

当搜索遇见AIGC： 京东的千人千面素材生成实践

Jason Li

京东零售视觉与AIGC部负责人



- 1、电商的发展历程回顾
- 2、电商2.0会是什么样子
- 3、千人千面的商品素材生成
- 4、我们能为商家做些什么

📍 北京

QCon

全球软件开发大会

会议时间：4月10-12日

- 大模型赋能 AIOps
- 越挫越勇的大前端
- 云上业务架构演进
- 多模态大模型及应用
- 海外 AI 应用创新实践

📍 北京

AiCon

全球人工智能开发与应用大会

会议时间：6月27-28日

- 端侧智能
- AI Agent
- 多模态大模型
- 金融+大模型

📍 上海

QCon

全球软件开发大会

会议时间：10月23-25日

- AI Agent
- 大模型训练推理
- 端侧 AI
- 搜推深度融合
- 数智金融

4月

5月

6月

8月

10月

12月

📍 上海

AiCon

全球人工智能开发与应用大会

会议时间：5月23-24日

- 通用大模型
- AI Agent
- 垂直领域应用
- 模型可解释性

📍 深圳

AiCon

全球人工智能开发与应用大会

会议时间：8月22-23日

- 模型效率与部署
- 多模态大模型
- 大模型安全
- 智能硬件

📍 北京

AiCon

全球人工智能开发与应用大会

会议时间：12月19-20日

- 通用大模型
- 智能硬件
- LMOps
- 具身智能

极客邦科技 2025 年会议规划

促进软件开发及相关领域知识与创新的传播



参会咨询



查看会议



电商的发展历程回顾

传统交易阶段

20世纪60年代以前



物物交换→货币交易、纯线下交易



电商技术萌芽阶段

20世纪70年代~1990年



EDP/EDI技术奠基、无纸贸易出现



互联网电商起步阶段

1991年~2004年



因特网出现、电商平台诞生（亚马逊/京东/阿里）、线上交易启动



电商1.0时代

2005年~2022年



移动电商普及、货架电商、内容电商、流量竞争、千人千面的推荐搜索



电商2.0时代

2023年以后



大模型、具身智能、3D/XR、极致的个性化体验、极高的购物效率

- 1、电商的发展历程回顾
- 2、电商2.0会是什么样子
- 3、千人千面的商品素材生成
- 4、我们能为商家做些什么



电商2.0会是什么样子

智能供需匹配



从“人找货”到“货找人”，精准预测并满足用户潜在需求

大模型需求预测、多维度用户画像、主动商品推送、场景化推荐

高效供应链与物流



供应链全链路智能优化，物流配送速度与精度大幅提升

供应链动态调度、物联网库存监控、无人机/无人车配送、最后一公里优化

极致个性化体验



从千人千面的搜索推荐到千人千面的商品素材

视觉AIGC技术、多模态大模型商品理解、用户圈层分析技术

全流程AI服务



AI贯穿售前到售后，提供高效、自然的7*24全周期服务

多模态智能客服、AI售前咨询、自动化售后处理、智能问题解决方案

沉浸式购物体验



突破二维局限，打造可感知、可交互的虚拟购物场景

3D商品建模、XR虚拟试穿/试用、虚拟购物空间、沉浸式场景搭建

- 1、电商的发展历程回顾
- 2、电商2.0会是什么样子
- 3、千人千面的商品素材生成
- 4、我们能为商家做些什么



千人千面的商品素材生成 (1/12)



户外功能需求型买家

核心关注点：材质性能、面料科技、实用功能，如防水性、透气性、防风性、耐磨度等



外观穿搭需求型买家

核心关注点：服饰整体美观度、设计风格及穿搭适配性，如颜色搭配、版型剪裁、款式是否符合潮流

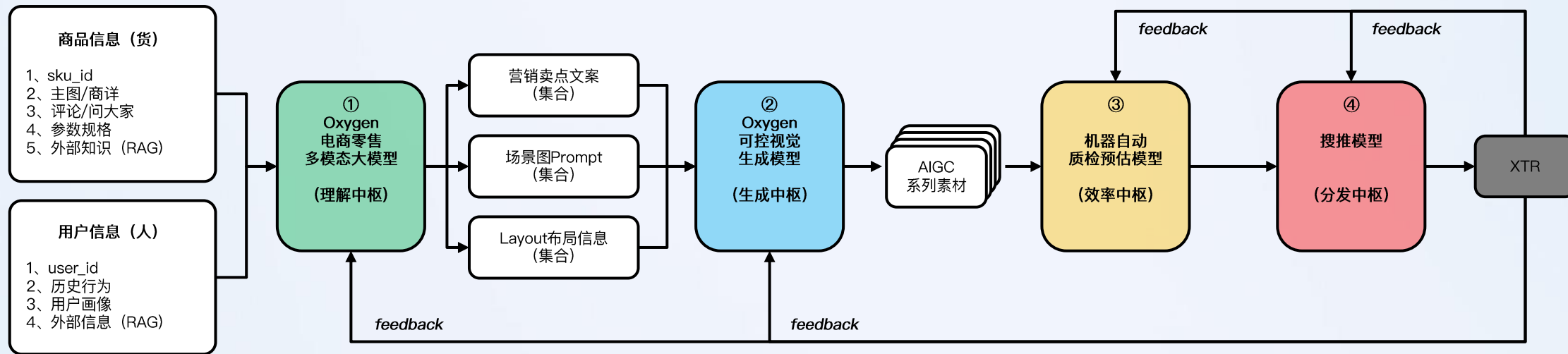


价格敏感型买家

核心关注点：价格高低、优惠力度及性价比，如是否为全年最低价、是否低于历史促销节点价格、是否有叠加折扣等



千人千面的商品素材生成 (2/12)

