

数字出版人热议 DeepSeek

■中国出版传媒商报记者 张馨宇 采写/整理

人工智能最重要的作用是业态升级与创新

■宋吉述(江苏凤凰出版传媒股份有限公司总经理)



凤凰传媒在AI技术应用于出版流程方面进行了多方面的投入和尝试。探索开发了“凤凰智灵”人工智能应用综合服务平台,集成10多种AI能力,提供了30多个与出版相关的智能应用,如智能问答、办公、出版及营销等服务,一定程度上提升了工作效率。相对于出版社,凤凰旗下的学科网对人工智能应用更加广泛深入。学科网目前为全国4万所学校、超6000万会员提供教育资源服务,日均更新内容资源2万套左右。在其运营中,AI降本增效的应用场景包括替代搜索引擎的AI工作助手、提升销售系统智能属性、版权服务及客服应用、辅助设计师创作、知识库问答、代码编写和SQL编写、文本处理和数据整理等多个方面。此外,公司还探索开发了教育领域的专业模型建设,如“AI小博士”。从凤凰传媒来看,AI实现了对出版社部分生产的改进提升,但总体上应用尚未普及深入,流程变化不大,专业化数字公司使用人工智能技术较多,更倾向于生产流程的变革。

【怎么看】我近期关注并使用DeepSeek(以下简称“DS”),还安排相关员工对其进行研究,特别是与凤凰传媒教育出版相关的中小学教育试题解答、试卷等教育资料生成等方面。总体测评来看,DS在小、初、高各学科的解题能力比较高,正确率90%以上。2月11日,我们打造的凤凰智灵平台正式部署了DS系统,员工可以方便地利用其生成内容,并计划开发针对中小学教育的专业模型服务。

我认为DS对于出版业的影响是非常复杂的。首先是挑战。随着其内容生产能力的提升,部分传统出版内容会被取代,出版的知识传播作用也会被削弱,出版的形式也可能发生很大变化。在未来愈加碎片化、精准化的内容知识服务中,很难想象现有图书形式的存在。其次是机遇。一方面,可以提高生产效率,另一方面,也是更重要的,是其作为一种新的技术革命,带来了新的文化服务形式和新的文化消费需求,也就是为出版业的产品创新、服务创新、市场拓展提供了可能,也就是“破圈”的机

遇。所以,如同近些年发展起来的大数据、VR/AR、区块链等数字技术一样,人工智能最重要的作用不是提高出书效率,甚至不是卖好书,而是业态升级与创新。

【怎么用】与DS公司合作,单体出版社是很难形成优势的。就出版社引以为傲的内容来说,DS已具备丰富海量的知识,而大部分出版社规模不大,出版门类多而杂,内容积累不多,很难形成内容优势。要发挥好出版社的优势,至少要从两个方面强化:一是专注于细分门类,加强专业化、体系化内容建设。做数字化产品与出书不同,图书讲究单品质量,即使比较零散、门类杂一点问题不大,但数字化产品高度依赖持续性积累与运营,需要海量内容支撑。二是垂直领域的运营能力。大模型虽然功能强,但对于每个细分领域来说都会有缺陷和不足,垂直化市场也需要专业化服务,这方面也不是互联网企业和通用产品的擅长。所以,出版社应该发挥自身在某些领域、区域的影响力,注重持续服务,深耕专业化市场。

此外,在人工智能产品建设运营中,还要注意两个内容相关难题:一是版权管理。当前人工智能训练中所使用的内容已经突破了原创性内容,不少模型都在使用二度加工、三度加工后的内容,版权溯源非常困难。但出版业数字化发展迟缓的主要原因不是盗版问题,而是数字内容不足不精,缺少市场吸引力,以及运营能力不足等。所以,总体不能因噎废食,应该积极拥抱新技术。二是内容合规。出版社做专业模型,服务于专业市场的前提是内容合规,应发挥好出版社的内容辨别与审核作用。当然,还要注重技术手段的应用与创新。

若与DS团队合作,期待DS技术在数字出版领域重点优化三方面:一是专业化内容的生成能力。DS与其他模型一样,通用性的内容生成较好,细分领域则缺乏专业性。二是内容合规性的审核。内容服务的基础是合规,不仅要尽量降低幻觉问题,更要杜绝非法内容。这固然可以通过加强审核来实现,但更重要的是技术手段,是模型本身对幻觉内容、非法内容的审核与过滤。三是行业性特色服务。例如针对教育,根据其特殊需求,提供评测等相应服务。这些专业能力建设可以基于和专业公司例如出版社等的合作而形成,期望通过合作形成专业化、精准化产品与服务。

