

# 从科研诚信到成果转化服务实践思考

演讲者：张秀梅 文章整理：万芸



**张秀梅简介:**万方医学网创始人兼总经理，管理学博士，研究员，中国科技信息研究所硕士生导师。在信息组织与信息构建、数字图书馆与数字出版、生物医学领域的知识组织与知识服务系统构建和产业化发展等领域深耕。主持或参加国家省部级科研项目（课题）十余项。

在《第三届医学研究与发表高峰论坛》中张秀梅博士应邀做了题为“从科研诚信到成果转化服务实践思考”的精彩演讲，本次演讲中张博士从中国科研诚信现状问题、科研诚信建设及学术不端的治理工具、中国科技成果转化现状问题、科技成果转化服务体系和信息服务实践的理论研究五个方面进行阐述。

近年来国际期刊出现了大规模撤稿事件，这引起了国内政府相关部门的重视。2021年6月3日，国家卫生健康委开设医学科研诚信专栏，宣传科研诚信政策法规，定期、不定期地查实并公开通报科研诚信案件调查处理结果。自2000年以来，中央各部委先后出台了约100项科研诚信与学术规范相关政策意见、法律法规及管理辦法。据不完全统计，截止到2020年底，全国31个省市自治区均出台了科研诚信相关政策，加强科研诚信建设实施意见，加强学风道德与学

术规范意识，科研诚信被提到了一个新的历史高度。

诚信和学术环境的治理是一个长期的过程，每个学术共同体的人员都有不可推卸的责任，引领科学研究走向正确的道路。围绕这个现状，作为国家科技部旗下的企业，万方数据也承担了很重要的相关情报数据处理的支撑工作。围绕科研诚信建设及学术不端推出了治理工具—科研诚信服务集。这个科研诚信服务集分为两个大的版块：科研诚信与学术规范通识教育平台和文献相似性检测服务。其中通识教育平台会提供教育教学、政策法规宣传、学术不端治理和成果共享展示。文献相似性检测服务版块会提供学术科研、科研管理、学术出版、人才评定和教育教学等量化内容。这些服务的底层是需要有海量数据池，通过对数据资源整合、规范化加工与处理，然后用高水平算法工具库共同匹配起来支撑上述的服务集。

针对文献相似性检测系统，前端是对数据规范化处理，包括制定标准规范、特征数据加工和技术开发。然后还需要形成不断延展和扩充的海量全文比对资源。有了资源和算法的基础后形成一个查大同、识小异，基于快速准确、科学全面和安全可靠的一套服务体系，面向不同的机构和个人应用场景提供服务。文献相似性检测系统可以从3个维度进行高度概括：即目前万方检测比对资源总量达4.87亿篇，形成了3类共8种场景化服务和面向5大用户群体提供多角度多种方式的检测服务。

科研诚信建设及学术不端的治理工具分成事前、事中和事后三个大场景。事前通过学习成长和考试考评来提升对诚信的认知。在事中提供监测和预警，在

DOI: 10.14218/MRP.2022.S016

从事科研活动过程中全流程提供贯彻科研诚信能力和素养提升的方法。在事后通过调查和处理等工具帮助各种场景包括职称评定、成果的转化等场合的应用。

诚信问题严重制约我国科技成果转化。如果缺乏诚信，将导致所有的宏观设计功亏一篑。科技成果转化不仅仅是科学研究、技术交易和技术服务的问题，也是政策、评估体系、诚信建设的问题。目前我国科技成果转化的痛点非常严峻，国家在多个层面包括立法等进行引导。供需信息匹配难，数据膨胀、保障体制不健全，技术壁垒的保护等问题都存在。科技成果转化的要素需要宏观管理体制和国家政策环境，但最重要的是需要有值得信赖的科学家做出值得信任的研究成果。

围绕上述问题，万方数据推出了科技成果转化服务体系，把文献数据和实体数据进行抽提、识别关联后形成产业、技术、项目、人物和项目数据库。基于这些成千上万的底层数据后构建了科技成果数据中台服务。中台服务除了实现多来源成果数据的聚集与融合，数据加工、发布和服务需求的汇集；数据采集与治理过程的统一管理。之后基于深度学习技术，包括以文献计量、数据挖掘、文本情感分析、实体关系分析以及统计、对比、评估服务，形成数据分析服务。此外还可以针对海量数据实现毫秒级检索速度，高安全性及高性能数据服务，帮助企业实现大数据高效管理。最后还可以设计灵活的数据输出，数据供应能力，支持在线、离线、接口等多种服务模式。因此，万方数据围绕所有数据的中台服务已经形成了数据聚合、数据分析、数据管理和分发服务体系。

有了上述服务体系后搭建了整体技术框架。资源数据池可以有科技成果数据、转化活动数据和相关产业数据。在底层是包括数据封装、用户画像、场景构建、检索引擎、关联挖掘和知识图谱的基础技术支撑。技术支撑之后对数据进行封装，开发各种应用接口提供服务。在运营管理体系中可以面向专家、高校科研院、企业、服务和投资机构提供成果信息和数据服务。

在服务体系下的成果转化工具可以概括为以下几

个方面。(1) 精准对接可以帮助成果和需求方进行智能匹配、背景调查和统计监测。还可以提供培育孵化服务包括成果筛选、评估和创新成果培育。(2) 实现智能匹配。基于成果生态数据池，以需求为导向，形成围绕技术、人才、团队等要素，面向需求找技术、招才引智、政策辅助参考等场景提供的描述性大文本需求快速匹配等服务，提高转化效率，分析技术转移风险。(3) 形成区域性创新成果转化的统计监测和产业图谱。(4) 基于海量数据和模型方法形成成果评估以及成果竞争力评价。在成果评估过程中有很多优势：有完善的专家信息库，协助快速建立评审团队；形成标准化与个性化相结合的指标体系，辅助管理者确立项目评估标准；公平的专家回避机制，完善科研诚信体系建设；可追踪的评估状态，加强项目监督评估过程；实现实时链接辅助评估小工具，实体画像能力辅助专家在线评审更轻松。成果竞争力评价通过确定评价对象、研发分析评估模型、科研整合与统计最终给出科研竞争力评估报告。通过综合、客观评价，了解产品定位，发现差距。受国家委托万方数据还建立了国家科技成果转化共享服务平台。

产业端也不能丢掉理论研究。最后张博士还分享了万方对信息服务实践的理论研究的探索和思考。基于大数据的科技咨询服务的平台研发项目提出的研究内容主要是如何以数据资源、关键技术为支撑，搭建数据驱动的科技咨询服务平台，形成数字化咨询产品，服务于咨询机构、咨询客户两类用户？有了大数据模式服务的需求和探索，最终希望形成基于大数据的科技咨询服务模式。该模式的应用主要是面向企业技术创新咨询和区域创新创业科技咨询应用。万方医学还推出了生物医药及大健康产业促进大数据平台，形成了贯穿高新区发展新经济的全过程，从 360 度为生物医药和大健康产业生态建设赋能。可以实现产业生态、企业培育、产业促进和园区管理，同时可以帮助园区形成大屏展示。在这个大的产业平台中还可以进行运行监测包括产业企业监测、药品品种监测和企业运行监测，形成企业培育，帮助其形成有核心竞争优势的创新型的服务体系。