

互联网数字报智能体技术探索

谢继亮¹ 魏明震¹ 刘 潇² 袁振德¹ 孙星海¹ 苑胜明¹

XIE Jiliang WEI Mingzhen LIU Xiao YUAN Zhende SUN Xinghai YUAN Shengming

摘要

针对传统互联网数字报资源在利用大语言模型（LLM）进行智能分析时存在的接口不统一、交互标准缺失及集成难度高等问题，文章提出一种基于 MCP（model context protocol）的标准化连接接口构建方案。旨在通过定义统一的通信协议与数据格式，实现 AI 模型与异构数字报资源的无缝对接。在技术实现层面，深入研究了数字报内容的抓取、分析与重构机制，并提出了基于二维视觉语义的报纸版面智能理解方法。接口设计涵盖了栏目检索、文章列表获取及结构化内容提取等核心功能模块。实验与应用场景分析表明，所提方案成功构建了包含报社生产、行业监管、读者服务及社会治理在内的多维应用框架，有效促进了数字出版与人工智能技术的深度融合，为传统媒体的智能化转型提供了新的理论依据与技术路径。

关键词

智能体；大语言模型；MCP 协议；互联网数字报

doi: 10.3969/j.issn.1672-9528.2026.04.030

0 引言

目前主流数字报平台（如人民网）提供数字版栏目导航、栏目文章导航和文章详情展示等功能（如图 1 所示），但未提供访问数字报结构化内容的标准接口。目前 LLM 客户端仅能访问到通过网页抓取采集到的碎片化的数字报资源，没有最大化发挥数字报资源的有效传播和利用。根据《2024 年中国数字报业发展白皮书》论述，AI 技术在报业的渗透率不足 30%，目前市场迫切需要各数字报资源方为 LLM 提供标准化的访问接口。

随着人工智能技术不断推进，大语言模型（LLM）在语言理解与生成方面的能力已经相当突出^[1-2]。与此同时，互联网数字报——作为传统报业在数字时代的一种重要形态，积累了体量庞大的新闻资源。但两者之间还缺少一条真正畅通的连接路径。AI 技术要想充分利用数字报的资源，并不容易；这种割裂的局面，也在客观上限制了两者实现更深层的协同。

针对这一问题，本研究提出一个基于 MCP 的统一连接接口。其目的在于，在大语言模型与互联网数字报资源之间构建起一条可实现无缝集成的通道，从而把数字报在 AI 时代潜在的、尚未被充分挖掘的价值进一步释放出来。

1. 潍坊北大青鸟华光照排有限公司 山东潍坊 261061

2. 大众报业集团 山东济南 250014

[基金项目] 山东省重点研发计划（山东省科技型中小企业创新能力提升工程）项目“基于国产操作系统的数字出版系统”资助（2023TSGC0710）

1 互联网数字报智能体系统设计

1.1 设计思路

MCP（model context protocol）本质上是一个标准化协议，用来支撑大语言模型与外部系统之间高效、稳定的通信^[3]。定义了一套统一的通信格式与数据结构，核心作用在于确保智能体与互联网数字报服务之间的信息交换足够准确。换句话说，通过 MCP，原本松散的数字报数据可以转化为智能体能够直接处理的结构化形式，为双方动态交互打下基础。

在这个系统中，互联网数字报智能体扮演的是核心交互单元的角色，集成了 MCP 客户端与自然语言处理能力，并借助本地或远程的大模型服务。主要完成两项任务：

（1）基于 MCP 技术对用户口语指令进行解析，并据此生成针对互联网数字报 MCP 服务的精确调用；

（2）对返回的数据做业务逻辑层面的分析，再结合提示词工程，最终输出具备可操作性的结果或汇总报告。

1.2 关键优势与设计理念

（1）标准化与解耦

MCP 协议在这里的角色更像是智能体与数字报服务之间的一层“中间件”。避免了智能体不得不与不同数据源进行逐一、定制化对接的麻烦，形成一种“一次对接、多端复用”的机制，类似于 USB-C 的扩展思路。