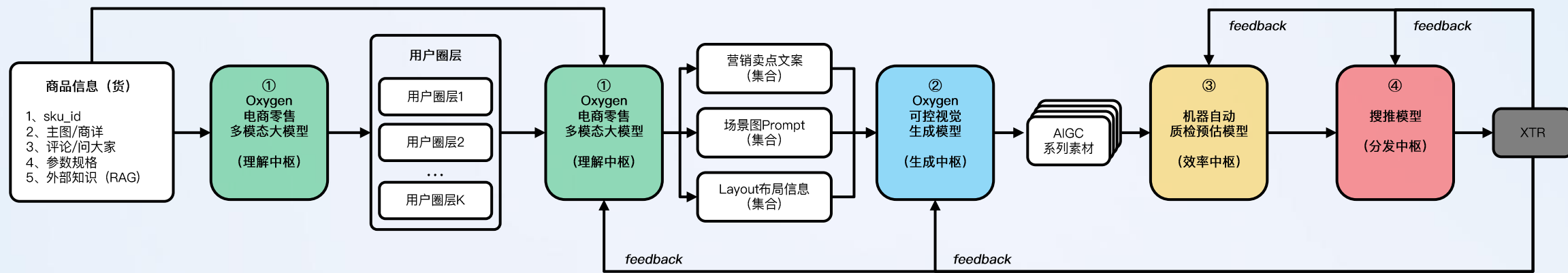


理想：个性化商品素材生成的完整技术框架

现实：推理效率是瓶颈，个性化素材生成是推理资源与业务收益的trade-off



千人千面的商品素材生成 (3/12)



一个可行的演进路径：千人千面 »» 千人百面 »» 千人十面

千人千面的商品素材生成 (4/12)



京东京造
美式黑咖啡粉
<https://item.jd.com/100139632942.html>



人群1: 健身爱好者



人群2: 办公室白领



人群3: 备考学生

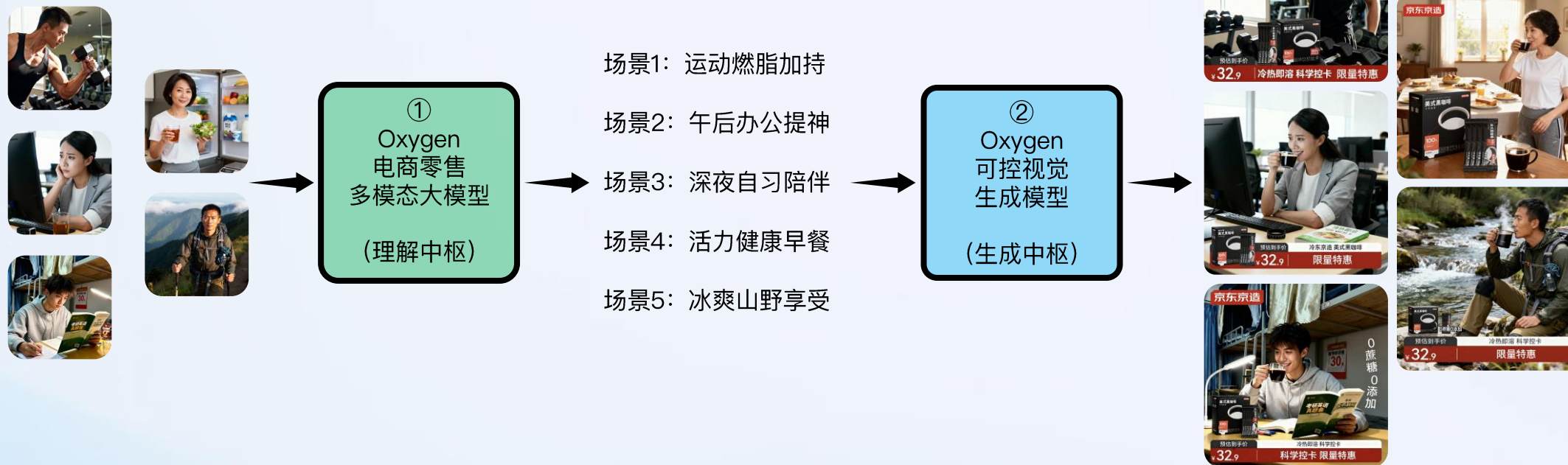


人群4: 控糖减脂人士

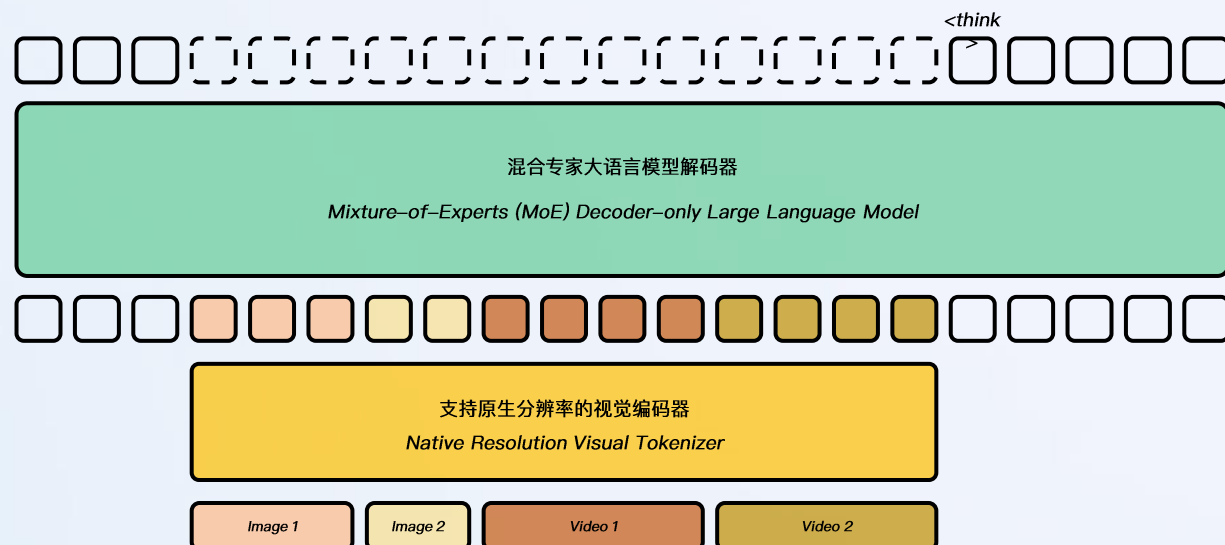


人群5: 户外爱好者

千人千面的商品素材生成 (5/12)



