

推动科研论文语义评价体系建设

2021年8月2日国务院办公厅发布《关于完善科技成果评价机制的指导意见》，特别提出要利用大数据、人工智能等技术手段，开发信息化评价工具。2021年8月3日国家自然科学基金重大项目公布“大数据驱动的科技文献语义评价体系研究”的年度招标选题，同步提出基于语义内容创新科技成果评价工具和模式。可见，完善科技成果评价机制，从原理机理、方法技术、场景应用等层面优化科技成果评价理论框架和方法体系，已成为当前学术评价乃至科研管理工作亟待突破的重要课题。

目前，我国科研论文量化评价基本沿用以文献计量指标为主导的评价体系，基于引文指标来评价学术影响力、基于网络指标来评价社会影响力，总体上属于形式评价范畴，这种模式导致科学评价中过度强调论文发表数量、被引频次、期刊影响因子等外在量化指标，没有深入到论文内容层面，从全文内容数据、引用数据、关注数据、使用数据等进行语义分析，忽视了从科学性、价值性、创新性、规范性方面进行评价。

所以，推进科研论文语义评价，需要在建立科研论文语义评价的资源体系基础上，探讨语义评价的方法体系，形成语义评价的解析流程，探讨科研论文评价中的贡献、创新、价值和影响的表达形式和识别方法，继而完善知识单元及其关系的图谱构建，结合科研论文评价目标、要素和场景进行应用示范和验证。首先，需要充分利用各类信息资源及其使用数据，建设跨行业、跨部门、跨地区的知识库、需求库、案例库和评价工具方法库；其次，将自然语言处理与文本挖掘技术应用到文本内容的自动分析与语义理解，对论文篇章结构进行自动解析，抽取科研论文中表现创新性、实用性、重要性的知识单元；再次，通过语义关联计算，从文本、图片、翻译到释义、观点、思路间的相似性，对论文进行相似检测，辅助论文结构规范性评价和论证逻辑性评价，发现科研论文中的新知识、新价值和新贡献、新突破；最后，建立科研论文评价数据服务平台，提供语义关联的知识图谱，实现对知识单元影响力、创新力的可视化展示，支撑同行评议，优化学术评价行业生态。

学术评价从文献计量、替代计量走向语义计算，从外在到内在，从文本到语义，从实体概念识别到语义理解对标，需要将语义关联、知识组织与科学评价相结合，将大数据、人工智能技术融入到科学计量学中，从外在统计评价转向知识单元评价，识别其中的新发现、新原理、新方法，测度文献规范性、文本相似性和观点创新性，革新和完善现有文献计量评价体系，形成定量与定性、形式与内容相结合的语义评价体系，开创信息化、语义化、智能化评价工具，丰富知识计量理论方法与工具体系。

■ 曾建勋