户的电话，根据预设的脚本进行产品介绍和优惠活动通知；客户服务方面，它们可以提供即时响应，解答客户关于账户信息、交易状态或产品服务的疑问，确保客户获得及时的帮助和支持，增强客户满意度。这些自动化操作不仅提升了工作效率，还保证了服务流程的一致性和规范性。

电销助手 Agent，融合了通义大语言模型、语音识别与语音合成等多项大模型技术，具备强大的数据处理和交互能力，它能够理解人类语言，模拟真实对话场景，完成电话接听、信息推送、客户咨询解答、意向筛选等一系列任务。

需求分析

全天候、不间断的客户支持服务，确保服务的时效性和质量

理解与响应客户需求

要求对话体验接近自然人类的交流，在对话过程中，如果客户提出问题或打断话语，机器人能够实时调整其对话策略，灵活应对各种复杂情况，保证对话的连贯性和效率。

效率与客户体验提升

根据不同客户的需求提供定制化的沟通和服务，从而增强顾客的满意度和忠诚度。

持续的优化

不断优化其对话模型，以更好地适应不同客户的需求和市场变化。

金融行业合规

如何在复杂的消费金融场景中准确识别违规行为？如数据泄密、人工客服不可控。

典型案例

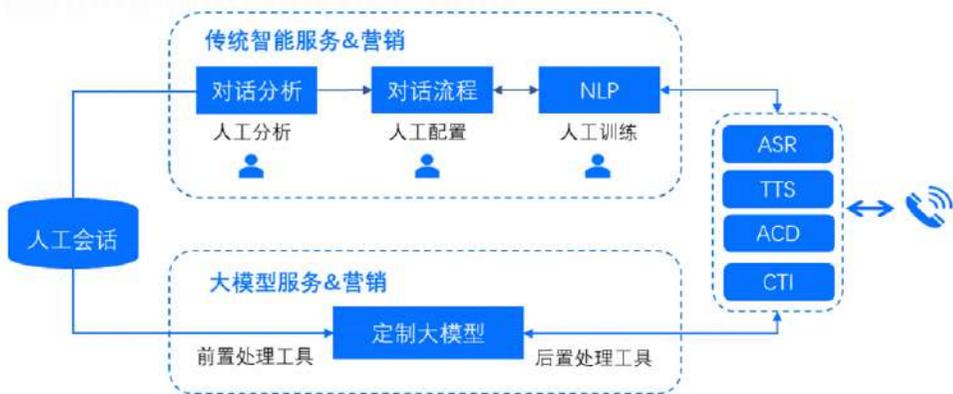
模力时刻

Qwen-72B

处理 基于用户历史通话记录，以及画像标签，生成个性化外呼话术。

效果 营销外呼接通率提升 30%。

某知名银行信用卡中心面临客户增长瓶颈，传统营销方式效率低下且成本高昂。智能外呼机器人被引入后，采用基于大模型的人机对话系统架构，实现意图识别、意图预测、用户画像、上下文理解等核心能力，为客户提供“千人千面”的精准智能服务。



通过外呼机器人智能批量外呼，针对潜在用户群体进行个性化话术沟通。凭借语音识别与自然语言处理能力，外呼机器人不仅完成了大量初步筛选工作，还准确识别出客户的潜在需求与意向等级。对于高意向客户，机器人即时流转至人工坐席，确保跟进的时效性和专业性。短短数月内，该银行信用卡新增激活量显著提升。

价值分析

📞 效筛选意向客户

自动拨打大量电话，对潜在客户进行快速筛选。通过大模型精准识别客户意向，筛选出高意向客户，为后续跟进提供有力支持。这一功能极大地降低了人工筛选的成本和时间，提高了营销效率。

👤 个性化营销沟通

机器人能够根据客户的画像和历史数据，进行个性化话术沟通。例如，针对银行信用卡中心，机器人可以分析历史客户数据，构建精准用户画像，然后智能批量外呼，针对潜在用户群体进行个性化营销。这种个性化的沟通方式不仅提高了客户的满意度，营销效果提升 30%。

🗣️ 提升客户体验

智能外呼机器人支持多种语言自由切换，能够跨越语言障碍，为全球客户提供无障碍的智能客服体验。同时，机器人还能够将通话内容转化为文字，方便人工进一步了解查看。这种便捷、高效的沟通方式不仅提升了客户体验，还增强了企业的全球用户触达能力，平均通话时长相比传统机器人提升 100%，交互轮次相比传统机器人提升 83%，客户满意度提升 28%。

客户经理绩效助手

Agent 概述

各分支行、员工的绩效关注点各不相同，想查询超出标准报表的指标需要向科技人员提出书面需求，并在取数过程中与科技人员反复核对数据口径、取数逻辑。业务人员很难及时获取数据，科技人员也要付出大量精力以应对。

基于大模型的智能问数助手让业务人员无需掌握编程技能，通过自然语言对话即可查询出数据库中的数据。也可通过自然语言对话的方式自助生成报表，美化样式，添加查询，调整数据，辅助分析。

需求分析

面对业务部门繁多的取数需求，科技人员疲于应对

业务部门会针对自己的绩效指标细节产生大量取数需求，需要有权限并掌握编程技能的科技人员配合抽取。导致大量科技人员需要挤出时间完成取数

取数口径理解差异，反复抽取核对

业务人员与科技人员对数据的理解存在差异，导致数据抽取结果偏离业务人员预期，需要双方反复核对口径并取数校验。

业务人员需要依赖科技人员制作报表

将频繁使用的指标构建固定格式报表，是解决频繁取数的有效方法。但业务人员只有在科技人员的支持下才能创建报表对外发布，报表的修改也需科技人员支持。

对业务数据的分析依赖业务人员经验

只有有经验的业务人员才能做好数据分析工作，新员工缺乏分析助手辅助。

典型案例

模力时刻

Qwen2.5-32b+Lora

处理 理解客户意图，将客户提问转写为绩效指标查询语句。

效果 意图识别准确率 90% 以上，指标查询准确率 80% 以上。

QwQ-32b

处理 理解客户意图，通过推理逻辑解为客户分析绩效指标问题及提升建议。

效果 意图识别准确率 90% 以上，回答内容采纳率 50% 以上。

某银行基于通义大模型及 QuickBI，使用智能小 Q 功能实现了基层员工通过自然语言对话的方式每日自主问询业绩指标。并通过自然语言对话的方式搭建各分行绩效报表，并辅助生成分析报告。



1 员工自主业绩查询

一线员工每日自主问业绩，时效可达 T+1，粒度可至末级产品，范围涵盖业务规模、FTP 利润/EVA、风险。

2 分行业绩报表自主搭建

在标准报表模板的基础上，分行管理人员可根据自己分行特点与需求通过对话方式调整数据元素、图表对象及报表格式

3 归因分析辅助

对各项数据都内嵌常见归因分析，辅助一线业务人员快速定位问题。

价值分析

☑ 业务人员自主问数，无需通过科技人员

业务人员以自然语言的方式向大模型问数用数。交互过程中，即便是金融专业知识、行话俗语也可被准确理解并生成相应的取数动作。取数用数无需再提请取数流程，无需求助于科技人员编写 SQL 程序。

👤 大模型辅助归因分析，让新员工快速上手

客户经理绩效问数 agent 内置了金融行业常见的绩效数据分析场景，可快速应用于归因分析、客户画像分析、客户商机挖掘等业务场景。使得新员工也可快速上手过去需要具备专家经验的工作。

☑ 业务人员自主掌控取数口径，无需与科技人员反复核对

基于大模型的多轮对话能力，业务人员可在问数过程中边取数边查看样本数据，逐步确认数据口径、约束条件直至最终确定准确的数据范围。无需再与科技人员反复解释核对。

☑ 业务人员自主创建报表，无需依赖科技人员制作

通过大模型的智能报表搭建 agent，业务人员可以自然语言的方式指导 agent 搭建数据分析报表并发布。无需通过科技人员或自行学习 BI 报表技能。

智能问数助手

Agent 概述

金融行业经过多年的从底层技术到上层应用到数据基础设施建设，基本已经实现了数据统一建设及管理。如何让业务人员更好的自主完成数据分析和应用是金融行业下一阶段的重点探索的目标，业务人员可以按照个人的分析意图准确的找到数据、聚合数据、分析数据，有效改变传统的业务与科技之间过度依赖“要数工单”的数据应用结果传递模式。

智能小 Q 是基于大模型打造自然语言对话式数据分析智能应用。让业务人员无需掌握编程技能，通过自然语言对话的方式便可自主获取数据及数据归因分析，实现数据应用的即问即得。为用户带来全新的数据分析、数据洞察的操作体验，极大提升数据分析效率。

需求分析

数据智能分析

借助智能化数据分析工具功能仅需要将数据应用需求通过自然语言输入的方式能够自动完成数据计算代码生成、数据计算以及应用结果可视化呈现，并且能够对数据结果进行提炼总结，以一句话的形式表达出智能数据分析结果的主题答案。

驾驶舱补充问答

驾驶舱作为金融行业经营管理的重要基础设施，智能数据分析应该能够深度集成到驾驶舱，管理人员在查看驾驶舱内的指标卡片过程中，如果对某些指标存在疑问或者临时想查看某个关键指标，可以通过对话的方式完成数据指标即时答复及呈现。

数据特征洞察

对于数据分析的结果能够进行深层次的洞察以及归因分析，尤其在数据出现大幅波动的情况下能够智能识别出造成波动的根本原因。

典型案例

模力时刻

Qwen2.5-32b+Lora

处理 理解客户意图，将客户提问转写为数据查询语句。

效果 意图识别准确率 90% 以上，数据查询准确率 80% 以上。

某银行基于智能小 Q 构建智能问答的数据分析体系，围绕银行资产负债管理、分支行经营赛、绩效考核马等场景实现面向各机构领导的智能数据应用模式。同时，结合 QuickBI 可视化仪表盘的能力打造面向领导驾驶舱的智能问数模式，充分满足行方将传统驾驶舱与 AI 数据智能分析有效结合的应用体系。



应用效果:

1 智能问数分析

基于行内多种场景的数据集面向管理部门以及业务团队进行灵活的对话式数据分析。

对于用户提出的问题、智能小 Q 结合通义大模型完成 NL2SQL 解析转换，并对结果数据进行提炼总结进一步提升数据面向业务人员对可理解性。

2 驾驶舱智能分析

将智能问数的能力嵌入到银行的经营驾驶舱内。

在经营驾驶舱内无法穷举所有指标卡片并进行可视化展现，因此将智能问数嵌入到驾驶舱内，能够对于没有呈现出来的指标进行快速的即问即答，提升驾驶舱数据内容服务的全面性。



3 数据洞察分析

报表数据洞察异常识别：面向驾驶舱一键开展数据洞察以及异常点分析，对于存在大幅异常波动的指标能够进行逐层下钻找到造成异常的根本原因。

智能问数波动归因分析：支持对问数的结果一键开展波动归因分析快速找到造成波动的主要原因，业务人员可以按照需求能够灵活定义归因的时间范围、拆解维度等信息，满足个性化数据归因分析。



价值分析

通过智能问数，银行行业可以更高效地利用数据资源，提升对数据价值的洞察力，释放数据的潜在价值。

提升效率

通过自然语言交互，快速获取数据洞察，减少人工查询和分析的时间。

降低门槛

非技术用户也能轻松完成灵活、自主的数据分析。

加快决策

智能问数可以通过对大量复杂数据的快速分析、异常洞察，为决策提供支持。

智能搭建助手

Agent 概述

金融机构已构建了企业级报表平台，并开发了庞大的报表体系。但一线分支机构、员工对报表的需求多种多样，想创建自己的定制报表需要向科技人员提出书面需求，并在报表搭建过程中与科技人员反复沟通核对数据口径、加工逻辑、报表样式。耗时耗力，牵扯了业务、科技人员大量精力

基于大模型的智能搭建助手让业务人员无需掌握 BI 报表搭建技能，通过自然语言对话即可自助生成报表，制作数据图表，美化样式，添加查询。

需求分析

面对业务部门繁多的报表需求，科技人员疲于应对

业务部门在生产经营过程中会产生大量取数、报表搭建需求，需要有权限并掌握编程技能的科技人员配合抽取。导致大量科技人员需要挤出时间配合完成。

数据、报表理解差异，反复核对沟通

业务人员与科技人员对数据、业务逻辑、报表的理解存在差异，导致最终结果经常偏离业务人员预期，需要双方反复核对沟通。

业务人员需要依赖科技人员维护报表

业务人员只有在科技人员的支持下才能创建报表对外发布，报表的修改维护也需科技人员支持。导致报表更新时效慢。

典型案例

模力时刻

Qwen2.5-32b+Lora

处理 理解客户意图，将客户提问转写为报表构建指令。

效果 意图识别准确率 90% 以上，报表构建行动准确率 80% 以上。

某银行基于通义大模型及 QuickBI，使用智能小 Q 功能实现了业务人员通过自然语言对话的方式方式搭建各分行绩效报表，并辅助生成分析报告。



1 大幅降低定制报表门槛

业务人员通过自然语言对话的方式即可完成报表搭建，无需掌握编程技能也无需技术人员支持。

2 业务人员自主把控调整报表

在标准报表模板的基础上，分行管理人员可根据自己分行特点与需求通过对话方式调整数据元素、图表对象及报表格式。

3 归因分析辅助

对各项数据都内嵌常见归因分析，辅助一线业务人员快速定位问题。

价值分析

👉 业务人员自主搭建，无需通过科技人员

业务人员以自然语言的方式搭建定制报表。交互过程中，即便是金融专业知识、行话也可被准确理解并生成相应的动作。取数用数无需再提请取数流程，无需求助于科技人员编写代码。

👉 业务人员自主掌控报表逻辑，无需与科技人员反复核对

基于大模型的多轮对话能力，业务人员可在报表搭建过程中边取数边查看样本数据，逐步确认数据口径、约束条件、计算逻辑直至最终确定报表样式。无需再与科技人员反复解释核对。

👉 业务人员自主维护报表，无需依赖科技人员制作

业务人员可以自然语言的方式指导 agent 修改已有报表并重新发布。无需通过科技人员即可即时维护。

数据资产问答

Agent 概述

在大数据时代，金融客户在数据分析的过程中往往面临找数难、用数难的问题。业务人员不了解数据资产结构，往往需要数据技术人员的支持，按照 IT 建设的模式提出数据加工需求或者取数需求。

在大模型技术的支撑下，智能小 D 将大模型、数据资产、业务人员进行融合组织，将数据检索过程从“被动查询”转变为“主动发现”，让业务人员通过对话方式找表、找指标、找 API、知晓表使用场景及使用方法等。

需求分析

现有的数据分析场景中，往往以被动支持的方式满足业务需求，这个过程存在业务数据支撑耗时较长，投入成本高等问题。

业务人员数据能力缺

业务人员往往难以准确评估数据的完整性、一致性和可靠性，缺乏对数据采集逻辑、清洗规则及潜在偏差的识别能力。同时面对技术术语（如数据模型、字段定义）时，业务人员难以将其映射到实际业务场景，导致需求表达模糊或分析结论偏离业务目标。

找数难、用数难

科技当前提供的数据资产服务，查找效率低，理解成本高，定位数据难；同时传统架构依赖单一技术元数据（如字段类型、存储位置），缺乏高效的检索能力和语义关联分析，导致跨源数据查询响应慢、资源消耗高。

线下互动，时间周期长

业务人员难以直接使用数据，过度依赖技术人员解读数据逻辑，形成沟通瓶颈，每次需求需协调数据工程师、分析师等多角色，平均耗时 2-3 天，且易因需求传递偏差导致返工。

典型案例

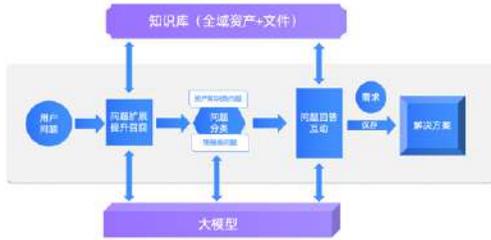
模力时刻

Qwen2.5-32b+Lora

处理 理解客户意图，将客户提问转写为数据知识查询指令。

效果 意图识别准确率 85% 以上，数据知识回答准确率 75% 以上。

在金融客户，基于通义千问大模型，并结合全域数据资产、文件构建成企业级的数据资产知识库，同时进行专家微调，最后将大模型产出内容输出到 ChatBot 里面，实现数据资产的智能问答。总体技术流程如下图所示：



1 用户问题输入

系统识别策略类问题还是资产问答类问题。例如 1、我是业务负责人，今年营销活动目标是 MAU 提升 20%，数据能帮我做什么？ 2、我是产品运营人员，我想了解**产品的销量指标定义是什么？

2 大模型推理

基于通义千问基模，以 RAG 形式进行推理。

3 全域数据资产整合

- 1、整合企业内核心知识资产，包含业务模型、业务流程、组织架构，元数据模型、专有知识、词汇、术语等。
- 2、整合企业所有结构化、非结构化数据，包含结构化：表、指标、标签、报表、数据集、API。文档：PPT、Word、Excel、Txt、PDF。媒体：音视频等。

基于大模型高效实现数据资产问答，满足业务人员找表、找指标、找 API、问表使用场景及使用方法等需求场景，找数、用数上极大提升了效率。



价值分析



提高理解效率

使用数据资产问答助手，能够对数据资产进行多维度的刻画，并丰富业务属性，加速业务人员对数据信息的吸收，提供数据的确信与理解能力。

提效降本

业务人员可借助数据资产问答助手，高效查找到自己业务想用的数据资产以及解决方案，业务人员也能够更加自主的进行数据的分析，极大的提升数据分析的效率，同时也释放部分科技人员解答数据问题的成本。

存款分析助手

Agent 概述

存款增长及成本控制是银行经营中的重要工作。分支行管理人员及一线人员都希望能有银行经营分析方面的专家每天针对本行 / 自己的存款业务进行分析，指出当前面临的问题及潜在商机。

银行存款分析助手智能体便是基于阿里云通义系列大语言模型，将银行经营分析专家知识进行编排，形成银行存款经营的数个典型分析方案。通过交互对话的方式，辅助引导银行人员存款分析的工具。

需求分析

银行存款增长及波动分析，存款商机挖掘需要有经验的业务人员耗费较多精力进行专项分析，而且分析结果通常会因人的经验差异而不同。

依赖经验人员

银行存款分析商机挖掘是一项十分依赖专家经验的工作。培养一名能从数千乃至上万数据信息中抓住关键信息洞察存款业务中隐藏的问题，进一步分析出潜在客户商机的数据分析人员需要数年时间。合格的分析专家在每一家银行都是稀缺人才。

人工分析效率低

存款业务涉及产品、客户、地域、市场利率、同业竞品等多方面，各项数据少则万成千上万条，多则过亿。且每家银行各系统记录的数据都是散落于各个角落，需要经验人员把数据归集构建，然后靠人工分析将耗费巨大的人力资源及时间。

难以统一标准化

分析人员对当前业务状况的判断结果很大程度上会受其个人经验、过往案例的影响，甚至分析人员工作时刻的状态也会导致结果差异。

典型案例

模力时刻

QwQ-32b

处理 理解客户意图，通过推理逻辑解答客户问题。

效果 意图识别准确率 90% 以上，回答内容采纳率 70% 以上。

银行存款分析助手是构建于银行现有的数据集市之上。通过 agent 编排能力把传统分析模型、报表与大模型分析 agent 相结合，让银行现有数据、模型继续发挥能力。同时通过调用大模型能力对现有数据分析进行补充，获得更优效果。

1 数据层

数据层依托于银行已构建的核心交易、客户、产品、管会、区域市场统计等数据集市，并对存量数据集市做适当修改以适应大模型数据处理量大但算力需求也十分突出的特点。

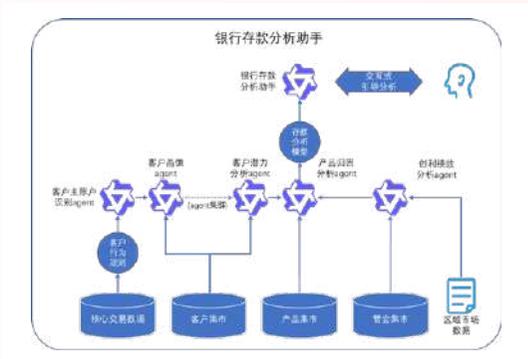
2 分析层

分析层是对数据挖掘、分析的主要层级。首先对银行已构建的模型代码、分析报表进行检视，评估现有分析方法的待优化部分，其中可通过大模型提升能力的则构建相应的 agent。如客户画像 agent、主账户流出追踪 agent 等。通过将传统分析方法与大模型相结合，提高分析的准确性并拓宽覆盖面与深度。

3 应用层

应用层负责将存款业务分析思路、结果向一线业务人员传授。存款分析是一项较为专业的工作，很难要求所有一线业务人员都熟练掌握。通过存款分析助手 agent 相关的流程编排，引导一线业务人员的用数及分析思路，让分析结果能被相对容易理解并应用。

某银行基于通义大模型，使用点金平台工具搭建了银行存款分析 agent。实现了在十多分钟内完成一家分支机构的存款分析并生成报告。该 agent 将全行数据分析专家共同总结出的分析思路及流程步骤通过智能体编排的形式固化在了 agent 流程中，使得任何人都可以获得专家经验支持，自主操作便可获得标准化结果。



价值分析

专家经验总结沉淀

通过大模型 agent 编排构建的过程，是把专家经验总结固化沉淀在计算机中的过程。其中每一个大模型 agent 并非是替代传统小模型、分析过程、业务数据，而是通过大模型的能力拓展存量模型与数据的智慧边界。大模型 agent 让过往沉淀的专家经验得到沉淀与传承，为存量科技能力注入新的活力。

高效分析

通过大模型 agent 编排批量处理海量的数据信息，避免了过往靠专家人工判读分析的漫长过程。实现高效的存款业务分析。

标准化分析

通过大模型 agent 编排使得银行存款业务分析成为标准化的流程，避免了人为评估业务状态，挖掘客户需求商机时的随机性与偏差。使得全行所有机构存款业务的评价都沿用统一的标准及体系流程。

贷款分析助手

Agent 概述

贷款业务是银行经营的核心业务。对贷款的分析与管理是规模、风险、效益多方面平衡的艺术，只有具备丰富经验的专家才能胜任该项工作。如果能把银行贷款专家的经验及典型案例总结归纳为标准化分析典范并让一线人员都能得到指导和帮助，将会大幅提升银行一线贷款人员的业务素质。

银行贷款分析助手智能体是基于阿里云通义系列大语言模型，将银行贷款专家知识进行编排，形成银行贷款分析的数个典型模式。通过交互对话的方式，辅助引导银行人员贷款分析的工具。

需求分析

银行贷款业务经营分析需要经验丰富的专家耗费较多精力进行专项分析，而且分析结果通常会因人的经验差异而不同。

依赖经验人员

贷款经营分析及商机挖掘是一项十分依赖专家经验的工作。培养一名能从成千上万数据信息中抓住关键信息，能从数据中洞察业务问题、挖掘潜客，并能将客户需求与本行产品匹配，能平衡风险与收益的贷款经营专家是漫长且可遇不可求的过程。经验丰富的贷款专家在每一家银行都是稀缺人才。

人工分析效率低

贷款业务涉及客户与业务、风险与效益多方面。相关的各项数据少则数百项，多则过万。每家银行各系统记录的数据都是散落于各个角落，需要经验人员把数据归集构建，然后靠人工分析将耗费巨大的人力资源及时间。