（2）动态扩展性

如果后续需要增加新的数字报数据源或者功能模块，只需对 MCP 服务的功能列表进行扩展即可，智能体自身的核心逻辑几乎不用改动。这种设计明显降低了系统在后续维护上的成本。



人民日报图文数据库（1946-2026）

人民日报 2025年06月18日 星期三 上 一期 下 一期

01版：要闻	02版：要闻	03版：要闻	04版：要闻
05版：评论	06版：要闻	07版：要闻	08版：广告
09版：理论	10版：经济	11版：政治	12版：文化
13版：社会	14版：综合	15版：国际	16版：民生
17版：国际	18版：产经	19版：消费	20版：副刊

- 在中亚，中国汽车如何赢得口碑（经济聚焦）
- 不断探索更“懂”消费者的产品（编辑手记）
- 一条东星斑背后的补链强链探索
- 夏日赏飞瀑
- 5月境内外汇供求总体平衡
- 全国铁路7月1日起实行新的列车运行图
- 绿色产品走俏海外
- 广西已拥有技能人才863.75万人
- 本版责编：白之羽 韩春瑶 林子夜

第10版：经济

PDF下载

图 1 《人民日报》数字版展示图

2 架构设计

互联网数字报智能体的整体系统架构，由三个核心模块构成：大模型服务层、智能体交互层，以及互联网数字报 MCP 服务层（如图 2 所示）。

2.1 智能体交互层

用户接口模块。承担“把人的指令收进来，再整理成系统能懂的样子”的工作。具体来看，既提供 Web 界面、命令行工具，也开放 SDK，支持文本或语音这两种自然语言输入形式。简而言之，无论采用文字输入还是语音交互方式，都

能接收，并把指令做一层格式化处理。

命令解析模块。核心任务包括：MCP 协议转换——把自然语言命令映射成标准化的 MCP 调用指令，得明确服务地址、功能编码和参数集；权限校验，看用户是否有权限访问数字报资源；错误处理，遇到无效指令或者格式不对的情况，不能直接崩掉，需返回较友好的提示。

结果生成模块。主要是接收 MCP 服务返回的原始数据。处理过程会结合业务提示词模板，包括数据清洗、结构化转换、多源信息融合分析。最终根据实际需求，自动生成包含表格、图表或关键摘要等形式的可视化报告。

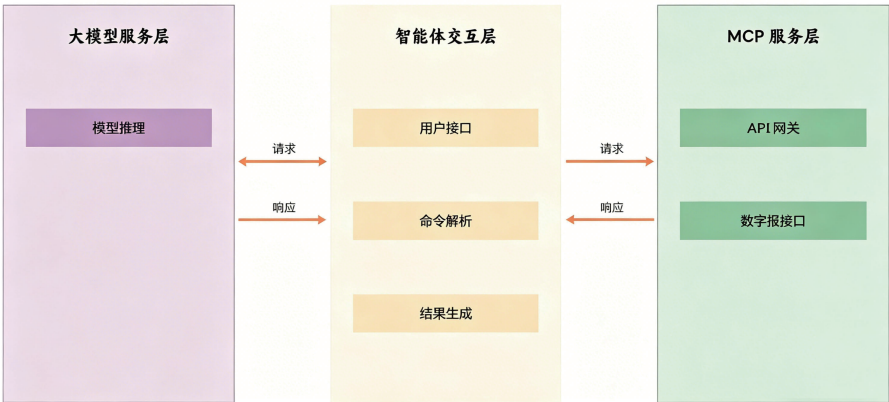


图 2 架构设计图

2.2 互联网数字报 MCP 服务层

API 网关可以理解成所有数字报资源的统一入口，负责请求路由，即根据智能体的指令，把请求分发到对应的数字报数据源。高并发的时候需做负载均衡，对流量做一定的控制。同时包括熔断保护，如其中一服务器出现问题，能自动切断非核心服务的调用，避免整体被拖垮。

