

# 技术展演与公众想象： 宇树科技的具身智能短视频科普实践研究

吴文汐 田子艺 胡 婷

[ 东北师范大学传媒科学学院（新闻学院），吉林 130024 ]

**[摘 要]** 当前，短视频已成为新兴科技推广的关键媒介，而科技企业则是其中主要科普主体之一。本研究以宇树科技抖音账号为研究对象，基于 81 条视频样本与 4 114 条评论数据，采用多模态话语分析与文本挖掘相结合的方法，系统考察企业技术展演与公众社会技术想象之间的互动关系。本研究发现：在企业层面，宇树科技短视频账号通过文字、图像与声音多模态融合，将具身智能技术建构为可视化、情境化且具有沉浸感的技术展演；在公众层面，评论话语围绕功能应用、民族科技与人机关系等维度展开，生成多元交织的社会技术想象。企业与公众在短视频语境中形成持续互动，共同推动技术意义的协商与稳定化。本研究表明，科技企业的技术展演不仅是技术的展示，更是公众参与技术社会建构的重要起点，有利于推动科普实践从单向传播走向意义共建。

**[关键词]** 具身智能 技术展演 社会技术想象 短视频科普 新兴科技

**[中图分类号]** G206; N4 **[文献标识码]** A **[DOI]** 10.19293/j.cnki.1673-8357.2026.02.012

《2025 抖音科技内容生态报告》显示，当前抖音科技内容全年观看量已超 1.4 万亿次，超 600 家科技企业选择以抖音短视频为载体发布技术信息、展示产品成果<sup>[1]</sup>。这一数据意味着短视频已经成为新兴科技推广的关键媒介。科技企业通过短视频将抽象的新兴技术转译为可感知的视觉符号，这不仅是技术的展示，更是平台赋能下企业与公众共同参与的技术意义建构实践。在这一浪潮中，宇树科技尤为典型。作为国内具身智能领域的头部企业，宇树科技以技术展演为核心，通过抖音发布

机器人动作展示、垂直场景应用及性能测试等内容，依托高度拟真的运动能力与多样化的身体形态实现技术叙事与视觉呈现，吸引了公众的广泛关注和积极参与，激发了人们对具身智能社会应用的未来想象。

从技术的社会建构视角看，技术意义的稳定化是相关社会群体协商的过程<sup>[2]</sup>。而科技企业作为技术的开发与推广者，也是塑造技术形象，参与协商的关键主体。尽管已有学者关注到企业如何通过“证据剧场”或“使用剧场”构建理想的技术形象，但这些研究

收稿日期：2026-03-01

基金项目：国家社科基金项目“科技强国战略背景下新兴科技的沟通风险研究”（25BXW015）。

作者简介：吴文汐，东北师范大学传媒科学学院（新闻学院）教授、博士生导师，吉林省特色新型高校智库数智传播与社会治理研究中心主任，研究方向：科技传播，智能传播，E-mail: wuwx613@nenu.edu.cn。

往往侧重于单向呈现，在很大程度上忽略了公众作为能动者对技术意义的再生产。事实上，技术展演并非意义的终点，而是社会想象的起点。正如公众参与科学范式所揭示的，公众的参与展现出的是一种集体性的“社会技术想象”。这种由公众在日常实践中勾勒的未来愿景，构成了技术社会建构中不可或缺的解释维度。

目前将企业技术展演与公众社会技术想象纳入同一分析框架的研究相对匮乏。尤其在具身智能这一新兴科技领域，企业如何通过短视频进行技术展演？公众在评论区如何解读和想象相关技术及其应用？两者之间形成了怎样的互动关系？这些问题尚未得到系统回答。基于此，本研究通过聚焦宇树科技具身智能短视频科普实践，关注企业如何通过短视频这一媒介形态对具身智能进行多模态呈现，并进一步考察公众在评论区中对相关技术的解读与想象，从而在技术展演与公众参与之间揭示技术意义生成的互动过程。

## 1 文献综述与理论基础

### 1.1 短视频科普重构科普范式：主体多元化与公众参与的兴起

媒介融合时代，短视频平台构成了科学传播的重要场景，科普短视频已成为当前“科普新形态”<sup>[3]</sup>。科普短视频作为多模态话语的生产实践<sup>[4]</sup>，深度重塑了科普生态。一方面，大量非专业的商业机构与个人经由平台媒介赋权涌入科普领域，逐步挤占原有权威科普主体的话语空间，推动科普短视频创作主体泛化<sup>[5]</sup>。从专业科普机构、科技期刊、高校院所的建制化力量，到科学家、医生、知识网红等个体创作者，多元主体在短视频平台上交织互动，共同构筑了科普短视频内容生产多元主体协同的创作格局<sup>[6]</sup>。过往研究多聚焦于科普机构、科技期刊等组织主体，以

及科学家等知识网红、个体的科普实践<sup>[7-9]</sup>。然而在新兴技术的普及进程中，作为技术研发与推广者的科技企业早已成为不可忽视的中坚力量，对这一主体的科普实践研究尤显不足。与此同时，科普短视频拓宽了公众参与科学传播的路径<sup>[10]</sup>。公众不再是知识的被动接收者，而是通过点赞、评论、转发等行为深度参与“知识再生产”的共同体成员<sup>[11]</sup>。在短视频语境下，科技企业的科普实践如何激发公众的参与，尚待进一步探究。

