

<https://doi.org/10.3799/dqkx.2022.862>



大数据和地质信息学能促进地质学定量化进入新阶段吗？

吴冲龙^{1,2,3}

1. 中国地质大学计算机学院, 湖北武汉 430074
2. 中国地质大学地质信息科技研究所, 湖北武汉 430074
3. 自然资源部基岩勘查区矿产资源勘查工程技术创新中心, 贵州贵阳 550081

地质学定量化是地质科学发展的必然趋势。它是系统地采用数学方法、数学模型和计算工具,对地质现象、地质过程和地质规律进行系统的挖掘、演绎、推理和表达,而非数学方法和工具的零散、简单应用。地球物理学、地球化学、数学地质学和地球动力学等分支学科的形成和发展,使地质学从早期纯粹的现象描述和定性分析,先后进入了“物质科学”和“能量科学”的范畴。近期,随着地质信息科技的兴起,地质学又进入了“信息科学”范畴,但从总体上看,仍然属于定性的描述性科学。数字化转型、大数据和“第四范式”(Gray and Szalay, 2007; Mayer-Schönberger and Cukier, 2013),将是推进这一定量化进程的快速通道。

1 问题提出

阻碍地质学定量化快速推进的主要因素有 3 种:(1)地质作用和地质过程所涉及的时空域巨大,初始条件复杂、影响因素众多,其发展演化轨迹存在确定性(机械性)、不确定性(随机性)和确定随机性(混沌),以确定随机性为主;(2)地质体主要部分深埋于地下,存在着结构信息不全、关系信息不全、演化信息不全和参数信息不完的特点;(3)地质科学研究和勘查工作的数据采集、处理方式,虽然有

定性、定量和半定量的,但以定性为主;其数据类型有结构化、非结构化和半结构化的,但以非结构化(文字、图件)和半结构化的为主。这就导致地质学的认知存在着极大的不确定性。

2 方法途径

随着各种微观超微观高新测试手段和对地观测技术相继出现,所获取的数据类型越来越复杂、维度越来越高、数量也越来越庞大,以至于地质数据呈现出一种爆炸的态势,有了多源、多类、多量、多维、多尺度、多时态和多主题特征,成为名副其实的地质时空大数据(吴冲龙等,2016)。地质科学问题本身所具有的高维度、高复杂性和高不确定性特征,也要求地质科技工作者必须考虑对整个对象时空域内的全体数据,进行跨界广度聚联、融合与深度挖掘。这种状况同时也给出了全面利用这些大数据获取更多隐秘信息,以便深刻地认识地质体、地质现象、地质过程和地质规律的可能性。

地质勘查工作的数字化和信息化转型是地质学定量化的基础。目前,一门研究地质信息本质特征及其运动规律和应用方法的崭新边缘学科——地质信息科学正在形成(吴冲龙等,2014)。其核心是地球信息学或称地质信息学(GeoInformatics)。从 20

基金项目:黔科合战略找矿[2022]项目(Nos.ZD003,ZD004)。

作者简介:吴冲龙(1945—),男,教授,博士生导师,主要从事矿产资源勘查和地质信息科技领域的研究与教学工作。E-mail:804077427@qq.com

引用格式:吴冲龙,2022.大数据和地质信息学能促进地质学定量化进入新阶段吗?地球科学,47(10):3913—3914.

Citation: Wu Chonglong, 2022. Can Big Data and Geoinformatics Promote the Quantification of Geology to a New Stage? *Earth Science*, 47(10): 3913—3914.

世纪 80 年代中期开始,部分研发人员在该领域进行了长期的开拓性研究,初步建立了该学科的理论体系、方法论体系和技术体系框架,并且以三维可视化平台为基础、以点源地矿数据库为核心,通过数字勘查和“玻璃国土”建设,实现了资源、环境勘查—评价—预测全流程计算机辅助化和三维可视化。地质信息科技和大数据方法的形成发展,突破了地质时空数据统合应用的难关,给地质学量化发展提供了机遇和可能,对于突破采样随机性和样品空间狭小、大量非结构化和半结构化数据无法利用,以及缺乏可靠的作用机理、因果关系和动力学模型,仅凭少量观测数据和旧模式进行判断、预测等限制,无疑有极大的好处。

因此,需要努力实践地质勘查的数字化转型、大数据理论方法和科学研究的“第四范式”,改变单纯追究因果关系的定量演绎推理科学观,建立以揭示关联关系为先的新科学观(Tolle *et al.*, 2011),研究和建立多源多类异质异构地物化遥大数据融合和挖掘的理论体系、方法论体系和技术体系。这是推进地质学量化的合理途径。

3 发展前景

这种基于大数据和科学研究第四范式的核心科学体系、战略愿景和战略目标,在一定程度上代表了大数据时代的新地学观和地球科学的发展方向。

在数字化转型、大数据和“第四范式”支撑下,有可能突破上述三大因素的阻碍,加快地质学量化进程。进行地质时空大数据统合利用,涉及一系列理论、方法和技术问题,包括:建立地质时空大数

据的一体化空间参考体系,开展数据空间基准、时态、尺度和语义的一致性处理,以及跨界数据的广度聚联、融合和深度挖掘;探索对各类静态地质勘查数据进行集成化、结构化、可视化转换,并与各类分布式动态地质观测数据进行一体化存储、管理和应用;研究地质时空大数据存储管理、智能计算和云服务技术。

在 20 世纪,地球物理学和地球化学引领了地质学发展;在 21 世纪,地质信息科学(地质信息学)将引领地质学发展。与社会科学量化相似(Watts, 2007),在地质信息学、大数据和科学研究“第四范式”引领下,地质学也将进入全面量化阶段。

参考文献

- Gray, J., Szalay, A., 2007. eScience—A Transformed Scientific Method, Presentation to the Computer Science and Technology Board of the National Research Council. Mountain View, CA.
- Mayer-Schönberger, V., Cukier, K., 2013. Big Data: A Revolution that Will Transform How We Live, Work, and Think. Houghton Mifflin Harcourt, Boston.
- Tolle, K. M., Tansley, D. S. W., Hey, A. J. G., 2011. The Fourth Paradigm: Data-Intensive Scientific Discovery. *Proceedings of the IEEE*, 99(8): 1334–1337. <https://doi.org/10.1109/JPROC.2011.2155130>
- Watts, D. J., 2007. A Twenty-First Century Science. *Nature*, 445(7127): 489. <https://doi.org/10.1038/445489a>
- 吴冲龙, 刘刚, 田宜平, 等, 2014. 地质信息科学与技术概论. 北京: 科学出版社
- 吴冲龙, 刘刚, 张夏林, 等, 2016. 地质科学大数据及其利用的若干问题探讨. *科学通报*, 61(16): 1797–1807.