

人工智能对媒体行业技能与未来就业的影响：基于机器学习和网络分析

（整个媒体行业分析结果）

论文正文主要聚焦内容生产类媒体行业，分析的是人工智能技术对内容生产类媒体行业的技能与未来就业的影响。媒体行业是一个更为复杂的系统。支撑内容生产的其它职业的变化也会最终影响媒体内容生产相关职业的发展。因此，全面地刻画整个媒体行业应对人工智能技术的挑战也具有重要意义。在本附件中，补充呈现了对整个媒体行业的影响，兼顾整个媒体行业的图景。主要从以下几个方面展开：

1、职业数据

整个媒体行业的一共包括了 17 种媒体职业。数据来自美国标准职业分类数据库和职业信息网站的分类，27-3000（Media and Communication Workers）和 27-4000（Media and Communication Equipment Workers）定义了媒体行业所涉及的 17 种职业。前者（27-3000）具体包括的职业有 10 种：广播电视播音员、公共广播系统和其他播音员、新闻评论员、记者、主编、公共关系专家、技术文档写作者、文案、广告脚本撰写员、诗人，文学、戏剧、歌曲等作家、口译和翻译。后者（27-4000）具体包括的职业有 7 种：音视频设备技术人员、广播技术员、无线电设备操作员、声效工程技术员、摄影家和摄像师、视频，电视和电影生产制作者、电影，视频编辑。这两大类职业涵盖采访、编辑、写作、拍摄、剪辑、播报各个环节。

2、研究方法

- （1）使用 Frey 和 Osborne（2017）在研究中构造出的 70 种职业的数据作为训练集，基于 2018 年的 O*NET 数据重新训练机器学习模型，得到了 17 种媒体职业的计算机化水平。
- （2）利用显示性指数比较优势算法计算得到媒体行业的核心技能、未来技能。
- （3）利用多元线性回归分析，探讨了计算机化水平、受教育水平、技能的认知水平、职业类别变量与职业薪资之间的关系。

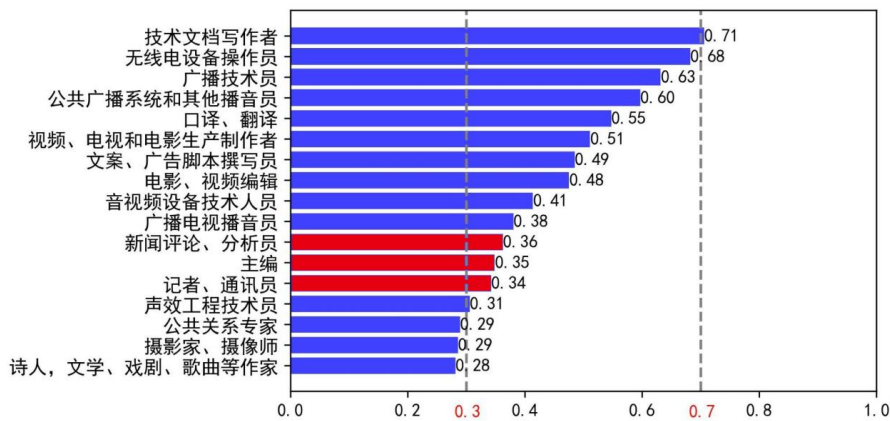
3、研究发现

3.1 媒体行业的计算机化水平

通过机器学习模型计算得到了 967 种职业的计算机化水平，该值由 0 到 1，值越大表示计算机化水平越高、越容易被机器替代。以 0.3 和 0.7 为分界值将职业划分低风险、中风险、高风险三类。图 1 展示了媒体行业的未来计算机化水平。红色代表内容生产类媒体