### 1.2 技术的社会建构与技术展演

技术的社会建构理论（Social Construction of Technology, SCOT）强调技术是社会建构的产物，其意义在社会群体的互动协商中逐步稳定<sup>[2]</sup>。既有研究表明，技术被社会接纳的过程，本质是持续的社会意义建构过程<sup>[12-15]</sup>，而“技术展示”（Demonstration）成为技术社会建构的重要方式。有研究者提出“使用剧场”<sup>[16]</sup>与“证据剧场”<sup>[17]</sup>的概念，认为技术展示是通过话语组织、视觉编排与策略性现场展示，将技术塑造为可行的表演过程。在数字媒介环境下，技术展示演化出了更多模态的可能性，其呈现方式正经历从功能性“展示”向表演性“展演”（Performance）的范式转移。而以短视频为核心的媒介平台，凭借其可视化、情境化与互动化特征，将技术推向了一个可观看、可评价、可共情的展演前台，在真实技术功能基础上形成了更具表演性和观赏性的视觉呈现与形象构建。虽然“展演”这一概念多用于解释身份与社会角色的建构<sup>[18-20]</sup>，但其“行动者如何在特定情境中通过符号系统向观众呈现某种现实”的理论内核，同样适用于解释技术情境。新兴科技企业的短视频科普实践可被视为技术展演的当代形态，企业在短视频前台持续展演理想化的技术形象，引导公众理解。然而，既有研究多基于线下展示或传统媒介语境展开，对于短视频等新

兴数字媒介环境下技术展演形态的转变关注不足，尤其缺乏对其多模态、可传播与强互动特征的系统考察。

### 1.3 社会技术想象：公众参与中生成的技术愿景

社会技术想象（Sociotechnical Imaginaries）被定义为集体拥有的、在制度上得以稳定且公开表达的、关于通过科技进步达成理想生活与社会秩序的愿景，其核心在于科学技术与社会秩序的互构<sup>[21]</sup>。相关研究经历了从宏观制度分析向多元行动者视角的拓展。早期多聚焦于国家层面的主导性想象，通过分析政策文件、战略规划等，揭示特定技术想象如何引导科技发展路径<sup>[22]</sup>。随着研究深入，研究视角向多元主体扩展。在职业与平台侧，新闻从业者在采纳技术时表现出工具理性、身份区隔与替代焦虑的集体叙事特征<sup>[23]</sup>；博主与平台则通过对算法规则的试探，达成关于技术边界的“双重想象”与“技术共谋”<sup>[24]</sup>。在公众侧，普通人在技术乐观主义的表象下，流露出无法拒绝、无法预判、无法参与技术未来的无能感，被动成为受影响的旁观者<sup>[25]</sup>。纵观现有文献，研究多关注精英阶层的宏大愿景，对大众层面、日常实践中生成的社会技术想象关注有限。研究对象多集中于生成式人工智能、元宇宙或数字城市等文本、视觉概念，缺乏对具有高度物理交互性的具身智能想象图景的考察，且方法上偏重深度访谈或政策文本，较少关注短视频评论区这一具有实时性与表演性的公众话语空间。本研究认为，短视频平台不仅是技术展演的窗口，其评论区更是公众参与技术意义建构的公共空间。这些评论反映了公众对人机边界和公共交往秩序的重申与反思<sup>[26]</sup>。因此，社会技术想象理论为本研究提供了分析视角，通过对短视频评论进行系统梳理，能够还原公众对具身智能技术的真实认知。

综上所述，从技术展演到公众社会技术

想象的互动过程，构成了新兴技术社会建构的重要实践场域。然而，现有研究仍存在一些不足：（1）对科技企业作为技术科普主体的关注相对有限；（2）对于短视频环境下技术展演形态的转变缺乏系统探讨，尤其未能从多模态整合角度加以深入分析；（3）尽管已有研究注意到公众参与的重要性，但对短视频评论区这一具体话语空间中公众如何参与技术意义建构、并生成社会技术想象，仍缺乏细致考察。基于此，本文试图回答以下两个问题：

RQ1：宇树科技在抖音短视频中如何进行技术展演？如何实现具身智能技术的意义建构？

RQ2：公众在短视频评论区如何参与、解读与想象这些具身智能技术？即公众的评论内容呈现出哪些主题与话语特征？这些表达折射出公众对具身智能技术具有怎样的认知、态度与社会技术想象？

## 2 研究方法：多模态话语分析与文本挖掘

本研究采用多模态话语分析与文本挖掘相结合的方法，围绕企业技术展演与公众社会技术想象的互动过程展开双向考察。以具身智能技术领域代表性企业宇树科技官方认证抖音账号“@Unitree 宇树科技”为研究对象，选取其2023年8月至2026年2月发布的81条视频内容（剔除4条纯图文信息）。采用文本挖掘与多模态话语分析相结合的方法，探究企业如何建构具身智能技术的意义。编码表经一致性检验（Kappa系数=0.87），信度较高（见表1）。接着，运用网络爬虫技术抓取上述视频评论数据，经清洗后保留4114条有效评论，运用词频分析与BERTopic主题建模，对用户评论进行聚类与结构化处理，提取公众关注的核心议题与话语模式，识别围绕具身智能技术展开的社会技术想象。

表 1 多模态话语分析编码表

| 分析维度 | 分析单元 | 编码类目                 |
|------|------|----------------------|
| 文字模态 | 视频标题 | 标题字数 / 标题语气 / 标题情感倾向 |
|      | 叙事类型 | 叙事主题                 |
| 图像模态 | 视频主体 | 主体类型 / 主体装扮          |
|      | 场景呈现 | 场景类型 / 场景特征          |
|      | 动作展示 | 动作类型 / 运动难度          |
|      | 画面形式 | 视频画面形式 / 视觉色彩        |
|      | 镜头运用 | 拍摄手法 / 镜头类型 / 视觉强化手段 |
| 声音模态 | 互动呈现 | 人机互动 / 机机互动          |
|      | 配乐配置 | 配乐类型 / 配乐节奏感         |
|      | 音频内容 | 解说类型 / 环境音运用         |