难以统一标准化

分析人员对客户的判断结果很大程度上会受其个人经验、过往案例的影响，甚至分析人员当时的状态也会导致结果差异。

典型案例

模力时刻

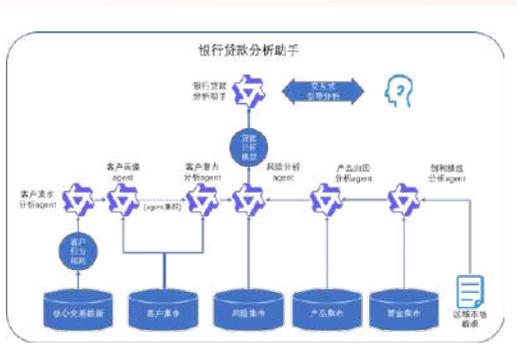
QwQ-32b

处理 理解客户意图，通过推理逻辑解答客户问题。

效果 意图识别准确率 90% 以上，回答内容采纳率 70% 以上。

某银行基于通义大模型，使用点金平台工具搭建了贷款经营分析 agent。实现了在十余分钟内完成一家分支行的贷款业务分析，并返回指标结果及生成报告。该 agent 将全行有经验贷款专家共同总结出的分析思路与典型问题方案通过编排的形式固化在了 agent 流程中，使得任何人都可以获得专家经验支持，自主操作便可获得标准化结果。

银行贷款分析助手是构建于银行现有的数据集市之上。通过 agent 编排能力用领域专家经验设计的分析流程与范式把传统分析代码、模型与大模型分析 agent 相结合，让银行现有数据、模型继续发挥能力。同时通过调用大模型能力对现有数据分析进行补充，获得更优效果。



1 数据层

数据层依托于银行已构建的核心交易、客户、风险、产品、管会等数据集市，并对存量数据集市做适当修改以适应大模型数据处理量大但算力需求也十分突出的特点。

3 应用层

应用层负责将贷款业务分析思路、结果向一线业务人员传授。贷款分析是一项较为专业的工作，很难要求所有一线业务人员都熟练掌握。通过贷款分析助手 agent 相关的流程编排，引导一线业务人员的用数及分析思路，让分析结果能被相对容易理解并应用。

2 分析层

分析层是对数据挖掘、分析的主要层级。首先对银行已构建的分析思路流程、模型代码进行检视，评估现有分析方法的待优化部分，其中可通过大模型提升能力的则构建相应的 agent。如，客户流水分析 agent，客户画像 agent 等。通过将传统分析方法与大模型相结合，提高分析的准确性并拓宽覆盖面与深度。

价值分析

专家经验总结沉淀

通过大模型 agent 编排构建的过程，是把专家经验总结固化沉淀在 AI 知识库的过程。其中每一个大模型 agent 并非是替代传统小模型、分析过程、业务数据，而是通过大模型的能力拓展存量模型与数据的智慧边界。大模型 agent 让过往沉淀的专家经验得到沉淀与传承，为存量科技能力注入新的活力。

高效海量处理

通过大模型 agent 编排批量处理海量的业务信息，避免了过往靠专家人工判读分析的漫长过程。实现高效的贷款分析过程。

标准化分析

通过大模型 agent 编排使得贷款分析成为标准化的流程，避免了人为评估业务价值、违约风险，发现客户需求商机时的随机性与偏差。使得全行各机构对业务的评价都沿用统一的标准及体系流程。

客户经营分析助手

Agent 概述

银行客户经营分析助手汇总客户相关数据并深入分析，从客户交易行为、业务结构、波动趋势、创立贡献、交易往来关系圈等方面入手。依据大模型流程编排中内嵌的专家经验路径，以标准化、高效自动化的流程构建客户交易流水画像，分析客户需求特征、偏好、价格敏感性，并判断客户潜力、目前主账户行等商机信息。

需求分析

客户经营分析及商机挖掘需要有经验的业务人员耗费较多精力进行专项分析，而且分析结果通常会因人的经验差异而不同。

依赖经验人员

客户经营分析商机挖掘是一项十分依赖专家经验的工作。培养一名能从数千乃至上万数据信息中抓住关键信息并构建出客户画像，进一步分析出客户需求和洞察商机的客户经理需要数年时间。合格的客户经理在每一家银行都是稀缺人才。

人工分析效率低

客户相关的各项数据少则数百条，多则过万。每家银行各系统记录的数据都是散落于各个角落，需要经验人员把数据归集构建，然后靠人工分析将耗费巨大的人力资源及时间。

难以统一标准化

业务人员对客户的判断结果很大程度上会受其个人经验、过往案例的影响，甚至业务人员当时的状态也会导致结果差异。

典型案例

模力时刻

QwQ-32b

处理 理解客户意图，通过推理逻辑解答客户问题。

效果 意图识别准确率 90% 以上，回答内容采纳率 70% 以上。

某银行基于通义大模型，使用点金平台工具搭建了客户经营分析 agent。实现了在数分钟内完成一个客户的综合分析，并返回结构化打标结果及生成报告。该 agent 将全行有经验客户经理共同总结出的判读经验规则通过编排的形式固化在了 agent 流程中，使得任何人都可以获得专家经验支持，自主操作都可获得标准化结果。

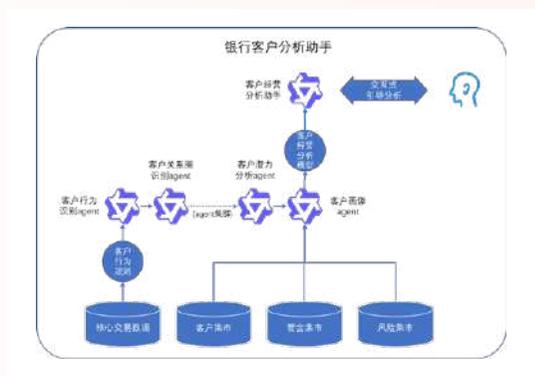
银行客户经营分析助手是构建于银行现有的数据集市之上。通过 agent 编排能力把传统分析代码、模型与大模型分析 agent 相结合，让银行现有数据、模型继续发挥能力。同时通过调用大模型能力对现有数据分析进行补充，获得更优效果。

1 数据层

数据层依托于银行已构建的核心交易、客户、管会、风险等数据集市，并对存量数据集市做适当修改以适应大模型数据处理量大但算力需求也十分突出的特点

2 分析层

分析层是对数据挖掘、分析的主要层级。首先对银行已构建的分析思路流程、模型代码进行检视，评估现有分析方法的待优化部分，其中可通过大模型提升能力的则构建相应的 agent。如，客户行为识别 agent，客户画像 agent 等。通过将传统分析方法与大模型相结合，提高分析的准确性并拓宽覆盖面与深度。



3 应用层

应用层负责将客户画像分析思路、结果向一线业务人员传授。客户画像分析是一项较为专业的工作，很难要求所有一线业务人员都熟练掌握。通过客户经营分析助手 agent 相关的流程编排，引导一线业务人员的用数及分析思路，让分析结果能被相对容易理解并应用。

价值分析

专家经验总结沉淀

通过大模型 agent 编排构建的过程，是把专家经验总结固化沉淀在计算机中的过程。其中每一个大模型 agent 并非是替代传统小模型、分析过程、业务数据，而是通过大模型的能力拓展存量模型与数据的智慧边界。大模型 agent 让过往沉淀的专家经验得到沉淀与传承，为存量科技能力注入新的活力。

高效海量处理

通过大模型 agent 编排批量处理海量的客户信息，避免了过往靠专家人工判断分析的漫长过程。实现高效的海量客户画像构建。

标准化分析

通过大模型 agent 编排使得客户画像分析成为标准化的流程，避免了人为评估客户价值贡献，发现客户需求商机时的随机性与偏差。使得全行全体客户的评价都沿用统一的标准及体系流程。

小微经营户潜在客户挖掘

Agent 概述

近些年，小微企业及零售客户中的经营户、个体工商户成为银行重点拓展的目标群体。这类客户有公司账户与企业个人账户混用，财务与风险需从两个主体通盘评估的特点。一线客户经理需要从大量小微企业、零售经营客群中挖掘出有潜力、风险可控的优质客户。

银行小微经营户潜在客户挖掘助手汇总客户相关数据并深入分析，从客户交易行为、业务结构、波动趋势、创立贡献、交易往来关系圈等方面入手。依据大模型流程编排中内嵌的专家经验路径，以标准化、高效自动化的流程构建客户交易流水画像，分析客户需求特征、偏好、价格敏感性，并判断客户潜力、目前主账户行等商机信息，并从中挖掘出优质的小微经营户潜在客户。

需求分析

优质小微经营户潜在客户挖掘需要有经验的业务人员耗费较多精力进行专项分析，而且分析结果通常会因人的经验差异而不同。

依赖经验人员

客户分析挖掘是一项十分依赖专家经验的工作。培养一名能从数千乃至上万数据信息中抓住关键信息并构建出客户画像，进一步分析出客户需求和洞察商机的客户经理需要数年时间。合格的客户经理在每一家银行都是稀缺人才。

人工分析效率低

客户相关的各项数据少则数百项，多则过万。每家银行各系统记录的数据都是散落于各个角落，需要经验人员把数据归集构建，然后靠人工分析将耗费巨大的人力资源及时间。

难以统一标准化

业务人员对客户的判断结果很大程度上会受其个人经验、过往案例的影响，甚至业务人员当时的状态也会导致结果差异。

典型案例

模力时刻

Qwen2.5-32b

处理 客户交易行为识别并打标。

效果 行为标签准确率 90% 以上。

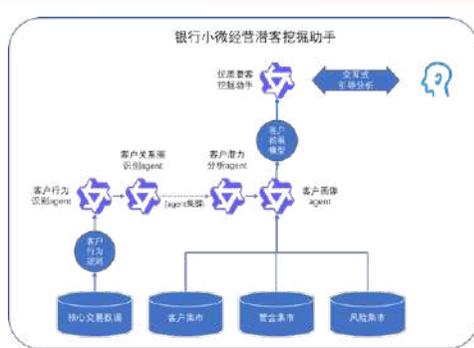
QwQ-32b

处理 通过推理逻辑解分析客群特征并挖掘潜在客户。

效果 潜在客户清单采纳率 50% 以上。

某银行基于通义大模型，使用点金平台工具搭建了小微经营户潜客挖掘助手。实现了在数分钟内完成一个客户的综合分析，并返回结构化打标结果及生成报告。报告主要结果为潜力贷款客户清单、潜力存款 / 理财客户清单、客户他行主账户推断、客户经营关系圈推断等内容。

该 agent 将全行有经验客户经理共同总结出的判读经验规则通过编排的形式固化在了 agent 流程中，使得任何人都可以获得专家经验支持，自主操作都可获得标准化结果。



银行小微经营户潜客挖掘助手是构建于银行现有的数据集市之上。通过 agent 编排能力把传统客户标签画像、客户挖掘模型与大模型分析 agent 相结合，让银行现有数据、模型继续发挥能力。同时通过调用大模型能力对现有数据分析进行补充，获得更优效果。

1 数据层

数据层依托于银行已构建的核心交易、客户、管会、风险等数据集市，并对存量数据集市做适当修改以适应大模型数据处理量大但算力需求也十分突出的特点。

2 分析层

分析层是对数据挖掘、分析的主要层级。首先对银行已构建的分析思路流程、模型代码进行检视，评估现有分析方法的待优化部分，其中可通过大模型提升能力的则构建相应的 agent。如，客户行为识别 agent，客户画像 agent 等。通过将传统分析方法与大模型相结合，提高分析的准确性并拓宽覆盖面与深度。

3 应用层

应用层负责将客户分析过程挖掘结果向一线业务人员交互式传授。客户挖掘分析是一项较为专业的工作，很难要求所有一线业务人员都熟练掌握。通过小微经营户潜客挖掘助手 agent 相关的流程编排，引导一线业务人员的用数及分析思路，让分析结果能被相对容易理解并应用。

价值分析

专家经验总结沉淀

通过大模型 agent 编排构建的过程，是把专家经验总结固化沉淀在 AI 知识库的过程。其中每一个大模型 agent 并非是替代存量数据及传统分析过程，而是通过大模型的能力拓展存量数据与模型的智慧边界。大模型 agent 让过往沉淀的专家经验得到沉淀与传承，为存量科技能力注入新的活力。

高效海量处理

通过大模型 agent 编排批量处理海量的客户信息，避免了过往靠专家人工判读分析的漫长过程。实现高效的海量客户分析挖潜。

标准化分析

通过大模型 agent 编排使得小微经营户潜客挖掘成为标准化的流程，避免了人为评估客户潜在商机时的随机性与偏差。使得全行全体客户的评价都沿用统一的标准及体系流程。

理财潜客挖掘

Agent 概述

理财投资客户是各金融机构全力争夺的客户群体。如何在激烈竞争的市场中挖掘出潜在客户进行精准营销、个性化服务，助力金融机构理财业务降本增效是所有金融机构的关注点。

理财投资户潜客挖掘 agent 借助大模型以及大数据平台，根据金融机构给定的规则、样例和私域知识库，构建全量或请求的客户画像，评分并筛选潜客，提供个性化投资推荐和沟通策略。

需求分析

金融机构在理财投资客户展业时面临诸多痛点：

市场竞争需求

理财市场竞争激烈，拉新成本愈发高昂。但各金融机构在经营中积累的庞大存量客户群体却未被高质量服务覆盖，从存量客户中精细化运营客户，发掘培育优质潜客的工作往往被忽视。

精准营销需求

金融机构间的产品和服务同质化严重，且疲劳轰炸的全量覆盖费时费力，需要更精准高效的营销方式来提升客户转化率。传统的营销方式难以精准定位目标客户，造成大量营销和人力资源浪费。

个性化服务需求

随着投 C 端客户理财意识和知识水平的提高，对个性化投资服务的需求日益增长。

数据驱动决策需求

金融机构均积累了大量的数据，但如何有效利用这些数据进行决策成为难题。

典型案例

模力时刻

Qwen2.5-32b

- 处理** 客户交易行为识别并打标。
- 效果** 行为标签准确率 90% 以上。

QwQ-32b

- 处理** 通过推理逻辑解分析客群特征并挖掘潜客。
- 效果** 潜客清单采纳率 50% 以上。

某金融机构创建 AI 智能体来协助客户经理跟进客户服务情况，并根据客户数据筛选最值得联系的用户名单，列出联系事项、产品推荐和投资理财方案以及对应的话术建议。

效果：

1 盘活存量客户

理财投资户潜在客户挖掘智能体帮助金融机构快速的在客户大盘数据中分析客户并筛选出 10% 增量高质潜在客户。

2 精准识别客户

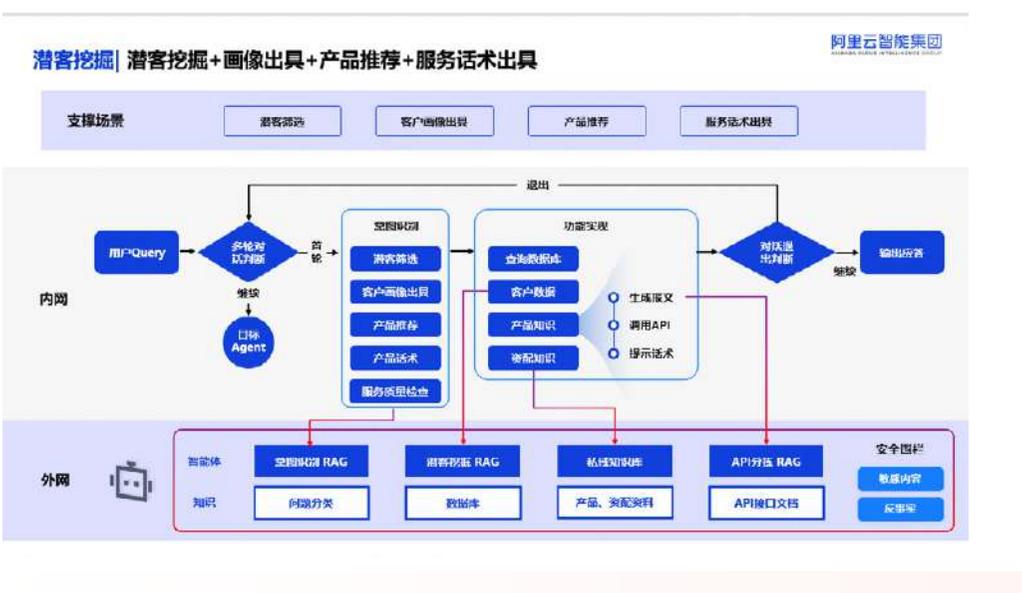
理财投资户潜在客户挖掘智能体通过整合分析海量客户数据，精准识别不同客户群体的投资偏好、风险承受能力和投资目标，实现精准筛选、匹配产品、触达并提供个性化话术提高营销转化率，降低营销成本，让金融机构业务人员提升人效。

3 精准识别客户

理财投资户潜在客户挖掘智能体通过深入分析客户数据，为客户量身定制产品推荐和投资方案，满足客户个性化需求，提升客户满意度和服务体验。

4 基于数据的辅助决策

智能体通过对数据的整合与分析，并打通业务人员的使用界面，为他们提供基于数据和机构私域知识的决策支持，帮助他们制定更科学高效的客户开发和服务策略。



价值分析

业务增长

帮助金融机构挖掘更多潜在客户，提高客户转化率，从而增加业务收入。通过精准的潜在客户挖掘和个性化服务机会提示，吸引更多客户选择本机构的理财服务。

成本降低

减少人工筛选潜在客户的工作量并避免业务人员对于部分客户的遗忘，降低人力成本和时间成本，提高营销效率，避免无效营销投入和打扰客户，实现资源的优化配置。

客户体验提升

基于客户的各类行为数据挖掘其服务倾向，并生成个性化、专业化的投资服务建议，增强客户对金融机构的信任和认可，提升客户口碑，并提高客户的业务转化率和转介绍率，形成良性的业务循环。

风险管理

在潜在客户挖掘和投资推荐过程中，充分考虑客户的风险偏好和承受能力，有效降低投资风险，并在金融机构私域知识和安全护栏内出具合规话术。

代发户促活助手

Agent 概述

代发客户是银行零售客户中重要的主账户潜力客户。但并非所有的代发客户都会将代发行作为自己的日常主账户使用，很多代发户仍停留在代发工具阶段。如何将代发户促活并培养为优质主账户客户是银行客户运营的一项重点工作。

代发客户促活助手基于模型以及大数据平台，在深度分析代发客群数据，构建精准客户画像的基础上，结合银行私域知识库中的产品和过往活动方案，为不同的代发客群客户定制营销策略。实现银行以数据驱动营销、准确结合行内产品与服务，复用过往成功案例，并可以实时进行策略优化的目标。

需求分析

金融机构的在对代发客户群体展业时一般面临以下痛点：

市场竞争需求

金融市场竞争激烈，各银行都在争夺代发客户资源，不仅要拓展新客户，更要留住老客户并促进其活跃。

多样化素材生成

制定创意营销方案后，制作宣传素材成为重资源消耗的任务。

围绕实时热点的特色营销需求

银行代发客群规模庞大、特征多样，且每年需要针对特定热点制定营销策略，但由于热点众多，信息获取不足，依靠个人经验和网络信息的方式效率低下，缺乏快速响应市场变化的能力，难以满足要求。

典型案例

模力时刻

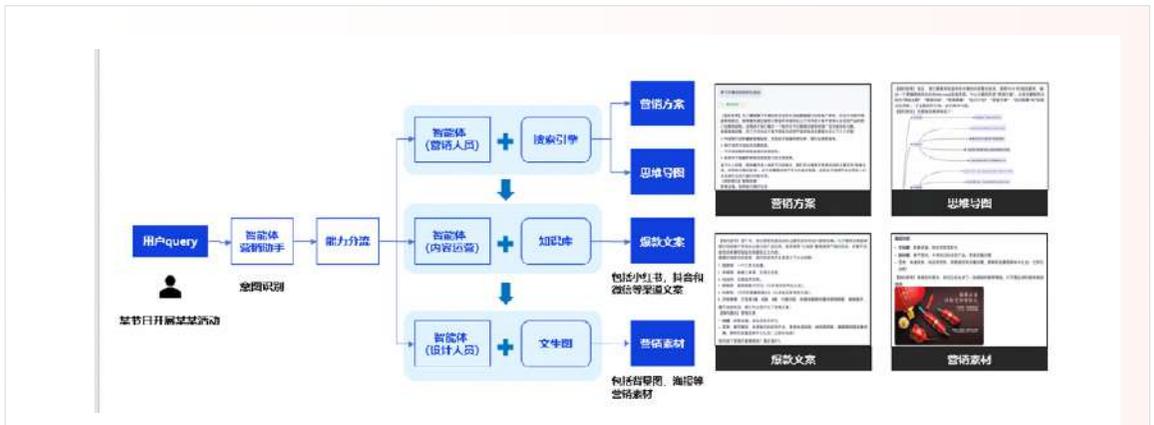
Qwen2.5-32b

- 处理** 客户交易行为识别并打标。
- 效果** 行为标签准确率 90% 以上。

QwQ-32b

- 处理** 通过推理逻辑解分析客群特征并挖掘潜客。
- 效果** 潜客清单采纳率 50% 以上。

某银行创建代发客户促活助手来协助客户经理对代发客群进行分析，并从知识库和互联网资料，根据目标代发客群的画像来出具营销策略，并形成营销方案和海报。缩短了客户经理的工时消耗，并提升了服务客群的精准度。该智能体使用 ReAct 的范式进行设计，可以从用户的反馈和确认信息中优化输出，并反复修改达到用户的预期效果。



1 构建客户画像

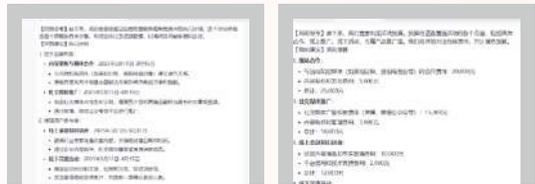
从客户基础属性、交易行为、身份推断、产品偏好、利率敏感性等方面构建客户画像，为客户运营提供基础信息。

2 形成营销方案

根据客户画像、营销思路、当前活动规划出具具体执行方案。包括行动计划、成本估算、预期增长、风险预案处理等。

3 生成营销素材

代发客户促活助手通过对行内代发客群过往的方案学习，互联网热点素材的分析，根据营销方案和目标客户的审美需求，为营销方案快速生成配套素材



价值分析

👤 挖掘存量代发户优质潜客

代发客户促活助手可以帮助银行洞悉客户行为，精准发力，提升客户粘性和忠诚度，在竞争中占据优势。

🔗 辅助营销策略

代发客户促活助手通过精准画像把握、银行私域知识库利用、互联网热点获取和本地化特色营销策略制定，帮助银行实现在不同热点下对不同客户群体的差异化创意营销，提高业务人员的营销效率，并降低营销成本。

👤 辅助客户运营

代发客户促活助手通过辅助生成营销策略级配套素材，助力银行快速启动客户运营活动。在激烈的行业竞争中提升代发客群的活跃度、资金留存率、产品与服务渗透率并深化与目标客群的客情关系。

知识问答

Agent 概述

智能知识助手是专为金融行业设计的智能搜索与问答解决方案，旨在解决海量数据管理和复杂业务流程中的知识获取挑战。通过集成大模型技术、检索增强生成（RAG）技术、思维链（COT）推理能力，以及先进的语义理解与深度语义匹配算法，该助手能够迅速从文档、法规、研究报告等多源异构数据中提取关键信息，并提供精准且上下文相关的答案，充分满足用户对信息精确度和时效性的需求。

基于通用文档解析模型，智能知识助手能够将企业内部大量的非结构化文档转化为数字化资产，并结合大语言模型技术，让用户可以通过自然语言快速而准确地从中获取所需信息。

需求分析

金融机构通常有大量企业财报、行业研究报告、审计报告等各种文档，涵盖 word，PDF，扫描件等各种格式，以及照片、截图等各种图片，这些数据过往由于缺乏有效的处理手段，致使这些海量的数据难以发挥价值；传统工具对自然语言提问理解能力不足，难以精准识别客户需求；数据检索与处理效率低，影响决策的时效性；服务和产品缺乏个性化，无法满足多样化用户需求；以及营销和业务流程中缺少智能辅助支持，导致人力投入大、响应速度慢。

