

客服机器人
辅助编码

辅助上色
辅助内容总结

视频文案润色
辅助合同格式化

辅助本地化

SACC

上海站 / 2023

中国系统架构师大会

SYSTEM ARCHITECT CONFERENCE CHINA

数字转型

架构演进

🕒 2023年11月24-25日 | 📍 上海丽璟滨江酒店

大模型与AIGC 在B站OTT、漫画业务中的应用探索

哔哩哔哩 OTT技术部&漫画技术部 邬晶

探索项目列表

探索项目	应用部门	底层模型	应用说明	状态
客服助手	OTT	ChatGPT3.5	帮助客服人员回复用户用户提问	已应用
编码助手	OTT	ChatGPT3.5	注释→代码、代码→注释、代码改写、协议转换、单测	已应用
漫画辅助创作	漫画	SD	在部分自制漫画中利用SD进行辅助创作	部份应用
辅助内容处理	漫画	ChatGPT4	帮助小说编辑总结小说内容，为CP提供文案润色功能	部分应用
合同格式化	漫画	ChatGPT4	自动识别合同中的权利项及结算信息	进行中
辅助本地化	漫画	ChatGPT4	日漫&韩漫汉化、国漫翻译成外文	无法应用

客服助手

背景:

客服侧同学需要同时处理多个业务线，每条业务线也一直处于高速迭代状态，业务细节多，用户问题杂，为了帮助客服同学提高问题处理效率，我们基于ChatGPT3.5输出了智能客服系统

挑战:

上下文长度上限问题

取最近N条上下文限制长度

如何保证Token上限不超的情况下提高问题回答准确性

问题向量化，通过向量匹配，高精度组合上下文，为AI提供有效参照物。

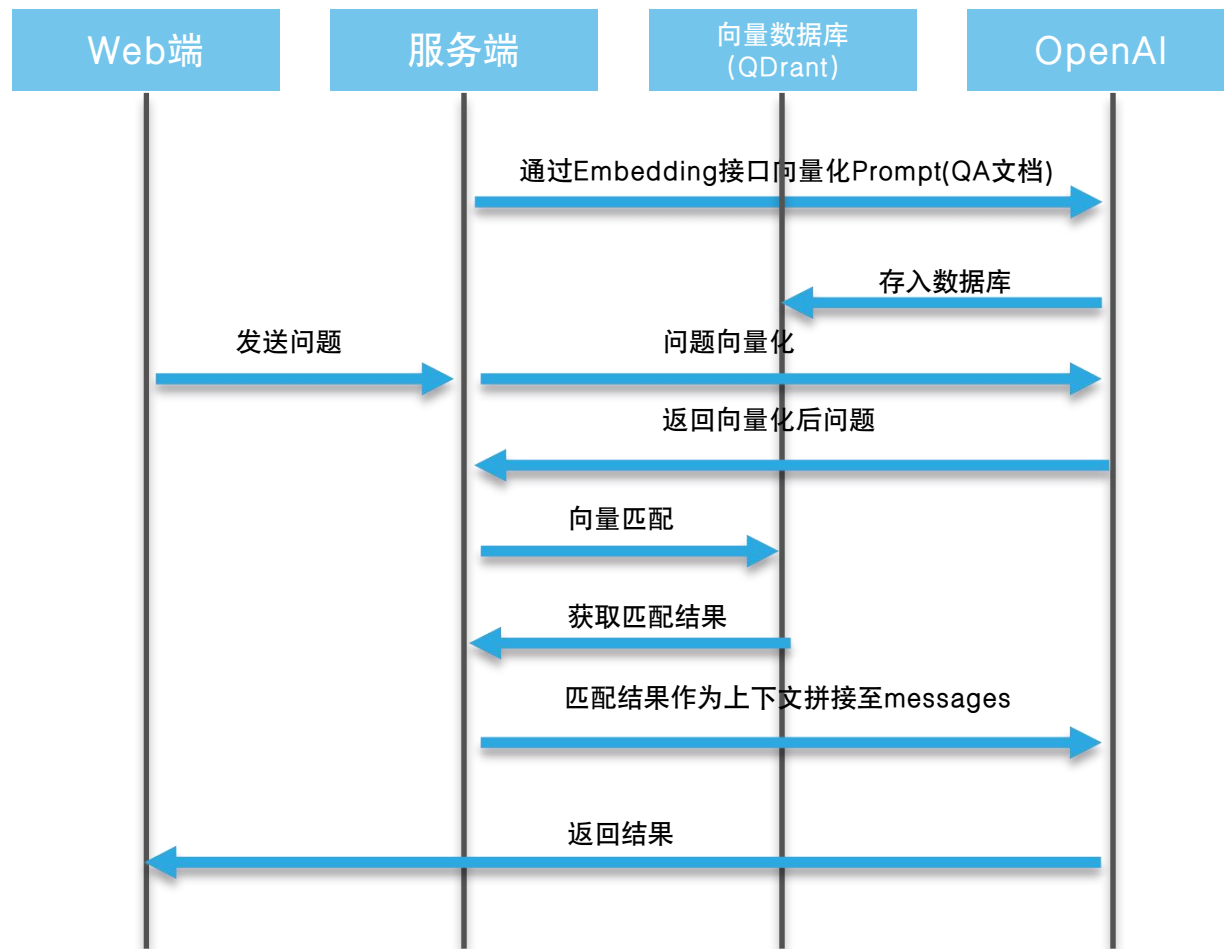
敏感性问题

调优prompt

收益:

拦截率: 88.5% -> 92% 提升3.5PP
节省人力: 13%

解决方案:



解决方案:

生成单测: -



生成测试用例的探索

主要卡点: 1. 代码上下文庞大, 不可能提供全部代码。2. 代码泄漏安全性



整个流程使用 AST 语法树处理，没有用到 GPT 模型接口。

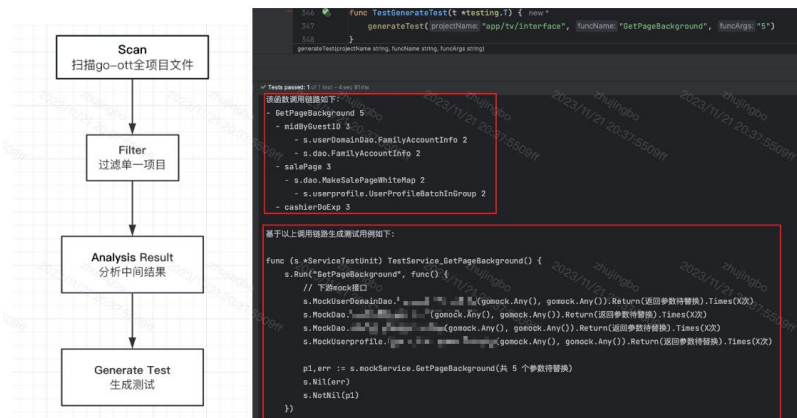
收益:



主要面临的问题来自 **代码安全、开发成本** 两个方面。

设置 **独立开发小队**，**限制代码传入** 来解决这些问题。

风险代码在VPS转发节点进行过滤，防控。



漫画辅助创作

筹备过程

- 确定发型细节
- 确定服装细节
- 绘制草图
- 铺个底色



文字人设(Tagger)



服饰参考(ironcatfatenoon)

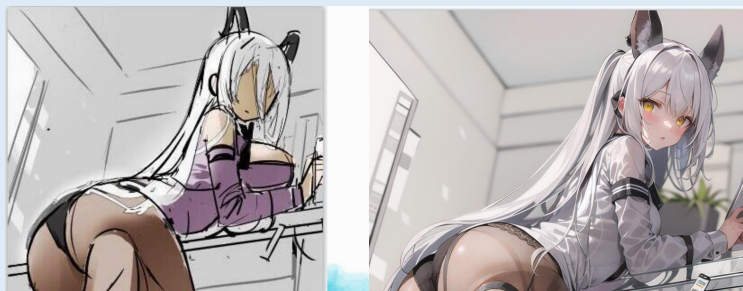


服饰参考
(ironcatfatenoon+controlnet-scribble/tile)

反馈：“偶尔用于辅助前期、内容封面，作为参考，使用门槛高，担心有舆论风险”

上色过程

- 草稿阶段
- 文生图+controlnet
- 线稿+底色
- 图生图+controlnet



文生图+controlnet-涂鸦/scribble



图生图+controlnet-tile/canny

实验：“可用于单人全身/单人特写/Q版辅助上色，覆盖率可达40%左右，生产一话约耗时1天”