接口服务模块则是具体去实现数字报资源的访问逻辑。按内容类型分

成新闻检索服务，支持按日期、版面、关键词过滤；数据聚合服务，主要进行跨报刊专题报道的整合；历史档案服务，用来回溯过期报刊的查询。

2.3 大模型服务层

主要提供自然语言理解和逻辑推理的能力。在智能体里，支撑几个具体场景。比如模糊指令的精确化——“昨天重要新闻”这类新闻，大模型将其转成“2023年10月1日头版头条”的明确查询。复杂查询也由其来拆解，像多条件组合的检索任务，会被分解成一步步执行。最后，技术性数据要转成更易懂的自然语言输出，进行人性化解读。

3 智能体访问数字报资源流程

互联网数字报智能体的核心工作流程，大致分为五个阶段。具体流程如图3所示。



图3 智能体访问数字报过程示意图

3.1 交互示例的流程拆解

例如，用户输入“请汇总2023年国庆期间《人民日报》要闻版关于科技创新的报道”。

大模型接到指令后，第一件事是解析关键参数。时间限定在2023年10月1日到7日，来源是《人民日报》，版面锁定要闻版，主题围绕“科技创新”。这些提取出来的信息，被转化为MCP指令，格式类似这样：<service:newspaper_query, params:{source:RM,date_range:[20231001,20231007], section:headline, keyword: 科技创新}>。

接下来进入调用环节。智能体通过MCP协议把这个指令发给对应的服务端，API网关先做权限验证，验证通过后路由到《人民日报》的专题数据库接口。返回JSON格式原始报道数据，包含标题、摘要、发布时间这些字段。收集到数据后，智能体再调用大模型做三层处理。先去重，把重复的报道剔除；然后按“政策发布”“成果展示”“专家访谈”三个类别做主题聚类；最后给每类报道生成200字左右的总

结摘要。

最终输出采用markdown格式。结构包括：报道统计——共检索到8篇相关报道，标题和日期以列表形式呈现；分类摘要——按三大主题做结构化总结；延伸建议——给出推荐的相关专题报道调用指令。

整个流程其实揭示了一个核心点：智能体在这里扮演的是“人机协同中枢”。MCP协议负责打通数字报资源的动态访问，大模型则用来提升自然交互的体验，最终输出的结果是能直接支撑决策的信息。

4 数字报抓取分析重构技术在MCP服务构建中的应用研究

搭建一个可用的数字报MCP服务，有两项基础工作绕不开。一是定义一套标准化的数字报描述信息体系，核心是明确元数据规范和数据的组织逻辑；二是建立一个可扩展的版面信息数据库，为后续存储、检索、深度应用打好底层基础。

实践中，一个比较高效的技术路径是对互联网上的公开数字报资源做定向抓取，再结合深度分析和结构化重构。这么做的好处很明显：既能保证数据来源够广、够新，也能通过技术手段把原始信息精准解析出来，转化成标准化格式，给MCP服务的功能落地铺路。

4.1 互联网数字报的信息构成与抓取边界

互联网数字报本质上是纸质报纸的数字化翻版，信息结构呈现明显的层级化特征，大致可以拆成三个维度。

一是报纸的基础属性信息。比如出版日期、报名、版次编号这些全局性的元数据。这类信息主要用于资源的分类归档和唯一标识。

二是版面层级。每个版在数字化过程中会被切分成若干个“版面区块”，这些区块是版面视觉布局和信息组织的基本单元。

三是落到区块内容上。按内容形态分，有文字区块、图片区块、表格区块三类。文字区块承载正文和标题，图片区块包含新闻图片和对应的说明文字，表格区块则以结构化形式呈现数据密集型内容，比如统计数字、人员名单等。

三个维度加在一起，构成了数字报抓取和分析的完整对象范围，也是后续所有技术处理的靶点。

4.2 核心信息的智能识别与解析

4.2.1 文本类信息的结构化处理

对于以HTML形式存在的文本内容，本文采用“标签解析+特征提取”的双层框架。标题识别这块，一方面借助HTML标签里的语义标识，比如<h1>到<h6>、特定的class属性，另一方面结合自然语言处理的“标题层级规则”——引题通常是补充说明性质的短句，主题是核心概括，副题则是延伸信息。通过这两层，实现引题、主题、副题的自动拆分和定位。