千人千面的商品素材生成 (6/12)



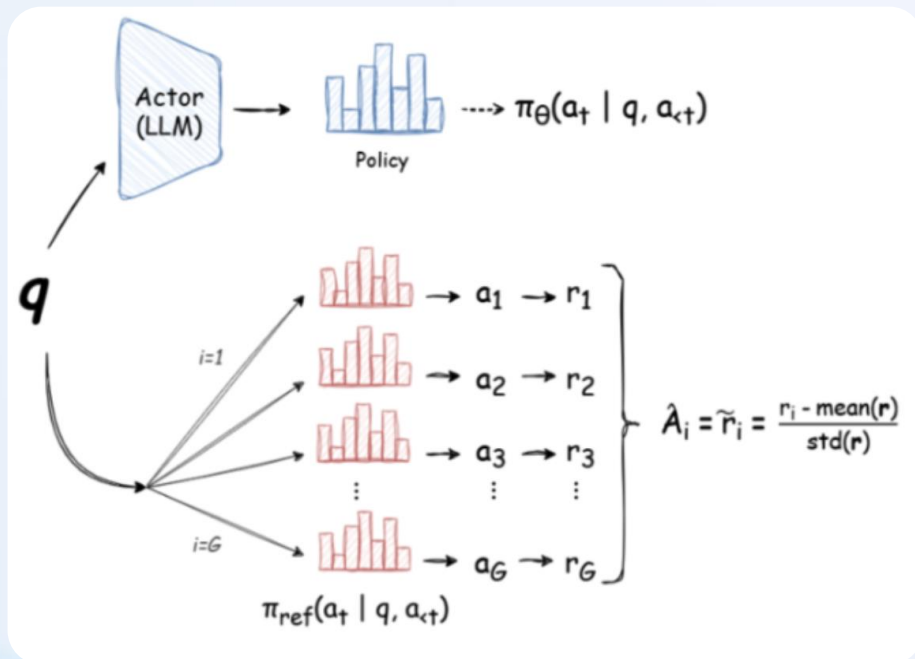
训练任务设计

通用任务	电商零售任务
文档解析 Document Parsing	商品分类 Product Classification
图表表格解析 Chart and Table Parsing	商品属性问答 Product Attribute Q&A
视觉定位 Visual Grounding	商品卖点生成 Product Selling Point Generation
文字检测识别 Optical Character Recognition	营销文案生成 Marketing Copy Generation
视觉问答 Visual Question Answering	商品合规性评估 Product Compliance Assessment
多模态推理 Multimodal Reasoning	商品价格合理性判断 Product Price Rationality Judgment
图像描述 Image Captioning	促销活动效果预测 Promotion Effect Prediction
视频理解与定位 Video Understanding & Localization	尺码信息解读 Size Information Interpretation

电商零售多模态大模型，通用数据+领域数据，激活领域推理能力同时保持通用能力是难点

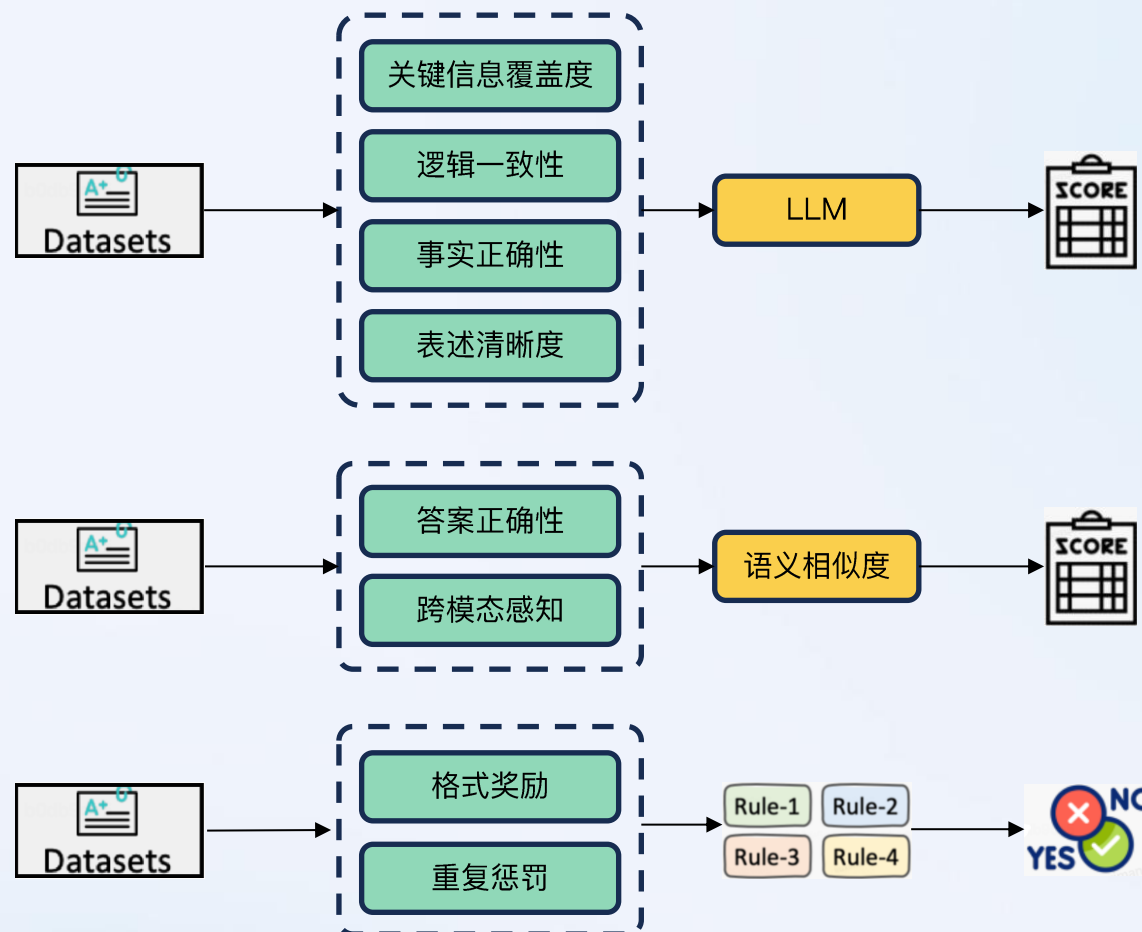
千人千面的商品素材生成 (7/12)