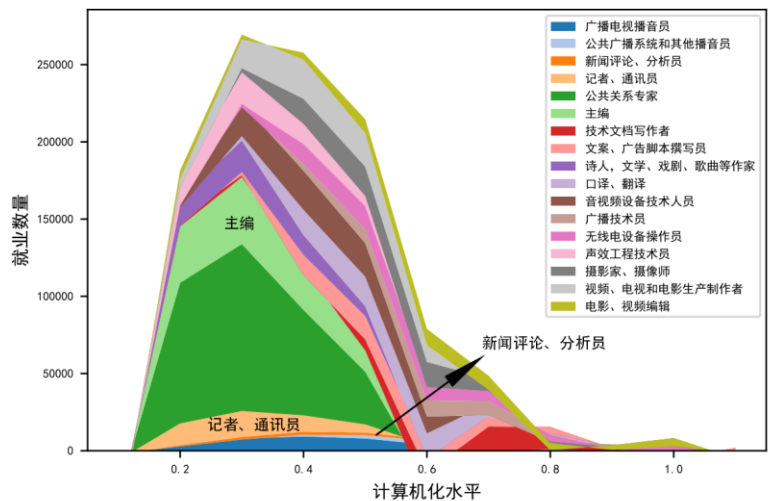
行业，蓝色为其他媒体行业。媒体行业的 17 个职业有一个职业（技术文档写作者）处于高风险水平（计算机化水平高于 0.7，占比 5.9%）。除此之外，13 个职业处于中风险水平（计算机化水平介于 0.3 至 0.7 之间，占比 76.5%），3 个职业处于低风险水平（计算机化水平低于 0.3, 占比 17.6%）。

图 1. 媒体行业未来计算机化水平



对媒体行业的就业数量进行统计，94%的媒体从业者处于中低计算机化水平。具体如图 2 所示：高计算机水平区间（计算机化水平 ≥ 0.7 ）的就业人数为 50350 人，占比 6.0%；中计算机化水平区间（ $0.3 < \text{计算机化水平} < 0.7$ ）的就业人数为 447530 人，占比 53.8%；低计算机化水平区间（ $\text{计算机化水平} \leq 0.3$ ）的就业人数为 333800 人，占比 40.2%。

图 2. 媒体行业未来计算机化水平与就业数量



与处于同样风险水平的其它行业相比，媒体行业更为安全。具体而言，在高风险类别的职业中，媒体行业计算机化水平低于平均水平（ $t(166) = 15.483, p < .001$ ）；在中等风险类别的职业中，媒体行业依然低于平均水平（ $t(594) = 6.777, p < .001$ ）；而在低风

险类别的职业中，媒体行业计算机化水平高于平均水平（ $t(204) = 17.404, p < .001$ ）。

在 Frey 和 Osborne（2017）看来，低风险、中风险、高风险区间意味着被人工智能技术所取代的先后顺序。其中，高风险职业将会相对较早被人工智能技术所取代显然。Frey 和 Osborne（2017）认为高风险职业被替代可能需要十到二十年。在第一波被人工智能技术所替代的媒体行业只有一个（即技术文档写作者）。因此从短期来看，内容生产类媒体行业的就业仍然较为安全；但是在更长远的未来，其他类型媒体行业将面临较高的替代风险。媒体从业者需调整自身技能来应对未来计算机化挑战。

3.2 媒体行业的核心技能

媒体行业的技能注重认知理解和表达沟通。技能的 O*NET 值帮助我们解媒体行业内部的核心技能。从 O*NET 数值来看，媒体行业内部最为核心的技能依次是获取信息（0.83）、英语语言（0.81）、电脑交互（0.81）、通讯与媒体（0.78）、创造性思维（0.76）、更新和调用相关知识（0.75）、口语表达（0.74）、语言理解（0.73）等能力。

表 1. 媒体行业与全行业 O*NET 水平均值 T 检验结果

技能名称	内容生产类 媒体行业 O*NET 均值	其他媒体行 业 O*NET 均值	全行业(除 媒体行业) O*NET 均值	内容生产类媒体 行业与其他媒体 行业 T 检验结果	整个媒体行 业与全行业 T 检验结果
信息获取	0.92	0.81	0.8	4.43***	10.05***
英语语言	0.96	0.78	0.66	3.71**	31.97***
电脑交互	0.77	0.81	0.63	1.4	5.51***
通讯与媒体	0.91	0.76	0.32	4.07**	93.17***
创造性思维	0.73	0.77	0.59	0.81	33.04***
更新、运用相关知识	0.73	0.75	0.68	1.86	5.05***
口语表达	0.78	0.73	0.67	1.88	17.84***
口语理解	0.77	0.73	0.68	2.65*	15.07***
与上下级或同事沟通	0.78	0.72	0.74	1.86	4.02***
建立和维护人际关系	0.74	0.72	0.67	0.61	13.05***