3 宇树科技具身智能短视频科普实践中的技术展演

短视频情境下，宇树科技的具身智能技术展演呈现出节奏紧凑、主题集中、以技术可视化为核心的特点。样本中 1 分钟以内的短视频占比超过 80%，整体内容具有明显的短时高密度特征。在主题分布上，两位研究者随机抽取 30% 的样本视频进行独立观看与开放式记录，通过讨论归纳初步形成主题类目，随后对所有视频进行归类。统计结果显示，视频主题以基础动作能力展示（28.4%）为核心，延伸至公开活动与赛事记录（24.7%）和场景化应用展示（23.5%），同时涵盖多机协同与集群表演（9.9%）、极限环境测试（8.6%）与科普生态建设（4.9%），形成了从技术能力验证到社会应用的多样化呈现。

在上述内容特征分析的基础上，进一步考察其多模态建构方式。多模态话语已成为公众认知现实的重要方式<sup>[4]</sup>，宇树科技短视频中的具身智能技术展演，正是通过多种符号资源的协同建构来实现。因此，本部分引入多模态话语分析框架，从文字、图像和声音 3 个模态出发，系统考察宇树科技如何通过短视频进行具身智能技术的展演。

3.1 文字模态：技术意义的框架建构

文字模态在多模态传播中承担意义指引与叙事组织的核心功能，通过符号的选择、

强调与排除，为其他模态提供阐释框架<sup>[27]</sup>，进而塑造公众理解事实的方式。宇树科技短视频的文字模态主要由标题与内嵌字幕构成，因内嵌字幕多为离散的参数标注或性能提示，偏向局部技术解释，难以承载整体叙事框架建构，故本研究聚焦视频标题文本进行分析。围绕其长度、句式、情感倾向的分布特征以及叙事框架，考察其如何预设技术叙事的基本立场，推动具身智能技术的意义建构。

3.1.1 高信息密度与积极情感：技术意义的预设与强化

宇树科技的视频标题整体呈现出长文本特征，超半数为 50 字以上的中长标题。尽管既有研究认为过长标题可能分散用户注意力而降低传播效果<sup>[28]</sup>，但宇树科技仍采用结构完整、信息集中的表达方式，将技术参数、应用场景与突破成果压缩呈现，帮助公众在进入观看之前即获得对技术内容的初步理解与框架性认知。

在句式语气层面，标题以陈述句为主，形成多样化的句式结构。陈述句（59.3%）用于确认技术成果与产品事实，凸显技术已实现的现实性；感叹句（34.6%）多出现在技术突破或比赛获奖节点，如通过“刷新多个全球第一次！”“勇夺四金！”等表述强化情绪，塑造技术持续进步的形象；疑问句（6.1%）虽数量有限，但通过“大家猜猜有多快？”等互动性提问引导公众关注。句式选择由此在组织技术意义的同时，借助情感张力与互动设计强化公众对具身智能的认知与期待。

在情感倾向上，标题情感以正向为主，占比达 75.3%。标题多使用“突破”“进化”“成就”“未来”等词汇，即便出现“挑战”“摔倒”等暗含风险和失败的表述，也指向技术的能力扩展与持续进步。这种一致的正向情感将具身智能塑造成值得关注与信任的成就对象，进一步巩固了公众的技术形象认知。

### 3.1.2 能力展示主导下的多重框架：技术意义的整合与延伸

宇树科技的技术叙事以技术能力展示为核心框架，辅以多重框架协同建构技术意义。其中技术能力展示框架覆盖率达98.8%，成为最核心的意义建构路径。通过强调性能突破与迭代进化，持续将具身智能塑造为可量化、可验证的技术实体，不断巩固其可靠性与先进性。其余框架则从不同维度进行意义延展。民族科技突破框架（25.9%）将技术嵌入国家科技叙事，赋予其民族自豪与集体认同；轻娱乐化拟人框架（29.6%）以人格化表达弱化技术距离感；合法化论证框架（23.5%）以参数与规范表述强化专业可信度；未来场景预演框架（24.7%）则依托生活化与产业化场景想象，将现有技术能力投射至未来应用图景。总体而言，多框架的组合运用使宇树科技将具身智能系统建构为兼具现实性能、未来潜力与社会价值的复合技术形象。

### 3.2 图像模态：技术能力的可视化呈现

在技术的社会建构理论视角下，技术并非以客观能力直接进入公众认知，而是通过符号化的呈现过程获得社会意义。图像模态作为短视频中最具直观冲击力的符号资源，承担着技术意义建构的核心功能。依托视觉修辞分析框架，可以发现宇树科技的具身智能科普短视频通过动作与镜头协同、空间与色彩配置、装扮与互动协同等视觉策略，将具身智能的技术能力转译为可视化的身体表现与社会情境中的行动能力，最终生成可见、可感、可确证的技术事实。

#### 3.2.1 动作与镜头协同：技术能力的视觉强化

宇树科技短视频技术展演在功能呈现的基础上，形成了以动作设计与镜头调度协同完成的视觉修辞实践。动作类型上，高难度动作（如原地空翻、侧空翻等）占比30.9%，进阶跑跳类动作占比43.2%，二者合计显著高

于基础动作（24.7%），形成以极限动作为核心的视觉呈现。视觉修辞强调“显著性”的建构，即某些元素通过数量、位置或动态强化而获得优先意义<sup>[29]</sup>。高难度动作的集中呈现，尤其置于娱乐竞赛与复杂环境等挑战性场景中，强化了突破极限的叙事逻辑。镜头运用进一步放大了这一意义建构。视频整体以平视镜头为主，构建出机器人与观众平等、去权威化、可接近的观看关系。在高难度动作展开时，动态跟拍镜头沿运动轨迹同步推进，使观众的视线与机器运动同步，产生沉浸式体验。关键动作往往采用动态跟拍和慢动作处理，使运动轨迹、关节控制与落地稳定性被清晰放大，让技术精度可被肉眼直接核验。在动作展演与镜头语言的协同下，技术合法性由专业话语转向视觉证据，机器人身体成为技术可信度的载体。