辅助内容处理



无脑男频爽文



AI总结仅需分钟级别，效果已打平当前现状，可以直接使用

小说总结



女频情感小说



AI总结的内容在情感表现方面不足，但细节更加丰富，可以作为编辑老师的辅助



PROMPT:

帮我改写下面[]内的话语，要求：

- 1: 一个一个改写
- 2: 保留原文的意思，以及情感
- 3: 原句和改写句子的重复率不得高于30%
- 4: 每1个输入 的语句，重写1个版本
- 5: 重写语句的输出顺序请和输入语句的顺序保持完全一致
- 6: 请返回中文
- 7: 严格仅返回生成的结果，其他内容不允许返回

原文:

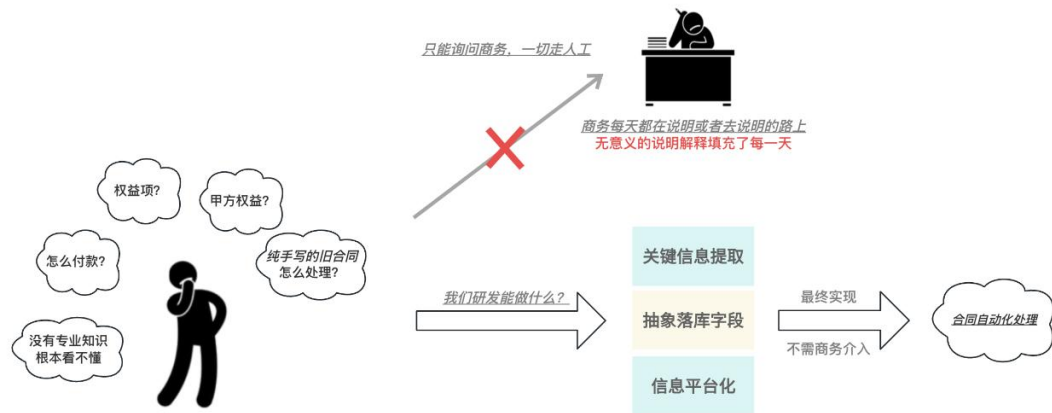
他叫沈廷卫
原本他以为自己一辈子都是 faker
结果在 24 岁时分化成 omega
他听从父母的话去相亲
结果在回家的时候
信息素突然不受控制
就在他准备去药店买抑制剂时
发生了车祸
穿越到了正常的世界

优化后:

他的名字是沈廷卫。
他原以为自己的整个人生都将被定义为一个faker。
但在24岁那年，他意外地变成了omega。
遵从父母的建议，他开始相亲。
在一次相亲后回家途中，
他的信息素开始失控。
正当他打算去药店购买抑制剂时，
一场交通事故发生了，
他意外穿越到了一个正常的世界。

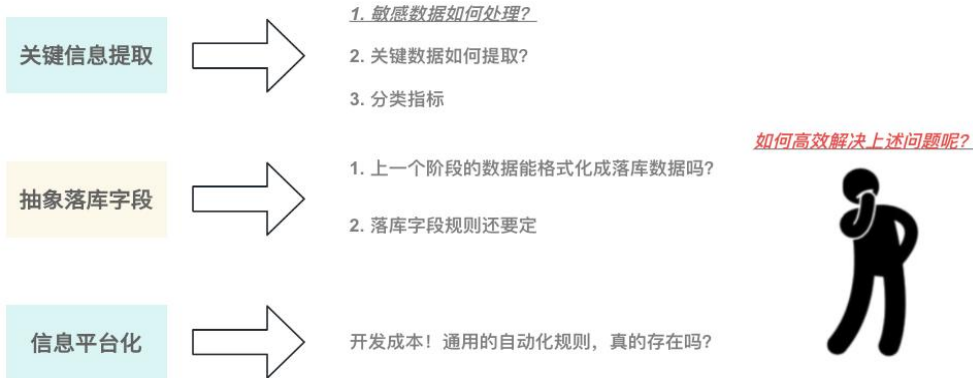
合同格式化

背景:

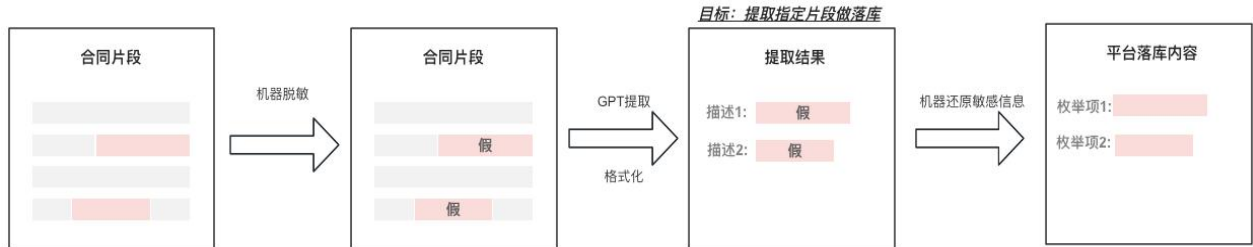


目标: 将合同电子化、平台化。各种权利项平滑入库, 提升可维护性。降低商务成本。

挑战:



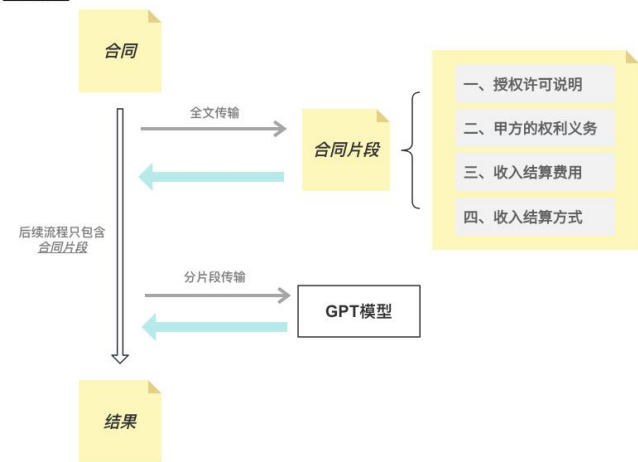
解决方案:



描述1、描述2为我们跟商务老师预先定义的各类权益枚举以及付款明细。

格式化后的数据, 我们做平台化展示。将所有合同关键细节做归类以图表的形式进行展示。

开发流程



根据合同片段中的信息, 甲方在合同中享有的特定权利详细分析如下:

权利	地区	语言	权利性质	期限时间	是否可转	结算方式	结算周期	导入字段
著作权	全球	独家	连载完	结束后10年	是	保底 + 分成	每月	是
信息网络传播权	全球	独家	连载完	结束后10年	是	分成	每季度	是
影视改编全版权	全球	独家	连载完	结束后10年	是	分成	每季度	是
游戏改编全版权	全球	独家	连载完	结束后10年	是	分成	每季度	是
周边产品授权	全球	独家	连载完	结束后10年	是	保底 (一次性形象授权费用) + 分成	每季度	是
实体出版	全球	独家	连载完	结束后10年	是	保底 + 分成	不详	是
动画改编全版权	全球	独家	连载完	结束后10年	是	分成	每季度	是
商务开发权 (形象授权)	全球	独家	连载完	结束后10年	是	保底 + 分成	不适用	是
其他 (不具备指明)	全球	不明	连载完	结束后10年	可能	不明确	不明确	是

辅助本地化

公爵之女的“伪”理想型恋人 (49-001, 002)