基于GRPO的强化学习



图片出处: <https://pub.towardsai.net/group-relative-policy-optimization-grpo-illustrated-breakdown-explanation-684e71b8a3f2>

Reward策略设计



电商零售多模态大模型后训练范式



千人千面的商品素材生成（8/12）

输入信息

商品标题：得力高120CM办公家用保管箱 电子密码锁 科技布内饰 弘雅AE70-M 黑色

一级品类：电脑、办公

二级品类：办公设备

三级品类：保险柜/箱

价格：¥2770

商详页卖点提取：电子密码锁；科技布内饰；高安全性能；双重解锁方式；防震防盗；纳米级科技布软包材质；金融级内载芯片；全钢材质；18道磷化防锈；C级锁芯；震动错码报警；支持两种解锁方式；环境温度适应性强；抗静电性能；连续开启耐久性。

用户评价摘要：质量好；做工好；材质精细；颜值高；性价比高



训练数据

问题：根据商品信息，预测显示屏是否支持触控操作？

思考步骤：

第一步：理解问题并提取关键图文信息

首先，商品标题中提到“电子密码锁”，主图展示了一个带有数字按键和明显的发光区域的面板。商品详情图进一步显示了保险柜的外观和面板细节。规格参数中虽然没有直接描述显示屏类型，但提及“密码”“电子”等关键词。

第二步：关联已知知识或规则

通常，现代电子密码锁保险柜会配备可视化操作界面，便于输入密码和显示状态。常见的显示屏类型有LED数码管、LCD液晶屏以及触摸屏。高端产品往往采用触摸屏，尤其是带有数字按键和发光区域的面板。

第三步：基于商品信息逐步推理出LED触摸屏控制的过程

从主图可以看到，面板上有数字按键且无实体按钮，显示区域有发光效果，这符合LED触摸屏的特征。商品卖点中提到“电子密码锁”“双重解锁方式”，结合图片，说明其操作界面是电子化且支持触控。高端定位（如“金融级内载芯片”“科技布内饰”）也暗示了更先进的显示和操作方式。商品的尺寸和价格也支持其为高端产品，通常会配备LED触摸屏控制。

第四步：验证逻辑合理性

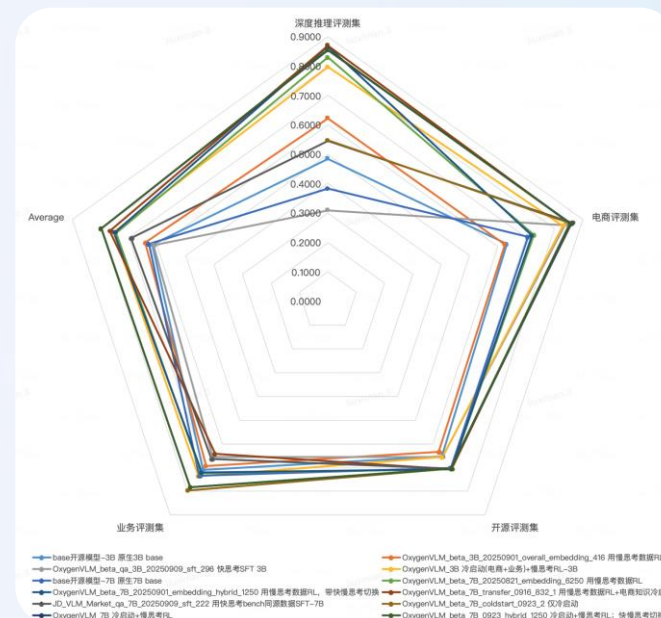
结合所有信息，商品的定位、图片中面板的设计、操作方式描述以及高端配置，都指向该保险柜配备的是LED触摸屏控制，而不是传统的机械按键或仅有数码管显示。逻辑链条完整，推理合理。

