

AIGC对出版产业链主体会计核算的影响及对策建议

焦建玲 桑广成 桑博涵 张雪亮

摘要：人工智能生成内容(AIGC)在出版行业的应用使得出版产业链主体更加丰富多元、出版业生产范式和内容供给模式得以改变和创新,也给出版产业链主体内容数据流转的会计核算带来影响。本文围绕AIGC辅助创作的语料采集和内容输出两阶段,梳理相关会计准则和会计政策依据,在界定数据版权归属的基础上,以出版产业链的延展为边界,探讨AIGC辅助创作引起的出版产业链相关企业会计核算问题并提出相关建议。

关键词：人工智能生成内容(AIGC);出版企业;版权归属;会计核算;数据资产

中图分类号：F235.92 **文献标志码：**A **文章编号：**1003-286X(2025)13-0044-05

人工智能生成内容(AIGC)通过机器学习、自然语言处理等技术,能够自动生成文本、图像、音频、视频等多种形式的内容,降低了内容创作的门槛且释放创作能力,支持数字内容与其他产业的多维互动,融合渗透辐射较多领域从而带来新业态、新模式。AIGC辅助创作延长了出版产业链,并给产业链主体内容生产的会计核算带来新问题。其涉及的出版产业链主体主要包括:大模型公司、语料预训练公司、数据加工公司、出版企业、版权人等,以上主体既可能存在身份上的重合关系,如拓尔思信息技术股份有限公司(以下简称拓尔思)自己建设语料库并开发大模型;也可分属不同的法律主体,如中文在线集团股份有限公司(以下简称中文在线)为其他大模型提供语料数据集。从版权看,AIGC辅助创作内容在语料采集和内容输出阶段均存在版权归属争议,这直接影响到出版企业资产和成本费用的确认与计量。本文围绕数据输入和内容输出,遵循会计准则及相关会计政策,在界定数据版权归属的基础上,以出版产业链的延展为边界,探讨AIGC辅助创作引起的出版产业链相关

企业内容生产会计核算问题并提出相关建议。

一、AIGC辅助创作内容产品涉及的版权权属问题

对训练数据版权的清晰界定是出版业大模型规范应用的前提,也是会计核算的关键。在我国,大模型在语料采集和内容生成等环节仍存在版权争议,未达成统一共识和标准。

(一)语料采集阶段

该阶段需解决将享有著作权的作品加工成供大模型使用的语料是否构成侵权问题。根据美国智库Epoch研究,到2028年,互联网上所有高质量文本数据将被使用完毕,将出现数据荒。AI研究专家和从业人员都认为,内容商(特别是传统出版商)价值崛起,传统纸质出版物都是专业人员精心写作和编辑的,具有规模庞大、质量较高、类型丰富、数字化加工基础较好、产权清晰、导向正面等特点。我国当前面临训练语料尤其是高质量有版权的中文语料不足的困境,

作者简介：焦建玲,河北经贸大学副教授;桑广成,河北出版传媒集团有限责任公司,正高级会计师,税务师,全国高端会计(后备)人才企业类十四期学员;桑博涵,复旦大学博士研究生;张雪亮,河北出版传媒集团有限责任公司。

阻碍了我国 AI 应用进程，主要是由于将有版权的语料用于训练存在侵权风险。训练 AI 智能体所使用的语料库是否获得了合法的使用权，不仅关乎平台自身的合法运营，还涉及原创作者的权益保护及整个内容创作生态的平衡。对此问题，目前有如下观点：

第一种观点认为，作品以数字化形式被复制于语料库中的行为属于复制权规制范畴，在此基础上可采用“合理使用”和“法定许可”两种合法路径。“合理使用”使用人无须经过著作权人许可、亦无须向其支付报酬，多用于科学研究。鉴于语料库产业链条上的企业多为商业机构，采用“合理使用”空间不大。“法定许可”允许使用人不经著作权人许可使用其作品，但应向著作权人支付相应报酬。报酬价格标准制定及由谁分配、如何分配等问题需先行解决。第二种观点认为，在当前的著作权框架下，语料训练阶段的复制行为没有侵犯著作权。理由在于受版权保护的作品用于人工智能训练（通常使用文本与数据挖掘技术 <TDM>）不同于传统复制：一是作品被打碎和分解为碎片化的“语料”，不再是独立使用的对象；二是属于非传播性复制、非特定性表达、非作品性使用；三是使用时间不同于传统复制权有授权期限限制，一般需要长期使用。