处理浮动标题时，传统依赖标签的方法常常失效。我们的做法是引入“文本特征驱动”的识别方法。统计文字段落的字数密度、字体类型（宋体、黑体）、字号大小这些视觉特征，构建浮动标题的特征模型。这类标题通常表现为字数少、字体粗、字号大于正文。再结合上下文语义关联，就能把那些没有明确标签标识的浮动标题准确识别出来。

4.2.2 人物信息的语义挖掘

通讯员、摄影记者等信息的提取采用关键词匹配结合语义上下文验证的方法。首先构建身份标识关键词库，涵盖“通讯员”“摄影”“摄”“文”等标识词，再通过全文检索定位文本中对应的候选位置。

后利用依存句法分析，对关键词周围的文本做语义解析。关键是要判断它是否符合“身份关键词+人名”的语法结构，比如“通讯员 张三”“摄影 李四”。同时要排除干扰，比如“摄影器材”里的“摄影”非人物信息。即可将通讯员和摄影记者区分，准确提取出姓名，为版权追溯和作者关联提供数据支撑。

4.2.3 表格信息的结构化转换

表格处理采用“模型提取+规则优化”的端到端方案，主要解决两个技术难点：表格结构还原和合并单元格处理。

结构提取环节引入 TableNet 这个深度学习模型。它基于卷积神经网络对表格图像（或 HTML 渲染后的表格区域）做特征学习，能精准检测表格线框和单元格边界，输出原始表格的行列结构和单元格位置信息，相当于先搭出一个结构骨架。

合并单元格的处理则靠定制规则引擎。分析单元格的位置重叠关系、内容关联性——合并单元格里通常放的是统一的分类标题，制定识别规则。比如同一行或同一列中连续且没有边界线的单元格，就判定为合并单元格，并自动生成映射关系。

最后把处理完的表格结构转成标准 JSON 格式，里面包含行列数、单元格坐标、合并规则、单元格内容等完整元数据。这样就完成了从“可视化呈现”到“结构化数据”的转化，后续数据分析、跨表格比对都能直接调用。

4.3 转版文章的跨版面关联

数字报里经常出现“下转几版”“上接几版”这类标识。一篇文章因此被切分到不同版面。阅读连贯性打了折扣，信息完整性也受到影响。

本文设计的关联技术，出发点为“关键词定位+语义相似度匹配”。先通过关键词检索，定位转版标识所在的文章片段，确定源版面和目标版面。接着提取转版处前后的文本片段——通常是转版前最后 100~200 字，再加上目标版面候选文章的开头 100~200 字。

两段文本语义相似度采用余弦相似度计算，其核心假设为：同一文章上下文间的语义关联度显著高于不同文章。通过设定相似度阈值（如 0.6），将超出阈值的内容判定为同一

文章的转版接续文本，结合版面编号进行双重验证，即可实现跨版面内容的自动拼接，恢复文章完整性，为用户提供无缝阅读体验与完整文本数据。

综上，数字报的抓取、分析与重构，其核心目标是将互联网中非结构化的数字报资源与标准化 MCP 服务对接。依托层级化抓取框架、多维度智能识别技术及跨版面关联方法，可实现从原始内容采集到结构化重构的全流程自动化处理。

该体系有效解决了数字报信息碎片化、格式不统一、内容关联不完整等核心问题，既为 MCP 服务提供了高质量、标准化的底层数据支撑，也为后续知识图谱构建、内容智能推荐、版权管理等深度应用开发奠定了技术基础。

5 报纸版面二维视觉语义的智能理解与应用

传统数字报纸系统能做到的，是获取版面基础视觉元素，主要集中在文章坐标信息和标题展示信息。

文章坐标信息用于记录版面内标题区、正文区、图片区等模块的空间属性，包括具体位置、尺寸大小、轮廓形态、分栏结构，以及图片与对应图说的位置关联关系。标题展示信息则用于记录标题的层级布局，如引题、主题、肩题的组合方式，以及颜色、字号等版式样式属性。