答案：支持LED触摸屏控制

多模态CoT数据可视化样例：商品属性推理

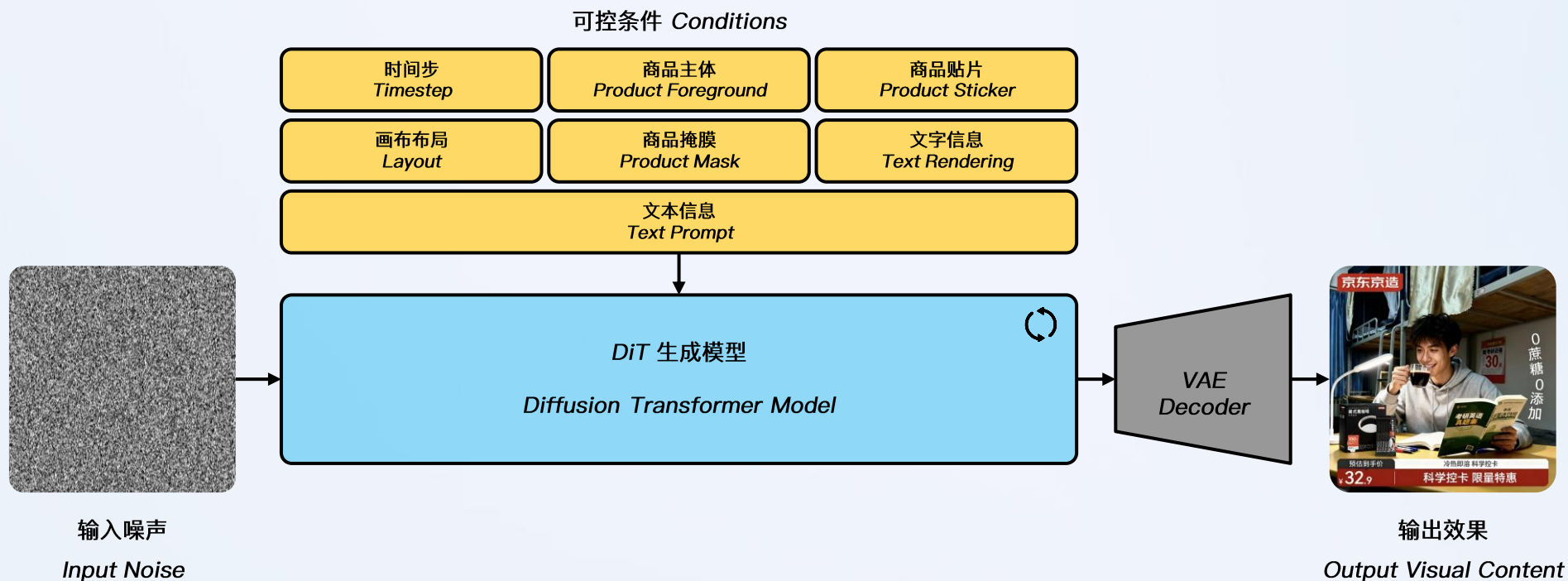
千人千面的商品素材生成 (9/12)

模型版本	评测集				Average
	深度推理评测集	电商评测集	开源评测集	业务评测集	
base模型-3B	0.4845	0.6317	0.6523	0.7103	0.6197
OxygenVLM-3B	0.7965 +0.3120	0.8304 +0.1987	0.6552 +0.0029	0.7382 +0.0279	0.7551 +0.1354
base模型-7B	0.3820	0.7071	0.7022	0.7339	0.6313
OxygenVLM-7B	0.8530 +0.4710	0.8576 +0.1505	0.7041 +0.0021	0.7836 +0.0497	0.7996 +0.1683



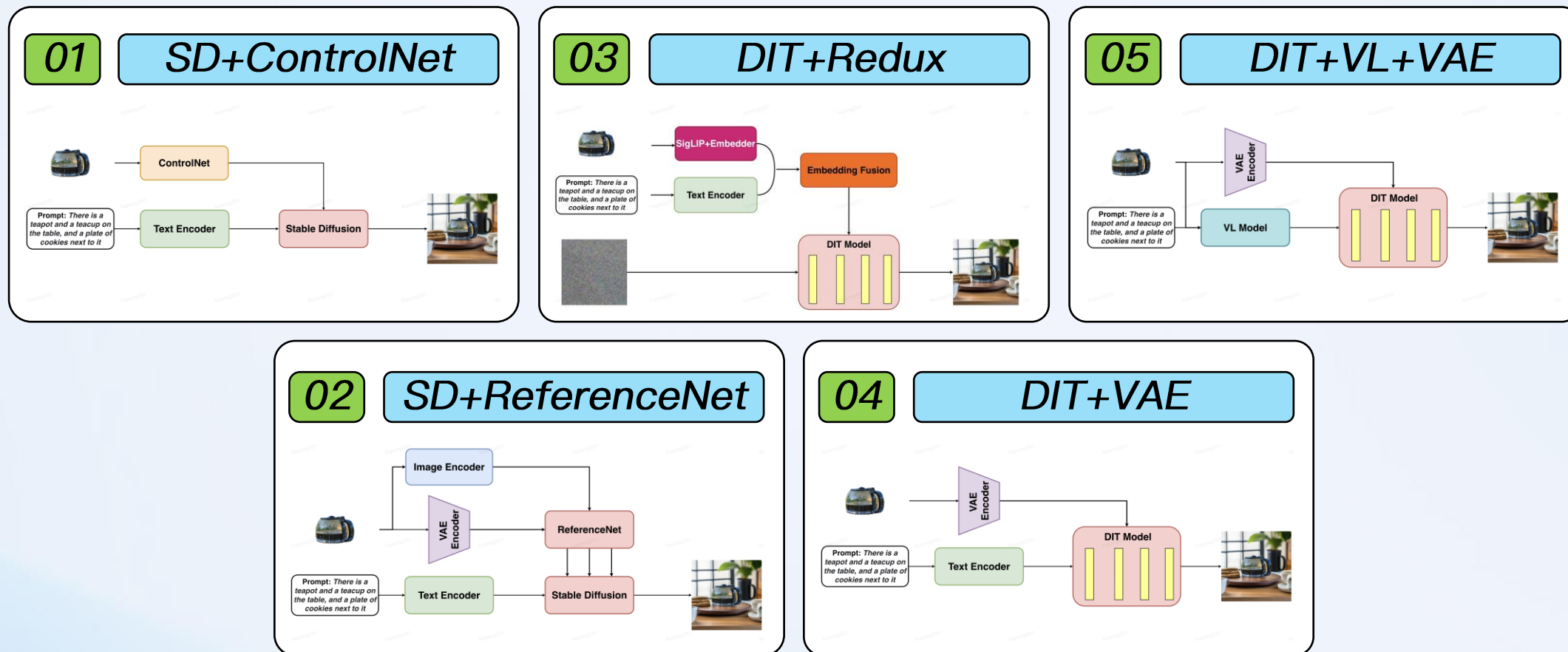
电商零售多模态大模型OxygenVLM多评测集定量评估效果

千人千面的商品素材生成 (10/12)