注： $p < 0.05^*$ ， $p < 0.01^{**}$ ， $p < 0.001^{***}$

同时，本文计算了内容生产类媒体行业的 O*NET 均值、其他类型媒体行业的 O*NET 均

值以及全行业（除媒体行业）O*NET 均值。T 检验发现：内容生产类媒体行业对信息获取、英语语言、通讯与媒体、口语理解的重视程度显著高于其他媒体行业。所有的媒体行业对这十个核心技能的要求显著高于媒体行业以外的其它行业（T 检验结果见表 1）。

显示性比较优势指数(RCA 值) 更好地衡量了媒体行业与其它行业相比所具有的核心技能。从显示性比较优势指数(RCA 值) 来看，媒体行业与其它行业相比更有竞争力的核心技能依次是创造性思维、原创性、电脑交互、写作（Writing）、思维流畅性、写作表达（Written Expression）、与组织外部人员沟通、阅读理解、写作理解（Written Comprehension）、口齿清晰。本文计算了内容生产类媒体行业的 RCA 均值、其他类型媒体行业的 RCA 均值以及全行业（除媒体行业）RCA 均值。T 检验发现：内容生产类媒体行业对原创性、写作、与组织外部人员沟通、阅读理解、写作理解、口齿清晰的要求显著高于其他类型媒体行业。所有媒体行业对这十个技能的要求显著高于其它行业（具体 T 检验结果见表 2）。此外，与其它行业相比，媒体行业对写作的要求更高。除此之外，我们还可以计算学科知识对于媒体行业的相对重要性。采用显示性比较优势指数(RCA 值) 计算得到媒体行业最为重要的学科知识依次是艺术、传播与媒体、历史和考古学、电信通讯、哲学和宗教、计算机与电子学、社会学和人类学、市场销售、外语。

表 2. 媒体行业与全行业显示性比较优势指数水平均值 T 检验结果注：

技能名称	内容生产类 媒体行业 RCA 均值	其他类媒 体行业 RCA 均值	全行业(除媒 体行业) RCA 均值	内容生产类媒体 行业与其他媒体 行业 T 检验结果	内容生产类媒 体行业与全行 业 T 检验结果
创造性思维	1.35	1.41	1	0.63	45.53***
原创性	1.58	1.35	1	2.35*	44.97***
电脑交互	1.31	1.38	1.01	1.15	6.65***
写作	1.59	1.27	1.01	2.54*	2.90*
思维流畅性	1.4	1.29	1	1.22	38.67***
写作表达	1.52	1.24	1.01	2.87*	32.48***
与组织外部人员沟 通	1.59	1.21	1	4.47***	25.57***
阅读理解	1.49	1.17	1.01	4.47***	28.99***
写作理解	1.44	1.17	1.01	4.08**	27.47***
口齿清晰	1.4	1.18	1.02	3.07**	32.18***