带来技术平权与公共技术能力提升

■赵海清(中国音像与数字出版协会出版融合工作委员会副主任、中新金桥集团公司总裁)



冲击和改变,目前看,DS超越其他大模型技术的方面在于其算法优势、技术开源、低成本应用、中文处理能力强等。

出版业大模型技术的发展和应用,将会由通用大模型转向专业大模型(或称为垂域大模型),满足出版业在内容生产、版权保护、精细化、专业化、个性化服务等方面的实际需求。随着技术的持续发展,大模型技术在智能编辑与创作、沉浸式阅读体验、知识发现和知识管理、教育出版与个性化学习、文化输出与国际传播、智能营销与用户互动等方面会出现更多细分的应用场景。

DS已经成为全民关注的现象级技术体系,我们更多关注和评判的角度的是其技术特性、使用门槛、可持续应用能力以及与业务的契合性等。我们的技术跟踪从2024年12月DeepSeek-V3的发布开始,其性能与GPT-4o和Claude Sonnet 3.5等顶尖模型相近,完全开源而且训练成本很低。

之前我们在公司所主导的“融合出版和数字教材”“融合采选”“知识图谱和知识发现”“数字内容发行”等相关数字出版应用技术平台中已经陆续接入了其他开源大模型技术能力,目前也已经完成了DS的部署和对接,并正在进行各种功能的效果评估和优化。未来我们也希望能够与这些DS大模型企业有更直接更深度合作。

DS在数字出版内容创作、校对、营销等环节会带来很多改变,这种改变首先体现在大幅提高传统工作的效率、质量和效果,其次我们还要观察是否会进一步带来颠覆性的变化,这种变化可能会彻底改变行业原有的经营管理思维、业务流程、人员能力要求和交付模式等。

【怎么用】与DS公司合作,是大模型底层算法技术基础、预训练数据与应用场景、用户服务方面的协作及优势互补。这一轮DS带来的技术平权和公共技术能力提升,让业务端可以更专注在优质内容创作和场景服务方面发力,而不用过多考虑技术支撑能力。因此合作各方可以充分发挥自己的优势,共同推进产业变革。

应用DS这样的大模型技术,可以有效地提升生产创作、内容合规、质量管理的工作效率,也可以高效率跟踪和监测版权运用状态,在具体操作方面,有着很大的想象和提升空间。但是在生产创作方面,过度依赖AI能力,可能会存在较为严重的内容同质化问题,影响出版

业核心竞争力。另外,出版业要能够应用大模型技术实现高质量高效能的生产创作和知识服务,就必须聚集更多的优质内容作为统一的数据基座,这很可能涉及版权权利合法运用等问题。

甚至在拓展国际市场方面,通过应用DS相关技术,可低成本快速实现中文作品自动翻译,显著降低海外用户对优质内容的发现和使用门槛,从而降低本地化成本并加速内容输出效率。同时,基于深度学习的智能分析能力,可精准匹配海外用户偏好,提升内容触达率和用户黏性。我们从2023年起开展了海外业务,今年也希望通过DS的技术加持,扩大中文电子书在海外的销售市场规模。

目前来看,AI技术变革必定会加速出版业的数字化和智能化转型,传统的数字出版也会从原有的数字化转型阶段全面快速向智能化方向发展。随着AI技术的应用,我们的生产工具和生产方式都在发生着本质的变化,数字出版也会更加关注用户的个性化服务需求,这将极大影响原有的业务形态和流程,这个变化是全面的。随着这个变化,首先要解决新的世界观和思维方式的形成,要习惯大模型驱动的新质生产力所带来的工作流程、作业方式、多模态内容输出、内容交互、数据分析、个性化服务、消费方式等各方面会产生的深刻变化。其次,是重点关注并把握数据确权和数据资产管理、数据赋能和版权运营、场景设计和产品服务、用户交互和隐私保护等方面带来的大量发展机遇。

为古籍整理编校和古籍活化提供全新解决方案

■苏瑞欣(古联(北京)数字传媒科技有限公司古籍实验室主任)