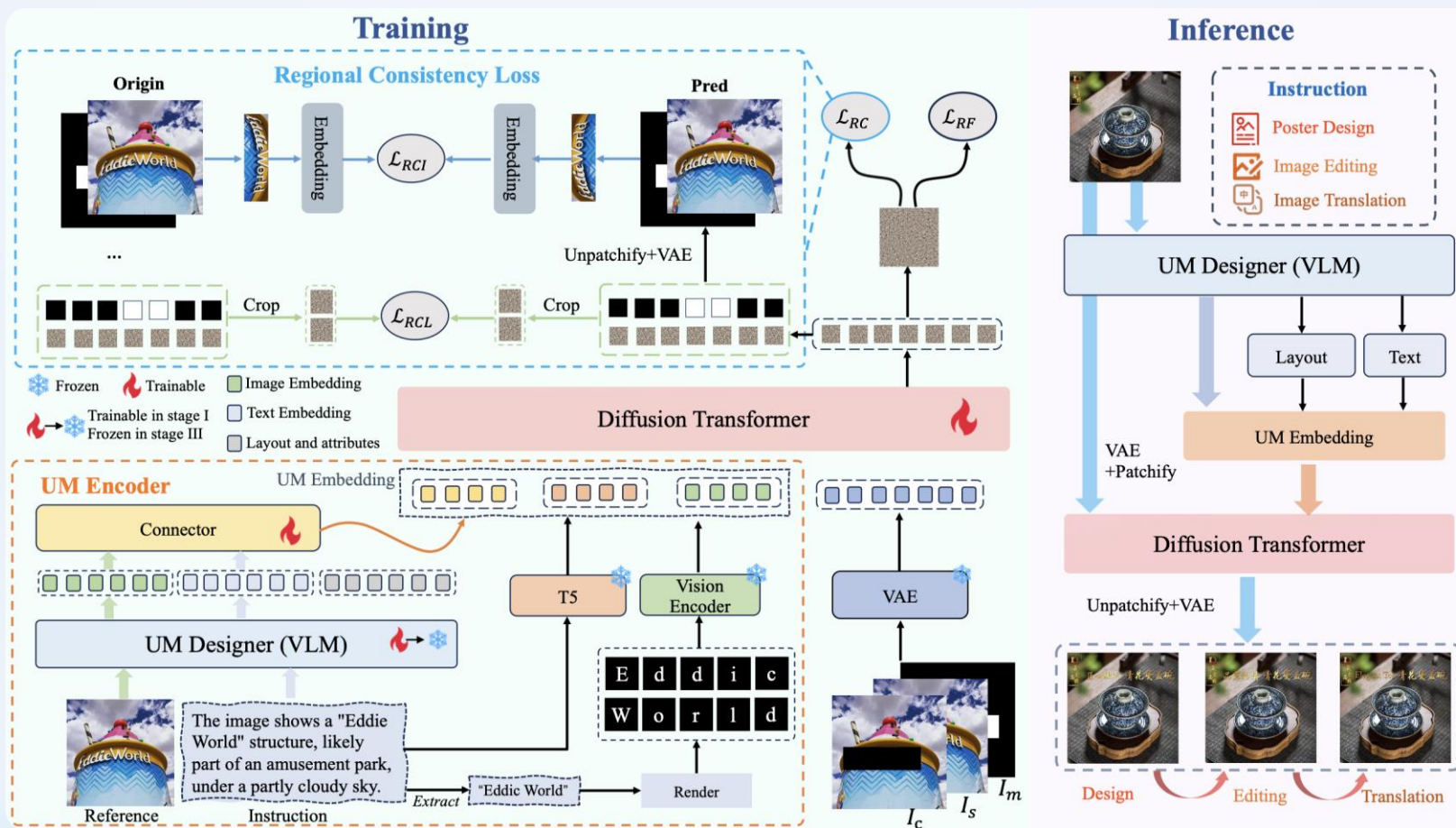
可控视觉生成模型，商品细节保持是技术难点，业务级端到端可用率低是行业现状

千人千面的商品素材生成 (11/12)



可控视觉生成模型近两年技术框架演化路径（不考虑Unified Model）

千人千面的商品素材生成 (12/12)



融合参考图/Layout/Prompt/文字的多条件统一生图框架

UM-Text: A Unified Multimodal Model for Image Understanding and Visual Text Editing

- 1、电商的发展历程回顾
- 2、电商2.0会是什么样子
- 3、千人千面的商品素材生成
- 4、我们能为商家做些什么

痛点一

商品多，干不完

- 京东数百亿动销商品
- 店铺平均4,000+ sku
- 自营采销人均负责10,000+ sku

痛点二

预算少，考核难

- 竞争激烈，竞对激进价格战
- 全店商品作图成本 v.s. 降价促销
- 10~20元/图 * sku数 * N个版本

痛点三

活动多，变化快

- 商品、店铺、平台促销活动丰富
- 采销精细化运营能力精力不足
- 投入产出比存在较大不确定性

从京东采销视角看商家痛点

● 我们能为商家做些什么 (2/10)