针对这些痛点，基于大模型技术的解决方案，通过强大的语义理解能力和 RAG 检索生成技术，可高效支持复杂场景需求，如实时市场分析、智能投资策略生成、客服话术推荐和精准营销内容制作，助力企业在知识管理、投研投顾和客户服务等方面实现全面升级和创新。

典型案例

模力时刻

Qwen-72B

处理 理解客户查询意图，对查询语句进行改写、扩写，召回相关知识片段，生成问题答案。

效果 查询意图识别准确率 90%，知识召回准确率 90%，问题回答准确率 90%。

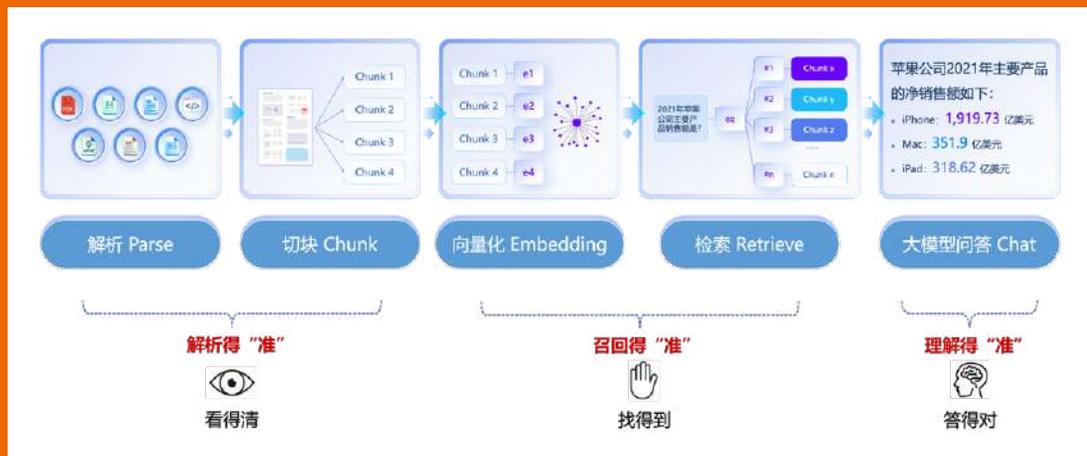
文档解析

处理 精准识别与提取非结构化文档信息，转化为数字资产。

效果 文档结构识别率 90%，关键信息提取率 90%。

企业智能知识助手，在多家企业客户落地，包括券商、银行、基金以及其他垂直行业客户。对于企业客户尤其是金融机构来讲，“准确”是最大的要求，因此，基于大模型的智能知识助手的回答，也要准确。智能知识助手的工作流程如图所示，首先，各种类型的非结构化数据会进行解析，识别里面的元素块，比如标题、段落、表格、图表、图片、页眉页脚等，然后会结合自研文档解析和 OCR 模型，将各类型元素块内容提取出来，输出为 JSON、HTML 等结构化数据，然后进行向量化并存储在向量库中。用户在使用智能知识助手时，系统会根据用

户的问题从向量库中检索相关文档块，发给大模型，最终由大模型把收到的问题和文档块进行理解并生成回答。整个过程可以简单分为“解析”-“召回”-“回答”三个步骤，要想最终准确的回答，三个环节每一步都要精准，解析得准可以比喻为“看得清”，召回得准可以比喻为“找得到”，理解得准可以比喻为“答得对”。



价值分析

基于大模型的企业智能知识助手，可以帮助金融机构实现海量非结构化数据的解析及知识提取，一步实现了非结构化数据的数字化、智能化和资产化。

提升信息获取效率

智能知识助手能够快速从大量的文档、法规、研究报告等多源异构数据中提取关键信息，并提供精准且上下文相关的答案。这大大缩短了用户查找和理解所需信息的时间，提高了工作效率。

优化客户服务体验

无论是内部员工还是外部客户，都可以通过智能知识助手获得即时的支持和解答。对于金融机构而言，这意味着能更有效地响应客户需求，提供个性化的服务建议，从而提高客户满意度和忠诚度。

促进合规管理

在高度监管的金融领域，智能知识助手可以帮助企业自动监测并理解最新的法律法规要求，确保所有操作符合规定，减少违规风险。

强化知识管理与共享

将分散在各种文档中的重要信息进行系统化整理和抽取，有利于构建企业内部的知识库，促进知识的有效积累、管理和传播，增强团队协作和整体竞争力。

增强决策支持能力

通过集成先进的自然语言处理技术和深度语义匹配算法，智能知识助手可以为管理层提供基于最新数据和深入分析的洞察，帮助做出更加明智和及时的决策。通过对行业报告、竞争对手动态等非结构化数据的分析，企业可以获得有价值的市场情报，辅助新产品研发和战略规划，加快对市场变化的响应速度。

复杂信息抽取

Agent 概述

面向金融机构复杂信息抽取的专家智能体，例如合同文件、信批文件、尽调报告、票据表单、报表、凭证等场景；服务金融机构中后台等运营部门，提升数据提取效率和精准度，提高数据资产积累能力和知识沉淀，辅助业务审核校验。

采用多模态大语言模型技术，结合 NLP 小模型双路校验。支持用户可视化配置抽取参数、优化参数抽取效果、提供结果确认 / 编辑功能以及丰富的 API 输出形式。

需求分析

非结构化文档信息抽取面临多重挑战，包括处理文档的多样性与复杂性、克服自然语言的灵活性和模糊性、精确识别实体及其关系、获取足够的标注数据以支持模型训练、有效利用上下文信息理解文本含义，以及确保处理过程中隐私和安全的保护。这些难点要求结合先进的自然语言处理技术、机器学习算法及领域专业知识来解决，以便准确高效地从非结构化数据中提取有价值的信息。

文档格式多样，抽取困难

金融机构中后台业务涉及大量合同、报表、回执单等多种格式文档，传统方法难以适应其复杂性。

人工抽取效率低、易出错

人工从海量文档中抽取参数耗时费力，且易因疲劳、疏忽导致错误。

参数配置复杂，缺乏灵活性

传统 NLP 抽取算法的上线周期长，优化节奏慢。

典型案例

模力时刻

Qwen-VL

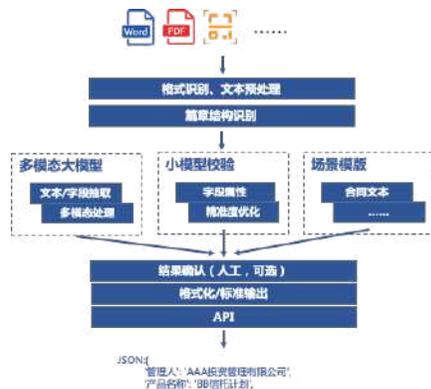
处理 多模态大模型理解文档板式与内容，调用信息抽取、分类等一系列下游任务 API。

效果 复杂信息提取准确率 90%。

某股份制银行资产托管部，托管合同智能抽取，基于大模型能力，支持用户灵活配置抽取参数；针对不同的文档类型，如产品合同、托管协议、补充协议等，用户只需输入简单的参数抽取信息，即可精准抽取所需参数。复杂信息抽取是指利用先进的技术和算法来解析、理解和分析文档内容的一系列过程。它主要涵盖三个核心方面：首

先是解析，即能够处理多种格式的文档，并将其转化为易于计算机理解的结构化形式；其次是理解，这涉及到构建文档的语义表示，使用预训练模型以及多模态大模型来深入理解文档的内容；最后是分析，涵盖了诸如版面识别、信息抽取、分类、问答等一系列下游任务，旨在最大化地发掘文档中的潜在价值。

文档智能技术方向的技术点和细分领域很多，属于 NLP 和 CV 交叉的领域，在文档智能建设过程中相关挑战也很多。复杂信息抽取 Agent，通过大小模型结合对段落文本信息、空间布局信息、表格信息、图像信息等多样化的信息模态智能处理，可以实现精准的信息理解和结构化数据提取。



价值分析

提升处理效率

传统上，处理合同和票据等非结构化文档需要大量的人工操作。通过自动化的信息抽取技术，金融机构可以快速准确地从这些文档中提取所需的信息，极大地提高了工作效率，减少了处理时间。减少了对人力的依赖，降低了因人为疏忽导致的错误风险。这不仅节省了人力资源成本，还提高了数据的准确性，确保了后续分析和决策的基础更加稳固。

增强风险管理能力

合同和票据等非结构化文档中包含着大量的关键信息，通过高效的信息抽取，金融机构能够及时识别潜在的风险因素，比如违约风险或欺诈行为，从而采取预防措施进行有效管理。

加速审批流程

对于业务申请中的相关材料，信息抽取技术可以帮助金融机构更快地完成业务评估和审批过程，金融机构能够提供更高效的服务响应，改善客户体验，增强客户忠诚度。

推动智能化决策制定

基于从非结构化文档中提取的数据，结合大数据分析和人工智能算法，金融机构可以获得深入洞察，辅助做出更加明智的战略决策，例如投资组合优化、市场趋势预测等。

智能写作

Agent 概述

随着数字化转型的加速和信息量的爆炸式增长，企业和组织在文档撰写方面面临着前所未有的挑战。传统的文档撰写方式不仅耗时费力，而且难以保证内容的一致性和准确性，特别是在需要严格遵守法规和行业标准的情况下。与此同时，团队协作的需求日益增加，如何高效地整合多方意见并确保文档版本的清晰管理也成为一大难题。

在这样的背景下，大模型技术为文档撰写带来了革命性的解决方案。其强大的自然语言处理能力、深度语义理解和跨领域知识整合功能，使得自动生成高质量文档草稿、自动化语法检查及风格一致性维护成为可能。同时，大模型还支持复杂的合规性检查，确保所有文档符合最新的法律法规和内部政策要求。利用大模型进行文档撰写已成为现代企业提高生产力、增强竞争力的重要手段。

需求分析

金融机构经常面临需要撰写大量专业报告的场景，金融场景文档创作通常面临以下业务痛点：一是海量历史数据的提取与整合效率低下，手动操作容易导致信息遗漏或误差；二是文档内容的专业性与准确性要求高，需符合行业标准并使用精确术语；三是多场景写作风格的灵活调整较难，难以快速适配个性化需求；四是数据安全与生成内容的可控性不足，存在信息泄露与内容偏差风险。基于通义千问大模型智能写作 Agent 通过先进的 NLP 算法和安全管理机制，精准解决以上痛点，为用户提供高效、安全、专业的文档创作解决方案。

典型案例

模力时刻

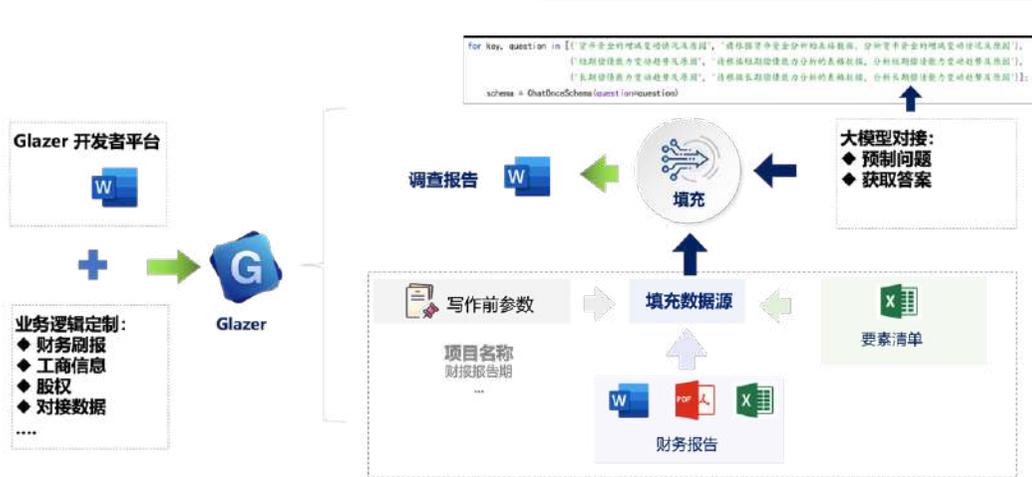
Qwen-72B

处理 按照配置的文档格式规范和行文风格，调用各类数据 API，生成专业报告。

效果 撰写时间缩短 70%~80%，业务人员报告处理效率提升 2~3 倍，错误率降低 80%。

智能写作助手已成功在多家金融机构落地应用，覆盖了银行、券商等多个领域。作为专业的报告撰写助手，它能够连接多种数据源，并利用 AI 模型自动提取相关财务数据，精准地整合到报告中。在此基础上，借助大语言模型的强大分析能力，基于提取的信息进行深入解析，为业务人员提供有价值的决策支持。此外，根据客户的实际需求，智能写作助手还支持对生成的报告进行定制化审核，既可以人工介入审核，也可以依据具体的业务逻辑实现自动化审核流程，极大地增强了业务审核过程中的效率和准确性，提供了强有力的支持。这一解决方案不仅提升了报告撰写的效率和质量，同时也优化了审核流程，确保了信息的准确性和合规性。

- 支持 AI 语义理解，智能生成内容框架；
- 支持自动完成公式计算、数据更新和内容撰写；
- 支持募集说明书、债券存期报告等各类复杂业务文档模板定制；
- 支持对接数据库数据、非结构化文档数据，可溯源定位到原始文件；
- 支持一键导出文档，撰写内容高亮批注，保留痕迹。



价值分析

提升文档撰写效率

金融机构需要处理大量的专业报告和分析材料，大模型可以自动生成高质量的初稿，减少从零开始撰写所需的时间，使团队能够更专注于内容审查和策略制定。

增强合规性管理

金融行业受到严格的法规监管。大模型可以帮助识别文档中可能存在的合规风险，并根据最新的法律条文自动调整内容，确保所有提交的文件符合相关法律法规的要求，降低违规风险。

提升确保信息准确性

通过自动化提取和整合来自多种数据源的信息，大模型可以保证数据一致且准确无误。这减少了人为错误的可能性，并确保了文档的专业性。

优化人力成本

自动化处理低价值重复性任务（如财报摘要、客服话术、产品说明书），释放专业人员精力聚焦高价值工作（如策略制定、客户关系管理）。

- 撰写时间缩短 70%~80%
- 业务人员报告处理效率提升 2~3 倍
- 错误率降低 80%

AI 程序员

Agent 概述

各类 AI 编程工具层出不穷，但大多数工具往往只能完成单一的任务，如代码补全或语法检查。AI 代码工具已从辅助编码升级到面向复杂任务的工程级协同编码。

通义灵码 AI 程序员使开发者在编程时能够更加专注于逻辑与创意。AI 程序员具备多文件代码修改（Multi-file Edit）和工具使用（Tool-use）的能力，可以与开发者协同完成编码任务，如从零开始的业务需求开发、复杂问题修复、单元测试的批量生成等，将重复性和繁琐的编码任务交给 AI 处理，提高整个开发过程的效率和质量。

需求分析

AI 编程工具功能比较单一，只能处理一些简单的问题，因此对新一代的 AI 编程工具备有以下功能：

缺陷分析

开发者全程只需一键确认，即可快速完成缺陷修复代码和代码提交的全过程。

需求开发

针对从 0 到 1 的开发任务，用户也只需输入自然语言描述需求，即可自动完成理解需求、拆解开发任务、编码、测试和修复等全链路开发过程。

单元测试

根据当前代码变更、单个或多个代码文件等批量进行单元测试生成，并自动进行编译、运行和报错修复，提升测试覆盖率。

典型案例

模力时刻

QwQ-plus

处理 根据用户的业务需求描述，AI 程序员自动完成理解需求、拆解开发任务、编码、测试和修复等全链路开发过程。

效果 开发效率提升 30% 以上。

AI 程序员与开发者一起围绕开发任务进行编码工作时，开发者需要输入清晰的需求描述和要求，AI 程序员可以为开发者在工程内进行多个文件的修改，开发者审查并接受相关代码变更即可。

1 AI 编码从片段级到多文件级

代码助手的能力已经从单文件简单的片段级别注释生成、单测生成、代码优化等，进化到多文件级别的编码任务，例如缺陷修复、需求实现、批量测试用例生成、多文件代码评审、批量代码重构、三方依赖升级等。



2 端到端完整编码任务

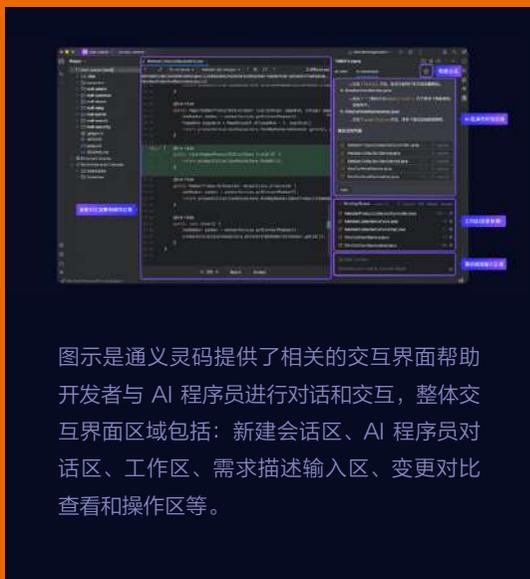
开发者只需要输入准确的需求和上下文，AI 可自主完成从需求理解、任务规划、代码生成、Diff 生成全过程。开发者无需从零开始编码，而是基于 AI 生成结果完成任务。

3 更多的复杂步骤实现自动化

AI 可以进一步实现上下文自动感知、编程工具自动使用、自动功能验证、自我反思迭代等自动化能力，进一步释放开发者的生产力。

4 更流畅自然的人机交互模式

更加类人对话的交互模式，从用户输入、过程干预、结果采纳过程均为人机协同，更融入开发者完成一个完整研发任务的编码链路。



价值分析

📁 工程级编码任务

AI 程序员具备多文件代码修改和工具使用的能力，可以与开发者协同完成编码任务，如需求实现、问题解决、单元测试生成、批量代码修改等。

🚀 提升开发效率

AI 程序员能够迅速生成产品原型，并在短时间内完成从 0 到 1 的业务需求开发，极大地缩短了开发周期。自动生成高质量的单元测试用例，确保代码质量的同时节省了大量的手动编写测试的时间。

✅ 提高代码质量

自主进行缺陷分析、制定修复计划，可快速完成缺陷修复代码和代码提交，帮助开发者高效修复缺陷，轻松解决繁杂的修复任务。

需求助手

Agent 概述

传统软件开发需求编写是一个详细且系统的过程，其中需求文档内容多且描述复杂，容易出现业务人员和技术人员理解不一致的情况。

需求助手借助大模型文档理解能力，以智能化的方式解析非结构化文本，提取关键需求信息，进行需求理解、编写、条目化和业务流程图生成等任务，并将其转换成清晰、可执行的需求文档。需求助手帮助业务人员和技术团队更高效地沟通，提升沟通效率和业务需求理解能力。

需求分析

在金融行业快速发展的背景下，传统的需求管理方式面临着以下挑战：

需求沟通不畅

由于业务和技术人员之间的专业背景差异，往往导致需求传达不清或误解，影响项目的顺利推进。

迭代更新频繁

随着市场变化和技术进步，需求经常需要调整，传统的手工修改方式耗时费力，容易遗漏关键点。

需求文档冗长难懂

大型项目中，需求文档可能非常庞大，包含大量细节，难以快速定位重要信息，增加了理解和维护的难度。

缺乏直观的业务流程展示

文字性需求文档难以直观地表达复杂的业务流程，使得开发团队在实施过程中容易出现偏差。

典型案例

模力时刻

QwQ-plus

处理 大模型根据需求进行需求理解、编写、条目化和业务流程图生成等任务，并将其转换成清晰、可执行的需求文档。

效果 需求文档编写效率提升 80% 以上。

需求助手通过对需求的理解和已有需求知识库内容，进行需求拆解生成需求条目，具体流程如下：

1 需求理解

解析并整理了来自不同部门的多样化需求，确保了所有利益相关者的需求都被充分考虑，形成了统一的需求文档，大大提高了项目的启动速度。

2 需求知识获取

检索需求知识库，获得与本次需求相似内容。

3 需求条目化应用

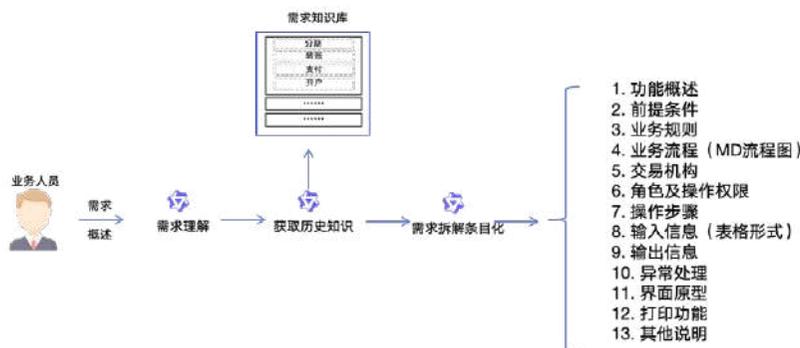
通过对需求文档进行条目化处理，每个需求项都有明确的责任人和优先级，使开发团队能够更加专注于高优先级的任务，提高了资源分配的合理性。

4 业务流程图生成应用

自动生成了包括开户、转账、贷款审批等多项核心银行业务在内的流程图，帮助新加入的成员快速上手，减少了培训成本。



上图显示为银行清算用户记账明细查询场景输出内容



价值分析

提升需求沟通效率

通过智能化的需求理解和编写功能，确保业务和技术人员之间的有效沟通，减少误解和返工，加快项目进度。

优化需求文档结构

将复杂的需求文档条理化，使得重要信息易于查找，增强了文档的可读性和可用性，方便后期维护和扩展。

增强业务流程可视化

自动生成的业务流程图不仅有助于内部团队的理解，也可以作为对外展示的重要资料，促进了跨部门间的协作。

支持敏捷开发模式

面对频繁的需求变更，需求助手能够迅速适应新的要求，保证需求文档与实际业务保持一致，助力企业采用更加灵活的开发方法论。

降低培训成本

对于新员工或外部合作伙伴，通过直观的业务流程图可以快速了解业务运作机制，降低了培训时间和成本。

促进知识积累与复用

系统自动记录的历史需求和流程图成为宝贵的资产库，为后续类似项目提供了有价值的参考，避免重复劳动。

编码助手

Agent 概述

程序员进行代码开发过程中经常会遇到需要大量相似性编码，时间投入价值低

主要依赖手动编写代码、调试和测试。但随着项目规模的增大、开发周期的紧张，传统开发模式往往面临以下挑战，重复性的代码耗费大量时间和精力、调试和维护过程中花费大量时间定位问题修复 bug、复杂的跨平台兼容性的问题。

基于通义大模型的智能编程辅助工具，提供行级 / 函数级实时续写、自然语言生成代码、单元测试生成、代码优化、注释生成、代码解释、研发智能问答、异常报错排查等能力，助力开发者高效、流畅地编码。

需求分析

开发者在日常编码过程中面临的多种挑战，这些问题直接影响了开发效率和代码质量：

重复性劳动

开发者需要频繁编写重复代码，这不仅耗时而且对提升项目价值有限。

单元测试的缺乏

缺乏编写单元测试的习惯导致代码质量不稳定，并增加了后期维护的成本和难度。

文档和注释负担

为保证代码可读性和可维护性，开发者需要撰写详细的注释和文档，但这增加了额外的工作量。现有代码中缺少足够的注释，使得快速理解和阅读代码变得更加困难。

学习成本和技术债务

使用新的 SDK、接口或编程语言时，查找相关信息和文档不方便，增加了学习成本和上手时间。技术债务积累严重，影响了代码的可读性和维护性，同时也提高了新加入成员的理解难度。

调试和资料检索效率低下

遇到运行异常时，开发者需要花费大量时间搜索相关资料进行排查，且这些资料可能不准确或不可靠。技术资料分散，检索这些资料的成本很高，特别是对于不熟悉的技术框架，这进一步加大了写出高质量代码的难度。

这些问题直接或间接地影响了开发效率和代码质量，要求采取有效措施来缓解这些问题，如采用自动化工具减少重复工作、加强单元测试实践、优化文档管理流程、以及建立高效的知识分享和检索机制等。