从人类视觉认知角度，这些客观视觉信息会形成“头版头条”“通栏大标题”“整版文章”等直观认知。该认知本质上是由版面视觉语言所触发的主观语义判断，能够直接反映报纸内容的重要性层级与版面编排意图，可视为报纸版面二维视觉语义的核心内容。

然而，传统处理技术仅能采集与存储客观视觉参数，难以实现从数据到语义的转化，无法提取“头版头条”这类主观语义信息。这导致版面中大量丰富的二维视觉数据未能得到有效利用，也限制了数字报纸系统智能化服务能力的提升。

针对这一技术瓶颈，本文提出了一套基于 MCP 智能体的二维视觉语义理解方案。具体而言，通过构建相对精细的提示词工程，将传统系统原本就能获取的文章坐标信息与标题展示信息，一并输入到大语言模型当中。MCP 智能体在这里扮演了关键角色——它负责整合视觉层面的空间布局信息，再配合大语言模型本身的语义理解能力，最终实现对“头版头条”“通栏大标题”“整版文章”等带有明显主观判断属性的版面要素进行精准抽取。有了这些信息，系统就可以自动为数字报纸内容生成对应的语义标签。

正是这些标签的生成，让数字报纸系统的智能化升级有了可落地的抓手。借助大语言模型对自然语言查询的理解能力，现有系统在面对“请问今天的头版头条是什么？”“今天头版以外重要的新闻有哪些？”这类问题时，便能够给出准确回应。

从宏观的角度看，这项工作的意义在于一次有效转化：将传统数字报纸系统中长期处于沉睡状态的二维视觉数据，

真正盘活为具备实用价值的语义信息。与此同时，信息检索的效率与用户的交互体验也得到了明显改善。可以说，这为数字报纸走向智能化，开辟了一条新的技术路径。

6 MCP 服务接口功能设计

接口设计遵循数字出版行业标准化规范，确保数据交互的兼容性与规范性^[4]。

6.1 按日期和报纸名称获取栏目

这一接口输入指定日期和报纸名称。MCP 服务返回该报纸当天所有栏目的 ID 列表。接口用途是为后续获取具体栏目下的文章提供一个入口，方便用户快速定位到自己感兴趣的内容区域。

6.2 按报纸名称、日期、栏目获取文章列表

在上一步获取栏目 ID 列表的基础上，传入指定栏目 ID，接口即可返回该栏目下的全部文章 ID 列表，便于用户精准获取特定栏目内的文章。系统需构建高效索引机制，依据日期、栏目与文章间的关联关系，实现对应文章 ID 的快速检索。

6.3 按时间范围、主题分类、关键词检索文章

支持用正则表达式来匹配关键词。

6.4 获取文章结构化信息

输入文章 ID。接口返回报道日期、所属栏目和文章的结构化信息。

结构化信息包含几个部分：

(1) 文章位置：文章在版面中的位置，本身就带有信息价值。分为左上、左中、左下、右上、右中、右下等。通过提示词说明位置和信息的对应关系，大语言模型就能理解文章在版面中的重要性权重，辅助对文章内容的分析。

(2) 标题：包含标题位置及宽高、排版方向，以及引题、主题、副题各自的内容和字体字号。大语言模型拿到这些细节，能更准确地理解文章主旨。

(3) 正文：包括正文位置及宽高、栏数，以及正文结构化内容。后者包含通讯员信息、浮动标题及字体字号、一般内容及字体字号。对正文的处理不只是提取文本，还要分析其在版面中的布局结构——通过版面分析算法确定位置、宽高、栏数等信息。对于结构化内容，要区分浮动标题和一般内容，提取属性，以结构化的形式提供给大语言模型，便于深入分析和处理。

(4) 图片：包括图片摄影记者、图片位置、宽高、图片在正文中的锚定信息，以及图片说明的位置、宽高、字体字号。大语言模型利用这些信息，能更好地理解图片所传达的内容，增强对文章的整体理解。