国外对受版权保护的 TDM 做法不一。美国采用“转换性使用”拓展了合理使用的边界；欧盟对受版权保护内容的任何使用都必须获得相关权利人的授权，除非适用相关的版权例外和限制；日本对 TDM 持开放态度，没有对实施主体和实施目的进行限制。从美国 17 位作家起诉 OpenAI 案、法国对谷歌大模型训练数据侵犯版权开出 2.5 亿欧元罚单可以看出，美国和欧盟对于将 TDM 用于商业用途的行为需要支付版权人报酬。从国内出版企业角度来看，除买断版权和已经超过著作权保护期的作品加工用于 TDM 无风险外，仍处于授权许可期内和超过授权许可期的作品，由于 TDM 是否属于复制权规制范围尚不明确，将其用于 TDM 会产生侵权风险。

（二）内容输出阶段

内容输出阶段需解决两个问题：一是人工智能生成作品是否拥有著作权；二是如拥有著作权，权利人是谁。关于 AIGC 能否纳入著作权法的保护范畴，理论界形成了“可版权说”“不可版权说”及“其他权属说”三类不同观点。“可版权说”认为 AIGC 具有独创性，应受著作权法保护。“不可版权说”认为 AIGC 不应纳入著作权法的保护范畴，因为这些内容并非由人类创造，而是由计算机模型和算法从海量

数据中撷编而成，缺乏独创性和原创性。“其他权属说”认为 AIGC 应当受到一种新的权利属性的保护，从而平衡生成者和使用者的权益。从实际案例来看，对于人工智能生成作品，美国版权局 2025 年 1 月最新裁定认为人类对 AI 素材进行选择、协调和编排形成的独创性整体（如拼贴）可类比汇编作品受版权保护，中国法院倾向于将 AI 视为工具且生成作品具备“独创性”与“人类智力投入”要素时有著作权，案例为 2024 年北京互联网法院首例涉及 AI 绘画大模型训练著作权侵权案判决。如 AIGC 可拥有著作权，其著作权归属同样是一个备受关注的课题。理论上存在以下几种可能：一是属于 AI 研发者；二是属于 AI 使用者；三是 AI 研发者和使用者共享；四是不授予版权而采用其他方式保护。实践中，DeepSeek、OpenAI、商汤商量、豆包、Kimi 约定内容输出版权归属于使用者，文心一言和科大星火则约定由公司及相关方保留相关权益，Suno 约定付费客户享有生成歌曲的所有权、免费客户歌曲所有权由平台享有。

二、AIGC 辅助创作内容产品的会计处理政策依据

语料采集阶段的会计处理一般应纳入《企业数据资源相关会计处理暂行规定》（财会〔2023〕11 号，以下简称《暂行规定》）的范畴。数据资源入表需满足“经济利益很可能流入企业”且“成本能够可靠计量”两个核心要件，且应有明确的业务模式，有客观证据表明该业务模式具有可行性、能够产生效益。满足资产确认条件的数据资源根据持有目的的不同可分为两类：一是自用的数据资源（包括对外提供服务、授权转让使用权），按《企业会计准则第 6 号——无形资产》（以下简称无形资产准则）核算；二是日常持有以备出售的数据资源（转让所有权），按《企业会计准则第 1 号——存货》（以下简称存货准则）核算。不满足资产确认条件的数据资源可按《企业会计准则第 14 号——收入》（以下简称收入准则）等处理。如符合无形资产确认条件，还需明确资本化和费用化的划分标准。如当数据采集时高管层未明确具体业务模式，经济利益流入可能性很小，相应采集成本费用化较为合理。当数据采集时高管层计划应用于某类业务模式、有过成功的案例、经济利益流入可能性较大且大于成本支出时，数据采集费用应资本化。实际工作中可采用“预期收益折现法”评估语料库价值，结合历史销售数据预测未来收益，作为资本化依据。企业一般在初次数据采集时将相关支出费用化，在业务模式明确后的某一时点将数据采集费用及

加工分析支出资本化。以数据换数据、技术换数据等特殊情
况取得的数据,按《企业会计准则第7号——非货币性资产
交换》执行。

内容输出阶段产生的图书、图像、音频、视频等成果,
虽然存在著作权争议,但其会计处理与有著作权的作品理
论上应保持一致。政策依据主要有四种,一是存货法,依
据为《新闻出版业会计核算办法》(财会[2004]1号)、《电
影企业会计核算办法》(财会[2004]19号)及存货准则,适
用于传统纸质图书出版;二是无形资产法,依据为无形资
产准则,适用于数字出版;三是存货或无形资产均可,依
据为《深圳证券交易所创业板行业信息披露指引第5号——
上市公司从事互联网游戏业务》(2019年修订)、《深圳证
券交易所创业板行业信息披露指引第6号——上市公司从
事互联网视频业务》(深证上[2019]673号),适用于游戏
和视频业务;四是当期损益法,依据为收入准则,适用
于版权分成的图书出版和数字出版。