典型案例

模力时刻

Qwen-coder

处理 根据当前代码文件及跨文件的上下文，生成行级 / 函数级代码、单元测试、代码优化建议等。

效果 代码采纳率 30% 左右。

QwQ-plus

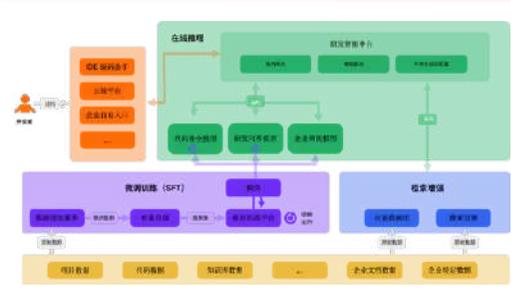
处理 根据海量研发文档、产品文档、企业研发知识，生成研发问题答案。

效果 提升搜索文档、查询知识效率 50% 以上。

通义灵码引用 Qwen-Coder 大模型，提供代码智能生成、研发智能问答、AI 程序员等功能，具体功能描述如下：

1 代码智能生成

经过海量优秀开源代码数据训练，可根据当前代码文件及跨文件的上下文，生成行级 / 函数级代码、单元测试、代码优化建议等。沉浸式编码心流，秒级生成速度，更专注在技术设计，高效完成编码工作。



2 研发智能问答

基于海量研发文档、产品文档、通用研发知识、阿里云的云服务文档和 SDK/OpenAPI 文档等进行问答训练，答疑解惑，帮助轻松解决研发问题。左下图是通用问答界面：

3 AI 程序员

具备多文件代码修改和工具使用的能力，可以与开发者协同完成编码任务，如需求实现、问题解决、单元测试用例生成、批量代码修改等。

价值分析

☑️ 跨文件感知让代码生成更贴合业务场景

客户端和模型层同步优化针对代码跨文件上下文感知能力，生成代码更加贴合当前代码库的业务场景，代码采纳率在 30% 左右。

☑️ 适配多 IDE 的原生设计，符合开发者使用习惯

适配 IDE 原生主题、交互模式，让开发者感受到 IDE 的原生视觉和交互体验，更符合开发者使用习惯。

☑️ 研发问答、文档 / 代码搜索能力，打造沉浸式编码

无需切换工具，IDE 客户端内即可满足开发者编码场景中的技术资料检索、技术难题解答的诉求，打造开发者沉浸式编码体验。

单元测试生成助手

Agent 概述

在现代软件系统开发过程中，代码的单元测试是确保代码质量和系统稳定性的重要环节，但单元测试面临功能覆盖不全、测试用例重复性多、浪费时间长、单测代码不更新等问题。

单元测试生成助手通过大模型技术，同时学习各种不同的训练测试框架和历史单元测试示例，通过与开发 IDE 客户端代码插件融合，提升代码测试的覆盖度，提升测试开发效率。

需求分析

单元测试用例测试框架比较多学习成本高、测试用例重复性多费时又费力、业务代码变化导致用例无效等问题，测试人员通常需要手动提取需求定义测试要点，而人工编写的测试用例往往面临着以下主要挑战：

测试框架多

当前存在多种单元测试框架，如 JUnit/Mockito (Java)、unittest/pytest (Python)、Mocha (JavaScript) 等，每种框架都有其独特的特性和使用方法。对于初学者或需要在多种语言间切换的开发者来说，掌握这些框架可能需要投入大量的时间和精力。

测试用例重复性多

在大型项目中，往往需要编写大量的单元测试用例来覆盖所有的代码路径和边界情况。这些测试用例在编写和维护过程中可能会产生大量的重复代码，尤其是当测试逻辑相似或测试数据相近时。

测试用例保鲜难

在软件开发过程中，业务逻辑和代码实现是不断变化的。当业务代码发生变化时，原有的单元测试用例可能不再适用或无法通过测试，这被称为测试用例的“脆弱性”。

典型案例

模力时刻

Qwen-plus

处理 选中指定的函数，通过灵码单元测试功能，自动生成对应的单元测试用例。

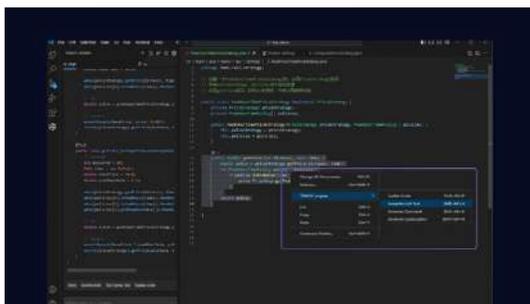
效果 代码单元测试用例覆盖度超过 70% 以上，生成单元测试用例效率提升 50%。

单元测试生成助手集成到 IDE 插件，根据函数代码段即可生成对应的单元测试。



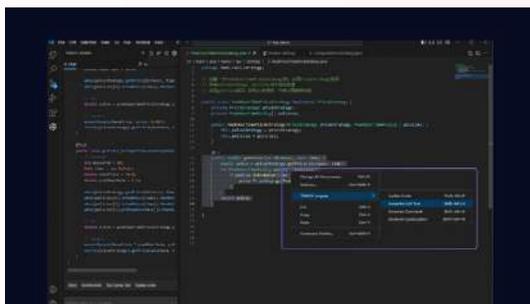
1 在 IDE 代码开发过程中

选中对应的函数内容，右键即可出现“Generate Unit Test”即可。



2 单元测试生成助手

会根据项目中的测试框架、开发语言等信息自动生成对应的测试代码，根据代码生成质量确定是否采纳测试代码。



3 满意的单测代码

如您对当前生成的单测代码不满意（需要使用特定的单测框架或者生成更多的单测方法），只需在问答区中的输入框输入相应的文本进行追问即可，或者也可以单击预置的单测追问 Tag，如示例中追问的 Tag：Retry、Use Mockito、Use Spring Test 和 Explain code，进行追问，直到生成满意的单测代码为止。

价值分析

单元测试助手的价值主要体现在提高软件质量和可靠性，确保代码在修改或重构后仍然能够正常运行。单元测试的优势包括：

提高代码质量

通过发现代码中的错误和漏洞，从而提高代码的质量和可靠性。

提高开发效率

在开发过程中及时发现问题，减少开发周期和成本。

便于重构和维护

确保代码在重构和维护过程中不会出现新的错误和漏洞。

有助于团队协作

作为团队成员之间的沟通和协作工具，提高团队的协作效率和质量。

测试用例生成助手

Agent 概述

在现代软件系统开发过程中，传统的测试用例编写需要大量的人工参与，由于需求文档的复杂性和多样性，会面临需求多、测试用例生成耗时长、功能覆盖不足等问题。

测试用例生成助手通过大模型技术，结合历史资产库学习大量的历史测试数据和领域知识，生成更准确、全面的测试用例。大模型测试用例生成助手自动化处理软件开发中的关键环节，提供了智能化的需求切片、测试要点提取、场景化测试用例生成的功能，不仅能够显著提高测试效率和质量，还能有效减少人工干预降低人为错误的风险。

需求分析

随着软件系统的复杂度不断提升，传统的人工编写测试大纲和用例的方式变得越来越难以满足快速迭代的需求。需求文档通常是测试流程的起点，测试人员通常需要手动提取需求定义测试要点，而人工编写的测试用例往往面临着以下主要挑战：

手动提取需求耗时费力

从复杂的需求文档中提取关键信息并定义测试要点是一个非常耗时的过程。

测试用例覆盖率不足

人工编写的测试用例往往存在覆盖不全、重复性高、场景不合理等问题。

典型案例

模力时刻

QwQ-plus

处理 大模型结合历史测试用例资料库，学习大量的历史测试数据和领域知识，生成更准确、全面的测试用例大模型测试用例生成。

效果 测试用例生成效率提升 80% 以上。

测试用例生成助手主要包括需求文档准备、测试文档库准备、大模型基于产品需求与文档库进行测试大纲和测试用例生成内容，最终人工确认。具体流程如下：

1 需求准备

产品需求人员提供产品需求文档，并上传到测试平台

2 测试知识库构建

基于历史测试文档构建历史测试知识库，并且进行案例数据的拆解工作。

3 模型解析

通过大模型调用进行需求问答切片、要素提取；自动解析需求文档，生成精准且全面覆盖功能、性能、安全等方面的测试框架；基于测试大纲智能生成符合需求文档的多样化测试场景，确保测试过程全面高效。

4 内容生成

根据测试大纲和测试用例输出，测试平台生成测试内容。

5 结果确认

测试人员可以人工审核每一个测试用例，确认用例的可用性。



价值分析

测试用例生成助手的应用为企业带来了多方面的价值提升：

提高测试效率

自动化生成高质量的测试大纲和用例，大幅缩短了准备阶段的时间，使测试团队能够在更短的时间内完成大量的测试工作。

增强测试质量

通过大模型的智能分析，系统能够识别出需求文档中可能被忽视的测试要点，确保测试大纲和用例的全面性和深度，减少了人工疏漏和错误的风险。

降低成本

自动化的测试准备和执行减少了对人工的依赖，降低了人力成本；同时，更高效的测试周期也缩短了项目的整体成本。

促进知识积累与复用

系统通过与历史测试资产库的结合，能够实时更新测试用例库，并为后续项目提供有价值的参考，避免重复劳动。

智能运维助手

Agent 概述

传统 IT 运维面临的问题多种多样，随着企业业务的增长和技术环境的变化，传统运维模式在面对大规模、高复杂度的系统时，往往存在响应速度慢、故障定位不精准等问题。

智能运维助手通过大模型技术、向量数据库以及知识库支持，将运维文档统一管理构建专业全面的运维知识库，另外基于智能体工作流技术构建常用故障诊断助手，根据运维人员的问题提供智能告警分析、运维知识库检索、场景智能诊断等功能，帮助运维团队快速定位并解决故障。

需求分析

随着金融行业 IT 系统的日益复杂化，传统的运维方式面临诸多挑战：

知识分散

运维规则和应急方案分布在不同的团队文档中（如数据库、操作系统、网络等），导致需要咨询或故障处理时必须询问多个团队。

缺乏实时支持

运维人员在查询文档内容时效率不高，在遇到紧急情况时，无法迅速获得有效的技术支持和解决方案，影响业务连续性和用户体验。

多种工具操作繁琐

日常工作中需要接触多种产品，每种产品都有自己的运维工具或界面。定位及处理问题时，运维人员需要手动对各产品的运维工具逐个进行操作，过程繁琐。

经验难以沉淀

分析和处理问题的经验通常以本地文档或口口相传的方式传承，学习成本高。

典型案例

模力时刻

QwQ-plus

处理 大模型对运维问题进行意图识别，结合各种运维知识库，生成运维问题答案。

效果 运维知识查询和处理效率提升 80% 以上。

为了克服上面的困难，智能运维助手将大量运维文档统一管理构建知识库，然后通过大模型方式提供了针对性强且高效的解决方案，满足了现代金融企业对于运维工作的高标准要求，具体流程如下：

1 运维知识库

汇总散落在各处的运维文档，包括产品文档、应急手册、问题案例、其他知识问答等，构建统一运维知识库。

2 问题意图识别

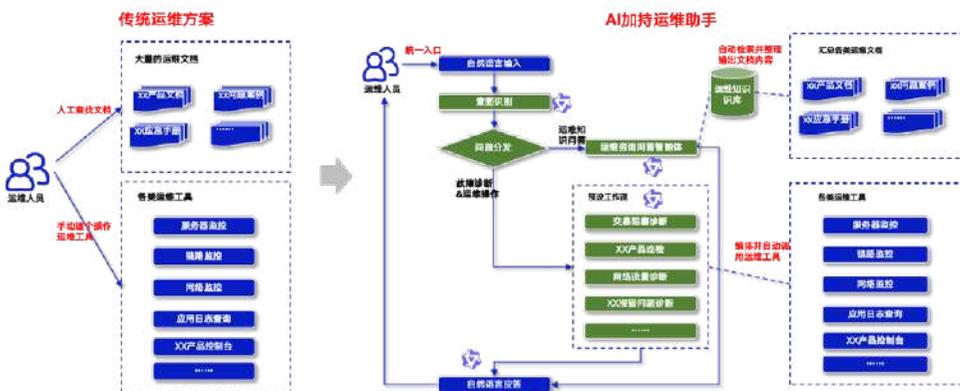
根据运维人员的问题输入判断是运维知识问答还是具体故障诊断 & 运维操作。

3 运维问答

运维知识问答场景，智能体会自动检索知识库内容，提供产品材料或应急手段。

4 故障诊断

沉淀了各种不同场景的诊断能力，包括业务交易阻塞诊断、产品巡检、网络流量诊断等能力，这些内置的故障诊断智能体可以直接对接到各种运维工，如服务器监控、链路监控、应用日志查询等等。



价值分析

大模型运维助手为企业带来了以下几方面的价值提升：

🔍 高效知识检索

各种类型运维文档汇总，实现对大量文档内容的管理能力，提升检索效率。

🛠️ 统一运维助手

实现统一的运维助手界面，通过大模型对话能力。

💡 运维经验沉淀

分析和处理问题的经验以诊断 workflow 形式沉淀下来，辅以高效的检索，极大降低了新人上手的门槛。

交互取证提额

Agent 概述

随着数字化时代的到来，金融支付服务行业的业务扩张和数据风控要求越来越高，要提高与观众的实时互动能力，并能根据客户提交数据进行快速核对和风控矫正，快速进行提额业务的处理，实现秒级或分钟级的业务办理，增加客户的信任和服务的粘合力，提高放款和转化。

为了适应 AI 智能时代发展需求，我们提供了交互取证提额的 Agent，能够根据不同的客户数据和需求，提供个性化的额度提升方案和建议。借助自动化流程和实时数据分析，实现快速审批和响应。通过高效的工具来汇总、分析和展示大量的客户数据，以快速做出更好的决策。确保数据分析和决策过程的透明度，提高客户信任度。

需求分析

核心业务痛点：

数据量大且复杂

金融机构处理海量的客户数据，包括信用评分、交易历史、财务报表等，传统的数据处理和可视化工具可能无法高效地管理和展示这些信息。

个性化需求

客户对于额度提升的原因和需求是多样化的，金融机构需要对每个客户进行个性化分析，但现有系统可能无法灵活应对。

实时性不足

额度提升的需求往往需要快速响应，但当前的人工审批和分析流程可能导致响应速度缓慢。

决策透明性和可解释性

金融决策需要有充分的理由说明，但复杂的数据和算法可能导致客户和内部员工难以理解决策逻辑。

典型案例

模力时刻

Qwen-VL

处理 在视频交互过程中，把视频拆帧成图像，使用 Qwen-VL 对其中内容识别。

效果 证件要素提取准确率 95%，画面理解准确率 85%。

Qwen-72b

处理 结合客户画像和历史通话记录，进行交互式问答。

效果 答案生成准确率 90%。

某消金公司 2025 年进行落地通过数字人和大模型的智能交互，引导客户提交相关固定资产证件或证书等证件，根据业务提额分成，与自动化流程结合，小额业务可以一键提额，大额业务获取真实数据，完成自动化流程审核，人工进行复核，可视化交互能力，更好地满足客户需求并提高运营效率。



通过大模型和数字人的智能技术完美的结局实际业务发展需要，并大大降低人工核验成本和用户满意度，利用大模型的自然语言处理能力，结合客户的历史数据与偏好，生成个性化的额度提升建议。结合大模型与流程自动化结合，自动处理额度审核的部分流程，包括客户数据验证、风险评估和建议方案生成。使用大模型生成简洁易懂的分析报告，帮助客户和决策者理解每个额度调整背后的逻辑和理由。

价值分析

📈 提高决策效率

通过大模型的高效数据分析能力，实现快速响应客户需求，缩短审批时间。

😊 增强客户满意度

个性化服务和快速响应提高客户体验和满意度。

↓ 降低运营成本

自动化流程减少了人工审核的时间和成本。

🏆 提高竞争力

通过先进的技术和透明化的决策流程，金融机构能够在市场中获得更好的竞争优势。

⚠️ 风险控制

更精准的数据分析能力有助于评估和管理信贷风险，确保额度提升的安全性和合理性。

视频交互陪练

Agent 概述

随着数字化时代的到来，金融服务行业的产品和服务迭代越来越快，需要业务人员快速跟上时代的发展，针对新一代产品进行有针对性的学习和培训，并根据不同的客户和产品特色进行训练，达到能够方便的解决产品知识的学习效果和适合产品的快速推广应用。

为了克服这些痛点，视频交互式培训 Agent 应运而生，旨在通过数字人和大模型的智能互动化的方式提升培训效果。该 Agent 采用先进的大语言模型技术，利用其强大的意图理解、话术生成、工具调用等能力，模拟实际工作场景，增强学习体验，通过个性化学习与实时反馈，帮助学员更深刻地理解知识并应用于实际工作中，从而有效提高学习效果。

需求分析

主要业务痛点：

知识更新滞后

专业服务人员需要频繁更新产品知识和政策，但有时缺乏时间和资源。

培训成本高

传统培训方式耗时耗力且费用高昂，包括场地、培训师以及材料。

个性化不足

一对多的培训模式无法满足每位代理人的个性化学习需求。

缺乏实时反馈

专业服务人员在面对客户时可能遇到无法即时解答的问题，影响销售效果。

典型案例

模力时刻

Qwen-72b

处理 以对话的形式，基于客户历史学习情况和真实作业场景，生成个性化提问。

效果 考题生成准确率 90%，评分准确率 98%。

某保险服务公司因业务迭代发展需要，业务人员需要灵活的学习方式，便于在工作闲暇中快速提高技能，根据经验水平和专长，提供定制化的培训内容；在实际销售场景中，能够迅速获得有关产品和政策的解答；具备模拟真实销售情境的功能，帮助代理提前练习应对策略。

通过数字人和大模型的智能交互问答，能够提供迅速且准确的知识更新，帮助代理人更快掌握产品信息，提高销售效率；根据业务的实际交互沉淀，还原业务交互链路，通过大模型进行智能交互问答和结果效果识别等；通过数字人模拟还原业务应用场景，让业务人员进行模拟演练，总结不足，制定定向学习方案；通过模拟陪练和实时反馈，提高代理人在实际销售中的表现。



价值分析

提升工作效率

通过大模型提供迅速且准确的知识更新，帮助代理人更快掌握产品信息，提高销售效率。

降低培训成本

减少对传统培训资源的依赖，节省时间和经费。

增强用户体验

通过个性化和互动式的学习，提升代理人的学习体验和满意度。

提高销售成功率

通过模拟陪练和实时反馈，提高代理人在实际销售中的表现。

数据驱动优化

基于数据分析，不断优化学习方案和培训内容，提升整体效果。

直播交互

Agent 概述

随着数字化时代的到来，金融服务行业的服务水平要求越来越高，要提高与观众的实时互动能力，并能根据观众偏好调整内容。通过创新互动方式提高用户参与度；采用技术手段优化流程，降低成本。提高内容生产速度，保持用户的新鲜感；同时真多不同种类的产品分享等，需要提供多语言支持和文化适应内容。通过大模型与数字人的直播交互融合，创建出一个或多个高度仿真虚拟形象，并通过实时在线直播的形式，在网络平台上向观众介绍、演示和销售商品或服务的一种电子商务活动。同时通过大模型智能交互能力，与数字人直播集成，推动运营最佳运营实践路径的平台，通过简单配置，即可一键完成开播内容准备，帮助商家高效创作直播内容。

证券：通过分享和直播，打造 IP 形象和特定人设，培养客户和引流获客；

保险：保险产品分享直播，小额按月保险的引导销售；保险代理人的私域朋友圈人设培养，获客引流；

需求分析

主要业务痛点：

内容同质化

大量直播内容缺乏新意，用户易于失去兴趣。

用户参与不足

用户与主播之间的互动有限，导致用户体验较差。

主播职业疲劳

真人主播面临高强度工作以及职业倦怠问题。

多语言和文化差异

跨语言、跨文化交流存在障碍，影响国际用户的获取。

技术成本高

高质量的互动技术实现成本高，包括硬件设备和技术开发。

典型案例

模力时刻

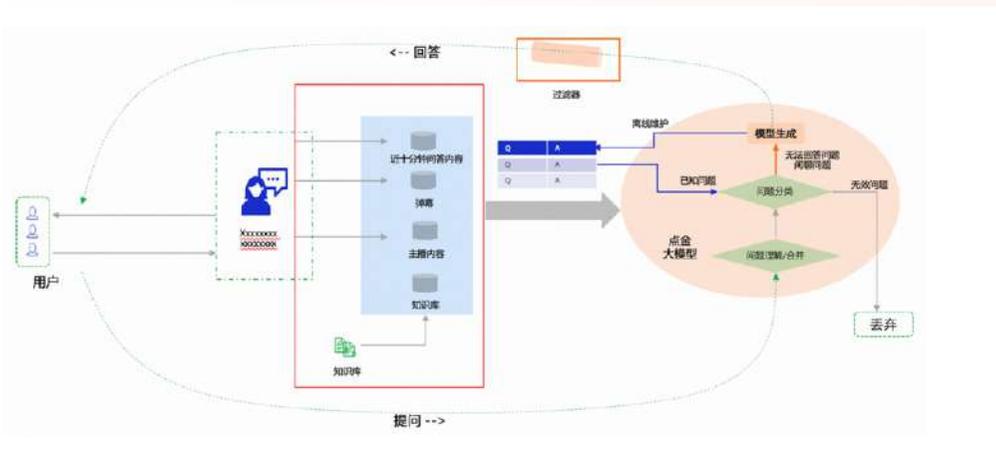
Qwen-72b

处理 以对话的形式，基于客户历史学习情况和真实作业场景，生成个性化提问。

效果 考题生成准确率 90%，评分准确率 98%。

保险代理人服务公司通过数字人直播交互，使用自然语言理解和交互能力进行实时字幕翻译和多语言支持；利用计算机视觉和语音合成技术创建高度逼真的数字人主播，降低真人主播的职业疲劳。结合情感识别技术，让数字人主播具备多样化的情感表达，增强观众互动体验，实现大模型驱动的聊天机器人功能，能够实时回答观众问题，提高互动性，结合用户画像和行为数据分析，为用户提供个性化的内容推荐。

在多媒体平台分享寿险产品信息，引导试用，再进行深入跟进转换，同时通过大模型进行快速且准确的知识更新，帮助代理人更快掌握产品信息，提高销售效率。同时在直播完后，通过大模型实现数据分析，掌握不同地区、不同文化的用户偏好，调整直播内容策略。自动调整内容表现的语言和风格以适应目标市场。



解决方案：

1 引入大模型智能交互能力

使用自然语言理解和交互能力进行实时字幕翻译和多语言支持。采用大模型生成内容，提高内容创作效率，比如自动生成脚本、视频主题和互动话题。

2 数字人技术优化

利用计算机视觉和语音合成技术创建高度逼真的数字人主播，降低真人主播的职业疲劳。结合情感识别技术，让数字人主播具备多样化的情感表达，增强观众互动体验。

3 智能互动系统

实现大模型驱动的聊天机器人功能，能够实时回答观众问题，提高互动性。结合用户画像和行为数据分析，为用户提供个性化的内容推荐。

4 跨文化交互设计

通过大模型实现数据分析，掌握不同地区、不同文化的用户偏好，调整直播内容策略。自动调整内容表现的语言和风格以适应目标市场。

价值分析

增强竞争力

通过个性化和高质量的内容与互动提高平台竞争力，吸引更多用户。减少对真人主播的依赖，降低人力和运营成本。

提高用户黏性

丰富的互动体验和内容定制增加用户的留存率。多语言支持和跨文化内容适应性有助于拓展国际市场。

提升盈利能力

通过提升用户体验和增加用户参与度，带动广告收入和打赏收益的增长。

视频分享

Agent 概述

随着数字化时代的到来，金融服务行业的业务获客难度越来越高，需要通过私域朋友圈或视频分享平台，根据不同专家或目标人设进行定向的业务分享，引导客户进入交流渠道或群组，进行业务的定向营销。