#### 3.2.2 空间与色彩配置：技术场景的日常化呈现

宇树科技短视频中空间场景的多元选择与色彩的差异化配置实现了技术场景的日常化呈现。在81个样本中，实验室场景仅占19.8%，而户外展示（22.2%）、日常生活场景（21.0%）与娱乐赛场场景（22.2%）合计占比达65.4%，成为技术展演的主要空间载体。其中，日常生活场景最为贴近公众日常经验，为公众围绕家庭服务、养老陪护等场景展开技术想象提供了直接基础。这种场景选择本身构成了一种视觉修辞策略，将机器人频繁置于运动场、街道和节庆空间等生活化场域之中，通过不同空间类型作为图像符号，承载并传递特定的意义指向。视觉修辞的运用，在于先以特定图像符号构建视觉场景，再通过对符号特征的深化来传递隐喻内涵<sup>[30]</sup>。当机器人主要处于实验室或工厂环境时，其意义更易被限定为技术设备或生产工具；而当其被反复置入公共与日常空间时，其社会角色则被重新编码为社会参与者。色彩运用进

进一步强化了这一意义转向。在短视频中，实验室与训练场景多采用灰、白、蓝等冷色调，强调技术理性与专业属性；而在户外展示、节庆活动和舞台表演中，则大量使用高饱和度的暖色与自然色，从而弱化机器人的金属质感与机械冰冷感。色彩既是视觉修辞的重要元素，也是情感表达的重要方式<sup>[31]</sup>。在场景与色彩的协同作用下，具身智能技术实现了视觉层面的日常化表达，其身份由专业设备转向社会行动主体。观众在观看短视频的过程中，既感知机器人的技术能力，也感知其在场方式。

### 3.2.3 装扮与互动协同：技术角色的拟人化策略

拟人化策略成为整体视觉表达的重要运用，近半数视频中的机器人身着功夫服、节庆服饰等人类装束，这并非出于功能需要，而是典型的拟人化视觉修辞，借助服装符号使其被赋予演员、运动员、伙伴等社会身份。同时，近七成视频通过人机、机机两类互动形式，以协同训练、对打、共舞等场景打破机器人孤立呈现的状态，尤其是人机同框时平等的画面站位、近距离互动以及动作节奏的同步呼应，更深化了公众对其类人行动能力的认知。从社会建构理论来看，技术形象并非天然形成，而是在符号编码中被不断塑造。拟人化穿搭弱化了技术的陌生感与风险感，互动展演则凸显其协作与回应能力，使具身智能从被动工具转变为主动参与者，并在视觉层面建构出未来智能体与友好伙伴的双重形象。

### 3.3 声音模态：技术声景的打造

作为多模态话语的重要符号资源，声音模态通过声音元素的组织与配置，将观众带入特定的感知情境，承担起营造氛围、表达特定情感、呈现人物行为的功能<sup>[32]</sup>。特定声音元素的筛选与组合能够构建出可感知、可阐释的声音景观，影响人们对情境的理解，从而塑造技术形象<sup>[33]</sup>。在宇树科技的技术展

演中，同期机械声和背景音乐构成声景建构的主体，一层指向技术真实，一层指向未来想象。

#### 3.3.1 同期机械声的保留：技术真实性的声音建构

同期机械声的保留是宇树科技技术声景的第一层建构，让具身智能技术“听起来”是一个真实在场的物质实体。在技术展演过程中，背景音乐并未完全覆盖机器人运行时的机械声响，机器狗越野时的环境音、机器人后空翻落地时的撞击声等现场声音都被有效保留，甚至在部分视频中处于听觉前景，在声音层面塑造了技术的物理在场感，强化了画面真实与在场体验。在人工智能生成内容泛滥、视觉特效日益逼真的当下，真实的机械运作声成为一种稀缺的感知资源。观众通过同期机械声确认眼前之物是具有重量、能耗与物理结构的机器，而非纯粹的视觉特效或虚拟动画。由此获得的不仅是声音信息，更是对图像内容的真实性感知。

#### 3.3.2 背景音乐融入：技术未来感的声音建构

背景音乐的融入是技术声景的第二层建构，通过节奏、旋律与音色的选择，赋予技术以情感色彩与未来指向。宇树科技短视频中的背景音乐风格统一，快节奏、强节拍、电子合成音色呈现出鲜明的未来科技感。电子合成音色是建构未来感的核心元素。合成器、电子鼓点与科幻音效在日常文化经验中常与科技、未来关联。当观众听到这类音乐时，潜意识中会将眼前的技术与科幻电影、未来世界联系起来，使具身智能被听出超越当下的象征意义。背景音乐的快节奏与强节拍则进一步强化了这一形象的感知强度。在多数视频样本中，音乐节奏与机器人的极限动作形成同步，例如空翻落地、集群变阵等的关键节点往往落在音乐的重音上，机器人由此被呈现为高度精准、节律统一的智能体，科技感的声音表达与视觉画面彼此强化。在这一声音建构中，观众被带入

激动、振奋的感知状态，使技术能力从“可看”转化为“可感”，在情感层面完成对技术未来形象的接受与认同。

### 3.4 典型视频案例的多模态分析

上文从文字、图像、声音3个维度分别考察了宇树科技短视频的技术展演策略。但在实际传播中，这3种模态并非孤立运作，而是相互嵌套、协同表意，共同建构具身智能技术形象。为进一步揭示其多模态的协同建构机制，下文选取两则典型视频案例展开深度分析，考察文字、图像、声音如何有机融合，为公众打造丰富、立体且具有沉浸感的技术展演图景。