三、AIGC 辅助创作内容产品的会计核算

目前语料库建设和内容生成过程中的著作权问题没有
解决,给出版产业链相关企业按拥有或控制来确认资产进
行会计核算带来困难。但人工智能相关业务已经产生,会
计需考虑对其进行准确核算。

(一) 语料采集阶段

会计核算的前提是对数据确权,国家网信办等七部门
发布的《生成式人工智能服务管理暂行办法》规定,AIGC
需使用具有合法来源的数据和基础模型,涉及知识产权的,
不得侵害他人依法享有的知识产权。这需要企业建立“数
据采集合规性审核流程”,要求上游数据商提供版权授权
链证明,并溯源至原始著作权人,规避侵权风险。如不
能上溯至源头应当至少溯源一级,查看上游数据商提供
的数据是否合法。

语料可划分为未经加工的原料数据和经加工的大模型
能够识别的成品数据集,原料数据需进一步加工才能成
为成品数据集,加工又可分为自行加工与委托加工。出
版产业链涵盖多个运营主体,各主体在内容创作、传播
与消费过程中扮演着独特角色,其业务特点各有不同,
会计核算也不尽相同。

1.大模型公司作为技术核心驱动者,专注于人工智
能大模型的研发与优化。其一般从网上收集开源数据,
同时从外部一揽子采购金融、政务、医疗等大量专业数
据集,也可能通过收入分成、技术换数据、用户上传内
容取得部分数

据。如2024年5月,OpenAI与新闻集团签署内容合作
协议,获取新闻集团旗下包括《华尔街日报》等十余家
媒体的内容并用于生成内容;DeepSeek从拓尔思采购
行业数据集。大模型厂商对于购买有限期间数据使用
权的,一般将其直接记入运营成本或物料消耗,对于买
断数据使用权的,一般将其记入无形资产。

2.语料预训练公司专注于收集、整理和预处理语料
数据,为大模型训练提供高质量的素材和预训练服务。
其与各类数据提供商合作,通过多点采购、网上爬取
可公开访问的开源数据等多种方式,广泛收集文本、图
像、音频等多种类型的数据,并进行清洗、标注和分
类等预处理工作。这些数据在供本企业垂直细分领域
模型调用的同时,还把数据集使用权或预训练数据及
软件出售给大模型公司,这类企业一般将全部或部分
数据集确认为无形资产(中文在线)或存货(北京海天
瑞声科技股份有限公司),为大模型企业提供数据预
训练服务的成本可记入企业履约成本。

3.数据加工企业一般为数据持有企业将电子数据、
纸质图书加工生成大模型能够识别的数据集,同时收
取加工费,这类企业对生成数据集不享有权利,此模
式类似于带料委托加工,企业对加工费和加工成本分
别确认收入成本即可。

4.出版企业已完成出版的纸质图书或融合出版图
书经著作权人授权许可(或买断版权),自己或委托加
工制作成语料数据集,供商业使用,可确认为无形资
产;未经著作权人授权许可,制作成语料数据集,供出
版社内部使用,由于能否取得著作权人授权不确定,
未来收益不确定,故将加工成本费用化较为适宜。此
外,还可区分“通用语料”与“专用语料”,通用语料
(如公有领域古籍)直接费用化;专用语料(如买断
版权的畅销书数据集)符合资本化条件时予以资本化。

5.AIGC服务提供者利用开源模型加上自有数据训
练形成新的模型应用。如科大讯飞股份有限公司利用
DeepSeek-Math模型和自己丰富的教育数据资源,与
DeepSeek共同推出AI数学辅导应用“星火助学”,
此时,新模型调试、数据训练等人工成本直接记入企
业运营成本较为适宜,自有或采购训练数据可视条
件予以资本化或费用化。

(二) 内容输出阶段

在AIGC辅助内容输出阶段,AI服务提供者对使
用者的授权分为收费-使用者享有权利、免费-使用
者享有权利、免费-使用者不享有权利、免费-服务
提供者和使用者共享权利等类型,由于这种输出内
容是否拥有著作权尚无定论,会计核算中可引入“人
机协同创作内容产品”明细科

