

中国管理科学研究院 现代教育研究所

中院培字[2024]第 018 号

关于举办“文生视频及多模态大模型核心技术开发与 应用研修班”的通知

各企、事业单位：

2024 年政府工作报告提出“发展新质生产力”，并将其列为年度十大工作任务之首。新质生产力正在重塑社会生产运行的方式，为企业开辟一条通向高质量发展的道路。新质生产力的提出加速了企业数字化转型步伐，人工智能技术作为核心新质技术，正在赋能企业数字化转型。从 2022 年大语言模型 ChatGPT 的问世，到今年年初出视频生成大模型 Sora 的推出，人工智能技术已发展到生成技术和多模态大模型阶段，成为每一位科技人才重点关注和研究的方向。

为使更多科技人才掌握更新的前沿技术，加快培养数字技术人才，根据《国务院关于推行终身职业技能培训制度的意见》提出的“紧跟新技术、新职业发展变化，建立职业分类动态调整机制，加快职业标准开发工作”要求，中国管理科学研究院现代教育研究所（<http://www.pdhib.org.cn>）联合北京龙腾智元信息技术有限公司特举办“文生视频及多模态大模型核心技术开发与应用研修班”。本次培训采用理论+实战培训模式。

一、时间安排：

2024 年 6 月 28 日 — 2024 年 6 月 30 日 北京（同时转线上直播）

（28 日报到发放上课材料，29 日-30 日上课）

二、培训专家

北京理工大学教授，博士生导师，副所长。国家重点研发计划评审专家、国家原 863 项目评审专家，北京市自然科学基金/重点基金评审专家、博士点基金/博士后基金评审专家，国际期刊 NLPR 主编，以及其它期刊的编委。受邀出版国家出版基金项目人工智能出版工程《人工智能-机器学习理论与方法》教材。

三、参会对象：

各省市、自治区从事多模态、大模型、人工智能、机器学习、深度学习、自然语言处理、图像处理、视频处理、文本生成、图像生成、视频生成、大数据、知识图谱等领域相关研究的企事业单位技术骨干、科研院所研究人员和大专院校相关专业教学人员及在校研究生等相关人员，以及对文生视频、多模态大模型研究感兴趣的广大爱好者。

中国管理科学研究院现代教育研究所

二〇二四年五月十六日

北京龙腾智元信息技术有限公司

二〇二四年五月十六日

四、培训目标

1. 掌握生成式人工智能核心技术。
2. 掌握多模态大模型训练技术。
3. 了解 Sora 的概念、主要功能与应用现状。
4. 掌握 Sora 的工作原理与流程。
5. 掌握 Transformer 模型训练技术。
6. 掌握扩散模型技术。
7. 掌握预训练模型训练技术。
8. 掌握 VAE 模型训练技术。
9. 掌握原生视频训练技术。

五、费用标准：

A 类、3880 元/人（含报名费、培训费、场地费、平台费、资料费、两天午餐费、A 类证书费）。

B 类、6880 元/人（含报名费、培训费、场地费、平台费、资料费、两天午餐费、A 类证书费、B 类证书费）。

注：住宿可统一安排，食宿费用自理

- 1、培训费由组织培训班的施教机构负责收取并提供培训发票。
- 2、5 人以上 9.5 折优惠，8 人以上 9 折优惠。
- 3、B 类证书费不享受优惠。
- 4、参加线上、线下培训学员均可享受视频录播回放权益，及本人再次免费参加线下同主题课程学习权益。

六、颁发证书：

A 类、参加相关培训并通过考核的学员，由中国管理科学研究院现代教育研究所颁发《**文生视频与多模态大模型开发与应用工程师**》（高级）岗位认证证书，可通过官方网站查询，该证书可作为有关单位专业技术人员能力评价、考核和任职的重要依据。

B 类、参加相关培训并通过线上考试的学员，由工业和信息化部教育与考试中心颁发《**人工智能机器视觉应用**》（高级）职业技能证书，可通过工业和信息化部教育与考试中心官方网站查询，并纳入工信部教育与考试中心人才库，该证书可作为有关单位职称评定、专业技术人员能力评价、考核和任职的重要依据。