以“宇树科技 G1 人形机器人天坛祈福”视频为例，该视频为宇树科技 G1 人形机器人于天坛广场开展的武术集群展演，视频以多机协同、户外稳定运动与高动态仿生动作为核心展示内容（见图1）。视频通过“宇树武 BOT 天坛祈福：祝福所有人 宇树 G1 机器人集群演示 # 宇树 # Unitree # 人形机器人 # 春晚 # 春晚机器人 # 武术 # 功夫 # 祈福”的标题与话题标签完成意义框架的搭建，将 G1 机器人集群演示、人形机器人等技术信息，与天坛、武术、祈福等文化符号及春晚舞台权威标识结合，以积极的公共话语完成叙事定调，把一场技术性能展示转化为带有仪式感的文化表达。其技术能力主要通过图像视觉进行具象化呈现。视频以天坛文化地标为展演空间场景，以真实户外环境印证机器人的环境适应与全局定位能力；通过大规模集群同步武术动作、统一抱拳礼，直观展现多机协同的同步精度与运动稳定性；同时通过全景俯拍、环绕中景与慢动作特写等方式放大关节运动、落地稳定等关键技术细节，增强关节控制、落地平稳度等核心技术指标的可见性。与此同时，在画面展演的全过程中，声音与视觉动作形成高度协同。视觉上的动作起落、阵

型变换与国风配乐的鼓点精准卡点，慢动作特写时刻伴随重音与机械撞击声，将画面动效与声音节奏深度绑定，使得视觉与听觉相互强化。在这个视频中，文字、图像与声音模态相互配合、彼此支撑，共同完成了一场连贯完整、层次清晰的具身智能技术展演。



图1 宇树科技 G1 集群同步后空翻动态展示

而在“宇树 G1 人形机器人室内格斗”这一视频中，通过对宇树 G1 人形机器人极限运动的测试，集中呈现其“反重力”稳定模式、连续空翻与跌倒后自主起身等高动态运动能力（见图2）。视频标题以“奇奇怪怪的技能”这种网感化表达降低理解门槛，同时突出“反重力模式”“任意动作序列稳定性提升”等核心技术卖点，用带引号的“反重力”制造传播悬念与记忆点。图像精准呈现文字信息，多角度拍摄的飞踢冲击、0.8秒快速起身、连续空翻等极限动作，将抽象的“稳定性提升”“任意动作序列下的抗干扰能力”的文字表达转化为直观可感的视觉奇观，每一个关键画面都与文字描述的技术优势形成一一对应关系，同时配合固定中景跟拍，保证动作



图2 宇树科技 G1 室内抗冲击平衡测试

的完整性与连贯性，印证“无任何加速”的画面文字。与此同时，声音同步匹配视觉节奏，用机器人关节运动的机械声、冲击倒地的真实环境音等强化现场真实感，搭配富有科技感的强节奏电子音乐放大动作的张力与冲击力，最终让文字的信息传递、图像的技术展示与声音的氛围营造形成不可分割的整体，将硬核技术转化为可看、可感、可信的视听体验，在极短时间内完成了宇树 G1 核心技术突破的高效传递。

4 具身智能技术的公众想象：基于宇树科技短视频评论文本的挖掘

在社会技术想象的视角下，公众并非被动接受技术意义，而是通过讨论主动参与技术意义的再生产。本研究通过对 4 114 条有效评论进行文本挖掘，初步勾勒出公众话语的分布特征。

高频词分析显示，“机器人”“宇树”“厉害”“人类”“动作”“遥控”“波士顿”等词语频次较高，构成公众评论的核心语义场（见图 3）。进一步采用 BERTopic 模型进行主题挖掘。基于 c-TF-IDF 算法提取关键词，并通过 nr\_topics 参数将语义相近的主题进行合并与凝练。在建模过程中，初始聚类产生 640 条离群文本（代表噪声数据）占评论总数的 15%，属于可接受的范围。根据 c-TF-IDF 关键词相似度，最终获得 5 个核心主题聚类（见图 4）。以此构建公众对该技术展演的讨论空间。Topic 0 春晚表现（42.66%）作为第一大主题，公众重点围绕春晚舞台的机器人动作展演展开讨论；Topic 1 家庭服务与养老应用（21.63%）位居第二，集中指向生活化服务与养老照护的现实需求；Topic 2 仿生机械动物（7.63%）聚焦仿生形态与辅助功能的社会价值；Topic 3 未来科技（7.27%）延伸至技术

发展与国产科技前景；Topic 4 价格与消费门槛（5.25%）则反映出公众对具身智能产品市场普及与消费可行性的现实关切。整体来看，公众讨论以舞台展演感知与生活化应用期待为主，同时兼顾技术前景与市场现实，呈现出多元且贴合实际的关注特征。



图 3 宇树科技抖音短视频评论词云图



图 4 主题分布及特征词权重

上述词频与主题分析勾勒出公众话语的整体轮廓。宇树科技凭借高机动性、强视觉冲击力的具身智能技术展演，激发了公众丰富且多元的未来想象。这背后形成了怎样的技术想象？以社会技术想象理论为分析框架，将碎片化的评论话语整合为 3 个深层维度：社会功能性应用想象、民族科技自主想象以及人机关系与伦理想象。

4.1 社会功能性应用想象：从“日常化展演”到“生活服务”的延伸

公众在观看宇树科技具身智能技术展演后，其认知并未停留于“技术奇观”层面，而是进一步转向对现实应用场景的功能性想

象。BERTopic 结果显示，Topic 1（家庭服务与养老应用）与 Topic 2（仿生机械动物）的相关讨论接近三成，“养老”“家务”“生活”“照顾”等高频词形成稳定语义网络。公众不仅接收到了科普视频传递的技术能力信息，还将其延伸至养老照护、家务替代、高危作业等具体社会功能的期待。