p<0.05*, p<0.01**, p<0.001***

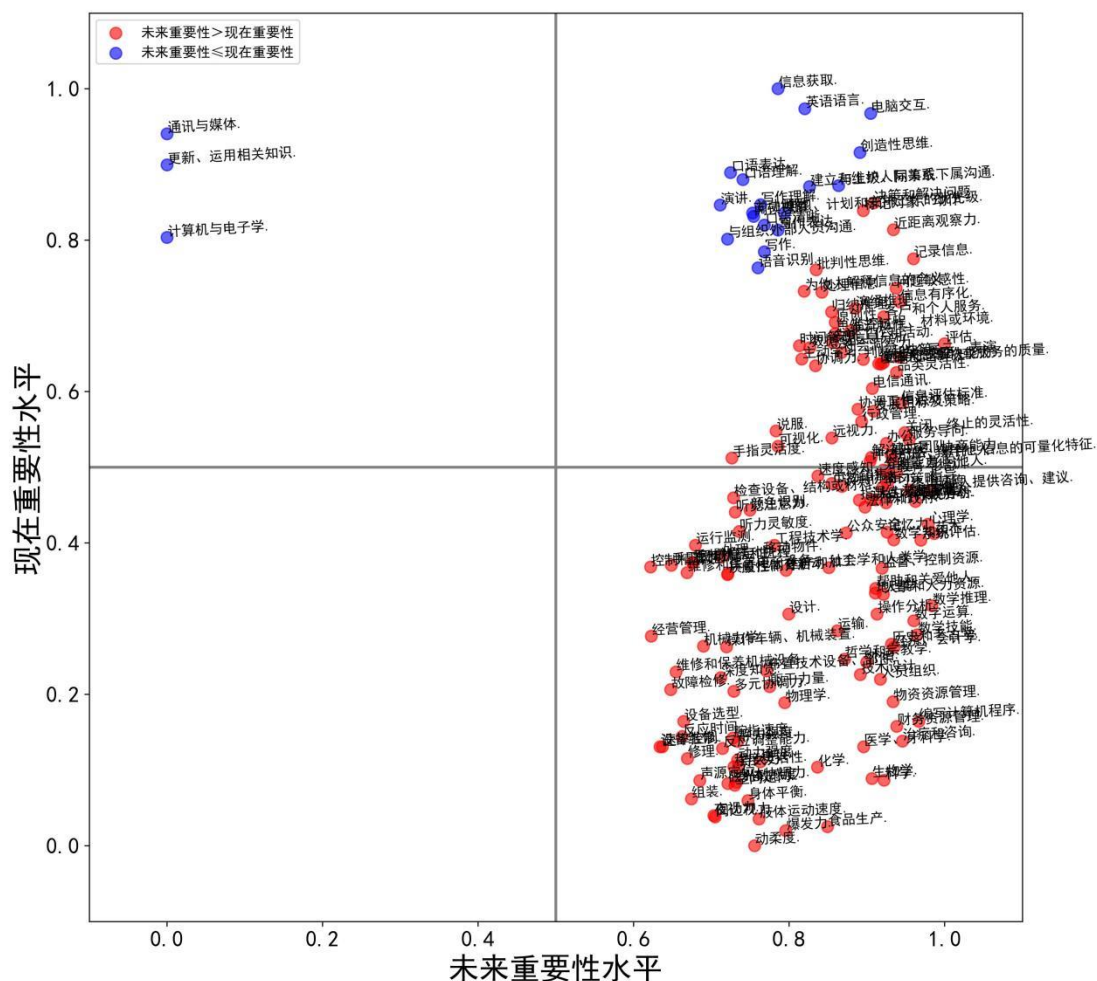
3.3 媒体行业的未来技能

本研究通过计算技能接近性，得到与每一个媒体职业接近性最高的技能。未来技能或新技能通常与现有技能相似或者互补，确保技能迁移的成本小而收益大。未来技能是未来重要性提升的技能。需要强调的是，未来技能的计算是为了挖掘新技能，但是这并不意味着现有的核心技能变得不重要。

图 3 展示了媒体行业现在和未来的技能分布情况。横坐标为技能接近性水平，显示

了技能在未来的重要性提升程度；纵坐标为技能的 O*NET 水平，衡量了技能在现在的重要性。每个点代表一种技能。对每一种技能的现在重要性水平和未来重要性水平进行了最大最小归一化。红色的点表示技能在未来的接近性大于现在的重要性，即重要性有较大提升的技能；蓝色的点表示技能现在重要性大于未来接近性，即职业现在所重视的核心技能。图 3 左上角的技能是在现阶段更重要的技能；右下角的技能是指在未来对于媒体行业来说重要性有所提升的技能。