旧版京点点AIGC内容生成平台

30+

覆盖京东零售30+业务场景，涉及商品发品、运营、营销等，根据不同场景类型提供内容素材

800,000+

支持800,000+京东商家，通过京点点一键生成店铺运营所需的各类商品素材

10,000,000+

基于京点点平台AIGC能力在京东零售各场景单日常调用量达10,000,000+次

95%+

内容生产效率提升95%以上，商品图像、营销文案、商品视频等内容素材制作从天级降至秒级

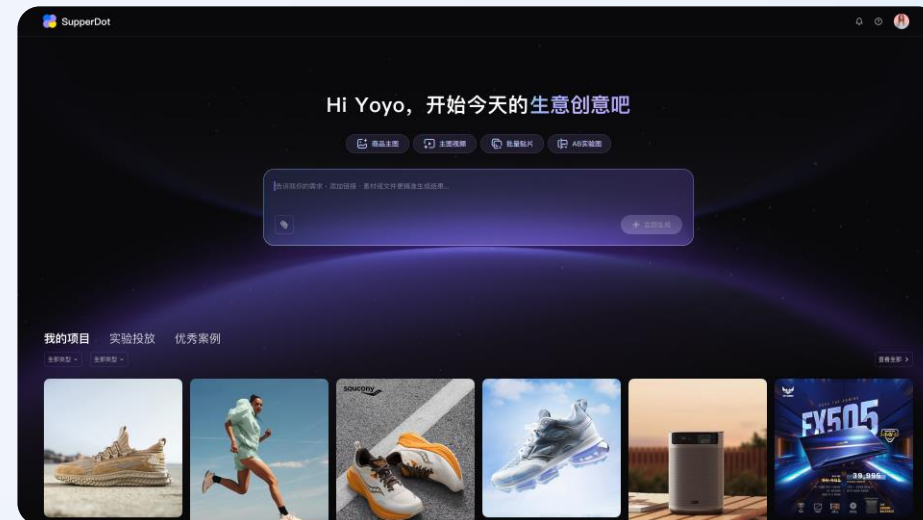
90%+

内容生产成本降低90%以上，商品图像、营销文案、商品视频等内容素材制作成本极大降低

● 我们能为商家做些什么 (3/10)

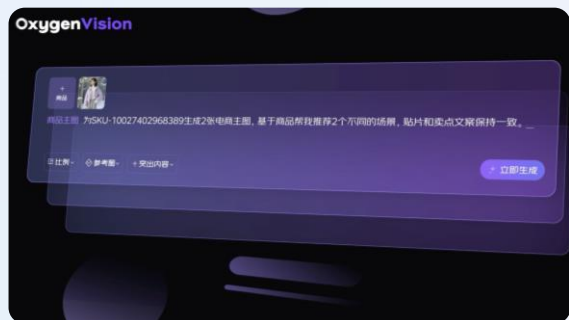


旧版京点点平台
GUI，工具箱式操作平台



焕新版京点点平台Oxygen Vision
LUI + GUI，自然语言交互操作平台

● 我们能为商家做些什么 (4/10)



对话式人机交互
纯自然语言交互形式



大模型规划与执行
拟人式分步骤有序执行操作



商品一致性保持
强多样性但不损失一致性



无缝接入AB实验平台
素材线上实验及经验总结

焕新版京点点平台Oxygen Vision的四个新特色

● 我们能为商家做些什么 (5/10)



● 我们能为商家做些什么 (6/10)

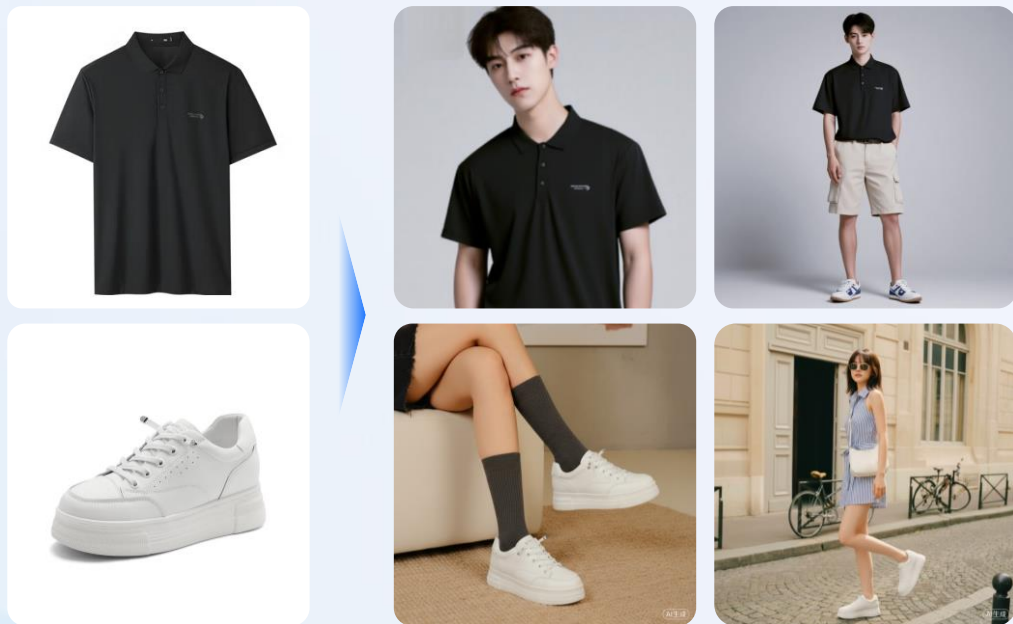
Agent员工上岗，让时尚品牌内容供给提效95%以上



焕新版京点点平台Oxygen Vision x 大时尚领域应用

● 我们能为商家做些什么 (7/10)