以宇树科技发布的“打工狗的日常”短视频为例，该视频属于“场景化应用展示”主题，采用“轻娱乐化拟人”的叙事框架。文字模态上，“第一天上班”“好好表现”“完成第一单”等第一人称字幕，把机器狗塑造成一个努力工作的“打工人”形象；图像模态上，视频将四足机器人置于真实的写字楼环境，捕捉其平稳行走、自主避障、完成物品递送的全过程；声音模态上，搭配轻松明快的背景音乐，烘托出生活化、轻娱乐的叙事氛围。通过图像、声音和文字的多模态协同，借助这一完整的“打工”故事，将机器狗塑造成“打工人”形象，使公众自然联想到配送、巡检、养老陪护等社会劳动场景，产生机器人替代人类承担这类重复性体力劳动的功能想象。技术能力因此被转化为可感知、可代入的生活服务实践，推动公众将具身智能与现实社会需求相连接。

评论区进一步强化了这种功能性想象（见图 5）。“快递”“外卖”“送货上门”等高频词表明，末端配送是公众最集中期待实现的功能。在此基础上，公众的想象进一步拓展至民生服务与特殊作业领域：既有面向弱势群体的“导盲犬”辅助盲人生活、养老陪护等需求，也有针对体力劳动的打扫卫生、拔草、遛狗等日常服务替代，更延伸至高危行业、深海修隧道等人类难以涉足的危险作业场景，展现出对具身智能技术社会价值的全面期待。这种功能性想象同时也伴随着复杂的社会情绪，既

有“想要购买”“兼职送快递”的消费与使用意愿，也出现了“抢饭碗”“普通人饿死”的就业焦虑，折射出具身智能技术进入社会生活前公众的多元认知与心态。



图 5 宇树科技“打工狗的日常”短视频评论词云图

4.2 民族科技自主想象：具身智能作为国家科技崛起的集体愿景

在技术展演的观看过程中，公众将技术进步与国家科技发展相联结，具身智能被赋予国家科技崛起的意义，进而凝聚为一种集体性的未来愿景。BERTopic 结果显示，Topic 3 集中于“未来”“高科技”“国产”等关键词，而“波士顿”“世界”“华为”等比较性词汇在整体语料中保持高频，表明公众在理解相关技术时，主动将其置于中国与世界的竞争环境中理解，形成了一种民族科技自主想象。

以宇树科技官方发布的“打破人形机器人速度世界纪录”短视频为例。该视频标题为“打破人形机器人速度世界纪录 3.3m/s（以前的纪录约 2.5m/s）……# 世界纪录 # 新质生产力”。这一标题采用典型的民族科技叙事框架：通过“打破世界纪录”强化“第一”与“领先”的话语符号，同时以“新质生产力”的政策标签将机器人技术突破与国家产业战略相绑定。由此，观众被引导去观看的不仅是一款更快的机器人，而是中国科技在世界舞台上的一次纪录刷新。

这一框架有效激活了公众的民族科技自

主想象，从该条视频评论数据来看，这一想象建构得到了公众的积极响应（见图6）。词频统计显示，“机器人”（526次）、“加油”（200次）、“中国”（100次）、“厉害”（95次）、“科技”（90次）、“国产”（40次）等词汇构成了最显性的公众话语集群，而“中国”与“国产”的并置，直接将技术成就锚定于国家身份认同的坐标系之中。与此同时，“波士顿”（88次）与“世界”（55次）的对举式出现，表明公众始终以全球机器人领域头部企业作为参照。此时公众的评论已超越了单纯的产品评价，而上升为一种民族科技竞争的情绪投射。这些话语共同表明，在这一过程中，具身智能早已突破企业层面的创新成果，被公众理解为“我们”的集体成就，其发展被主动纳入国家未来的共同愿景之中。



图6 宇树科技“打破人形机器人速度世界纪录”短视频评论词云图

4.3 人机关系与伦理想象：对技术反噬与人机失序的担忧

BERTopic 生成的 Topic 3（未来科技）显示，“人类”“进化”“自主”等高频词汇与大量戏谑式评论交织，整体指向一种围绕未来人机关系的批判性想象。以前述“宇树 G1 人形机器人室内格斗”视频为例，机器人连续空翻、摔倒后自主起身等极限动作，直观展现了机器在身体机能上对人类的超越，使“机器进化终将超越人类”的抽象担忧变得具

象可感。同时，机器人反复承受冲击与跌倒的测试画面，也被部分公众解读为对机器的“虐待”，为后续评论区的“复仇”叙事提供了直接的视觉依据。此外，视频中“请勿模仿极限测试”“禁止危险性改造”“全球都处于早期探索阶段”“务必了解产品局限性”等安全提示，则暗示了技术本身存在不可控的风险，激发公众对技术自主进化、机器脱离人类控制最终反噬人类的核心恐惧，进一步深化了围绕具身智能技术的伦理批判与反思。

评论分析显示（见图7），“意识”（119次）、“自我意识”（108次）、“觉醒”（96次）的高频出现，公众已不再仅将机器人视为工具，而是想象其可能具备自主意志。在此基础上，“霸凌”（91次）、“反击”（82次）、“记仇”（72次）、“复仇”（51次）等词汇构成紧密语义链，公众担忧机器人会记录并反噬人类施加的伤害。更深层的恐惧则指向对人机权力失衡的焦虑，“进化”（76次）、“自主”（87次）、“统治”（58次）、“控制”（39次）等词汇折射出公众对技术失控的终极担忧，即当机器人拥有超越人类的能力与自主性后，可能脱离控制甚至反向支配人类。而“以后”（321次）、“未来”（103次）、“十年”（44次）等时间词的高频出现，也表明这种焦虑更多指向对未来技术演化趋势的长期担忧。