图 3. 媒体行业现在和未来的技能分布



整个媒体行业未来技能的覆盖范围很广，既包括评估、评价等软技能，也包括编程和数学等硬技能。例如，未来技能的前十名分别是评估（Monitoring, 0.52）1、艺术（Fine Arts, 0.51）、数学推理（Mathematical Reasoning, 0.51）、为他人提供咨询&建议（Provide Consultation and Advice to Others, 0.51）、心理学（Psychology, 0.51）、

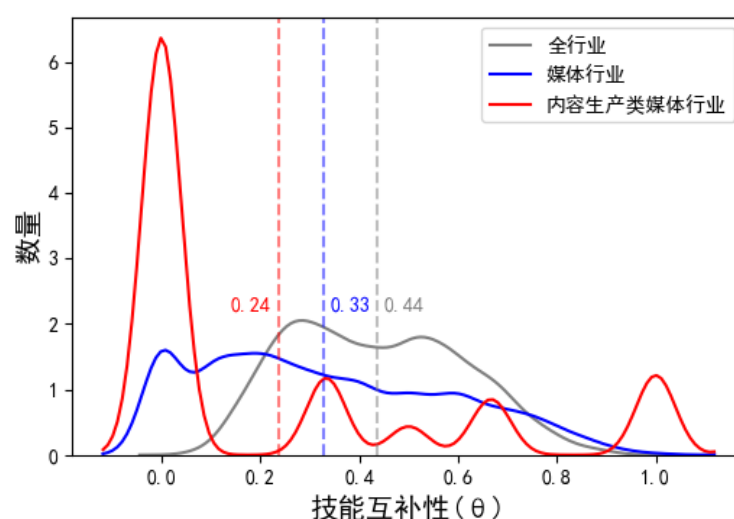
¹此处的“监测”是指评估自己、其他个人或组织的绩效，以进行改进或采取纠正措施。
<https://www.onetonline.org/find/descriptor/result/2.A.2.d>

系统评价 (Systems Evaluation, 0.51)、编写计算机程序 (Programming, 0.50)、数学技能 (Mathematics, 0.50)、系统分析 (Systems Analysis, 0.50)、数字运算的能力 (Number Facility, 0.50)。

3.4 媒体行业技能互补性

技能之间的互补性衡量了两个技能互补的程度，为人工智能时代媒体从业者的技能更新提供了方向。如图 4 所示，全行业技能互补性具有双峰分布的特点，而媒体行业的技能互补性低。通过双总体 T 检验，进一步证实了发现，媒体行业的技能互补性显著低于全行业 ($t(25758) = -39.588, p < .001$)。内容生产类媒体行业技能互补性显著低于整体媒体行业水平 ($t(25758) = -23.909, p < .001$)。技能的互补性低意味着该职业所需的技能更多独立出现，较少与其他技能成对出现。内容生产类媒体行业的技能互补性低限制了基于技能互补性的团队合作，同时也意味着不容易进行技能的迁移。