(5) 表格：包括表格位置、宽高、在正文中的锚定信息，以及表格说明的位置、宽高、字体字号。大语言模型获取这

些信息后，能更准确地处理表格数据和相关内容。

7 AI 模型与互联网数字报资源无缝集成后的应用场景

7.1 报社角度

对报社而言，互联网数字报智能体的应用推动内容生产与管理迈向智能化。在内容适配方面，大语言模型可通过 MCP 接口获取当期版面结构化信息，综合考虑手机、平板、计算机、电子纸等不同终端在屏幕尺寸、分辨率及交互方式上的差异，对内容排版需求进行分析，并依托语义理解与自然语言生成技术，对标题、正文、图片、表格等元素进行重组与优化，实现面向多终端的内容快速适配。例如，手机端可将长文自动分页展示、自适应缩小图片尺寸；电子纸设备则侧重优化文字版式，提升阅读舒适度，既改善用户体验，也拓宽传播覆盖范围。

在专题剪报生成方面，报社可借助数字报智能体检索历史数据，通过输入专题关键词，模型依据 MCP 接口提供的文章标题关键词、正文内容、主题分类等结构化信息，从海量历史报道中精准筛选相关内容。随后利用自然语言处理技术完成摘要提取与内容整合，自动生成专题剪报。此类剪报既可服务于内部记者编辑的选题策划与资料搜集，也可作为增值服务面向用户，满足其深度阅读与资料整理需求。

在智能选题策划方面，互联网数字报智能体通过 MCP 接口对历史数字报数据进行分析，挖掘热点话题传播趋势，如某类事件报道量的时序变化特征，并结合读者点击量、阅读时长等行为数据，自动生成选题建议，同时关联相关历史报道作为参考资料。

在跨媒体内容生成方面，基于数字报结构化文本信息，系统可将新闻报道自动转换为短视频脚本，提取核心事件、人物与数据要素；或调整表述方式生成口语化音频新闻，适配短视频与音频传播平台，实现内容形态的高效转化。

7.2 监管角度

从报纸主管监管部门的视角来看，该集成系统可提供更为高效的监管手段。依托大语言模型的智能分析能力，可实现对数字报出版情况的实时监测。通过获取报纸文章列表及具体内容，监管部门可精准掌握稿件发布状态，判断相关报道是否合规。结合大语言模型的敏感信息检测能力，可实现内容实时监测，对违规信息及时处置，有助于维护健康的新闻舆论环境。该监管模式兼具高效性与准确性，对规范数字报出版行为具有积极作用^[5]。

在此基础上，内容审核可进一步深化，不再局限于敏感词检索。借助大语言模型的语义理解能力，可识别具有隐喻性、隐含性的违规内容，通过上下文语义分析，判断文章是否存在隐性负面导向等问题。

在跨平台内容一致性监管方面，系统可实现更精细化的管控。通过 MCP 接口汇聚多家媒体数据，对同一事件在不

同数字报中的报道内容进行比对，核查是否存在信息冲突或误导性表述，并利用大语言模型的对比分析能力自动生成监管报告。

7.3 读者角度

对于读者而言，系统集成后可带来更丰富、更便捷的阅读体验。读者能够将当前阅读的报纸与其他报纸进行内容对比，快速获取特色报道，实现信息来源多元化。同时，系统依托大语言模型提供智能检索服务，支持读者精准查找某一主题的系列报道，满足个性化阅读需求。例如，针对某一热点事件，读者可检索该报在不同时期的相关报道，形成完整的信息脉络，清晰掌握事件发展过程与多方观点。

7.4 社会角度

对于社会人士而言，借助互联网数字报智能体检索报纸相关信息时，MCP 接口实现的无缝集成的方式，显著优化了信息获取体验。传统报纸信息搜索往往仅返回零散的单篇文章，而集成后的系统依托 MCP 接口获取的结构化报纸知识数据，可对搜索结果进行深度加工处理。例如，当用户搜索某一热点事件的相关报道时，大语言模型不仅会呈现相关文章列表，还会通过解析文章结构化信息，整合事件发展时间线、各方观点对比、相关数据统计等核心内容，并以结构化或可视化形式呈现，帮助用户快速、全面地掌握事件全貌。

