关于购买“北大法宝”数据使用许可服务的任务相关性、合理性和经济性说明

北京大学新闻与传播学院研究员王天夫承担国家重点研发计划课题“互联网新业态广告智能化监管新范式研究”（课题编号：2023YFC3305001，所属项目“互联网新业态违法违规广告智能监测和预警关键技术研究”）”。项目执行过程中，计划与北京北大英华科技有限公司签订《数据与信息使用许可协议》委托业务，从任务相关性、合理性和经济性三个方面说明如下。

一、委托业务的基本内容

根据协议，乙方（北京北大英华科技有限公司）授权甲方（北京大学）使用“北大法宝”法律法规数据库、司法案例数据库、行政执法数据库中广告、反不正当竞争相关条款与文书数据（发布日期 2014 年 1 月 1 日至 2026 年 12 月 31 日），其中广告违法行政处罚决定书数据量预估约 100 万条；乙方运用大语言模型和大数据工程技术，按甲方研究设计对原始数据进行特征工程、向量化、标签化等加工，交付关系型数据库和可供大语言模型加载的知识库，并提供不超过 1%的样例数据用于代码调试与算法验证；同时提供“北大法宝”正式账号 5 个（使用期至 2030 年 1 月 1 日），数据加工交付后提供 3 年运维服务。

二、任务相关性

本项委托业务与课题研究任务直接相关，是完成课题考核指标的必要条件。其一，课题研究内容 1.1“构建广告监管基础数据库”明确要求“融合执法大数据，对违法违规广告的类型特征进行总结并分类，形成违法违规广告共性特征数据库”，考核指标 1.1 要求形成广告监管基础数据库 1 套、覆盖直播带货、短视频广告、软文广告、算法广告等大类违法违规典型表现并形成特征编码、数据库记录不少于 10 万条且通过第三方测试。本项委托所提供的约 70 万条广告违法行政处罚文书，正是任务书所述“执法大数据”的核心数据来源，所交付的结构化、标签化数据库直接服务于该考核指标的达成。其二，课题创新点 1“人工智能驱动的监管基础设施建设”要求对广告监管的法律法规和违法违规典型案例进行内容分析，构造“广告内容—广告形式—类型特征—法律法规”逻辑链条，直接依赖本次许可的法律法规条文与案例文书数据。其三，研究内容 1.3 违法违规广告传播机理模型、1.4 智能监管

新范式均需以违法违规广告分类特征数据为基础，协议约定的违法分类体系与大语言模型结构化提取与之对应；协议同时许可数据用于大语言模型训练、微调与检索，与课题“以人工智能治理人工智能”的技术路线一致。

三、合理性

一是数据来源权威、具有不可替代性。北京北大英华科技有限公司是国内成熟、专业的法律信息供应商，“北大法宝”对全国省、市、区县各级市场监管部门（含原工商、质监、食药监、综合执法等）发布的行政处罚文书收录完整、更新及时，并覆盖《广告法》《互联网广告管理办法》《反不正当竞争法》等多部法律依据，课题组难以通过其他公开渠道获得同等覆盖度、同等时效且授权合规的数据资源。二是合作模式与课题进度匹配。协议约定签订后 20 个工作日内交付样本数据，数据提取次数不超过 10 次、可分阶段运算提取，运维期 3 年覆盖课题验收及成果固化阶段，与课题执行期（至 2026 年 10 月）和年度计划中数据库建设的时间节点相衔接。三是合规与风险安排完备。协议明确数据仅用于非营利性学术研究，约定双方遵守国家安全法、网络安全法、数据安全法、个人信息保护法等法律法规；乙方保证所提供数据不侵犯第三方合法权益并承担相应法律责任；原始数据知识产权归乙方、加工后交付物知识产权归甲方，权属清晰；甲方在完整收到数据和加工交付物后 15 个工作日内凭发票付款，并设有逾期交付、账号无法正常使用等违约赔偿条款，经费支付风险可控。

四、经济性

一是价格水平低、费效比高。二是显著低于自行建设成本，且无法解决数据合法授权问题。三是优于单纯商业采购。四是经三方比价择优确定，成交价格为市场最优。五是数据来源最直接、授权链条最短。

综上，本项委托业务与课题研究内容和考核指标直接相关，供应商选择与实施方案合理、合规风险可控，价格经济、费效比显著优于自行建设和单纯商业采购，符合国家重点研发计划资金任务相关性、政策相符性和经济合理性的管理要求，具备实施的必要性。

负责人：王灵丞

日期：2026.7.20