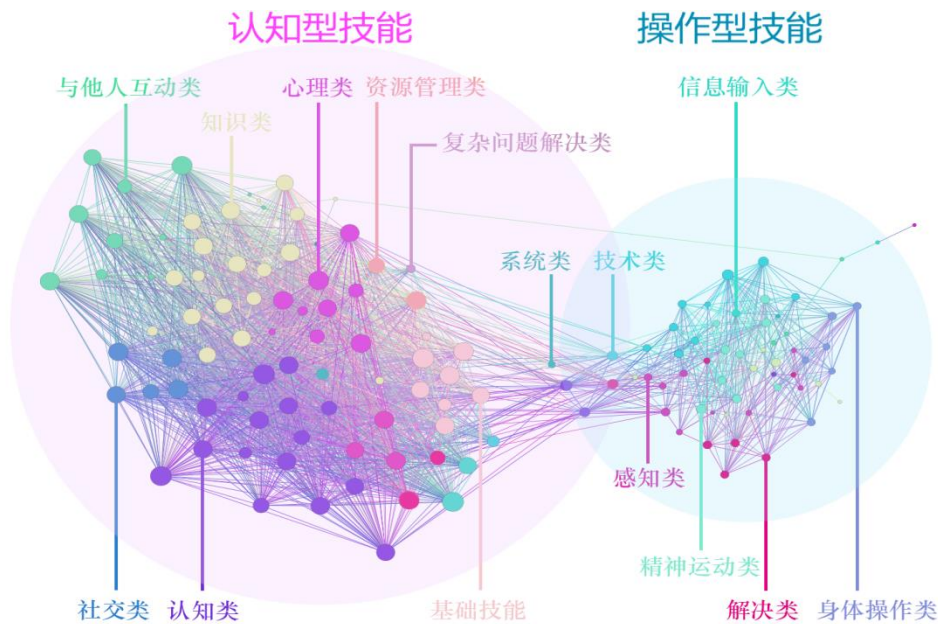
图 4. 媒体行业技能互补性 (θ) 分布



以 0.6 为分界线 (Alabdulkareem, et al., 2018)，将技能互补性划分高、低两个区间。相比于其它行业，媒体行业在高互补性区间里数值更高 ($t(5029) = 10.112, p < .001$)，而在低互补性区间里数值更低 ($t(20727) = -61.556, p < .001$)。为了进一步形象化理解媒体行业的技能特点，进行了网络可视化。如图 5 所示，每一个节点代表一种技能，两个节点之间的连边代表技能存在互补关系。为了展示更清晰，图 5 仅展示互补性大于 0.6 的连边 (Alabdulkareem, et al., 2018)。节点的大小代表该技能与其他技能互补程度，点越大，说明互补性水平越高。²

² Alabdulkareem, A., Frank, M. R., Sun, L., AlShebli, B., Hidalgo, C., & Rahwan, I. (2018). Unpacking

图 5. 媒体行业技能互补性网络



基于技能网络，有以下发现：首先，媒体行业的技能网络可以划分为**认知型技能**和**操作型技能**两个大的网络社区。其次，认知型技能的节点比操作型技能的节点更大，即互补性更强。与之契合，统计分析发现：认知型技能的互补性水平显著高于操作型技能（ $t(1935)=11.14, p<.001$ ）。技能的度中心性取值越大，该技能对其他技能的影响越大。结果表明：第一，认知型技能的度中心性高于操作型技能（ $t(159)=-3.491, p<.005$ ）。第二，技能的度中心性水平与技能互补性水平正相关（ $R=0.58, p<.001$ ）。例如，社会洞察、沟通交流等认知型技能同时具有较高的度中心性水平和互补性水平。

3.5 回归分析

本研究通过多元线性回归来分析职业薪资水平的影响因素。表 3 模型 1、2、3 展示的是全行业（940 种）的薪资对照情况；表 3 模型 4、5、6 展示的是人文社科类职业（371 种）的薪资情况，其中模型 1、2、4、5 对照组是媒体行业，模型 3、6 对照组是内容生产类媒体行业。首先，表 3 模型 1 展示了全行业薪资与教育水平、认知技能水平、计算机化概率之间的关系。回归发现受教育水平与薪资收入正相关（ $\beta = 0.001***, p < .001$ ）；职业的未来计算机化风险与薪资负相关，职业未来计算机化风险越高，薪资越低（ $\beta = -$

the polarization of workplace skills. *Science advances*, 4(7), eaao6030.

0.327***, $p < .001$), 认知技能水平与薪资相关关系不显著。为了比较媒体行业与其它行业薪资的差异, 表 3 模型 2 将媒体行业为参照组与全行业做对比。发现管理类、商业和金融运营类、计算机和数学类、建筑和工程类、建筑及采掘类、运输类职业薪资显著高于媒体行业。而媒体行业的薪资水平显著高于社区和社会服务类、教育和图书馆类、健康支持类、生活、物理和社会科学类、食品服务类、个人护理和服务类、办公室和行政类职业。进一步, 我们将内容生产类媒体行业作为参照组进行回归, 如表 3 模型 3 所示, 发现内容生产类媒体行业的薪资几乎低于其他所有职业。

