#### 一、语义大模型

语义大模型，是国内首个且唯一完成万亿参数全自主训练的、全国产、全尺寸、全开源的基础大模型。不仅斩获国际顶级赛事双赛道冠军，更获得通信学会年度十大科技进展，以及世界人工智能大会最高奖SALL之星。

在全尺寸中，目前已开源1B到115B参数系列模型。模型开源方面，国产集群训练模型下载量排名第一。数据开源方面，开源1T高质量数据语料，成为国内下载量最大的开源预训练数据集之一。助力开发者快速构建行业智能应用，更赋能金融、政务等场景的复杂决策。

目前语义大模型已成为平台化、标准化能力。在拟真人通话代接模型中，日均服务500W通电话；在某市政的民生诉求大模型中，7个场景取得第一。

以深圳市12345热线项目为例。在分类分拨采纳率达到90%+，智能受理的准确率达到85%+，整体话务受理能力提升20%。下面我们会通过一个视频来了解12345能力的具体流程展示。

同时包括用大模型进行党建公文写作也是一个实际使用场景。支持53+N类海量的写作场景，包括结构化专业文章写作以及高质量的长文生成，以及一键排版和金句优化。

软件工厂，是中国电信专门为开发者打造的基于对话实现代码能力赋能的新一代智能开发工具。我们通过一个小短片来了解一下是如何为前、后端开发赋能。

（播放到语言交互停止，1min25s）下面是关于该产品比较细致的一些操作说明，通过刚刚的案例演示主要聚焦在原来需要写代码的地方，可以用到交互式的自然语言或者是图片来帮助用户，快速实现进行相关页面的简易开发

[备用]

【在模型效果层面，在国际计算机语言学学会ACL以及其他多赛道中包揽冠军，并且在SuperClue榜单中，位于第一梯队开源榜单前三名。除语义理解和文本生成的能力，我们支持业内独有的超千万行超大表格的复杂推理能力，~~目前，已在四川省电信、officeAI、硅基流动等单位开展项目落地推广。~~我们直接在网页上体验，选择其中一个案例，提供2015-2023全国366个重点城市的空气质量数据，可以针对该百万行表格完成表格取数、数据分析统计已经完整分析报告的生成。~~直接生成分析报告。在报告中，会提供重点数据分析结论以及合适的图表展示形式，帮助用户快速掌握超大表格的信息，完成从简单数据查看到有效信息提取的过程。~~

同时实现分钟级快速创作的星辰慧笔能力，能够参考用户提供的文档格式、领域知识等进行‘标题-概要-全文’三段式的写作，通过交互式输入实现整体风格以及局部内容的调整优化。

目前可以直接登陆天翼AI开放平台进行使用和体验。】

#### 二、语音大模型

星辰语音大模型基于电信方言数据积累，目前是国内首家通过网信办算法+服务双备案的语音大模型，并且是业内首个支持中英+方言自由混说的语音识别大模型，目前已实现60种方言的识别，10种方言自由混说，比如上海话、四川话、粤语等。

在数据层面，万号智能客服每个月1亿通电话，利用这些资源，我们也建成了业内规模最大的多方言语音数据集；目前已规模落地于智能客服、智能家居、适老服务、反诈拦截等场景，比如中山的12345热线，省专共建的新模式，短时间内低成本拿到远超头部竞品的效果并顺利推广，以及稍后会演示到的贵州电信反诈外呼回访。

[备用]【同时在语音生成方面，打造高仿真的标杆音色、实现近乎零样本的精准复刻】

这边给各位领导做展示的为全双工语音交互：和传统的唤醒式的“小度小度”相比，支持实时双向对话、无需唤醒词即可随时打断追问并且能够实现超拟人情感的多方言语音生成。

追问“请介绍上海的景点”“请用粤语介绍上海的美食” 。

在接受到一个较为完整的指令以后，大模型的打断响应速度上还是很棒的。

接下来，我们用两个不同大模型训练出来的语音产品进行一场辩论赛，感受一下我们的Telespeech在语义理解以及打断响应速度的能力进展。我们的辩题是人工智能究竟是造福人类还是危害人类。正方是我们的telespeech，反方是我们市面上豆包。

下面有请我的同事

进行辩论你的辩方是造福人类，请阐述你的观点。因为时间原因我们语音大模型的内容就展示到这。

#### 三、视觉大模型

过去的十年之间，我们在TeleAI1.0的时代。在多模态理解大模型中，有很多像人脸识别、视频结构化的单点的技术。但是视频、语义、语音他们是彼此各自发展、没有互通，同时就 点对点（case by case ）的去做任务。

那么在TeleAI2.0时代，我们研发的新一代跨模态的视觉大模型，精准对图像、视频、语义进行多源处理。同时实现行业事件认知、动态知识的挂载下的视频图片的万物布控。