古联公司致力于中华古籍数据大语言模型的研发与应用,2024年主要从两方面推进:一是“籍合”古文大模型;聚焦古籍整理与研究的基础内容生产环节,开发了文白翻译、古籍注释、稿抄本OCR识别、标识引等专业任务模型,并研发古籍智能整理产品。其中,OCR模型已投入生产实践,并实现商业化。二是“中华物产大模型”:与南京农业大学合作,针对物产知识数据进行自动标注、知识抽取,并提供知识问答服务,探索古籍数据在产业经济领域的应用价值。

【怎么看】自2025年1月中旬起,我们开始关注DeepSeek(以下简称“DS”)。这一契机源于古籍整理中的实际需求:通用大模型对古籍中大量生僻字的支持不足,导致应用效果欠佳。在尝试了包括GPT-4o在内的多个模型后,我们发现DS在注释古籍例句

中的生僻字方面表现突出。

进一步测试发现,DS能够将古籍整理场景中的复杂问题转化为相关古籍内容的检索行为,根据古籍原文解答问题,而不是像大多数通用大模型那样仅能调用互联网已有的相关问题及说明。这一“思考”过程已非常接近人类专业研究者的解题思路。它还具有十分优秀的“文学写作”能力,能通过将联网搜索与深度思考相结合,完成对其“作品”的事实核查与上下文逻辑一致性检查,为古籍整理的编辑校对工作和古籍活化提供了全新解决方案。虽然测试更多同类问题后难免会观察到效果的波动,但在此类问题上已是当下国产大模型中应用前景最为乐观的。

【怎么用】基于此,我们计划利用DS的API或

让更多人可以参与到AI发展进程

■汤杨(龙源数字传媒集团执行总裁、文可AI首席执行官)

【怎么看】我从春节才开始关注DeepSeek(以下简称“DS”),先是国外媒体的评价,然后作为用户体验,也看了发布的论文了解其基本逻辑。这的确是中国创新的里程碑事件,很自豪也很受启发。

有人提出,DS是目前对于汉字理解最深刻的模型,在我看来,不如说是对人性理解最深的模型。之所以这次DS能破圈,可以从两个层面来说:站在用户层面,它能让表达非常差的人也能获得一个超过85分的结果。用户几乎不是很需要提示词学习、思维层级、表达能力、批判能力等,也能够获得一个好的答案。有技术突破,但是我认为也有很多产品逻辑层面的突破。比如DS会帮你做很多决定,风格、写作目的、细节程度,在没有要求的情况下,会尽可能完整,尽可能包括更多知识点。

我认为DS对出版业影响最大的是降低了各出版社的私有部署成本。出版社之前最担忧的是,在做AI应用的同时把内容喂给各个大模型,现在私有部署了就可以不用担心这一点。因此,在我看来,DS可以让更多出版商参与到数字出版的AI化进程。以前由于对数据外流的担忧,很多人想下场但又不敢下场,现在这个担忧获得了结构性解决,让更多中小型出版单位和传媒单位可以参与进来。

【怎么用】实际上,在关注到DS后,龙源数字传媒第一时间做了接入,并且部署了私有版

本。接入后,最大的好处是我们作为3000家刊社和出版社的合作伙伴,可以更好地保护合作方的知识产权。另外DS在创作型文字方面比目前市面上的大模型都更有优势,所以在文学方面会有AIGC内容出来。去年年底,我们推出的“一书一AI”出版社图书伴生智能体业务,也是今年集团的重点业务,借助DS,可以实现的功能多了一些。

值得关注的是,DS能够在模型层面就有了控制意识形态的能力,这非常重要,意味着大模型是有价值观的,就有能力向善。但其中也存在一些风险,比如如果开启DeepSeek-R1和互联网关联,会幻觉出非常多的细节和数据,这方面希望使用它生成内容,特别是知识内容的人需要注意进行事实依据的核实。

现在很多人都在谈与DS团队的合作,我认为应该让他们去做真正的尖端科学,商业化的部分充分利用最新的科学边界来实现,交给那些理解AI技术及其发展原则的企业去完成,和各产业理解AI、尊重AI发展原则的企业携手共同助力产业发展。出版、传媒、金融、医疗等领域都希望深度探索时有一个主要合作伙伴负责商业化,如此一来,也能够解决DS创始人梁文峰和团队构建的生态关系。这是一种后资本时代之后的新生态。



让出版社以最少资源获得较好AI能力

■付鹏(中信出版集团信息数据中心总经理)