为适应业务的发展要求，通过 AI+ 数字人或多模态的视频分享 Agent 应运而生，实现快速、高效且低成本的视频生成，提升内容的多样性和吸引力，根据用户数据生成定制化的视频内容。在社交媒体广告，通过生成具有吸引力的营销视频，提高品牌曝光率和用户互动；也用于展示产品功能和优势，比图文更能吸引潜在客户；也会为活动、会议和展览会制作宣传视频，吸引参会者，增加活动的知名度和影响力。

需求分析

核心业务痛点：

高制作成本

传统的视频制作涉及昂贵的设备、专业的团队和较长的制作时间。

内容创新不足

大量雷同的内容难以吸引观众注意力，无法满足不断变化的市场需求。

个性化不足

难以在大批量生产中实现个性化，导致无法精准满足不同客户群体的需求。

典型案例

模力时刻

Qwen-72b

处理 以对话的形式，基于客户历史学习情况和真实作业场景，生成个性化提问。

效果 考题生成准确率 90%，评分准确率 98%。

券商通过专业分析师在新加坡进行定向人群分型，并在多媒体平台发布，不仅可以做专业的股评分析，还可以做公众号或私域朋友圈进行分享转发，吸引客户的访问和订阅；也可以在不同国家和地区制作本地化营销视频，以便更好地渗透国际市场。同时 AI 智能化在营销中正扮演着越来越重要的角色，能够有效地传递信息并创造深刻的用户体验。

通过大模型驱动的视频生成，简单易用，它可以利用数字人 / 多模态和自然语言处理，生成视频脚本和视觉内容。也可以快速生成视频素材，如场景、角色、和对白等；集成 AI 工具自动化整个视频生产流程，包括脚本创作、视觉渲染和后期编辑。使用用户数据和市场分析，定制化生成符合用户需求和喜好的个性化视频内容。



价值分析

成本效益提升

降低制作成本和时间，实现规模化生产。

内容创新

引入先进的 AI 工具，大大提升创意和内容多样性。

用户参与度提高

个性化的内容能够更好地吸引和维持观众的关注。

市场响应速度

快速调整和响应市场需求，增加市场竞争力。

内容安全审核

Agent 概述

随着金融行业多媒体宣传（如视频、直播、图文）形式越来越多，金融行业多媒体内容合规审核面临金融行业法律法规复杂，人力消耗大且效率与准确率不足、审核结果一致性、行为风险识别难度高、实时性不足等问题，通过人工审核工作枯燥，人力投入高，难以支持对多媒体内容高效、准确的识别；

基于阿里云通义系列大语言模型及多模态大模型的技术方案，实现按照通用规则、营销规则、业绩回顾规则、直播和短视频规则对文图、视频进行内容高效审核，为金融用户提供智能、高效的内容合规审核解决方案。

需求分析

多媒体内容审核人工审核方式人力消耗大，不同的审核员审核结果存在一致性问题，尤其对于短视频、直播场景存在审核难度高等痛点问题。

法规复杂性

金融行业的法律法规不断更新，且非常复杂，要求审核人员时刻保持对最新规定的了解，这对人工审核来说是一项挑战。

效率与准确性

传统的人工审核方式可能耗费大量时间和人力成本，而且难以保证完全不出错，特别是在处理海量信息时。

一致性问题

不同审核员可能因为个人理解差异，在执行标准上存在不一致的情况，影响了审核结果的一致性和公正性。

风险识别难度

一些违规行为或风险点可能隐藏得很深，不容易被发现，特别是当涉及到复杂的金融产品或者交易结构时。

实时性不足

在快速变化的金融市场环境中，及时响应最新的市场动态和法规变化是一个难题。

典型案例

模力时刻

Qwen-VL-Max

处理 图片类型审核，采用 Qwen-VL-Max 多模态大模型，输入图片 + 规则提示词，产出规则审核结果。

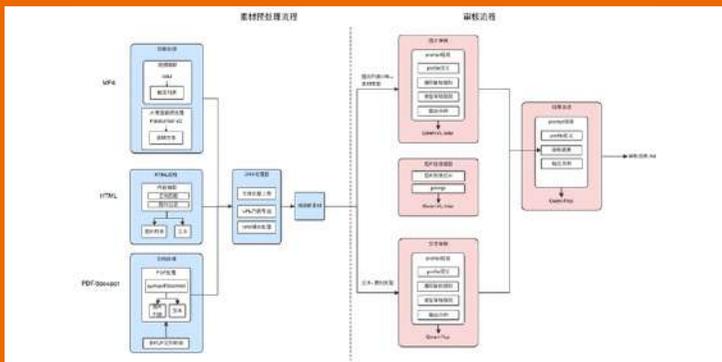
效果 提高审核效率 50%。

Qwen-Plus

处理 纯文本类素材，采用 Qwen-Plus 大语言模型，输入文字信息 + 规则提示词，产出审核结果。

效果 提高审核效率 50%。

在某期货公司，通过对 MP4 视频、HTML 文件和 PDF/doc/ppt 等异构数据预处理，基于通义千问 VL 模型和 Qwen-plus 模型匹配审核规则，对生成素材进行审核，生成审核结果；总体技术链路如下图所示：



1 异构数据预处理

● 视频数据：

- 1) 利用 OSS 结合 IMM 做视频截帧，将视频拆帧成为图片列表，存放至 OSS 中作为审核素材；
- 2) 音频使用百炼 Paraformer 转文本，文本保存至 OSS 中作为审核素材；

● 网页数据：

通过脚本解析，分类提取文本类信息和图片列表，保存至 OSS 中作为审核素材；

● 文件数据：

文件数据统一转 PDF，使用阿里云 docmind 或开源库 pymupdf 处理，得到图片列表和文本类数据；

2 大模型审核

- 图片类型审核，采用 Qwen-VL-Max 多模态大模型，输入图片 + 规则提示词，产出规则审核结果；
- 图片内容提取，提取图文信息，形成文字描述信息；采用 Qwen-Plus 大语言模型，输入文字描述 + 规则提示词，产出审核结果；
- 纯文本类素材，采用 Qwen-Plus 大语言模型，输入文字信息 + 规则提示词，产出审核结果；

价值分析

📖 持续学习

基于机器学习的大模型可以不断从新的案例中学习，快速更新其审核规则，以适应不断变化的法规环境。

📈 高效处理

大模型能够快速处理大量数据，大大提高审核效率，减少人工成本。

🚩 提高准确性

大模型可以通过算法优化来减少错误率，提供更加准确的审核结果。

🔗 一致性保障

大模型按照预设的算法和规则进行审核，可以确保审核结果的一致性，不受人为因素的影响。

🔍 深度分析能力

利用先进的自然语言处理技术和数据分析能力，大模型能更深入地挖掘潜在的风险点和违规行为。

制度撰写

Agent 概述

在金融行业中，企业必须遵循复杂、多变的监管要求和内部合规政策。随着金融业务的快速发展和市场环境的变化，金融机构需要不断更新和完善合规制度，以防范法律风险和维护市场信用。传统的合规制度撰写过程耗时且容易出现疏漏，无法迅速响应法规变更和业务创新带来的合规需求。因此，金融机构迫切需要一种智能化的工具，来辅助制度的撰写、编辑和更新。金融行业大模型合规制度撰写助手正是在此需求背景下应运而生，帮助企业提高制度撰写的效率和准确性。

基于阿里云通义系列大模型，为银行、保险等金融机构的合规和法务部门提供高效的制度撰写支持。制度撰写 Agent 能够对法规或制度所属的合规领域进行智能分类，使用户能够根据不同的业务类别进行有针对性的管理和查阅，针对新颁布法规能够深入分析法规内容，智能匹配受影响的内部制度，并提供废止旧制度或修订新制度的建议，能够有效提高合规管理的效率和准确性，还帮助金融机构更好地应对法规变化，降低合规风险。

需求分析

在金融行业的合规管理中，随着金融市场的快速发展和监管政策的不断变化，金融机构需要及时更新和调整内部合规制度以适应新的法规要求。然而传统的合规管理方式往往难以跟上法规更新的步伐，导致制度滞后于外部法规，增加了合规风险。其次，合规部门在处理大量复杂的法规文件时，常常需要投入大量的人力和时间进行手动分析和匹配，效率低下且容易出错。这些问题影响了金融机构的合规管理水平，制度撰写 Agent 能够智能分类法规、自动匹配内部制度并提供制度废止、修订以及撰写的建议，对于解决这些痛点具有重要意义。

典型案例

模力时刻

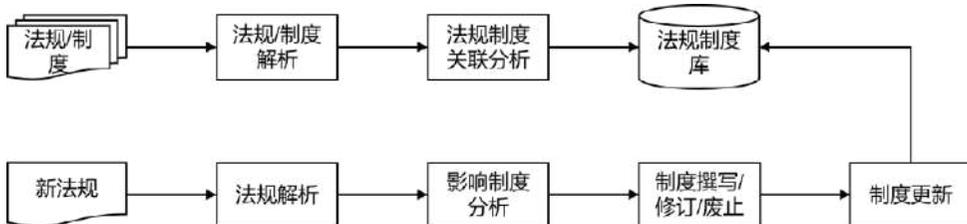
Qwen-72B

处理 自动解析法规内容、深入分析法规条款，判断对现有制度的影响，调用制度撰写 / 修订 / 废止任务 API，完成合规制度更新。

效果 合规制度撰写效率提升 50%。

制度撰写 Agent 的核心功能模块包括存量法规制度分析和增量法规分析。在存量法规制度分析模块中，Agent 首先读取并导入现有的法规和制度数据，运用自然语言处理技术对文本进行解析和切分，以提取出关键信息和结构化数据。随后，通过语义匹配和关键词提取等技术手段，分析法规与现有制度之间的关联性，识别出潜在的合规

风险点和法规制度关联关系。最后将分析结果存储在关系型数据库中，构建起一个全面的法规制度库，为后续的合规管理和决策提供支持。



在增量法规分析模块中，Agent 针对新颁布的法规进行读取。它会自动解析新法规的内容，并对其中的条款进行深入分析，以判断这些新法规将对现有制度产生何种影响，如是否需要废止某些过时的制度条款，或是对现有制度进行修订以适应新的法规要求。然后利用通义千问大模型的语言理解和生成能力，根据法规的影响类型及其对制度的具体要求，智能生成新的制度文本，或是提出修订和废止制度的具体方案。这一过程不仅提高了制度更新的效率，还确保了新制度与法规的高度一致性，降低了合规风险。

制度撰写 Agent 能够帮助金融机构及时、准确地响应法规变化，保持内部制度与外部法规的同步更新，从而提高合规管理水平，减少合规风险，增强企业的市场竞争力。通过将法规制度关联分析和通义千问大模型相结合，制度撰写 Agent 实现了法规与制度的自动化分析、匹配以及智能化撰写和修订，推动了合规管理工作的数字化和智能化转型。

价值分析

📄 提高合规文档质量

金融行业需要遵循严格的法律法规和内部政策，编写高质量的合规制度文件至关重要。制度撰写 Agent 可以确保这些文档基于最新的法规和最佳实践编写，减少人为错误，保证文档的专业性和准确性。

📈 提高制度制定效率

传统的合规制度编写流程通常耗时较长，涉及多个部门的协作和多次审核。使用制度撰写 Agent 可以显著缩短这一过程，通过自动化模板和内容建议快速生成初稿，使得团队能够更专注于审查和优化内容。

🔗 保持制度内容准确与一致性

金融行业监管环境复杂多变，合规要求频繁更新。制度撰写 Agent 能够实时整合最新的法律法规和行业标准，帮助金融行业及时更新其合规制度，确保所有政策文件始终保持最新且一致。

合规问答

Agent 概述

金融行业是一个高度复杂和严格监管的领域，涉及到大量的法律法规、合规要求和内部政策。随着金融市场的不断发展和创新，金融机构面临的合规挑战也随之增加。为了确保业务运营的合法性和合规性，金融机构需要及时全面地了解并遵循相关规定，减少因非合规操作而产生的风险和成本。

基于阿里云通义系列大模型，为银行、保险等金融机构的合规和法务部门提供高效的问答支持。它可以帮助金融机构快速、准确地获取合规信息、解答合规疑问，提高合规工作的效率和可靠性，同时也能辅助监管机构进行市场监察和执法，维护金融市场的公平和稳定。

需求分析

在金融行业，合规管理是确保企业稳健运营的关键环节，但传统合规工作面临着诸多痛点。合规部门在日常工作中经常需要解决大量关于公司内部制度规范的问题，这些工作重复且耗时，而且由于法规制度繁多且更新频繁，需要专业人员投入大量时间和精力去查找和解答。此外，外规的频繁修订和废止可能使得人员对法规或制度内容更新不及时，增加了合规风险。这些痛点不仅影响了合规工作的效率，还可能导致合规风险的增加。合规问答 Agent 这对这些问题提供了一种能够高效管理合规信息、快速响应合规查询需求的解决方案，以提升合规管理的整体效能。

典型案例

模力时刻

Qwen-72B

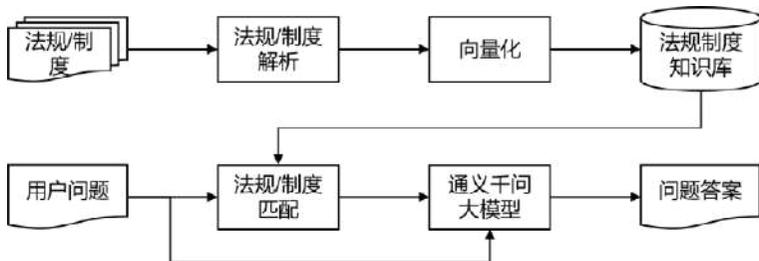
处理 理解客户查询意图，召回相关合规知识片段，生成问题答案。

效果 合规知识召回准确率 90%，合规问题回答准确率 90%。

合规问答 Agent 主要包括法规制度知识库构建和答案生成两个部分，知识库构建是基于客户自有法规数据和制度数据构建法规制度知识库，首先对各类型的法规制度数据进行读取，利用自然语言方法对法规和制度进行解析和切分，然后利用语义模型对切分后的数据进行向量化得到法规制度的特征向量，最后将法规制度及其特征向量存入关系型数据库和向量数据库构建法规制度知识库；在答案生成部分，首先对用户输入的问题进行向量化并通过语义相似度计算为其匹配相关的法规或制度信息，然后将匹配结果和用户问题通过通义千问大模型获取问题答案，最终将答案返回给用户。

合规问答 Agent 是一款专为金融机构合规部门设计的智能问答系统，其核心功能模块包括法规制度知识库构建和答案生成两大部分。在知识库构建环节，系统首先读取客户提供的法规数据和制度数据，这些数据可能包括法律法规、行业规范、内部规章制度等多类型文本。接着，利用自然语言处理技术对这些文本进行解析和切分，将复杂的法规制度内容拆解为更小的单元，以便于后续处理。然后，通过语义模型对切分后的数据进行向量化，将文

本内容转换为计算机能够理解和处理的特征向量，这些特征向量能够捕捉到法规制度的关键信息和语义特征。最后，将法规制度及其对应的特征向量存储到关系型数据库和向量数据库中，从而构建起一个结构化、可检索的法规制度知识库，为后续的问答提供坚实的知识基础。



1 在答案生成环节

当用户输入一个问题时，系统首先对问题进行向量化处理，将其转换为与知识库中法规制度特征向量相同维度的向量。然后，通过计算用户问题向量与知识库中法规制度特征向量之间的语义相似度，为用户匹配出与问题相关度最高的法规或制度信息。这一步骤能够帮助系统快速定位到与用户问题最相关的知识内容，为生成准确答案奠定基础。然后将匹配到的法规制度信息和用户问题一起输入到通义千问大模型中，利用其强大的语义理解和推理能力，生成针对用户问题的答案。通义千问大模型能够综合考虑法规制度的语义信息和用户问题的具体情境，生成符合逻辑、准确可靠的答案。最后，系统将生成的答案返回给用户，为用户提供及时、准确的合规咨询。

2 在技术方面

合规问答 Agent 融合了多种前沿技术，实现了对复杂合规信息的高效管理和检索。首先，在知识库构建环节，通过自然语言处理技术对法规制度文本进行深入解析和切分，能够更准确地提取出关键信息，为后续的向量化和检索提供更高质量的数据基础。其次，利用语义模型对法规制度数据进行向量化，使得文本内容能够以特征向量的形式存储和处理，这不仅提高了知识库的检索效率，还为语义相似度计算提供了可靠的依据。此外，通义千问大模型凭借强大的语义理解和推理能力，在答案生成环节发挥了关键作用，能够生成符合法规制度要求和用户问题情境的准确答案，有效解决了传统问答系统在复杂场景下答案不准确、不全面的问题。这些技术创新使得合规问答 Agent 在合规管理领域具有显著应用价值。

价值分析

增强合规管理

金融行业受到严格的监管，法规频繁更新。合规问答助手能够实时提供最新的法律法规信息，帮助金融机构迅速理解和适应新的监管要求，确保其业务操作始终符合规定。

降低违规风险

通过提供准确的合规指导，减少因不熟悉或误解法规而产生的违规行为，从而避免高额罚款、法律诉讼及声誉损失等潜在风险。

提升工作效率

传统查找合规资料的方法耗时费力，而合规知识助手能够迅速响应查询，提供精准的答案或指导，极大地提高了员工的工作效率，使他们能更专注于核心业务活动。

企业工商信息打标

Agent 概述

根据各个渠道获取的各类企业工商信息，由于经营范围或资质的变化，导致企业经营范围变化，会对企业投融资和对外服务信用，有一定的变化，需要快速针对获取的数据进行分析、加工、整理他们的信用状况，形成系列服务和产品，助力企业在生产经营活动中降本增效。为每个中小企业建立一份信用档案，中小企业可以认领、维护自己的信用档案，并在未来的招投标、融资、招聘等诸多场景下用于增信。

通过大模型 PE 构建企业工商信息打标 Agent，能够充分利用其强大的自然语言处理能力，有效解决了企业在工商信息数据提取中的多个痛点。它能够理解文本中的语义，从而处理指代词和相关信息，提高提取精确度。同时，大模型可以明确产品和服务的分类标准，减少提取逻辑的模糊性。此外，针对包含多项服务或产品的短句，它具备自动拆分能力，确保逐项准确提取。

需求分析

企业工商信息数据提取过程中面临多个主要痛点。首先，大模型仅允许从原始语句中进行结构化提取，无法基于经营范围推理生成字段值，这限制了其对指代词等的处理能力。此外，产品和服务的分类及提取逻辑往往模糊不清，甚至可能存在特定的提取条件，需要进一步明确。对于短句的处理，当其中包含多个服务或产品时，必须进行拆分并分别提取相应值。而在提取标准方面，“产品”字段仅限于物品名称的提取，而“服务”字段则需要完整匹配相关内容，导致提取过程更加复杂。

典型案例

模力时刻

Qwen-plus

处理 基于大段的文本描述，提取企业经营产品范围和提供服务内容。

效果 关键实体提取准确率 95%。

某征信公司通过对企业经营分类，快速构建和更新对外年报检索知识库，为金融从业者和投资者提供相关金融行为的依据，同时业务企业服务订阅和 API 服务调用带来非常大的日活和交易流量。主要处理流程如下：

1. 收集数据和标准样例数据；
2. 工程实现数据清洗，过滤特殊字符、无效词组、括号内容；
3. 大模型智能化提取和结构化格式输出；
4. 数据处理，删除无符合规则的内容，比如“提供”等动词、或者服务中“进出口服务、技术咨询”等空洞字眼。

业务处理效果：

1 输入数据企业经营范围

< 经营范围 > 生产配合饲料、浓缩饲料（凭有效的饲料生产企业审查合格证经营），生产添加剂预混合饲料（凭有效的添加剂预混合饲料生产许可证经营），养殖畜禽种苗（限分支机构凭种畜禽生产经营许可证经营）以及上述产品自销。饲料原料贸易、饲料技术研发、技术服务及进出口业务（涉及配额许可证管理、专项规定管理的商品按照国家有关规定办理）。（依法须经批准的项目，经相关部门批准后方可开展经营活动）</ 经营范围 >

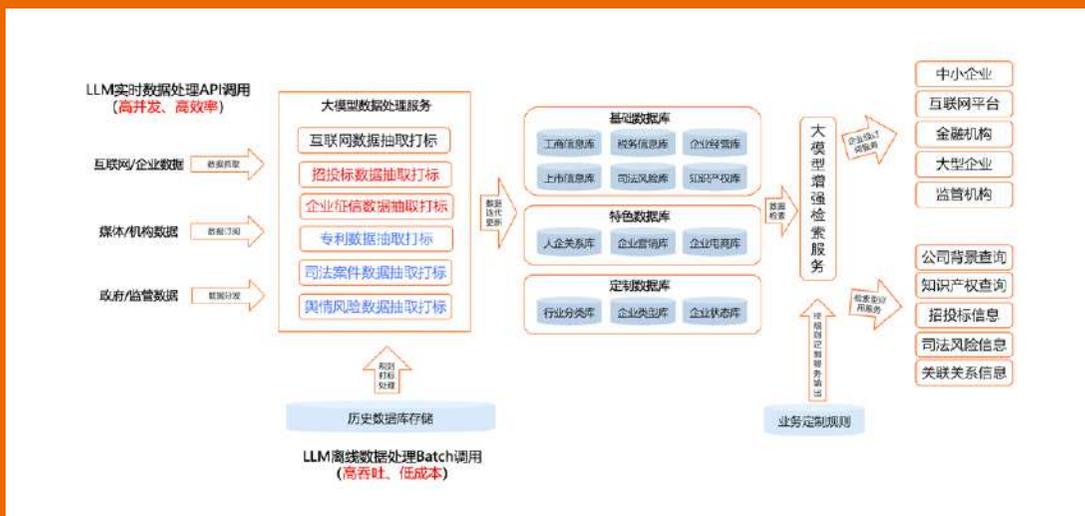
请对该数据进行分类和抽取

2 提取结果

```

{" 服务 ":
[" 养殖畜禽种苗 ", " 饲料原料贸易 ", " 技术服务 ", " 进出口业务 "],
" 产品 ": {
" 设计 / 开发 / 技术研发 ": [" 饲料技术研发 "], " 生产 / 制造 / 加工 ": [" 配合饲料 ", " 浓缩饲料 ", " 添加剂预混合饲料 "],
" 销售 ": [" 配合饲料 ", " 浓缩饲料 ", " 生产添加剂预混合饲料 ", " 畜禽种苗 "]
}}

```



价值分析

📈 数据提取效率

大数据部门可以利用大模型实现招投标和企业经营信息的自动解析与分析，减少人工工作量，提升分析效率和准确性。

🏛️ 监管与合规

监管机构可通过大模型实现对企业财报的自动化监控，及时发现异常和违规行为，保障市场公平和透明。

📊 投资分析与决策支持

投资机构利用大模型对大量企业财报进行快速解析和比较，辅助投资决策，提高投资回报率。

可信搜索

Agent 概述

近年来，随着互联网、云计算、人工智能等技术的飞速发展，大模型在金融行业广泛应用，但是受限于金融行业的特殊性，数据不可出域，同时内网、办公网对外网的封闭，因此在大模型应用场景下的数据和信息相对比较滞后，对于互联网上的实时信息和数据需要有一个入口提供，因此对于各大金融机构引入可信搜索服务的主要驱动因素在于提升金融大模型的实时性和全面性。

通过可信搜索在意图识别、策略优化和数据 rerank 上的差异化能力，打造高质量的通用搜索查询服务，帮助大模型应用获取最新的事实和数据，提升大模型实时问答、提高问答质量和响应速度，以满足不同场景的应用需求，从而提升自身服务质量和用户体验。