在技术上，在国际顶级AI竞赛中获得8项冠军，并通过信通院、银检中心、公安部三大权威认证，并且在多模态理解的权威榜单Opencampass等达到国际一流水平。

那么它也是基于中国电信8,000万路摄像头，这样的海量数据能够去实现的一项技术。

目前展示的是通过给我们的大模型提供《交通道路安全法》，我们将他喂进大模型，然后输入违规关键词，比如违规掉头，现在会看到将违规的视频画面截取出来，我们点开一个看，这辆车在前面违规掉头。

大模型可以做到自主学习并理解法条对应的语义知识，并快速在海量的视频库和视频流中识别并检索出对应的交通违规行为，并针对相应的违规行为提供治理建议和判别的法条依据。后续也可以在学校或公安场景中，作为一个辅助功能来投入使用。

目前，我们的万物布控正是搭载在星河AI巡检助手上，为城市治理提供智慧方案。星河AI巡检助手是中电信自研的智能巡检系统，融合多模态大模型技术，以六大模块支持消防、交通、市政等多个行业。

接下来我们会通过一个视频来了解该产品的能力展示。

（播放到四大核心功能48s停止播放）刚刚介绍了完整的业务闭环流程，下面就是该产品能力在消防安全委员办的相关文件的展示，因为时间原因我们就接着往下

**另一个是我们鉴伪大模型**

生成式AI的安全风险已经成为一个社会热点，有些黑色产业链，会利用AI技术，比如拦截摄像头，通过盗取他人身份证及照片，利用AIGC技术生成一段活体检验视频，直接通过身份验证，盗取钱财。

通过千万级的数据集以及50+伪造攻击算法库，我们在七月的WAIC大会上发布业内首个基于思维链的可解释性鉴伪检测能力，通过分析“为什么是假的”来进一步提高真伪判断的准确率，该能力获得国家三大认证以及国际顶会连续三年五冠军。

接下来让我们来直接在网页上现场体验一下我们的该项能力。这是一段典型的AI诈骗视频：画面中人物的面容、声音、动作都极其逼真，如果仅靠传统单模态检测技术，可能难以识破。但我们的鉴伪平台通过全模态关联分析，识别出伪造痕迹——人物面部边缘出现异常模糊、远处车流违反物理规律碰撞、人物服装随时间不合逻辑变化，同时音频存在机械性重复的克隆特征，、最终鉴伪平台给出非真实拍摄的视频结论。

目前上线到多款手机，实现视频通话场景的实时伪造检测。

#### 四、多模态大模型

刚刚展示的是多模态的理解能力，这边演示时多模态的生成能力，实际上大模型时代，最大的特点时从过去的判别模型走向生成模型，能进行图像、视频、声音等多种多媒体的高质量生成，已经广泛适用于广告传媒、文化转播以及影视制作等领域，也是最近人工智能发展的一大趋势。

在技术上，国内第一个创新提出基于storyboard技术路径生成视频的团队，我们的技术国际上最权威的视频VBench榜单打榜，一直排名第一。目前，又提出了 二阶段的图文生视频世界模型VAST，能实现对三维运镜进行控制、流程生成长视频连贯动作，以及音视频同步生成等技术亮点。

那么接下来我们通过一个短片来感受下音视频同步生成的能力。

#### 五、语音语义一体化

这里介绍的是业内首次提出的超语义语音学，因为我们知道语音不仅传达字面的意思，比如会通过语调、语速、停顿甚至一个咳嗽来表达更多的情绪态度和意图。虽然没有在文字中，却影响着听者，因此我们提出超语音语义学。那其中除了显性语义包括词汇、句法之类的以外，还包括语速、叹气的情感线索，背景声音、交流节奏的语境动态，以及重音、暗示的隐形语义。这些维度不仅对语音理解有所贡献，也对语音生成有所助益

下面展示的就是我们的超拟人对话大模型，也是前面提到的贵州电信反诈外呼回访的标杆案例。反诈一方面是用户端建起的安全防护；那么该项目从运营商侧我们也会主动出击，找出那些潜在的诈骗行为。

接下来介绍的是端到端场景的服务平台，平台融入了语音语义一体化对话技术，实现机器人与用户超拟人自由对话，系统能针对用户的提问与表述迅速精准回应，避免答非所问，

这套以对话和分析为主的系统除了应对反诈场景，还可以处理营销、催收催缴的多元业务场景，并实现业务闭环。

我们已在贵州省电信反诈中落地，每日呼出通话超1w通，其中识别中高风险的约100-200通；业务结果上从每周40个涉案量降低至20个，大幅提升反诈预防效果。

接下来，我们通过一个简短的通话来感受下用户接到回访的场景复现（研发打过来，和线上业务流程是一样的）领导感兴趣的，可以直接您和他对话体验