本单位近两年开始全面做AI布局,对于AI在出版过程中的应用作了深入探索,在AI技术应用上,我们针对选题、翻译、审校、营销等各个环节业务,进行了多项AI赋能的探索,从结果来看,AI确实可以为出版大幅度地提质增效。基于探索结果,我们自主投入研发了夸父AI平台,目前正在社内铺开使用。另外,在数字内容制作上,从图文内容、音频内容、视频内容、数字人等方面,中信出版也作了全面探索,从结果来看,AI可以很好地辅助数字内容创作,在某些领域,可以使创作效率提升1倍以上。可以说,AI技术变革已经使AI能够部分替代数字出版物的创作过程,并且可替代的比例也在快速提升。

【怎么看】我是在春节前关注到DeepSeek(以下简称“DS”)的,从使用情况来看,DS的表现不输于其他任何大模型,更惊艳的是它的深度思考功能,可以直观地看到DS的思考过程非常有逻辑性、非常智能。另外,DS的模型是开源的,目前中信出版已经在万象亿新智慧出版体系-夸父AI平台中部署了70B的DS模型,并且部署的算力需要非常小,这使我们可以用最少的资源获得比较好的AI能力,对出版社来讲非常友好。

从过去2年来看,AI大模型技术已经被广泛应用在内容创作领域,但是对于优质内容的输出来讲,质量是不够的,而DS通过更优秀的推理能力,结合思考能力,不仅可以写出通顺的文稿,更能够较好地梳理文稿结构,提出较为全面的策划角度。从内容创意的角度看,人机结合或将成为未来的趋势。在审核方面,我们认为AI是没办法完全替代人的,因此DS仍然只能起到一定程度的审核作用,并且DS也未能避免出现“一本正经地胡说八道”的现象,因此,对于未经人工审核的内容的使用是存在较大风险的,审核环节必须是人机协同,有编辑

参与确认的过程。在推广环节,通过DS生成图文及视频营销素材将会被运用得越来越顺畅,这在未来应该是所有公司都需要具备的基本能力。

【怎么用】中信出版对于技术的应用态度是开放的,会根据自身业务情况来评估如何与DS或类似企业合作。我们认为大模型技术公司,与应用大模型的企业之间,应该是各自做好自己的事情,从定位上来讲,出版企业应该聚焦于自身的内容内核,围绕自身出版内容的特点,建立创作、编校及营销体系。而DS会关注通用化的内容交互,多模态的交互,进一步提升推理能力。关于优势互补,我们认为适合AI的应用场景非常重要,也就是选择合适的抓手。从中信出版过去的探索来看,翻译试译、AI生图、营销视频几方面是业务用得比较多的,这些场景的共同特点都是,AI工具相对成熟,过往以人工内容制作为主,但对结果没有特别严苛的要求。但相对地,如内容审核,“理念性”设计等工作,目前在实际应用中褒贬不一。

若与DS团队合作,我们希望DS能够有更丰富全面的域信息收集能力,在长文稿的编撰处理方面能够做得更强,另外在多模态内容的创作方式上能够做得更成熟,同时希望在针对出版行业的智能知识库、智能出版体系建设等方面能够与其背后团队作深入探讨。

目前,DS在海外已有不小的关注度,类似DS等AI有较强的多语种转化能力,甚至有根据文化适应性做文本转化的能力,这些是非常好的工具,能够贴合场景和话题,对内容进行二次加工。在海外拓展方面,我们相信AI工具将提升效能,更精准地进行传播。中信出版一直以传播中华优秀传统文化、加强国际文化交流为使命,面对AI赋予的机遇,我们更是责无旁贷。



开源模型,优化现有的文白翻译、古籍注释、智能检索等技术,提升古籍整理出版智能化产品的性能,并寻求与其研发团队合作的渠道,开发新的产品。

然而,DS目前仍存在较为严重的“幻觉”问题,即使在联网检索的情况下,也可能输出引用页面中不存在的内容。这对出版行业而言是高风险问题。一方面非专业用户难以辨别AI生成内容的真伪,加之社交网络缺乏对AI生成文本的检测机制,可能导致虚假信息广泛传播,甚至形成错误的群体认知。另一方面,专业内容创作者和编辑无法信赖AI输出内容,需逐一核查引用链接,甚至在AI幻觉数据泛滥之后,可能需要耗费更多时间验证出版物中的信息。

为解决这一问题,我们建议探索大模型厂商与出版社的合作路径,通过利用DS等大模型的能力,激活经过“三审三校”和社会检验的高质量数据,提升大模型的可信度,从而实现技术与内容的双向赋能。

(下转第4版)