AI模特图-万物试穿



商品平铺图

生成商品场景图 (静态+动态)

某KA服饰品牌成交转化率提升**110%**以上，京东品牌自营店更换AI模特图商品销量**TOP1**，新品上新速度提升**95%**以上。

AI模特图-万物试戴

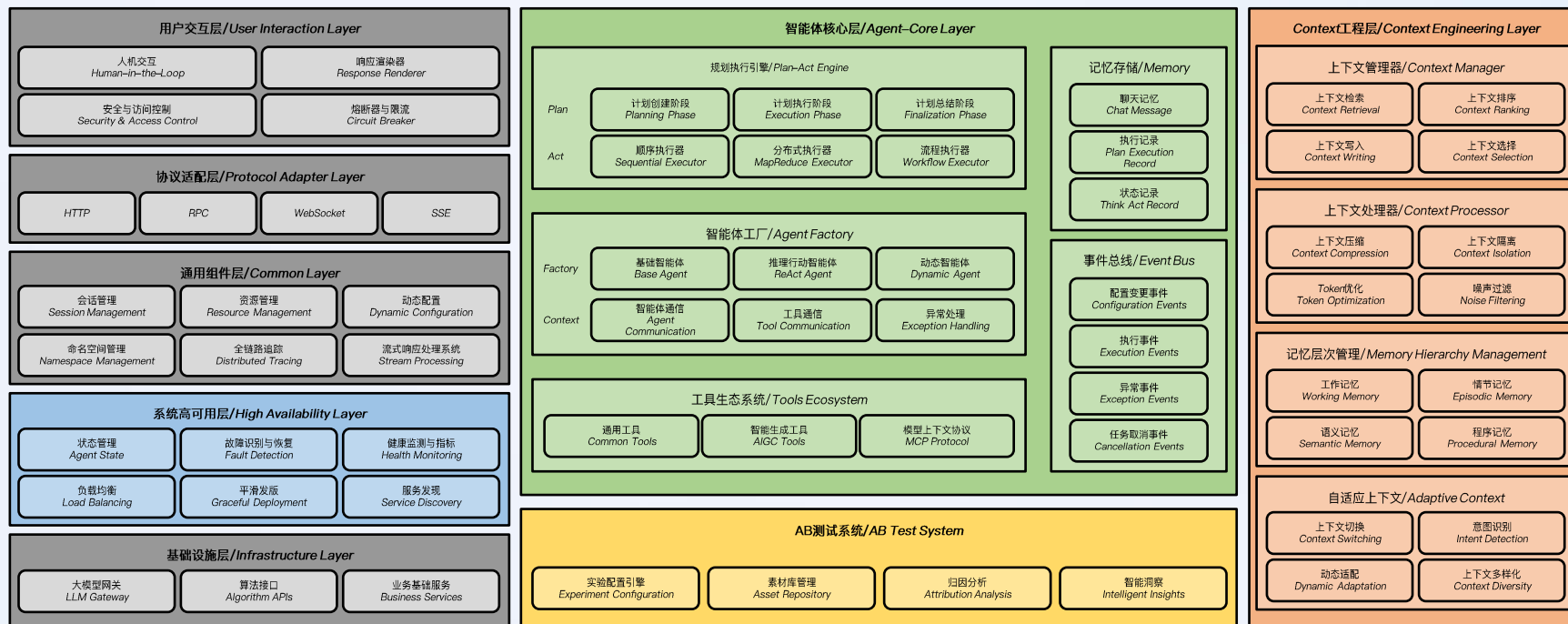


商品平铺图

生成商品场景图 (静态+动态)

某KA珠宝品牌成交订单量提升**360%**以上，成交GMV提升**320%**以上，成本较传统拍摄降低**90%**以上。

我们能为商家做些什么 (8/10)



所谓的智能体都是10%的大模型 + 90%的软件工程

我们能为商家做些什么 (9/10)



焕新版京点点平台多智能体协作流程图

● 我们能为商家做些什么 (10/10)

批量素材生成能力

- 支持指定店铺
- 支持上传列表

长短视频生成能力

- 支持主图短视频生成
- 支持营销长视频生成

经营效果直驱能力

- 支持点击类目标驱动
- 支持转化类目标驱动

外部平台支持能力

- 支持京东外商家
- 支持跨境电商商家

焕新版京点点平台二期功能规划预览

📍 北京

QCon

全球软件开发大会

会议时间：4月10-12日

- 大模型赋能 AIOps
- 越挫越勇的大前端
- 云上业务架构演进
- 多模态大模型及应用
- 海外 AI 应用创新实践

📍 北京

AiCon

全球人工智能开发与应用大会

会议时间：6月27-28日

- 端侧智能
- AI Agent
- 多模态大模型
- 金融+大模型

📍 上海

QCon

全球软件开发大会

会议时间：10月23-25日

- AI Agent
- 大模型训练推理
- 端侧 AI
- 搜推深度融合
- 数智金融

4月

5月

6月

8月

10月

12月

📍 上海

AiCon

全球人工智能开发与应用大会

会议时间：5月23-24日

- 通用大模型
- AI Agent
- 垂直领域应用
- 模型可解释性

📍 深圳

AiCon

全球人工智能开发与应用大会

会议时间：8月22-23日

- 模型效率与部署
- 多模态大模型
- 大模型安全
- 智能硬件

📍 北京

AiCon

全球人工智能开发与应用大会

会议时间：12月19-20日

- 通用大模型
- 智能硬件
- LMOPs
- 具身智能

极客邦科技 2025 年会议规划

促进软件开发及相关领域知识与创新的传播



参会咨询



查看会议

THANKS

欢迎试用焕新版京点点平台

oxygen-vision.jd.com