需求分析

当前，尽管金融机构在自己内部的向量知识库在知识召回方面表现出较高的准确度，但面对金融行业的快速变化和用户对即时信息的迫切需求，现有知识库的实时更新能力显得不足。

引入可信搜索服务的主要驱动因素在于提升金融大模型的实时性和全面性。尽管内部向量知识库在知识召回方面表现出较高的准确度，但面对金融行业的快速变化和用户对即时信息的迫切需求，现有知识库的实时更新能力显得不足。通过整合求是网、人民网、证监会指定信息披露网站巨潮资讯网等权威网站，外部可信搜索服务的集成能够提供即时的、动态的、可信的金融信息和资讯。基于可信搜索智能体，提供可信搜索服务交互和体验，提供互联网最新可信金融政策和市场动态，以提高大模型生成内容的准确性和深度，并满足一线业务人员对实时金融信息可用性、时效性的需求，从而提升用户体验与满意度。

典型案例

模力时刻

Qwen-72B

处理 理解客户查询意图，改写 Query、优化搜索条件并对搜索结果进行精排，实时返回最新互联网信息和数据。

效果 查询结果准确性 90%，查询服务可用性 99.5%。

某国内领先的综合性证券公司，致力于利用 AI 提升内部运营效率和客户服务质量。面临的挑战主要包括：营销话术生成时缺乏足够的素材，智能客服系统需要补充更广泛的通用知识，以及智能投研活动亟需实时的行业数据支持。这些问题凸显了在资料丰富性、知识广度及数据即时性方面的需求，以提升各应用领域的效能与智能化水平。

可信搜索服务可支持丰富的参数配置，灵活选择和过滤所需数据，以适应垂直业务场景；支持全文截取，适应内网安全限制，提供完整内容而非仅 URL 链接；支持高实时性，便于及时获取最新金融资讯，支持快速决策。提供精排的搜索结果，提高数据获取精准度，协助准确捕捉金融市场动态。通过整合可信搜索服务，大模型应用的效果得到了显著提升：

1 提升大模型实时问答能力

通过实时搜索功能，为大模型提供最新的信息和数据，增强其回答实时性问题的能力。

2 简化大模型应用开发

提供 Query 改写、搜索条件优化、精排等优化，降低大模型应用开发难度，减少模型调优工作量。

3 提高大模型问答质量

多源融合的方法直接提供高质量的问答结果，还为客户模型提供多模态校正和推理优化的基础，从而显著提升整体问答性能。

4 多维度数据源搜索集成

通过整合这些垂直领域数据源，为用户提供更全面、个性化的搜索体验，满足多样化信息需求。

5 加速大模型应用响应速度

提供流式接口，以最快速度输出原始内容，保障大模型数据应用的快速响应。



价值分析

提升信息准确性和时效性

大模型能够获取到最新鲜的信息和事实，确保提供的内容具有高度的准确性和时效性。这对于需要实时更新的应用场景尤为重要，如金融市场分析、新闻报道等。

丰富知识图谱

可以帮助大模型不断扩充其内部的知识图谱，使其涵盖更广泛的主题和领域。这不仅增强了模型的理解能力，还提高了回答问题的全面性和深度。

支持复杂决策制定

对于需要综合多种因素进行决策的应用场景（如智能投研），可信搜索服务提供了重要的参考依据。通过整合多渠道的数据资源，大模型能够帮助用户做出更为明智和基于数据驱动的决策。

增强用户体验

通过提供更精准、更丰富的搜索结果，可信搜索服务有助于提高用户的满意度和粘性。无论是改善搜索引擎的结果质量，还是提升智能客服的响应效率，都能有效增强用户体验。

个性化服务优化

大模型可以更好地理解用户的需求和偏好，从而提供更加个性化的服务和建议。

企业招投标信息打标

Agent 概述

根据各个渠道获取的企业招投标信息，对其进行分析、加工、整理，形成系列服务和产品，助力广大小微企业在生产经营活动中降本增效。为每个中小企业建立一份信用档案，中小企业可以认领、维护自己的信用档案，并在未来的招投标、融资、招聘等诸多场景下用于增信。

通过大模型 PE 构建处理招投标信息 Agent，能够解决数据的准确性，高质量的信息提取和分类需要对模型进行精细调整，以便理解不同类型的招投标文档。同时，系统的处理效率也极为重要，必须能够快速处理海量文档，以满足企业日常运营的需求。还需考虑各企业在信息提取和分类上的特定需求，关注不同的行业或文档格式，以便实现系统的定制化。

需求分析

在招投标文档处理中，主要面临的痛点包括格式和内容的多样性与复杂性，这导致了解析过程中可能出现大量非结构化数据，从而增加了解析的难度。此外，利用大型模型的有效性依赖于大量高质量的标注数据，缺乏足够的标注数据将可能导致模型性能不佳。同时，如果需要处理多语言信息，设计支持多语言的模型也增加了整体系统的复杂性。

典型案例

模力时刻

Qwen-plus

处理 基于原始的半结构化招投标描述，大模型按固定 JSON 格式提取出结构化专利描述。

效果 内容提取准确率 92%。

某征信公司通过对招投标信息提取和分类，快速构建和更新对外年报检索知识库，为金融从业者和投资者提供相关金融行为的依据，同时业务企业服务订阅和 API 服务调用带来非常大的日活和交易流量。主要处理流程如下：

1 信息提取

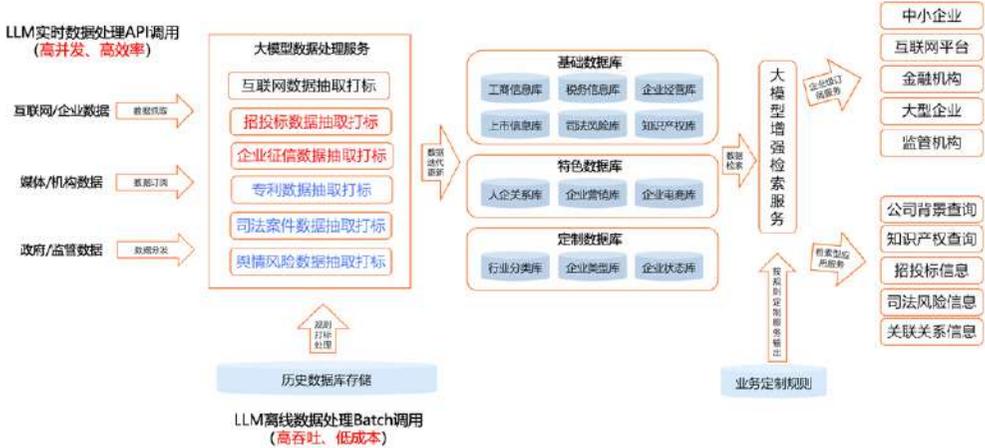
该公司利用大型模型进行招投标文档的句子级和段落级信息抽取，提取关键数据，例如投标公司名称、项目名称、投标截止日期、预算金额和技术要求等。

2 信息分类

在提取信息的基础上，进行多维度的分类，分类标准包括行业类别、项目规模、地区以及供应链相关性等。

3 数据预处理与清洗

对文本数据进行清洗和标准化，去除无关信息和噪声，以确保数据的准确性和可用性。



The screenshot shows a document analysis tool interface with four main sections:

- 超大文件快速读取和加载** (Large File Fast Reading and Loading): Shows a '招标文件' (Bidding Document) with a red seal.
- 结构定义灵活便捷Json** (Flexible and Convenient JSON Structure Definition): Shows a JSON structure for defining document fields like '项目名称', '项目金额', '投标人', etc.
- PE构建 Prompt业务定义和规则** (PE Construction Prompt Business Definition and Rules): Shows a complex JSON structure for defining prompts and rules for document analysis, including '项目信息', '投标人信息', '招标信息', etc.
- 结果验收灵活验证和结果标注测评** (Flexible Verification and Result Annotation Evaluation): Shows a list of results for document analysis, including '项目名称', '项目金额', '投标人', etc.

价值分析

📈 提高效率

自动化的信息提取和分类显著减少了人工审查和数据输入的时间，提升了企业的响应速度。

🔍 洞察挖掘

通过分析历年的招投标数据，企业可以获得市场趋势和竞争对手的策略洞察。

⬇️ 降低成本

通过减少对人力资源的需求，企业可以在数据处理上节省大量成本。

企业知识产权信息打标

Agent 概述

根据各个渠道获取的企业知识产权信息，进行分析、加工、整理他们的信用状况，形成系列服务和产品，助力广大小微企业在生产经营活动中降本增效。为每个中小企业建立一份信用档案，中小企业可以认领、维护自己的信用档案，并在未来的招投标、融资、招聘等诸多场景下用于增信。

通过大模型 PE 构建企业知识产权信息打标的 Agent，能够通过大模型数据处理能力提升数据质量、一致性以及安全性。大模型通过先进的自然语言处理技术，能够自动识别和清洗来自不同来源的知识产权数据，无论格式如何，均可实现高效的统一处理。这种智能化的数据预处理能力不仅能提升数据的一致性，还能有效提高数据的整体质量，为后续分析奠定坚实基础。与传统规则模型相比，大模型具备更高的灵活性和精准度，能够适应多样化的知识产权信息背景，并能根据实际应用需求进行动态调整和优化，从而实现更为精准的打标，提升信息处理的效果。

需求分析

企业知识产权信息打标的业务痛点主要体现在数据质量和一致性、传统规则模型的适应性以及信息安全与隐私保护三个方面。首先，知识产权数据来源广泛，格式各异，质量参差不齐，因此需要有效的数据清洗和统一处理方法。其次，虽然传统规则模型在处理速度上具有优势，但其灵活性和精准度不足，导致效果不尽如人意。最后，在处理企业敏感数据时，尤其是在跨境数据流动时，必须高度重视信息安全和隐私保护。

典型案例

模力时刻

Qwen-plus

处理 基于原始的半结构化专利描述，大模型按固定 JSON 格式提取出结构化专利描述。

效果 内容提取准确率 92%。

某征信公司通过对企业知识产权信息进行打标和分类，快速构建和更新对外年报检索知识库，为金融从业者和投资者提供相关金融行为的依据，同时业务企业服务订阅和 API 服务调用带来非常大的日活和交易流量。主要处理流程如下：

1 数据收集与整理

从各种来源收集企业的知识产权信息，如专利数据库、商标注册信息和版权记录。对数据进行预处理，包括去重、清洗和格式化。

2 大模型打标与优化

使用现有的大语言模型进行打标，以适应具体的知识产权信息提取任务。根据具体的任务需求，可能需要对模型进行业务定向分析，以改进其识别和分类能力。

3 信息提取

利用大模型自然语言处理技术从文本中提取关键信息，如专利编号、申请日期、发明人等。处理非结构化数据，并将其转化为结构化格式以便进一步分析。



价值分析

提高效率

自动化工具减少了对人工提取和分类的依赖，节省时间和人力成本。

增强竞争力

通过及时准确地获取和分析知识产权信息，企业可以更好地进行战略决策和竞争分析。

创新驱动

帮助企业识别未被充分利用的技术领域和潜在合作伙伴，促进创新。

风险管理

有助于识别潜在的知识产权侵权风险，减少法律争端。

财报解析

Agent 概述

根据财报解析企业财务健康状况和经营绩效，按照关注点进行年报和财报解析、打标和摘要总结，利用大模型技术在财报解析中的应用，可以有效解决现有的痛点和需求，实现自动化、高效化、智能化的财报分析，为企业和投资者提供强有力的支持。

通过大模型和金融专业数据解析能力构建财报解析 Agent，能够有效应对上述痛点。首先，它利用先进的自然语言处理技术，自动识别和分析财报中的复杂数据、文字说明和表格，从而实现精准的信息提取和结构化处理，显著减少手动解析所需的时间和可能的错误。其次，大模型具备强大的自适应能力，能够根据不同企业和年份的财报格式及内容差异，灵活调整解析策略，提升自动化处理的效率和准确性。此外，凭借其深厚的语言理解能力，大模型能够正确解读财报中的专业术语和财务指标，并深入分析其上下文关系，从而为企业提供更全面和深刻的财务洞察。这些优势使大模型在财报解析中成为一种理想工具，极大地提升了企业的决策支持能力。

需求分析

财报解析面临多个显著挑战。首先，财报通常包含大量复杂的数据、文字说明和表格，信息结构繁琐，手动解析不仅耗时且易出错。其次，不同企业和年份的财报在格式和内容结构上存在显著差异，这增加了自动化处理的难度。此外，财报中涉及的专业术语、财务指标及其解释，以及上下文关系，均需具备较强的语言理解能力，使得文本理解与语义分析变得更加复杂。

大模型为财报解析提供了一种高效且智能的解决方案，能够有效应对上述痛点。首先，它利用先进的自然语言处理技术，自动识别和分析财报中的复杂数据、文字说明和表格，从而实现精准的信息提取和结构化处理，显著减少手动解析所需的时间和可能的错误。其次，大模型具备强大的自适应能力，能够根据不同企业和年份的财报格式及内容差异，灵活调整解析策略，提升自动化处理的效率和准确性。此外，凭借其深厚的语言理解能力，大模型能够正确解读财报中的专业术语和财务指标，并深入分析其上下文关系，从而为企业提供更全面和深刻的财务洞察。这些优势使大模型在财报解析中成为一种理想工具，极大地提升了企业的决策支持能力。

典型案例

模力时刻

Qwen-plus

处理 基于企业发布的财报文件，大模型理解其内容并按指令 Prompt 生成结构化字段。

效果 内容提取准确率 95%。

某金融客户通过对年报财报解析，快速构建对外年报检索知识库，为金融从业者和投资者提供相关金融行为的依据，同时业务企业资讯 APP 带来非常大的日活和交易流量。主要处理流程如下：

1 多模态大模型的处理与信息提取

利用大模型对财报中的文字内容进行识别和分类，自动提取关键部分如利润表、资产负债表、现金流量表等。实体识别与关系抽取：通过模型识别财报中的关键实体（如公司名称、财务指标）及其关系，提取结构化数据。

2 数据标准化与结构化处理

通过大模型可以学习不同财报格式的特点，自动将非结构化数据转换为统一的结构化格式，便于后续分析。利用模型进行数据清洗，处理缺失值和异常值，保证数据的完整性和准确性。

3 语义理解与情感分析

大模型具备强大的语义理解能力，能够准确理解财报中复杂的语句和专业术语，提取有价值的信息。通过分析管理层讨论与分析（MD&A）等部分的文字内容，判断企业的经营状况和未来趋势。



价值分析

🕒 企业内部财务分析

企业财务部门可以利用大模型实现财报的自动化解析与分析，减少人工工作量，提升分析效率和准确性。

📋 监管与合规

监管机构可通过大模型实现对企业财报的自动化监控，及时发现异常和违规行为，保障市场公平和透明。

📊 投资分析与决策支持

投资机构利用大模型对大量企业财报进行快速解析和比较，辅助投资决策，提高投资回报率。

交互式培训助手

Agent 概述

随着数字化时代的到来，传统的培训方式正面临着激烈的变革需求，特别是在真实作业环境的模拟和学员实际参与度方面。传统培训属于“沉默式”自主学习，主要是提前准备好文档、ppt、知识库，依靠学员的自主性，按时保质完成学习和考试，过程中也无实时的反馈与指导。在时间碎片化的大背景下，学员很难有意愿抽出大段时间完成枯燥的学习。

为了克服这些痛点，交互式培训 Agent 应运而生，旨在通过智能化和互动化的方式提升培训效果。该 Agent 采用先进的大语言模型技术，利用其强大的意图理解、话术生成、工具调用等能力，模拟实际工作场景，增强学习体验，通过个性化学习与实时反馈，帮助学员更深刻地理解知识并应用于实际工作中，从而有效提高学习效果。

需求分析

当前培训存在以下几个主要痛点：

培训与真实工作环境存在差距

无法模拟真实作业环境，传统培训缺乏实际操作机会，影响学员的实用能力和工作表现。

培训灵活性不足

传统规则配置的培训方式难以快速适应行业变化，导致内容维护不及时，影响培训的有效性。

典型案例

模力时刻

Qwen-72B

处理 基于业务目标，生成培训考题，以对话的形式与学员互动。

效果 培训考题生成效率提升 100%。

交互式培训 Agent 通过自动出题 + 场景 / 人物设定，模拟真实作业场景，实现以下重点功能：

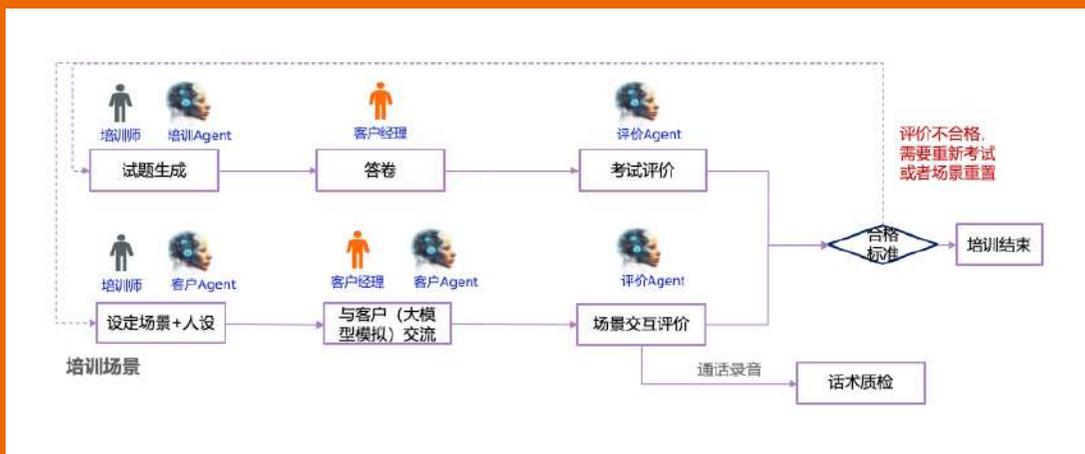
1 多场景人设 Agent 配置

通过大模型模拟客户特点，模拟真实工作场景，让学员在互动中进行实操训练，提升对知识的应用能力。通过小组讨

论、在线问答、角色扮演等多种互动形式，提高学员参与度和学习热情，增强知识的吸收和应用能力。

2 大模型自动出题

公司的业务知识和规章制度实时在变动，传统的人工出题方式耗时较长，导致时效性不足。有了大语言模型和多模态大模型的支持，在感知到知识源头发生变化后，实现分钟级知识变更，客户经理能看到最新的知识。



某金融机构在员工培训中应用了交互式培训 Agent，通过白屏化配置，模拟真实工作场景，增设互动模块，明显提升了学习效果和工作能力。培训师，5min 分钟内完成场景人设的设置，以及题目的自动生成；客户经理以对话交互的形式，与客户 Agent 进行沟通，完成相关任务与题目；还有评价 Agent，针对客户经理的话术实时给出建议，提升实战技能，实现边学边练。

价值分析

当前培训存在以下几个主要痛点：

📁 培训场景更真实

借助大模型角色扮演与 Agent 人设设置，能够模拟真实场景中不同类型客户的角色，让培训变得鲜活，也更能提升员工的实战能力。

📁 知识更新更及时

基于大模型对文本的理解与处理能力，实现自动出题、自助问答等功能，省去了做文字整理与出题等中间环节，让新鲜的知识第一时间可得。

课件生成助手

Agent 概述

在教育与培训日益数字化的背景下，课件制作的效率与质量直接影响教学效果。传统的课件制作流程往往依赖于讲师的专业经验，过程繁琐且缺少互动性。为此，基于先进的大模型技术，课件生成助手应运而生，旨在通过智能化手段提升课件制作的效率与效果，满足教师和学习者的多样化需求。

课件生成助手能够一站式解决 PPT 以及视频课件制作中的关键痛点，提供简便、高效和个性化的课件生成体验。通过集成丰富的功能特性，助力教师高效构建沉浸式的教学内容。

需求分析

目前针对多媒体素材的课件制作存在若干痛点：

专业依赖性

传统方法依赖讲师的丰富经验，使得新手教师难以快速适应课件制作。

耗时费力

大量的非结构化内容，例如：视频、PPT、甚至是图片，需要人先把里面的信息提取出来，再进一步提炼成课件。

典型案例

模力时刻

Qwen-VL

处理 识别 PPT、视频等多媒体课件中内容。

效果 知识提取效率提升 200%。

Qwen-72B

处理 基于培训提纲，生成文本为载体的学习课件。

效果 课件生成效率提升 100%。

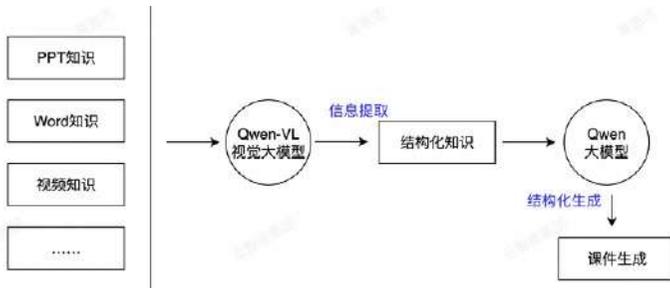
某金融机构试点课件生成助手，辅助讲师快速基于异构数据源生成课件：

1 多源异构信息提取

针对非文本类知识，可以使用多模态大模型完成信息提取。相比传统的 OCR 只能提取字符，Qwen-VL 视觉理解大模型还能理解字符内容，例如：视频的字母、图片的文字描述。因此，可以配上一段提示词，让大模型把视频拆帧、图像中的信息提取并以结构化形式表达。

2 课件生成

提取到结构化知识后，按照课件的章节段落、内容粒度、版面设计等要求形成 Prompt，送给大语言模型进行课件内容（文本）生成；如果需要生成多媒体课件，可以利用大模型 Function Calling 的能力，搭配使用外部工具，如图片生成、PPT 生成等，完成课件的制作。



如上图所示，是一个基于 PPT 生成演讲稿培训课件的示例。通过大模型加持的课件生成 Agent，课件生成时间从以前的周级别，下降到分钟级别，能更好的支撑讲师高效生成不同类型的课件，辅助讲师高效完成相关工作。

价值分析

基于多模态大模型的课件生成助手，极大的提升了课件制作的效率，进而提升培训效果：

📄 课件制作门槛降低

原本需要专业人士制作的课件，现在培训师可以通过白屏化工具一站式完成，显著降低了课件制作的成本。

📈 知识到课件的转化效率提升

以往有新的业务知识或规章制度，制作课件需要以周为单位；现在有了大模型的加持，新的知识最快可以分钟级的转换为课件，被大家获取到学习。

对练机器人

Agent 概述

在金融行业数字化转型的大背景下，传统的员工培训、实践训练方法面临诸多挑战，往往无法满足实际工作需求。一是实施成本太高，需要花费大量人力准备培训内容、考试、反馈、质检等；二是静态的出题培训考试无法模拟真实工作中多样化的场景，影响了员工在真实工作中的表现。

通过对练机器人模拟客户经理营销展业过程中的客户场景及人设，面向客户经理做真实场景的对练培训，核心目标是提升新人客户经理的营销话术及技巧。机器人包括模拟人设 agent 和教练评价 agent，通过评价 agent 对整个对练过程做把控和评价。

需求分析

目前传统的陪练方式存在以下痛点：

场景单一

缺少复杂和多样化的客户场景模拟，员工无法接受真实情境下的挑战。

反馈滞后

缺乏实时指导，错误的处理方式难以及时纠正，影响员工的技能提升。

评估不足

传统的评估往往未能全面反映员工的实际能力，缺乏针对性和灵活性。

典型案例

模力时刻

Qwen-72B

处理 基于陪练角色设定，和客户经理进行多轮对话。

效果 扮演 20+ 不同类型角色，真实演练客户经理服务场景。

金融机构的客户经理以展业营销为主，人员流动性高，新人往往缺少客户洞察和话术引导，转化率不高。通过对练机器人帮助客户经理快速提升展业话术技巧，提高营销转化率。

把一线优秀员工的服务案例沉淀到场景知识库中，既做了相关经验的沉淀总结，又可以利用大模型对练机器人，带动新人客户经理在同类型场景下进行对练实践，让客户经理对该行业有认知，随时随地利用碎片时间对练起来。