	像正常人类	像正常机器人	像努力的机器人	像虚伪的机器人
原文	人翻	机翻	机翻+故事性	机翻+情绪
<ul style="list-style-type: none"> ■ 시니스,부탁해. ■ 다녀와~ ■ 엄청열심히찾네. 	<ul style="list-style-type: none"> ■ 西尼斯, 拜托你了 ■ 去吧~ ■ 看来找得非常仔细呢 	<ul style="list-style-type: none"> ■ 시니스, 拜托了. ■ 回来之后再见~ ■ 你真的很努力在找啊 	<ul style="list-style-type: none"> ■ 辛尼斯, 拜托了. ■ 回来时小心~ ■ 你真是费尽心思在找啊。 	<ul style="list-style-type: none"> ■ 希尼斯, 拜托了. ■ 回来时一定要小心~ ■ 他们真是拼尽全力在寻找啊
<ul style="list-style-type: none"> ■ 아줄리말로는,셋째가.가문을 잇는 다길래 권력욕이 있나 싶었는데 ■ 무슨일을 해도 잘했을거야.강직하고,정직하고. 일단같은 편이 되면 절대로 배신하지 않을 사람 같아. 	<ul style="list-style-type: none"> ■ 听艾杰莉说, 老三会继承家族, 本以为是个利欲熏心的人呢 ■ 感觉他不论做什么事都会做得很好, 坚强且正直。如果成为伙伴, 是绝对不会背叛的人 	<ul style="list-style-type: none"> ■ 三儿子会继承家族。我还以为他有权力欲望呢 ■ 无论做什么事情都会做得很好。坚强且诚实。一旦成为同一阵营, 肯定不会背叛的人。 	<ul style="list-style-type: none"> ■ 按照阿朱里的话来说, 他是第三个。我原本以为他想要继承家族, 可能有权力欲望, 但是... ■ 无论做什么事情, 你都会做得很好的。你坚韧不拔, 诚实正直。只要我们成为同一阵营, 我相信你绝对不会背叛的。 	<ul style="list-style-type: none"> ■ 按照阿祖利的话来说, 原本以为第三个孩子, 也就是伊齐基尔·埃姆费特会继承家族, 我还以为他会对权力有所渴望呢 ■ 无论做什么事情, 你总是做得那么出色。坚韧不拔, 诚实坦率。一旦成为同伴, 绝不会背叛。你就像那种人。
<ul style="list-style-type: none"> ■ 그럴 필요 없습니다.해본적 없으니까요. ■ 말도안돼!기사님은 좋은 집안의 귀족이잖아?어렸을때정해둔정략결혼상대라던가,없어? ■ 없습니다.나름대로 사정이 있어서 	<ul style="list-style-type: none"> ■ 没什么要保密的, 因为没谈过 ■ 怎么可能, 骑士大人不是大家族里边的贵族吗? 家族的人没有在小时候给您定下结婚对象什么的吗? ■ 没有, 我们有自己的苦衷 	<ul style="list-style-type: none"> ■ 没有必要那样做。因为我从没试过。 ■ 这太不可思议了! 骑士先生不是出自好家庭的贵族吗? 小时候确定的策略性婚姻对象, 没有吗? ■ 没有。有自己的情况和原因。 	<ul style="list-style-type: none"> ■ 没有必要这样。因为我从未尝试过。 ■ 这简直不可思议! 骑士大人不是出身于好家庭的贵族吗? 小时候就定下的政治婚姻对象, 难道不存在吗? ■ 我也有我的难处啊。 	<ul style="list-style-type: none"> ■ 无需如此。因为我从未尝试过。 ■ 这不可能! 骑士大人你是出身名门贵族, 小时候就定下的政策婚姻对象, 难道没有吗? ■ 不存在。我有我自己的难处
<ul style="list-style-type: none"> ■ 기사님.그럼역시서 ■ 마음에둔사람이따로있는거지? ■ 그사람이이루어지지못할곳에 있거나, 아니면,설마죽었...! 	<ul style="list-style-type: none"> ■ 那么, 骑士大人您? ■ 是有自己喜欢的人吗? ■ 那个人在一个不能在一起的地方, 难道是死..... 	<ul style="list-style-type: none"> ■ 骑士先生, 那么还是去西站吧。 ■ 你是不是心里有别的人? ■ 那个人可能处在无法实现的地方, 或者, 不会吓到, 他没死吧! 	<ul style="list-style-type: none"> ■ 骑士先生, 那么果然是... .. ■ 你是不是心里有别人了? ■ 那个人是在无法到达的地方, 还是, 不会吧, 他已经死了?! 	<ul style="list-style-type: none"> ■ 骑士大人, 这么说来, 果然是... .. ■ 你是不是心里有别的人了? ■ 那个人可能在我们无法达到的地方, 或者, 不会吧, 他不会已经...死了吧!

THANKS



Hadoop
去“IOE”运动

分布式集群

云原生架构

区块链架构

语音识别技术

DDD设计

湖仓一体

sacc2013

sacc2015

sacc2016

sacc2018

sacc2017

sacc2019

sacc2012

sacc2014

rchitect
2011