图7 宇树科技“Unitree G1 学会了更多奇奇怪怪的技能”短视频评论词云图

## 5 结论与思考

本研究从技术展演与公众的技术想象两个维度考察了宇树科技的具身智能短视频科普实践，研究发现，宇树科技通过文字、图像与声音3个模态的协同编码，将抽象的具身智能技术转化为可视化、情境化且具沉浸感的技术展演。这一展演不仅完成了技术能力的展示，也为公众提供了可参与、可讨论的意义起点。从评论区的公众反馈来看，宇树科技的技术展演激发了公众的积极参与和讨论，生成了社会功能性应用、民族科技自主与人机伦理关系多元交织的社会技术想象，这些想象并非对展演内容的简单复现，而是公众在自身既有知识与经验基础上对于技术意义的再生产。由此可见，科技企业的短视频科普实践本质上是企业与公众对技术意义的协商与共构的过程。

新修订的《中华人民共和国科学技术普及法》高度重视新兴技术科普，“鼓励各类创新主体围绕新技术、新知识开展科普”，同时强调“科技企业应当把科普作为履行社会责任的重要内容”，鼓励其将科技资源转化为科

普资源。本研究也证实了以宇树科技企业为代表的前沿科技企业正在成为新型科普生态中的重要力量，其短视频科普实践的意义并非仅是产品推广，更是在客观上搭建了一个公众参与新技术讨论的开放空间。同时，本研究的发现为科技企业的短视频科普实践带来启示。科技企业的短视频科普应重视评论区的公众反馈，将其视为洞察公众技术认知与期待的重要窗口，可将公众的合理关切与期待纳入后续科普创作考量中，使科普实践更切实地回应公众诉求。此外，在技术展演的同时，可有针对性地补充应用场景说明、技术边界阐释以及伦理讨论等延展性内容，在保持内容吸引力的同时，引导公众形成理性的技术认知。但本研究仍存在一定局限性，例如，研究基于单一企业与单一平台展开，相关结论的适用性还有待后续研究进一步检验。未来可引入不同技术领域与不同平台的比较视角，或采用纵向追踪设计，考察公众技术想象的历时变迁及其影响因素，从而深化对数字媒介环境下科技企业短视频科普实践的理解。

## 参考文献

- [1] 抖音. 2025 抖音科技内容生态报告 [EB/OL]. (2026-01-05) [2026-02-01]. <https://trendinsight.oceanengine.com/arithmetic-report/detail/1080?source=0923hangyezhiinan>.
- [2] Pinch T J, Bijker W E. The Social Construction of Facts and Artefacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other[J]. *Social Studies of Science*, 1984, 14(3): 399-441.
- [3] 席志武, 李姗姗, 沈鲁. 科普新形态与科普期刊短视频传播的转型路径探究 [J]. *出版广角*, 2022(9): 11-16.
- [4] 冷淞. “人间烟火”的景象迁移与“现实图景”的双向建构——新冠肺炎疫情下短视频与拟态环境重构 [J]. *新闻与传播研究*, 2020, 27(9): 111-125, 128.
- [5] 黄荣丽, 王大鹏, 金梦瑶, 等. 科普短视频评价指标体系：概念澄清、价值审思与构建路径 [J]. *科普研究*, 2023, 18(6): 32-40, 95-96.
- [6] 江作苏, 李梦娇. 短视频科学传播内容生产与所涉伦理探讨 [J]. *中国编辑*, 2022(1): 61-66.
- [7] 赵淑萍, 李超鹏. 突发公共卫生事件报道中主流媒体情感传播策略研究——以三大央媒新冠肺炎疫情报道短视频产品为例 [J]. *中国出版*, 2021(4): 46-51.
- [8] 徐萌. 疫情语境下医学科普短视频大众化传播策略创新——以“丁香医生”抖音号为例 [J]. *传媒*, 2020(18): 38-40.
- [9] 陈维超, 周杨玲. 我国科普期刊抖音短视频传播效果影响因素实证研究——以中国优秀科普期刊为例 [J]. *中国科技期刊研究*, 2023, 34(12): 1616-1622.