此外，结构化报纸知识数据对跨领域研究与知识发现具有重要价值。学者及研究人员在开展社会科学、历史文化等领域研究时，可依托大语言模型，通过关键词、时间范围、主题分类等多维度检索数字报资源；模型会从结构化数据中提取有效信息，自动生成研究报告、数据图表等成果，为研究工作提供丰富的资料支撑，助力学术研究的创新与发展。这一优势同样适用于普通公众，结构化知识数据的普及方式，大幅提升了信息获取效率，不仅有助于个人积累知识储备，也能帮助公众更顺畅地理解社会热点与文化历史，推动知识在社会层面的广泛传播与共享。

在文化传承与知识图谱构建方面，通过 MCP 接口可从数字报中提取历史事件、人物、文化符号等核心要素，结合大语言模型构建“报纸知识图谱”。以某一历史时期的报道为例，可通过时间线形式梳理相关内容，同步可视化呈现人物关系网络，为历史研究、文化传播提供有力支撑，助力传统文化的传承与弘扬。

应急信息服务也是其重要应用场景之一。当突发公共事件发生时，互联网数字报智能体可通过 MCP 接口实时抓取多家数字报的相关报道，整合零散信息生成事件真相核查报告，通过辟谣澄清、梳理权威信源的方式，为公众提供准确、可靠的应急信息，助力突发公共事件的有序处置。

8 结论与展望

本研究提出一套基于 MCP 的大语言模型与互联网数字

报无缝集成方案。其核心思路是设计统一的连接接口，实现 AI 模型与数字报资源的深度打通。文章分别从报社、监管机构、读者及社会四个维度，探讨了该方案的典型应用场景。在实际落地过程中，仍面临数据安全保障、模型性能优化等亟待解决的关键问题。未来研究可从两方面推进：一是深化 MCP 协议研究，进一步提升数据交互的安全性与传输效率；二是持续优化大语言模型对数字报数据的处理能力，不断拓展应用场景，推动数字报资源与 AI 技术深度融合，为新闻传媒行业发展注入新动能^[6]。

参考文献：

- [1] BROWN T B, MANN B, RYDER N, et al. Language models are few-shot learners[C] //NIPS'20: Proceedings of the 34th International Conference on Neural Information Processing Systems. New York: ACM, 2020: 1877-1901.
- [2] OpenAI. GPT-4 technical report[EB/OL]. (2024-03-04)[2026-04-02]. <https://arxiv.org/abs/2303.08774>.
- [3] Anthropic. Model Context Protocol (MCP)[EB/OL]. [2026-04-02]. <https://modelcontextprotocol.io/specification/versioning>.
- [4] 中国新闻出版研究院, 全国新闻出版标准化技术委员会. 数字内容 API 接口规范 V2.1[Z]. 北京: 中国新闻出版研究院, 2022.
- [5] 国家互联网信息办公室. 网络内容生态治理规定 [Z]. 北京: 国家互联网信息办公室, 2019.
- [6] 中国新闻出版研究院, 中国报业协会. 数字报业人工智能应用指南 (2024 版) [Z]. 北京: 中国新闻出版研究院, 2024.

【作者简介】

谢继亮 (1974—)，男，山东巨野人，本科，高级工程师，研究方向：信息技术应用、中文信息处理。

魏明震 (1971—)，男，山东单县人，本科，高级工程师，研究方向：信息技术应用、中文信息处理。

刘潇 (1980—)，女，山东济南人，本科，初级工程师，研究方向：媒体融合、信息技术。

袁振德 (1978—)，男，山东寿光人，本科，高级工程师，研究方向：信息技术应用、中文信息处理。

孙星海 (1989—)，男，山东安丘人，本科，工程师，研究方向：多媒体出版、中文信息处理。

苑胜明 (1985—)，男，山东宁津人，本科，软件设计师，研究方向：多媒体出版、中文信息处理。

(收稿日期：2025-09-06 修回日期：2026-04-08)