1 场景知识库

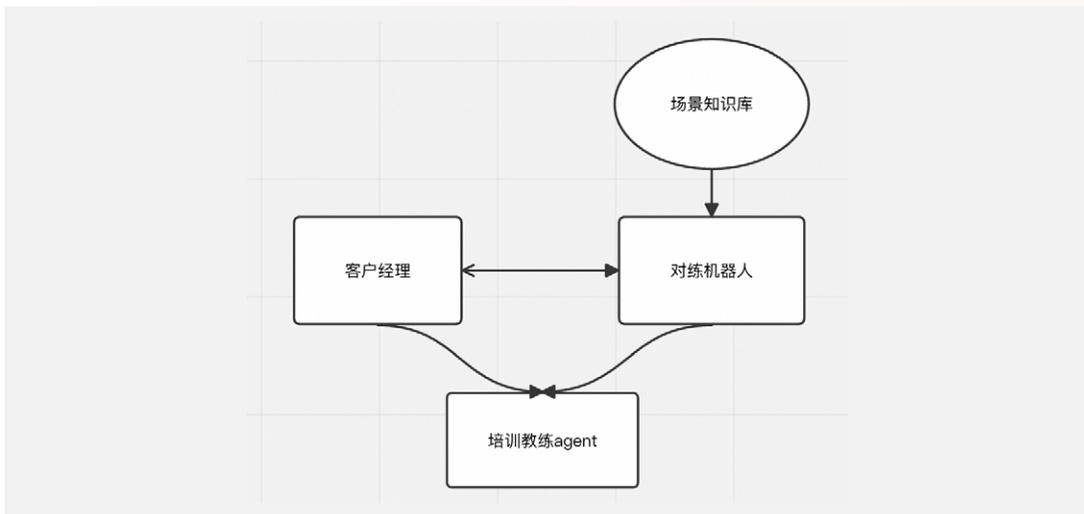
依赖行方沉淀的在营销展业过程中的代表性案例，包括相对应的客户行业、场景、客户的人设、优秀展业话术等。

2 对练机器人

客户经理的每次对练过程，从场景知识库选出一个案例进行对练。对练机器人通过大模型 Prompt，模拟场景中客户的人设（高冷、健谈、所在的行业背景、教育水平、关注点等），尽可能的反应案例中客户的人设，与客户经理展开对。

3 培训教练 agent

以第三方的角色，观察客户经理与对练机器人的对话。用途：纠正对练机器人的答复更贴切人设，整个对练过程的控制（中止、打断），以及在一次完整对练后对客户经理的话术进行评价，并给出相关建议。



价值分析

新人客户经理快速成长

通过教练机器人，员工在每次对练后能迅速获取详细的表现反馈，有效针对性地进行调整和进步。

沉淀行方知识与经验

通过把优秀员工服务案例沉淀到场景知识库中，能把行里的知识传承下去，从个体优秀转变成整体优秀。

模拟真人对话实践

对练机器人可以根据场景知识库中不同的案例，以及对应客户的人设，为员工提供了近百种场景的模拟，包括投诉处理、服务咨询等，员工能够在真实的对话中练习应对策略。

智能培训问答助手

Agent 概述

培训流程中一个重要环节就是问答，学员在学习过程中以问答的方式快速检索到所需知识。传统知识问答通常基于规则或统计模型，依赖于预先把培训的课程和教材，构建知识库，再通过查询这些结构化的知识来寻找问题的答案。这种方式受限于其知识库的大小和质量，生成的回答往往是固定的或模板化的，灵活性较差。而且问答链路当中意图识别等决策点需要预先枚举全部规则，配置周期长、成本高。

在 GenAI 时代，基于通义大语言模型及多模态大模型的技术方案构建智能问答助手，围绕专业类、知识类、工具类三大场景应用，实现培训内容构建、域内知识问答、智能搜推课程等功能，为企业带来生产力的提升和企业培训组织方式的重大转变。

需求分析

现有的传统问答助手搭建，需要进行过多的手动配置，存在人力成本高、耗费时间长、问答效果无法满足培训需求等问题。

成本高

知识库生成

需要依赖特殊的代码库完成文档的信息提取，不支持格式多，且提取效果不佳。

query 意图识别

依赖关键字判断用户意图，且传统的关键词库依赖人工手动枚举配置，成本过高。

时间长

完成知识库构建、意图对应的动作模版构建，需要消耗大量的时间成本，制作一份知识库消耗 2-3 周的时间。

效果不佳

- 手工制作知识库难以保证格式统一；
- 后续检索过程当中依赖关键词做召回，不相关信息多，存在错召 / 零召回现象；
- 用户语义理解困难，使用者体验不佳。

典型案例

模力时刻

Qwen-72B

处理 构建培训陪练知识库，大模型以多轮的形式和学员对话，快速问答。

效果 学员获取知识效率提升 300%。

某 TOP 企业培训公司：

知识库构建



基于通义千问语言大模型、通义千问多模态 VL 大模型、Docmind，提取异构培训教材信息，并向量化入库。工程链路如上图所示：

1 知识库管理

经处理的向量化数据存储于 ADB-PG 向量数据库中，应用系统在用户登录后，按企业 / 组织权限展示 / 召回相关数据，供用户选择。

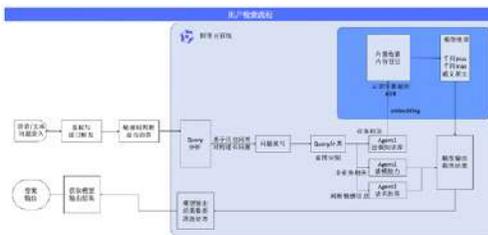
3 模型安全能力

使用 AES/RSA 做 query、回复的加密传输。

5 文档解析

- 图片使用 VL 能力解析
- 语音使用阿里云 paraformer / SenceVoice 模型解析
- mp4 经过策略抽帧后，将图片传给百炼解析
- 其他特殊格式依赖用户当前数据工程链路处理后，将图片等格式内容传给 VL 进行解析

问答助手构建



query 先经过客户侧敏感信息判断；而后经过意图识别 agent，通过基模能力判断业务需求，流转给下游不同的 agent 做响应；最后输出 agent 回复结果。工程链路如上图所示：

2 多模态文档处理

通过通义千问 VL 模型、文档智能的协作，对课件文件的文字内容、图片内容、表格内容，做充分的提取，形成 markdown 结构化数据。

4 意图识别 agent:

- 意图 1：业务相关问题，结合 RAG+ 基模回答
- 意图 2：非业务相关问题，闲聊，基模回答
- 意图 3：基模判断包含敏感信息，设置特定话术拒绝回答

公司采用上述基于大模型的知识构建与问答架构后，技术价值明显：

- 服务并发提升：大幅提升模型并发，大幅加速模型训练、模型推理、文档解析效率，解决原环境中中文档解析慢的问题。
- RAG 向量化增强：结合阿里云向量化模型能力，分词 + 向量双路召回，大幅提升召回率与问题回答准确率（依据 POC 测试数据）；ADB 底层权限管控多样化，可以针对企业 / 组织进行表 / 行级别的权限管控。
- Multi-Agent 应用架构搭建：结合多智能体编排能力，将 query 处理链路多样化，满足业务侧对不同分支剧本的要求。
- 意图识别精准：提升意图识别准确率至 95%，全面升级 RAG 入口用户交互体验。

价值分析

通过 AI 大模型问答能力构建企业分权限隔离的 AI 知识库，结合持续训练调优，形成企业专属知识大脑。借助企业知识库及大模型生成各类方案，高效传递并掌握知识，赋能业务运营过程，提升业务绩效。

使用 AI 大模型问答对话能力、丰富的多领域知识，特定或开放场景下提供模拟实战的、个性化的、实时互动的陪练体验，并根据个人的情况不断调整和拓展训练内容，高效训练快速掌握技能。

智能考试助手

Agent 概述

在数字化学习日益普及的背景下，传统培训场景考试的方式正面临着考试效率与评价客观性的双重挑战。考官和学员之间的互动性不足，考试准备和反馈过程流于形式，亟需一种高效率、智能化的考试解决方案。基于生成式人工智能技术，智能考试助手应运而生。使用大语言模型，自动从原始学习材料中抽取大量的题目—答案对，形成题库；进一步，出题人将考察要点组织成 Prompt，让大模型实现智能自动组卷；考试结束后基于大模型自动阅卷，提升考试全流程端到端效率。

需求分析

目前的传统考试存在以下痛点：

出题效率低且知识覆盖不足

传统出题依赖人工，效率低且同质化现象严重，难以满足多样化的考核需求。

出卷过程耗时且更新不够及时

考试内容难以与行业前沿知识同步，导致考核的时效性和相关性不足。

阅卷主观性高且效率低下

人工阅卷的主观性与差异性影响评分的公正性，导致考生的真实水平难以准确反映，且主观题阅卷强依赖人工，导致效率低下。

典型案例

模力时刻

Qwen-72B

处理 基于考察目标，大模型自动生成考题，根据学员情况组合成试卷，并自动批改。

效果 学员考试效率提升 200%。

使用大模型的智能考试 Agent，能够端到端提升考试的效率，如下流程所示：

1 智能出题

利用大语言模型 + 多模态大模型，分析培训内容及，参考相关知识点提示，自动生成考试问题；通过 Prompt 设定可调参数，确保问题的多样性和覆盖面。

2 智能组卷

一方面，大模型基于试卷的考核要点分类和权重，自动从题库中筛选考题进行组卷；另一方面，针对特定学员，可以结合其历史考试情况，优化试卷难度与题型组合，实现个性化出题。

3 智能阅卷

基于知识库和评分标准，大模型获取试卷题目及答案，进行自动评分阅卷，确保结果的尺度一致客观公正，减少教师的阅卷负担。



价值分析

大模型智能考试 Agent，在出题、组卷、阅卷三个方面，给考试这件事带来端到端的提升：

提升考题设计效率

通过使用智能考试 Agent，显著减少了教师在出题和组卷过程中所需的时间和精力，提高了考题设计的效率。

增强考试的有效性与公平性

智能考试 Agent 不仅帮助自动化了阅卷过程，降低了人为错误的可能性，还确保了评分的一致性和客观性，从而提升了考试的整体有效性和公平性。

促进教育个性化与智能化发展

解放教师从重复性劳动中后，使其能够更专注于了解学员的学习情况和需求，进而提供更加个性化的指导和支持。同时，也为教育领域的智能化发展奠定了基础，有助于进一步增强培训的效果和成效。

工作报告助手

Agent 概述

在企业日常办公中，高效的工作流程和精准的决策支持是提升机构竞争力的关键。工作报告作为日常运营和战略规划的重要工具，其生成过程往往涉及大量资料的整理、分析和总结，传统的手动撰写方式不仅耗时费力，还容易出现信息遗漏或偏差。随着金融业务的复杂化和数据量的激增，金融机构亟需通过智能化手段优化工作报告的生成流程。

工作报告助手是聚焦于办公提效的智能体，旨在简化和优化工作内容提炼与报告编写的流程。通过智能化的数据收集、分析及工作内容提炼总结，该智能体能够避免重复劳动和流水账式的报告展示，帮助用户快速准确地生成公文、工作总结等各类工作报告中标准化的内容，将更多工作精力放在真正有意义的思考和沉淀上。

需求分析

日常工作中工作报告编写主要业务痛点如下：

重复劳动多

用户经常需要手动整理相似的工作内容，耗费大量时间在数据收集和格式化上，导致工作效率低下，难以集中精力于更具创造性和战略性的工作任务。

难以保持一致性

由于团队成员各自为政，使用不同的工具和方法来记录工作进展，这不仅增加了整合信息的难度，也使得跨部门协作时难以维持一致的标准和风格。

信息不准

时政知识更新迭代快导致信息量大且复杂，而个人知识储备不足易出现漏判和误判，影响了报告的专业性和可信度。市场上常规校对软件多以普通差错为主，无法识别复杂语义环境下的错误，而公文一旦出错影响巨大，会导致社会舆情等问题。

典型案例

模力时刻

Qwen-72B

处理 大模型遵循报告撰写要求，基于原始文档和素材，生成模板化的工作报告提纲与具体内容。

效果 工作报告撰写效率提升，从 2 天变成 2 小时。

在某保险机构，基于通义千问大模型，并针对公文领域任务进行参数调优，实现企业内部知识的公文领域解决方案，具备智能审校、智能排版、智能写作、智能检索等模块能力。

1 智能审校

对于政治性差错、知识性差错、基础性差错等都有严格把关，平均准召率可达 85% 以上，准确率高达 99%，领先同行业。

- 政治性差错：重要固有表述差错，重要引用差错，政治专用词汇和术语差错，重要人名和职务差错，重大政治活动盛世表述差错，机构表述差错，地缘表述差错，敏感违禁表述提醒，时事信息表述差错等；
- 知识性差错：专名差错，一键替代差错，事实性差错，引用差错，行业术语差错，字母词差错，纪年年号差错，时间差错，自然科学知识差错，文史哲知识差错等；
- 基础性差错：字词差错，语法差错，标点符号差错，量和单位差错，数字用法差错等。

2 智能排版

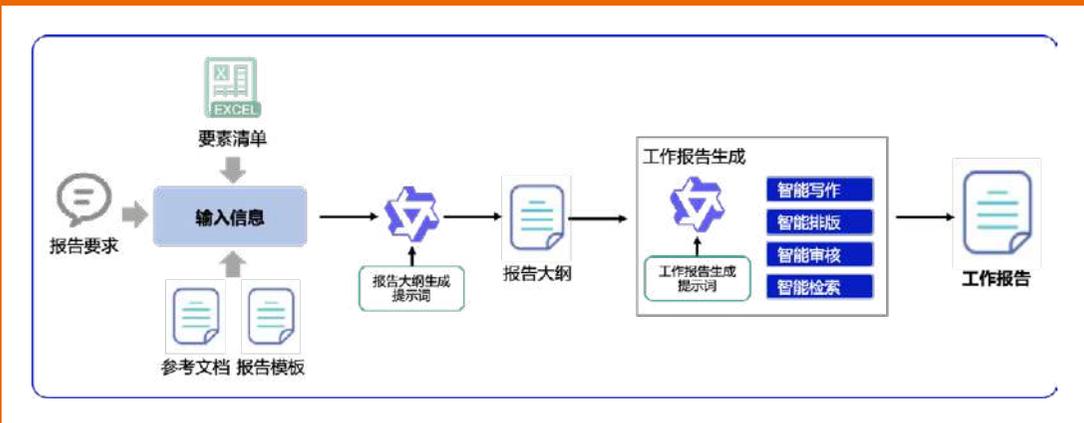
根据样例库中的样例来自动排版，也支持手工调整。

3 智能写作

自动生成全文、提纲、文件解读、问答助理，可完成续写、改写、扩写、提炼总结等工作。

4 智能检索

在支持连接公网的前提下，数据来源于中央级，国家级，省市级等 3000+ 单位官网，央媒党媒省媒，可以实时更新总书记，公文，政策，规章，法律等超 620 万数据；也可支持单位内部文档导入。



价值分析

🕒 专业数据，全链赋能

权威政务数据进行训练，集成多项能力，赋能写作全流程。

- 公文专家数据集：海量公文大数据支持；公文数据时效性强。
- 资料收集整理：写作思路拓展；生成后一键排版。

📁 私有化低成本

私有化信创环境部署，小参数模型，成本低训练推理。

- 私有化训练和部署：训练推理场景化定制；交付及实施成本可控。
- 更新维护训练方便。

🔒 内容安全

涉政内容把关，大模型输入输出全程监控。

- 审核引擎集成保障内容安全：通过网信办大模型备案。

OA 小蜜

Agent 概述

在现代企业环境中，员工经常会遇到各种与 IT 设备管理、行政服务以及人事政策相关的问题。传统的解决方式通常需要员工联系相关部门，等待人工客服解决问题，这种方式效率低下且耗费时间。基于此背景，对于一个能迅速准确回答常见问题、指导操作流程，并能够自动完成某些任务的智能化助手存在明确的需求。OA 小蜜可以提供帮助企业员工更高效地处理日常办公事务的解决方案。根据提问用户的问题进行意图识别，将对应问题的答案以及出处提供给用户。OA 小蜜不仅能够理解自然语言输入，还能够根据企业的具体需求进行定制化开发，提供个性化的服务和支持。通过集成企业内部系统和资源，它能够快速响应员工提出的关于 IT 支持、行政管理和人力资源等多方面的问题。

需求分析

主要业务需求：

自动化

减少人工干预，提高工作效率。

智能性

利用 AI 技术理解复杂查询并给出精确答案。

易用性

界面友好，易于所有员工使用。

可扩展性

可以根据不同企业的特定需求调整功能和服务。

典型案例

模力时刻

Qwen-72B

处理 办公场景中，理解员工 query 的意图，检索内外 部知识，调用接口完成任务

效果 员工意图识别准确率 90%，工具调用准确率 95%

在某金融机构，基于通义千问大模型，帮助企业员工解决办公过程中的高频问题，包括 IT、行政、人力资源等相关问题。当用户（企业员工）输入办公相关问题，先通过大模型识别用户意图，并判断是否涉及私域知识，如涉及，则调用私域 openAPI。通过调用办公智能体集群来实现对不同办公细分问题的支持。

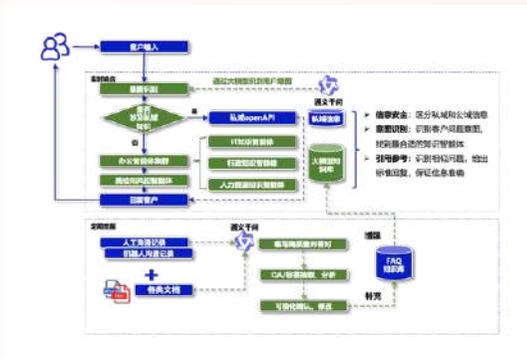
在 Agent 上线后，可以通过大模型实现对各类沟通记录和文档的“定期挖掘”，完成高质量问答对的自动化编写、标签抽取、分析等，补充 FAQ 知识库，达到对大模型知识库的持续增强优化。

具体支持以下功能：

总体技术链路如下图所示：

1 IT 支持

- 当员工询问“我的电脑什么时候到期需要更换？”时，“OA 小蜜”Agent 可以查询员工的资产记录，并告知具体的更换日期。
- 如果想知道“我还可以领用哪些 IT 配件？”，助手会检查库存情况和个人配额限制，列出可领取物品清单。
- 对于“公司 WiFi 怎么连接？”、“打印机怎么安装？”等问题，“OA 小蜜”Agent 提供详细的图文指南或视频教程帮助员工自行解决问题。



2 行政服务

- 员工可以通过“OA 小蜜”Agent 轻松预订会议室，查看可用时间段，并直接发送邀请给参会人员。
- 来访者预约也可以通过该平台完成，包括填写来访信息、选择访问时间和地点等步骤。
- 提供各园区的大巴乘坐时间表和路线图，方便员工规划出行。

3 人力资源

- “OA 小蜜”Agent 针对社保、公积金、商保等福利政策提供详细的解释说明，解答员工疑问。
- 它还能提醒员工重要事项，如保险续费期限、年度体检安排等。



价值分析

提升效率

自动化处理常规事务，让员工专注于核心工作，减少等待时间。

降低成本

减少对专职客服人员的需求，降低人力成本。

改善体验

提供 7x24 小时不间断的服务，增强员工满意度。

促进沟通

加强了部门间的协作，使得信息流通更加顺畅。

数据驱动决策

通过收集和分析用户交互数据，为管理层提供有价值的洞察，有助于优化业务流程和服务质量。

智能会议助手

Agent 概述

智能会议助手是一款基于通义千问大模型的智能解析与结构化处理的智能体工具，专为高效处理会议音视频内容而设计。它能够快速解读会议记录，提取核心信息，生成精准的会议纪要和摘要，同时支持全文语义搜索、智能导航及语音同步定位功能。通过结合自然语言处理与语音识别的前沿技术，智能助手显著提升信息转化效率，帮助用户在投研会议、路演、内部决策和客户沟通等金融场景中更高效地管理和利用会议信息。

该智能体旨在解决传统人工整理会议纪要效率低下、易遗漏关键信息的问题，助力用户捕捉重点、减少错误，为业务决策提供科学支持。它不仅能优化语音转录的准确性和可读性，还能通过通义千问大模型实现全文摘要、关键问答提炼及多场景下的个性化服务，全面赋能用户的会议管理与决策流程。

需求分析

在办公、投研、路演、内部决策和客户沟通等金融核心场景中，用户面临大量复杂的会议音视频数据，需快速获取关键信息以支持高效决策。传统人工整理会议纪要的方式效率低下，容易遗漏重要内容，且信息检索和分析过程繁琐，难以满足及时性和精准性的需求。此外，语音记录中的非标准表达和冗余信息使转录内容的可读性和实用性降低，进一步增加了用户的整理和分析负担。智能会议助手针对这些痛点，提供高效、精准的信息解析和管理能力，以满足用户对实时记录、摘要提取和语义搜索的迫切需求，帮助快速捕捉核心信息，降低人工干预的复杂度，提升业务决策效率。

典型案例

模力时刻

Qwen-72B

处理 会议场景中，基于原始会议文本，自动完成多种任务（摘要、总结、待办）的处理。

效果 生成信息准确率 90%，会议后信息处理效率提升 60%。

针对某头部城商行客户对于会议信息处理的多元化需求，该项目采用了产品与服务深度融合的模式，依托智能会议助手，实现了与银行内部的某三方商业办公系统实现无缝对接。智能会议助手能够高效地完成会议音视频文件的自动解析与优化工作，同时运用智能技术对会议文本数据进行深度分析，提供包括智能摘要、要点导航、AI 问答在内的全方位智能服务。可实现实时识别会议内容，区分发言角色，会议总结生成，提高会议记录效率。会议记录内容统筹管理，精准回溯。



智能会议助手可实现实时识别会议内容，区分发言角色，会议总结生成，提高会议记录效率。会议记录内容统筹管理，精准回溯。支持本地化部署，无需互联网支持，通讯采用加密算法保存数据，信息更安全。基于大模型的会议分析，会后自动整理会议纪要，提取会议关键词，支持对会议内容进行章节速览、发言总结、要点回顾、提取 PPT。

价值分析

提升会议内容处理效率

- 自动生成会议纪要和摘要：通过自动化生成精准的会议纪要和摘要，减少了传统手工整理的时间消耗，同时避免了因人工操作可能导致的关键信息遗漏问题。
- 减少手动干预：降低了对人工记录和整理的依赖，使得参会者可以更加专注于会议本身而非记录过程。

增强信息利用能力

- 快速捕捉核心内容：用户能够迅速了解会议的核心要点，无需花费大量时间去梳理冗长的会议记录，从而提高了信息吸收的速度。
- 支持语义搜索与智能导航：不仅限于关键词匹配，还支持基于语义理解的高级搜索功能，帮助用户更精确地定位所需的信息。智能导航则进一步简化了信息检索的过程，提升了用户体验。

优化信息检索便捷性与时效性

- 提高检索效率：借助先进的语义搜索技术，用户可以在短时间内找到想要的内容，极大地缩短了信息查找的时间。
- 增强时效性：确保重要信息能够被及时发现并应用于实际工作中，有助于加快决策速度。

流程助手

Agent 概述

在现代企业环境中，员工经常需要遵循一系列复杂的业务流程来完成日常任务。这些流程可能涉及到多个部门之间的协作、繁琐的手动数据录入以及对规则和政策的理解。传统的流程管理系统往往缺乏灵活性，不能很好地适应变化的需求或提供足够的用户友好性。

流程助手是基于通义千问大模型，并融合专有 AI 模型与工作流编排工具的企业级智能体应用。目的是帮助企业级客户充分利用企业专有知识库、协同办公系统、第三方信息系统等现有资源，快速、灵活构建流程助手，旨在帮助企业员工更高效地完成流程相关的任务，实现流程问答、流程发起、快速填写表单等能力。