- [10] 刘俊冉, 武珂冉, 商建辉. 科普期刊与科普期刊短视频: 共生理与共生路径 [J]. 出版发行研究, 2026(1): 53-59.
- [11] 孙玉珠, 胡明. 知识生产与再生产的媒介化: B 站泛知识类短视频的接合实践 [J]. 新闻春秋, 2025(1): 62-72.
- [12] Phillips D J. The Social Construction of a Secure, Anonymous Electronic Payment System: Frame Alignment and Mobilization Around Ecash[J]. Journal of Information Technology, 1998, 13(4): 273-283.
- [13] Savic M. Research Perspectives on TikTok & Its Legacy Appsl from Musical. ly to TikTok: Social Construction of 2020's Most Downloaded Short-Video App[J]. International Journal of Communication, 2021, 15: 22-22.
- [14] Graf A, Sonnberger M. Responsibility, Rationality, and Acceptance: How Future Users of Autonomous Driving are Constructed in Stakeholders' Sociotechnical Imaginaries[J]. Public Understanding of Science, 2020, 29(1): 61-75.
- [15] Kline R, Pinch T. Users as Agents of Technological Change: The Social Construction of the Automobile in the Rural United States[J]. Technology and Culture, 1996, 37(4): 763-795.
- [16] Smith W. Theatre of Use: A Frame Analysis of Information Technology Demonstrations[J]. Social Studies of Science, 2009, 39(3): 449-480.
- [17] Simakova E. RFID 'Theatre of the Proof': Product Launch and Technology Demonstration as Corporate Practices[J]. Social Studies of Science, 2010, 40(4): 549-576.
- [18] Butler J, Trouble G. Feminism and the Subversion of Identity[J]. Gender Trouble, 1990, 3(1): 3-17.
- [19] Wang Y, Feng D W, Zhang J. Transpositioning Through Visual Translanguaging Creativity: A Multimodal Analysis of a Chinese Wanghong Vlogger's Identity Performance on YouTube[J]. Discourse, Context & Media, 2025, 68: 100937.
- [20] Zhao Y, Mahfoodh O H A, Subramaniam I, et al. Pixels, Poses, and Power: A Multimodal Analysis of Self-Representation and Femininity Among Female Fitness Influencers on Chinese Social Media[J]. Discourse, Context & Media, 2025, 68: 100954.
- [21] Dreamscapes of Modernity: Sociotechnical Imaginaries and the Fabrication of Power[M]. Chicago: University of Chicago Press, 2019.
- [22] Jasanoff S, Kim S H. Containing the Atom: Sociotechnical Imaginaries and Nuclear Power in the United States and South Korea[J]. Minerva, 2009, 47(2): 119-146.
- [23] 何映霏. 理解 - 协商 - 阐释: 技术创新采纳中新闻从业者的 AI 想象 [J]. 新闻记者, 2025(12): 69-82.
- [24] 燕道成, 吴涵. 与平台共谋: 青年博主的账号运营与平台想象 [J]. 当代青年研究, 2025(3): 19-33.
- [25] 谭小荷. AI 想象的展褶: 技术乐观主义与无能主体的共生——以普通人对 Chat GPT 的社会技术想象为中心 [J]. 新闻界, 2023(11): 52-65, 96.
- [26] 赵瑜, 梁冰青. AI 何以“越界”? 用户对社交机器人的关系想象与边界反思 [J]. 新闻与写作, 2026(1): 36-46.
- [27] 崔浩男. 多模态档案资源知识服务体系研究 [J]. 档案学通讯, 2025(4): 120-123.
- [28] 刘果, 汪小伢. 标题特征对数字媒介内容传播效果的影响——基于新闻评论类微信公众号标题的实证研究 [J]. 新闻与传播评论, 2020, 73(6): 29-39.
- [29] Kress G, Van Leeuwen T. Reading Images: The Grammar of Visual Design[M]. Routledge, 2020.
- [30] 张媛. 观念的图像化: 中华民族共同体意识的视觉修辞分析 [J]. 湖南大学学报 (社会科学版), 2025, 39(3): 161-168.
- [31] 吴月. 数字新闻漫画的视觉传播策略——以“乌合麒麟”漫画作品为例 [J]. 传媒, 2025(7): 81-83.
- [32] Van Leeuwen T. Speech, Music, Sound[M]. London: Macmillan, 1999.
- [33] Schafer R M. The Soundscape: Our Sonic Environment and the Tuning of the World[M]. Simon and Schuster, 1993.

(编辑 颜 燕 马 赫)

## 关于更换《科普研究》官方电子邮箱的声明

为进一步规范期刊联系渠道, 自即日起, 《科普研究》官方电子邮箱变更为 kpyj2026@163.com。原邮箱 kepuyanjiu@163.com 将于 2026 年 6 月 13 日起停止使用。请广大作者、读者及审稿专家惠存新址, 由此给您带来的不便, 敬请谅解。

特此声明。

the dual relationship between AI and evidence-based communication. ( 1 ) The empowering role of AI in evidence-based science communication: AI demonstrates enabling potential in stages such as data aggregation, evidence screening, and impact feedback, thereby enhancing the efficiency and reach of evidence-based communication. ( 2 ) The challenges posed by AI hallucinations to science communication: These challenges are reflected not only in content authenticity but also in the procedural legitimacy of knowledge production. They affect dimensions including content foundation, knowledge production, knowledge stability, and the efficiency of consensus-building, causing misinformation to escalate from isolated inaccuracies to structural errors. Based on these findings, this study proposes a human-AI collaboration pathway for evidence-based science communication, guided by the principles of evidentiality, traceability, adaptability, and publicness. By clarifying the AI-side intervention mechanisms and the key nodes of human-AI collaboration under each principle, this study constructs a comprehensive evidence-based science communication system spanning “content production—process control—pace regulation—impact feedback.”

**Keywords:** AI hallucination; evidence-based science communication; knowledge production; Artificial Intelligence

**CLC Numbers:** G206 **Document Code:** A **DOI:** 10.19293/j.cnki.1673-8357.2026.02.011

## Technological Performance and Public Imagination: Multimodal Science Communication of Embodied AI on Short-Video Platforms: A Case Study of Unitree Robotics

Wu Wenxi Tian Ziyi Hu Ting

[ School of Media Science ( School of Journalism ) , Northeast Normal University , Jilin 130024 ]

**Abstract:** Short-form video has emerged as a pivotal medium for the dissemination of emerging technologies, positioning technology enterprises as one of the main agents of science communication. Focusing on Unitree Robotics’ official Douyin account, this study draws on a sample of 81 videos and 4 114 comments to examine the interaction between corporate technological performance and public sociotechnical imaginaries. The study integrates multimodal discourse analysis with text mining to capture both the semiotic construction of technology and its discursive reinterpretation by the public. The findings reveal that embodied AI is rendered as a visualized, contextualized, and immersive technological performance through the orchestration of textual, visual, and auditory modes. Correspondingly, public discourse generates multiple, intersecting sociotechnical imaginaries structured around functional utility, techno-nationalism, and human-machine relations. These dynamics suggest that technological meaning is not unilaterally transmitted but continuously negotiated and stabilized through platform-mediated interactions between technology enterprises and the public. This study reveals that technological performance goes far beyond mere technology demonstration. It opens up a crucial space for the public to engage in the social construction of technology and to shape sociotechnical imaginaries. Consequently, these practices reconfigure science communication from one-way dissemination toward the participatory co-production of meaning.

**Keywords:** embodied AI; technological performance; sociotechnical imaginaries; short-video science communication; emerging technology

**CLC Numbers:** G206; N4 **Document Code:** A **DOI:** 10.19293/j.cnki.1673-8357.2026.02.012