目核算这种权益的取得成本。随着AI的普及,可能会出现编辑作者化趋势,出版社可安排编辑利用AI创作适销对路的内容产品(形成职务创作),也可能通过外购或版权分成来获取类版权权益。内容输出阶段涉及的产业链主体包括出版企业、影视、动漫等行业。内容输出产品可能是单一产品出版,也可能是多模态融合产品出版。无论是单一出版还是融合出版,类版权权益是内容输出产品成本的重要构成内容,也是该阶段会计核算的重难点。下文根据出版类型及版权取得方式分述出版产品的会计处理。

1. 单一纸质图书出版。将外购类版权权益费用和其他出版成本列入生产成本,用该方法进行成本核算的企业几乎覆盖全部传统出版企业。对于自有创作团队利用AI进行IP创作研发,可将人工费用列入生产成本,也可列入开发支出,考虑到版权资产价值具有不确定性,可将研发成本全部费用化,这种方法常见于民营上市出版机构的图书研发中。

2. 单一数字产品出版。将外购类版权权益记入“无形资产”,将除版权外的出版成本全部记入“研发费用”。此方法主要适用于由互联网企业或通讯企业跨界进入出版领域的数字出版企业,为开展数字阅读业务而买断版权时使用。对于按销售数量、销售收入分成等版权分成的经营模式,一般采用当期损益法,将版权成本、渠道成本直接记入“主营业务成本”。目前数字阅读企业掌阅科技股份有限公司和中文在线均存在这种成本核算模式。

3. 多模态融合产品出版。由于多模态融合产品一般只采用授权许可、按次/期间收费、会员充值、出售拷贝等形式,不会出售所有权,故将其按无形资产核算较为适宜。对于自用和对外服务兼具的融合出版业务,参照国际会计准则,可将融合出版业务/项目成本核算划分为五个环节:一是将外购非分成版权记入“无形资产——版权”;二是将融合出版产品/项目立项后至发布前发生的选题策划、编辑、审校等费用记入“研发支出——预发布资产”;三是将数据库建设、数据的采集清洗整合分析及可视化加工等费用记入“研发支出——数据资源”;四是将研发支出金额按照费用化和资本化条件转入“管理费用——研发费用”或“无形资产——预发布资产/数据资源”;五是封面设计、版式设计、排版、图书纸张、印刷、资源卡等费用记入库存商品的制造成本。无形资产摊销及分成版权成本记入“主营业务成本”。如将版权内容开发为电视、电影、游戏、动漫等独创性内容的可将内容版权和预发布资产合二为一形成新的无形资产。

四、对策建议

(一) 构建以版权弱保护为核心的语料采集和内容输出法律制度

1. 对有版权的作品语料采集提供弱保护。算法、算力、数据是大模型发展的三个核心要素。当前,我国大模型除高端算力因芯片受限稍显不足外,算法和数据与美国差距不大,数据方面主要是欠缺高质量的中文出版业数据集。在目前国际上激烈的技术竞争格局下,促进我国生成式人工智能技术的发展、缩小与发达国家之间的差距应当成为首要的政策目标。故对受版权保护的中文出版数据进行弱保护可成为当前及今后很长一段时间内的政策取向。建议采用“免预授权+适当付费”的法定许可模式完善现有版权法律,形成语料库建设的收益分配机制。发挥版权集体管理机构作为版权费用支付的提存转付作用,以平衡著作权人、数据加工者、数据经营者、数据训练者、数据模型构建者之间的利益关系,从而使作品著作权人到数据模型构建者之间的授权链条清晰、完整,不出现中断。作价标准考虑统一按字数、图片数、视频时长制定,不宜太高,不宜区分名著与非名著、畅销书与非畅销书、高票房与非高票房。上述理想状态需要成熟、被著作权人广泛接受的著作权集体管理机制作为支撑,因此需进一步壮大著作权集体管理机构权能,为语料库使用作品的付酬提供强有力的支撑。

2. 对人工智能生成内容提供弱保护。建立以使用者为主要著作权人的内容输出法律制度,明确人工智能创作作品著作权归属有约定从约定,无约定或约定不明时归使用人。按照国家互联网信息办公室等四部门制定的《人工智能生成合成内容标识办法》,对人工智能生成的作品进行标识,以提高透明度。鉴于生成式人工智能并不具备人类的思想、情感,缺乏自我意识,其生成内容类似从大规模数据库中摘录并攒稿,是否具有独创性和创新性争议较大,且法律主体资格确认存疑,边际成本极低,短时间内可大量生成,存在同质化问题,若以狭义的著作权进行保护,必将对人类“作品”市场造成巨大冲击,我国可考虑通过设立新的邻接权对人工智能生成内容进行保护,权益内容以财产权为主,署名权为辅。将生成内容列入法定许可范围,允许合理使用,缩短保护年限,从内容发布起计算,建议保护期为5~10年。