从表 3 模型 2 可以看出整体媒体行业薪资在人文社科类行业中具备优势。为了比较这种差异, 筛选出了 371 种人文社科类职业进行多元线性回归。如表 3 模型 4 所示, 在人文社科类行业中, 受教育水平与薪资收入正相关 ($\beta = 0.001***$, $p < .001$); 认知技能水平与薪资正相关 ($\beta = 2.299***$, $p < .001$); 职业的未来计算机化风险与薪资负相关 ($\beta = -0.421***$, $p < .001$)。将人文社科类职业作为控制变量进行分析 (表 3 模型 5), 认知技能水平变得不显著, 其他结果与全行业结果类似, 媒体行业薪资虽然低于管理类、商业和金融运营类职业, 但是显著高于其他大部分人文社科类职业。回归模型 R 方为 0.82, 具有较高解释力。表 3 模型 6 将内容生产类媒体行业作为对照组与人文社科类行业做对比, 结果发现: 内容生产类媒体行业薪酬显著低于其他人文社科类职业。综上, 媒体行业的薪资水平显著高于大部分人文社科类行业, 低于理工科类职业; 在媒体行业内部, 内容生产类媒体行业是偏低的。高教育水平、低职业计算机化水平有利于获得更高工资。

表 3. 职业薪资的回归分析结果

变量	全行业			人文社科类行业		
	模型 1	模型 2	模型 3	模型 4	模型 5	模型 6
常数	10.631***	10.443***	10.007***	8.890***	10.17***	9.28***
教育水平	0.001***	0.001***	0.001***	0.001***	0.001***	0.001***
认知技能水平	-0.134	-0.007	-0.071	2.299***	0.453	0.308
计算机化水平	-0.327***	-0.298***	-0.300***	-	-0.193**	-0.179*
				0.421***		
职业类别						
管理类		0.460***	0.935***		0.462***	1.436***
商业和金融运营类		0.154**	0.631***		0.121*	1.096***
社区和社会服务类		-0.495***	-0.021		-0.461***	0.512***
法律类		0.029	0.503***		0.027	0.993***
教育和图书馆类		-0.324***	0.146***		-0.272***	0.690***
媒体（非内容生产类）			0.507***			0.986***
健康支持类		-0.305***	0.171***		-0.325***	0.647***
食品和服务相关		-0.370***	0.108		-0.434***	0.541***
个人护理和服务类		-0.365***	0.112*		-0.404***	0.571***
销售类		0.097	0.577***		0.035	1.015***
办公室和行政类		-0.104*	0.376***		-0.183***	0.792***
农林渔类		-0.061	0.412***			
计算机和数学类		0.286***	0.760***			
建筑和工程类		0.175***	0.646***			
生活、物理和社科类		-0.155**	0.315***			
医疗和技术类		-0.041	0.426***			
保护服务类		0.142*	0.617***			
建筑和清洁维护类		-0.119	0.354***			
建筑及挖掘		0.203***	0.675***			
安装维护和维修类		0.098	0.569***			
生产类		0.027	0.499***			
运输类		0.220***	0.692***			
R-square	0.59	0.78	0.79	0.53	0.82	0.84

注：p<0.05*, p<0.01**, p<0.001*** 职业类别为哑变量，模型 1、2、4、5 中职业类别变量的对照组是媒体行业，模型 3、6 中职业类别变量的对照组是内容生产类媒体行业

4、研究总结

研究发现媒体行业在未来十到二十年较安全，但长远来看会受到较大的技术冲击，内容生产类媒体行业的计算机化水平显著低于其他类型媒体行业。其次，技能的进步存在偏向。现阶段媒体从业者的核心技能以写作、信息获取、理解和表达等为主，而未来从业者则需要掌握数字化技能、培养数学思维，内容生产类媒体行业还需增强决策判断和服务导向等未来技能。未来就业的变化会受到技能偏向的影响。与操作类技能相比，认知类技能的度中性水平显著高于操作类技能的度中性水平，因此在职业转换上更具优势。最后，回归模型的结果表明教育水平越高、计算机化风险越低，职业收入越高。