需求分析

一个高效的流程助手需要满足：

即时响应

无论何时何地都能迅速解答关于流程的问题。

自动化流程

减少人工干预，提升效率。

简化表单填写

自动填充常用信息，减少重复劳动。

易于使用

界面直观，适合所有技能水平的员工。

集成现有系统

无缝对接企业现有的 IT 基础设施，如 ERP、CRM 等。

典型案例

模力时刻

Qwen-72B

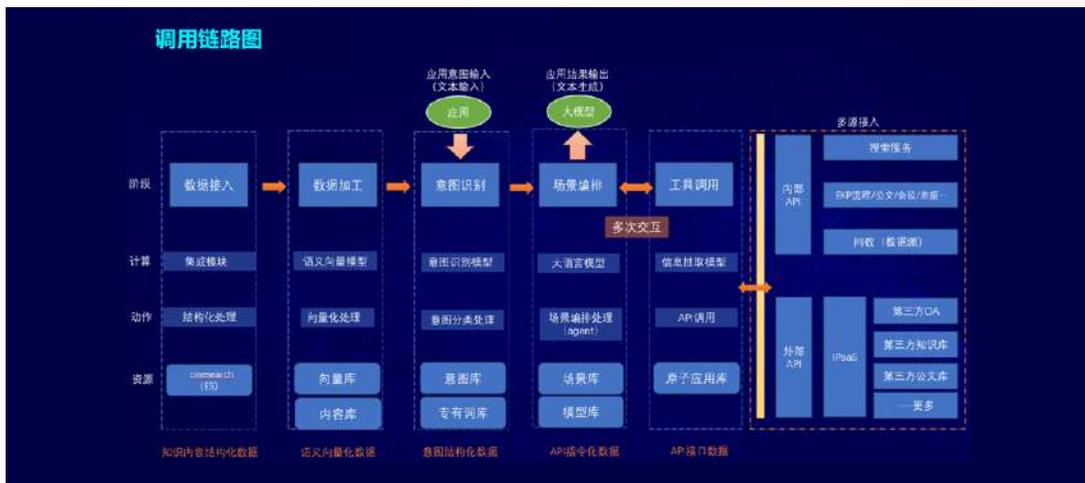
处理 大模型基于用户输入，理解其意图，利用 Function calling 能力调用对应 API。

效果 API 调用准确率 90%。

在某数科公司，基于大模型，进行用户意图识别、场景编排处理、工具调用分析等关键环节，完成流程助手 Agent 的搭建。

在意图识别环节，Agent 利用意图识别模型对用户输入进行分析，参考意图库和专有词库中的信息，精确分类用户

意图，为场景编排提供明确的方向。场景编排环节根据已识别的用户意图选择最合适的操作路径或场景，调用必要的 API 接口来执行具体的任务，支持从内外部多个不同的数据源获取信息，不仅增强了自身的功能范围，也提升了对企业复杂环境的适应能力。



该项目实现了流程的高效发起和知识的快速获取：

1 知识快速获取

为集团用户提供不同场景的知识问答服务、帮助用户找到知识、获取答案，盘活集团丰富的知识资产。

2 流程高效发起

通过流程助手，帮助用户以自然语言对话的形式，准确找到从集团大量的流程中找到所需流程，快速填写表单。

价值分析

引入基于大模型的流程助手为企业带来的价值是多方面的：

提升效率

通过自动化处理常规事务，减少了员工等待时间和重复性工作，使他们能专注于更有价值的任务。

改善体验

提供了 7x24 小时不间断的服务，增强了员工满意度；简单易用的界面设计提高了用户采纳率。

降低成本

减少了对专职流程管理人员的需求，降低了人力成本；同时，由于错误率降低，也减少了因误操作而产生的额外费用。

增强透明度

记录所有的流程交互数据，便于后续审查和审计，同时也为管理层提供了宝贵的洞察力，有助于优化业务流程和服务质量。

促进合规性

确保所有流程都严格按照规定执行，减少了人为疏忽导致的风险。

营销策划助手

Agent 概述

在竞争激烈的金融行业中，有效的营销对于吸引新客户、保持现有客户忠诚度以及推动产品销售至关重要。金融机构不仅需要要在市场上树立独特的品牌形象，还需要针对不同客户群体提供个性化的服务和产品推荐，以增强市场竞争力。然而，传统的营销策划过程往往耗时费力，需要大量的市场调研、创意构思和方案撰写工作。

基于大模型的“营销策划助手”Agent 是一款专为金融行业设计的智能营销策划智能体，旨在帮助金融机构更高效地完成营销策划工作。它集成了营销创意生成、营销方案生成和营销思维导图生成等功能，能够显著提升营销策划效率，降低人力成本。

需求分析

对营销智能助手有以下需求：

快速响应市场变化

及时捕捉市场热点，生成符合当下趋势的营销创意。

自动化方案生成

减少人工编写方案的时间，提高效率。

定制化内容创建

根据不同产品或服务的特点，量身定做营销材料。

可视化规划工具

以图形化方式展示营销思路，便于团队沟通和理解。

风险管理

识别潜在风险并提出应对预案，保障营销活动顺利进行。

典型案例

模力时刻

Qwen-72B

处理 大模型基于营销创意、营销素材、受众特点，生成营销文案与计划。

效果 提升营销人员准备营销方案与物料效率 50%。

在某银行项目中，大模型通过学习产品信息，结合营销热点，自动生成营销方案，思维导图。

1 营销创意生成

当市场部希望围绕“绿色金融”这一主题开展宣传活动时，营销策划助手可以根据当前的环保政策、社会关注点等因素，迅速生成一系列创新性的营销创意，如举办线上环保讲座、推出低碳投资组合等。

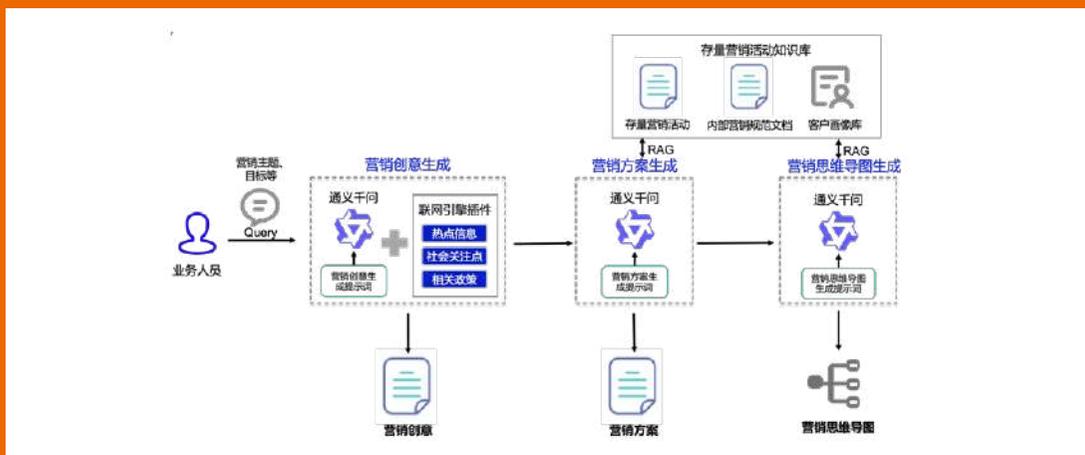
3 营销思维导图生成

在准备大型营销活动前，团队可以通过营销策划助手生成一张涵盖所有关键要素的思维导图，包括不同阶段的任务分配、资源需求、合作伙伴关系等，使得整个项目的结构一目了然，方便讨论和调整。

2 营销方案生成

对于一款新的理财产品，工具会综合考虑产品的特点（收益稳定性、风险水平）、目标受众（年龄层、收入水平）以及市场竞争状况，生成包含以下元素的营销方案：

- 营销主题：如“稳健理财，智享未来”
- 营销目标：设定具体的销售额、客户增长率等指标
- 营销策略：确定广告投放渠道、促销活动形式等
- 执行计划：详细列出每一步骤的时间表和责任人
- 营销文案：撰写吸引人的宣传语句和产品介绍
- 活动预算：估算整个营销活动的成本
- 风险点及预案：预见可能遇到的问题及其解决方案



价值分析

📌 提升创新能力

通过自动化的创意生成机制，激发更多的灵感，避免陷入固有的思维模式。

📈 提高效率

大幅缩短了从概念到实施的时间周期，使营销活动能够更快地上线。

📊 增强精准度

利用大数据分析和机器学习算法，确保营销方案更加贴合市场需求和消费者偏好。

👥 促进协作

思维导图等功能促进了跨部门间的交流与合作，提高了决策透明度。

⚠️ 降低风险

提前识别潜在的风险因素，并制定相应的应急预案，减少了不确定性和损失的可能性。

📊 数据驱动决策

积累大量营销活动的的数据，有助于深入分析效果，持续改进未来的营销策略。

营销素材生成

Agent 概述

在当今数字化营销领域，营销素材生成作为吸引新客户、保持现有客户忠诚度以及推动产品销售的关键工具，正扮演着越来越重要的角色。随着市场的发展，金融机构意识到仅仅依靠传统的广告和推广方式已难以满足客户需求，也无法在众多竞争对手中脱颖而出。为了在市场上树立独特的品牌形象，并针对不同客户群体提供个性化的服务和产品推荐，金融机构需要更加精准、高效的营销策略支持。

基于大模型的“营销素材生成”Agent 是一款专为市场营销人员设计的智能解决方案。它能够根据不同的营销场景自动生成高质量、多样化、个性化的营销内容。这些内容可以是文本、图像、视频等不同形式的素材，以满足企业在各个数字营销渠道上的需求。

需求分析

金融机构的营销痛点：

素材生成效率低

传统营销素材制作依赖人工设计，耗时长且成本高。

多渠道适配复杂

不同渠道（如邮件、短信、APP、社交媒体）对素材格式和内容要求不同，适配工作繁琐。

缺乏个性化推荐

难以根据客户画像和行为数据生成个性化推荐内容，导致营销效果不佳。

合规性要求高

金融行业监管严格，营销内容需符合合规要求，人工审核效率低。

典型案例

模力时刻

Qwen-72B

处理 大模型基于营销创意与目标，生成个性化的营销文案。

效果 提升营销人员准备文案效率 100%。

Wanx2.1

处理 基于用户提示词生成营销图片。

效果 设计师画图效率从 1 天到 1 小时。

在某银行项目中，“营销素材生成”Agent 根据某一营销活动的策划方案要求，结合企业存量营销素材的规范和风格，制作该营销活动的多渠道营销素材，解决传统方式下素材制作耗时长、难以实现个性化的问题。

1 营销素材生成

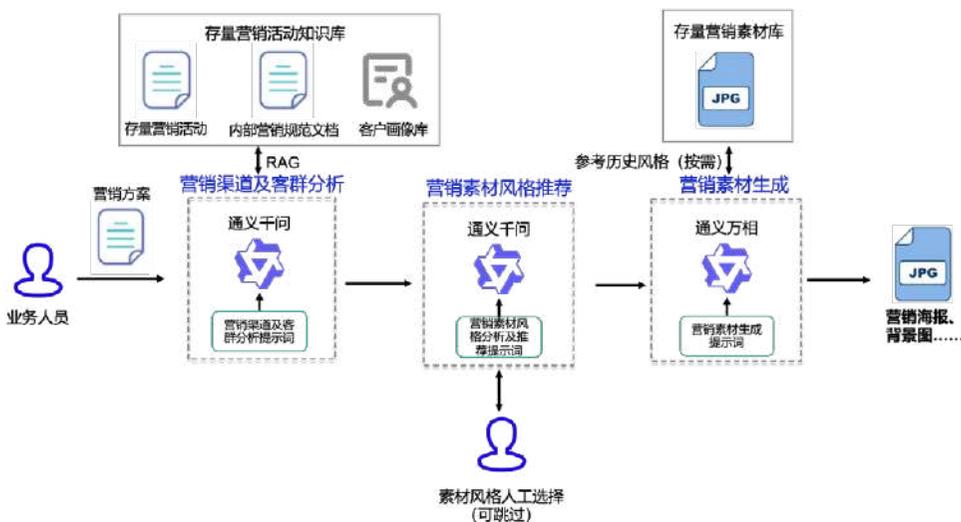
使用素材生成功能，输入产品信息、目标客户画像、营销方案、存量素材风格等信息，Agent 自动生成多套营销背景图、海报等营销素材。

2 多渠道适配

工具自动将生成的素材适配到邮件、短信、APP 推送、微信、抖音等渠道，确保内容格式和风格一致，确保素材在各渠道的显示效果。

3 智能推荐

基于历史素材及本次营销活动相关信息，智能实现素材风格的选择和生成。



价值分析

成本节约

减少了对外部设计师和文案撰写者的依赖，降低了人力成本。

竞争优势

借助先进的 AI 技术和大数据分析，使企业在激烈的市场竞争中保持领先。

用户体验

更贴近用户的个性化内容增强了品牌的吸引力，促进了更高的用户参与度。

灵活性与可扩展性

无论是小型创业公司还是大型跨国企业，都可以根据自身规模选择合适的服务方案，并随业务发展灵活调整。

营销文案生成

Agent 概述

在当今数字化营销领域，营销文案的质量和准确性对于建立信任和吸引客户至关重要。面对快速变化的市场环境和日益增长的客户需求，如何高效地生产出既能反映品牌价值又能触动客户心灵的营销文案，成为了金融机构提升营销效果、保持竞争优势的关键挑战。因此，探索更加智能化、个性化的营销文案生成方法，成为行业发展的必然趋势。

基于大模型的“营销文案生成”Agent 为金融行业需求提供智能解决方案，旨在通过 AI 技术帮助机构快速生成高质量、个性化的营销文案，并智能推荐最优的营销策略。该 Agent 集成了文案生成、多渠道适配等功能，能够显著提升营销效率和客户转化率。

需求分析

结合金融机构营销诉求和特点，营销文案有如下需求：

快速响应市场变化

传统文案创作依赖人工撰写，耗时长且成本高，应该随着金融市场和政策的变化，及时更新营销策略。

多渠道适配

不同渠道（如邮件、短信、APP、社交媒体）对文案格式和内容要求不同，适配工作繁琐。

个性化定制

根据不同客户群体的需求和偏好，定制化营销信息。

专业合规

文案必须准确传达复杂金融产品的信息，避免误导消费者，且符合金融监管机构的规定，如广告宣传不得夸大收益或隐瞒风险。

典型案例

模力时刻

Qwen-72B

处理 大模型基于营销知识库，生成不同渠道多样风格的营销文案。

效果 提升营销人员准备文案效率 100%。

在某银行项目中，打造基于大模型能够快速生成营销文案的 Agent 应用，触发渠道覆盖小红书、抖音、微信等，需要针对产品情况和渠道特色生成个性化文案。

1 营销文案生成

根据用户提供的信息和设定的目标自动创作出高质量、合规且具有吸引力的营销文案，能够理解并处理复杂的金融概念，确保生成的内容既专业又易于理解。

- 智能填充：通过输入少量关键信息（例如产品特点、目标受众），系统可以自动填充文案中的具体内容。
- 多语言支持：支持多种语言的文案生成，满足国际业务需求。
- 创意与优化：结合市场趋势和客户反馈不断优化文案内容，保持文案的新鲜感和竞争力。

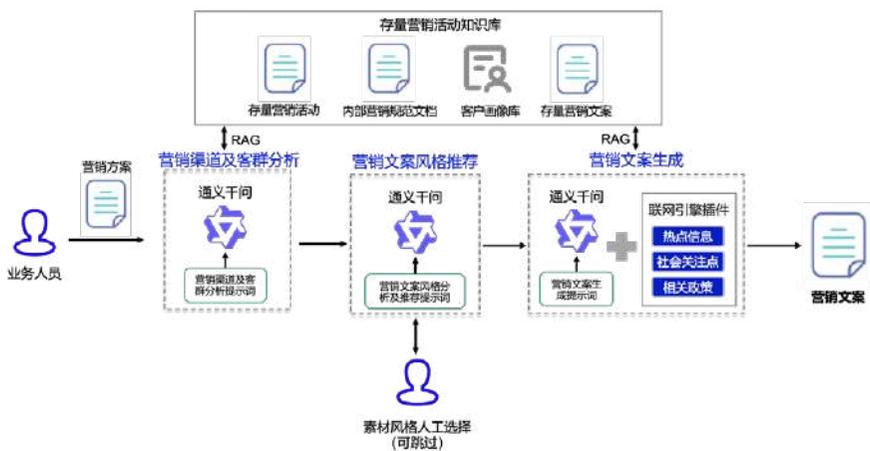
2 多渠道适配

使生成的营销文案能够无缝应用于不同的数字营销平台，包括但不限于银行网站、社交媒体、电子邮件营销、移动应用等，确保文案在各种媒介上的展示效果，并且保持品牌形象的一致性。

- 风格匹配：适应每个平台的用户群体和互动方式，调整文案语气和风格，以提高用户的接受度。
- 视觉元素整合：提供适合各平台的视觉设计建议，如图片选择、视频剪辑等，增强整体传播效果。

3 法规遵从检查

可以内置审核机制，确保文案符合当地法律法规要求，避免法律风险。



价值分析

📈 提升营销效率

通过自动化生成和智能推荐功能，大幅缩短营销文案的制作时间。

⬇️ 降低成本

通过自动化工具减少人工干预，降低营销成本，提升投资回报率（ROI）。

👤 增强客户体验

多渠道适配功能确保客户在不同平台获得一致的品牌体验。

🛡️ 合规与安全

自动检查文案是否符合金融法规要求，减少因违规带来的法律风险。

营销素材生成

Agent 概述

在当今竞争激烈的金融行业中，社群运营已成为金融机构提升客户粘性、增强品牌影响力的重要手段。随着社交媒体和数字化平台的普及，客户期望通过更加互动和个性化的方式与金融机构建立联系。社群运营不仅能够帮助金融机构实时了解客户需求，还能通过精准的内容推送和互动活动，提升客户参与度和忠诚度。然而，传统的社群运营方式往往面临内容单一、互动不足、数据分析不精准等挑战，难以满足日益多样化的客户需求。

基于大模型的“社群运营助手”Agent 能够在社群运营之初，依据具体需求高效自动生成定制化的运营计划。在执行过程中，大模型能够实时整合客户信息与过往沟通记录，智能推荐精准匹配的营销话术，以提升互动效果。运营活动结束后，AI 还能协助对客户进行细致标签分类，并沉淀优秀案例，为后续运营提供宝贵经验与参考。

需求分析

客户现有的社群运营面临三大痛点

事前计划耗时

每月都需要收集热点，生成运营计划，耗时耗人力，运营文案缺乏风格和创意。

事中响应迟缓

运营人员没有有效的实时分析和回复建议辅助，沟通过程影响营销效果。

事后分析不足

用户兴趣分析和案例总结不足，难以实现个性化运营，同时缺少案例培训新员工。

典型案例

模力时刻

Qwen-72B

处理 大模型生成运营计划，针对不同客群生成个性化营销话术案，参考优秀案例。

效果 提升私域社群运营人员工作效率 50%。

某银行在社群运营中，将业务流程分为事前、事中、事后三个阶段，并借助通义点金大模型提升运营人员的效率，具体流程如下：

1 运营前期

在运营前期的规划环节，主要通过“运营计划生成”智能体，结合运营分类等知识输入，提升制作运营计划表的效率。

2 客户引导

关键客户引导环节，大模型通过分析历史客户数据、实时分析聊天记录，识别出关键客户特征和偏好、自动分析对话内容，进而为每个客户群体提供个性化的跟进建议和辅助话术。

3 营销活动或客户服务后

在完成一次营销活动或客户服务后，社群运营智能体将负责总结聊天记录，并根据预设标签自动为客户和交互结果进行标记。这一步骤极大地减少了人工手动标注的时间成本，同时保证了数据的一致性和准确性。

4 长期运营过程中

最后，在长期运营过程中，社群运营智能体不断收集和分析成功的营销案例，从中提炼出有效的策略和方法，并将其整合进营销素材库中。这样不仅能持续丰富和完善营销资源，也为未来的活动策划提供了宝贵的参考。



价值分析

减少人工投入

流程自动化，提升计划效率，增强文案创意

提升营销效果

实时提供分析和建议，提升营销效果

降低人员门槛

自动分析总结案例，提升运营和培训效果

结语

Grok 向左、DeepSeek 向右；通用模型向前、推理模型向后。算力堆砌、极致工程优化、算法创新、高质量数据，都能够带来模型能力的不断提升，也推动了模型的 Scaling Law 法则从 Pre-Training Scaling 扩展至 Post-Training Scaling 和 Test-Time Scaling。技术进步带来的性能提升和推理成本下降，叠加开源带来的技术平权，使得金融机构能够以更低的成本使用更强大的模型。杰文斯悖论与梅特卡夫定律的双重作用，将促进生成式人工智能技术更好地融入金融行业，在助力金融行业高质量发展的同时，也必将推动人类社会迈向更高层次的智能时代。

附录

行业	一级分类	二级分类	页次	
银行	信贷	尽调报告生成助手	11	
		信贷授信方案助手	13	
		交易流水分析助手	15	
		企业股权 / 关联分析助手	17	
		财务报表分析助手	19	
		行业分析助手	21	
		信审资料查全助手	23	
		尽调报告审核助手	25	
	风控	信贷进件风控助手	27	
		风控智能助理	29	
		生成用户可疑交易报告	31	
	AI Native 手机银行	APP 零售业务助理	33	
		APP 信用卡服务助理	35	
		APP 理财业务助理	37	
		APP 养老咨询规划	39	
	证券	投研	研报快读	41
研报观点问答			43	
路演解析			45	
行业周报观点			47	
日报早评收评			49	
热门事件助手			51	
深度研报写作			53	
投行		股权激励助手	55	
		投行法规解读	57	
投顾		投顾资讯简报	59	
		理财产品问答助手	61	
		营销优先级判断	63	
智能运营		信披报告审核	65	
		营销物料审核	67	
保险		产品定义与销售	保险产品解读	69
			条款解析助手	71
	保险营销创作		73	
	保险产品搜索		75	
	保险产品推荐		77	
	保险产品问答		79	
	保险产品比对		81	

行业	一级分类	二级分类	页次
保险	核保核赔	智能预核保	83
		核赔辅助	85
		智能影像处理	87
	监管合规	条款智能核验助手	89
		条款智转助手	91
		对外披露审核	93
通用	智能客服	坐席通话质检	95
		意图识别	97
		智能语音导航	99
		机器人营销外呼	101
		机器人催收外呼	103
		客服知识构建	105
		客服知识问答	107
		数字人客服	109
		人工客服助手	111
		App 对话机器人	113
		智能工单助手	115
		智能投诉预警	117
		智能语音分析	119
		电销助手	121
		智能用数	客户经理绩效助手
	智能问数助手		125
	智能搭建助手		127
	数据资产问答		129
	存款分析助手		131
	贷款分析助手		133
	客户经营分析助手		135
	小微经营户潜在客户挖掘		137
	理财潜在客户挖掘		139
	代发户促活助手		141
	知识助手	知识问答	143
		复杂信息抽取	145
		智能写作	147
	研发助手	AI 程序员	149
		需求助手	151
		编码助手	153
		单元测试生成助手	155
		测试用例生成助手	157
		智能运维助手	159
	数字人	交互取证提额	161
		视频交互陪练	163

行业	一级分类	二级分类	页次
通用	数字人	直播交互	165
		视频分享	167
	内容审核	内容安全审核	169
		制度撰写	171
		合规问答	173
	信息检索与打标	企业工商信息打标	175
		可信搜索	177
		企业招投标信息打标	179
		企业知识产权信息打标	181
		财报解析	183
	培训陪练	交互式培训助手	185
		课件生成助手	187
		对练机器人	189
		智能培训问答助手	191
		智能考试助手	193
	办公助手	工作报告助手	195
		OA 小蜜	197
		智能会议助手	199
		流程助手	201
	营销助手	营销策划助手	203
营销素材生成		205	
营销文案生成		207	
社群运营助手		209	

【2025年2月】版本