(二) 强化出版产业链AI相关的合同和授权管理

业务合同是会计核算的依据,在使用AI大语言模型的过程中,合同和授权管理需得到高度重视,相关各方应确保

版权内容获得合法授权,明确各方的权益和责任,建立合同会审制度,财务、业务、法务、技术等部门应密切配合,夯实财务核算的基础,防范合同内容风险。对于多方合作的数据项目及智能合约,在项目启动前应依据规范签订详细的数据权利协议,明确各方数据资产属性及入表归属。另外还需加强法务团队建设,深入研究数据资产、著作权相关法律法规和国内AI侵权司法判例,运用系统思维和底线思维,确保在数据业务开展过程中的操作符合法律要求并有利于资产确权与区分。鉴于AIGC创作内容产品涉及的版权归属无明确法律规定,还需从会计核算角度防范侵犯著作权风险。例如有的网站公告禁止爬虫等程序抓取、下载或缓存网站的相关内容,如数据采集方未关注此类公告,则在利用AIGC技术生成内容出版的过程中可能会因为侵犯被使用作品的复制权、改编权、信息网络传播权等权利而落入到侵权困境当中。对此,财务部门应会同相关部门对数据来源不明确,可能导致潜在赔偿的人工智能输出作品计提侵权风险准备金,计提基数用人工智能输出作品产生的收入,计提比例以能覆盖侵权支出为宜。

(三) 出台以稳健性为主导的人工智能相关会计核算手册

2004年出台的新闻出版、电影企业会计核算办法均参

照工业企业制造成本法定制,不能体现文化行业的独创性、创意性和创新性,与其他会计制度对类似业务的会计处理存在分歧,建议财政部以无形资产为主(可选择单独设置一级会计科目或在无形资产下设置明细科目)兼顾当期损益法,将出版、影视、游戏、动漫等内容生产业务整合为一个会计核算制度,规范数据采集、数据加工、数据训练、内容输出等人工智能产业链各环节的会计核算,强化信息披露,以提高会计信息横向可比性。■

责任编辑 林荣森

主要参考文献

- [1] 王立梅, 王迁, 陶乾, 等. 人工智能语料库建设的法律问题[J]. 数字法治, 2024, (5): 28-46.
- [2] 隋明志. 论人工智能生成内容的著作权保护模式——以我国首例“AI文生图案”为例[J]. 出版与印刷, 2024, (4): 11-19.
- [3] 张王国, 李妍. 生成式AI在专业出版领域的应用及未来思考[J]. 出版参考, 2024, (9): 34-39.
- [4] 李益斌, 李浩洋. 欧美中人工智能监管规范比较研究[J]. 当代世界与社会主义, 2024, (5): 161-169.

图片新闻

中国会计学会金融会计专业委员会2025年学术年会在吉林举办



近日,中国会计学会金融会计专业委员会2025年学术年会在吉林财经大学举办,会议主题为“数智时代下金融创新与金融会计研究协同发展”。中国会计学会秘书长、财政部会计司一级巡视员刘光忠,吉林省财政厅党组成员、总会计师张言,中国会计学会金融会计专业委员会主任委员、中山大学教授谭劲松,中国会计学会金融会计专业委员会副主任委员、北京工商大学副校长毛新述,厦门大学教授曲晓辉,德勤有限公司中国副主席吴卫军,中国建设银行周志明,中国银河证券股份有限公司王淡森,安永华明会计师事务所专业业务部合伙人范勋,吉林银行金融研究中心张光

华以及来自全国81所高校、科研机构和实务界的专家学者参加了会议。主题报告环节围绕“中国自主会计知识体系构建”“会计的明天会怎样?”“商业银行准则应用中的一些问题和思考”“证券行业若干会计热点和难点问题分享”“数智时代会计人才培养”“人工智能在财会金融领域的应用进展”“IFRS 18惊鸿一瞥:GAAP为什么要求披露Non-GAAP”“人工智能在中小银行数智会计的应用与挑战”等进行了分享。本次会议遴选52篇优秀论文在8个平行论坛汇报交流,涵盖金融科技与会计数智化转型、绿色金融会计与ESG披露、数智时代金融会计理论创新、金融监管与会计信息披露协同等多个主题。

(本刊记者)