注：报到时请提交电子版彩照（大于 20KB，红蓝底皆可）、身份证复印件和学历证明复印件各一份。

七、注意事项

1. 指定报名邮箱：2196813614@qq.com。
2. 报名成功后，会务组在报到前一周发具体报到通知及行车路线。
3. 学员需自备能连网电脑一台。

八、具体课程安排

时间安排	课程大纲	
第一天 9: 00-12: 00 14: 00-17: 00	一、生成式人工智能技术	1. 生成式人工智能在计算机视觉领域的发展历史 2. 生成对抗网络 GAN 3. 可变分自编码器 VAE 4. 自回归模型 Auto-regressive Model 5. 扩散模型 Diffusion Model
	二、多模态大模型	1. 从 ChatGPT 到 Sora 的技术演进 2. 大语言模型 3. 视觉大模型 4 多模态大模型
	三、Sora 的概念、主要功能与应用	1. 什么是 Sora 2. Sora 的主要功能 3. Sora 的技术特性 4. Sora 应用场景（Sora 的个人应用场景 Sora 的企业应用场景）
	四、Sora 的工作原理与流程	1. Sora 的架构 2. Sora 的工作原理 3. Sora 使用的技术 4. Sora 的流程分析
	五、Transformer 模型	1. Transformer 模型的基本架构 2. Self-Attention 机制的原理与计算过程 3. Multi-Head Attention 的设计与作用 4. Positional Encoding 的实现方法 5. Rotary Positional Embedding 6. Transformer 中的 Feed-Forward Networks 7. Layer Normalization 的原理 8. Transformer 模型中的残差连接 9. Teacher Forcing 技术 10. 编码器与解码器的结构差异 11. 视觉 Transformer
第二天 9: 00-12: 00 14: 00-17: 00	六、扩散模型	1. 前向扩散过程 2. 反向生成过程 3. 网络架构 4. 参数化 5. 采样方法 6. Stable Diffusion 模型 7. Diffusion Transformer 模型
	七 、 CLIP : Contrastive language-Image Pre-training	1. CLIP 架构 2. 对比预训练 3. 数据集分类器创建 4. Zero-shot 推理 5. 提示词工程与集成
	八 、 VAE: Spacetime Latent Representation	1. Autoencoder 2. VAE 模型原理 3. 重参化 4. VAE 与 AE 的区别 5. Spacetime Latent Representation
	九、数据工程	1. Longer Detailed Captions: LLM 优化 Prompt 2. Re-captioning: Text-Video pairs 3. Native aspect ratios: 原生视频训练
	十、模型训练	1. Sora 训练流程 2. patchify—视频数据转换为图像块 3. Scaling Transformer 生成 4. Latent 转换为视频向量
	十一、行业实现	1. 文生视频 2. 文生图像 3. OpenSora
	十二、实操	文生视频代码分析

文生视频及多模态大模型核心技术开发与应用研修班报名回执表

单位名称 (发票抬头)							
纳税人识别号					发票类别	<input type="checkbox"/> 会议费 <input type="checkbox"/> 培训费 <input type="checkbox"/> 资料费	
通讯地址						邮 编	
填表人姓名		部 门		电 话			
姓 名	性 别	部 门	职 务	工作（研究）方向	学 历	手 机	邮 箱
希望通过培训解决的问题		问题一： 问题二：					
培训费		万 仟 佰 拾 元			缴费方式		
一、汇款	汇入指定账号： 户 名：北京龙腾智元信息技术有限公司 开户行：中国银行股份有限公司北京房山支行 账 号：328571911214						
二、微信、支付宝 (可刷信用卡、公务卡)		 					
是否需要住宿： <input type="checkbox"/> 是 <input type="checkbox"/> 否 入住日期_____，住宿天数_____天 标 准 间_____间，大床房_____间				参会带队学员签名：_____ 2024 年 月 日			
		关注微信公众号，了解更多资讯！					

联系人：王 源
 手 机：13718491258（微信同号）

电 话：010-81311930
 报名邮箱：2196813614@qq.com