

The Characteristics and Utilization Value of Geological Big Data

Jianlai Yao

Baoding Jiayu Software Technology Co., Ltd., Baoding, Hebei, 071000, China

Abstract

The rational development and utilization of big data will open up a new cognitive field, provide novel models for solving practical problems, promote the rise of new fields in the digital economy, and lead the innovation of human lifestyles. Digital geology, as the integration of mathematical geology and information technology, has become the “data science” of geological science, which uses data methods to deeply explore geology. Geological data science focuses on researching and developing the potential value of geological big data based on the characteristics of geological data and the needs of geological work. The paper briefly introduces the core characteristics of geological big data, such as mixing, sampling, causality, spatiotemporal, polymorphism, and diversity. It is suggested that when developing and utilizing geological big data, a database, method library, problem library, and model library should be established on the basis of the knowledge base, in order to collect, analyze, and deeply mine geological big data more targeted, and solve various practical problems to achieve precise geological research and application.

Keywords

geological big data; characteristics; economic value; mineral exploration

地质大数据特点及其利用价值

么建来

保定市佳宇软件科技有限公司, 中国·河北保定 071000

摘要

大数据的合理开发与利用将开启全新的认知领域,为解决实际问题提供新颖模式,推动数字经济新领域的崛起,并引领人类生活方式的革新。数字地质作为数学地质与信息技术的融合,已成为地质科学的“数据科学”,它运用数据方法深入探索地质学。地质数据科学专注于根据地质数据的特性和地质工作的需求,研究并开发地质大数据的潜在价值。论文简要介绍了地质大数据的混合性、抽样性、因果性、时空性、多态性和多元性等核心特点,并建议在开发利用地质大数据时,在知识库的基础上建立数据库、方法库、问题库还有模型库等,以便更具有针对性地收集分析、深度挖掘,从而研究地质大数据,并能解决各种实际的问题,实现精准的地质方面的研究与应用。

关键词

地质大数据; 特点; 经济价值; 找矿勘探

1 引言

随着现代科学信息技术的飞速发展,现在各行各业都已经进入了数字化信息科技时代,人们对于大数据的归集分析运用以及深度挖掘,这标志着大数据时代的到来。收集、获取、掌握、汇总、分析以及挖掘信息数据的能力已经成为衡量一个国家实力强弱的标志。大量的信息数据伴随着具体需求不同而决定其产生经济价值不同,进而效益也不同,大量有益的信息都是从大量数据中经过数据分析、深度挖掘出来的。数据信息随着时间持续地产生、不断地流动,进而形成了大数据信息,而数据成为当今每个行业和商业领域的重要因素。因此大数据正在改变着人们观察和考察世界的方式方法,正在引发一场非同寻常的思维革命,以前所未有的速

度引起各个不同领域的深刻变革,大数据除了将更好地解决社会、经济、国防、军事等问题,还有助于解决农业、房产、医疗、健康和地质等各种行业问题,越来越多的问题可以通过分析大数据来进行解决,形成以人为本的大数据战略。

在科技信息高度发达的今天,在地质方面的现代化和科技的创新都离不开信息化,数字地质推动地质勘测、矿体研究、环境保护以及生态修复的发展。大数据研究已成为继实验、推理和传统计算科学后的新兴科学范式,其应用对于提升各类预报能力并广泛服务于学术和经济领域具有重要意义。确实,大数据不仅为我们开启了新的认知维度,更提供了解决现实问题的新途径,从而催生了与实体经济并行发展的新兴经济领域,引领着人类生活方式的深刻变革^[1]。

2 数字地质时代

2.1 什么是地质大数据

地质大数据是指通过各种手段收集、整合和分析大量

【作者简介】么建来(1976-),中国河北唐山人,本科,工程师,从事大数据及相应软件开发研究。

地质学数据,包括地球物理、地球化学、地质勘探、地质调查等多种数据类型,以期获得更全面、更准确、更深入的地质学信息。地质大数据在地质中的成因分析、数学特征研究、地质体评价以及预测等方面都发挥着重要作用。地质大数据是数学地质发展的新阶段、延伸以及拓展。

地质大数据具有双重性,精准的数据能够带来财富,然而虚假的数据则可能引发灾难。在推进全面利用大数据找矿的过程中,强化地质大数据研究至关重要,这更有助于提升地质大数据的科学性、创新性以及可靠性。并采用与地质相关的技术处理数据,可以显著提升地质大数据的信息化水平。

2.2 “智慧地球”时代

在“智慧地球”的新阶段,地质找矿向“智慧找矿”升级,这些有些区别于早期的人工智能和专家系统找矿,这一科技创新的转变来源于数据分析的水平、信息技术的提取和运算技术的提升的显著进步。借助云计算技术,连接成矿观测系统、勘查系统以及地球动力学系统并利用人工智能技术,实现地质数据的自动获取、信息归集、综合分析和建模评价。最终目标是精准确认、圈定和评价成矿远景区、矿床乃至矿体,从而提升找矿效率和准确性。

随着成矿预测研究的日益深入,由于找矿信息数据不断丰富,使得找矿范围也逐渐缩小,靶区的细密级别也持续提升,勘探的风险也在逐步降低,这使得找到矿产资源成功的概率也逐渐增大,同时。在地质领域,地质中的数据异常按照不同尺度水平,可分为区域性异常、局部性异常、显微地质异常,这些异常的分类有助于我们更精准、精确地理解地质现象,为找矿工作提供有力支持。地质大数据的形式多种多样,因此需要针对每一种数据类型采用相应的处理方法。在综合分析时,标准化和归一化是关键步骤,旨在统一地质变量的量纲和数据尺度。有时,通过变换可实现地质变量的正态分布,或将非线性关系转化为线性关系。

2.3 地质大数据应用现状

全世界范围内各国也都在积极推动地质大数据的发展,德国不来梅大学的 Peter Baumann 等^[2]人研发了 Earthserver 大地球数据分析引擎,为“大地理数据”提供了开放、易操作、跨平台的访问方式。美国地质调查局在“大数据时代”背景下提出了核心科学体系^[3],以更好地研究和理解复杂的地球系统。此外,美国国家科学基金会还启动了“地球立方体”项目^[4]。

多个国家已经建立了各自的地质数据共享平台,如美国的 Geospatial One-Stop、加拿大的 GeoConnections 等,这些平台使得政府、企业、学术机构乃至公众能够迅速搜索、访问和接收地理空间信息,实现信息共享,从而做出更明智的决策^[5]。在中国,自然资源部地质信息技术重点实验室搭建的 GeoBDA 地质大数据技术研究实验平台^[6],能够满足地质领域多种数据需求。此外,我们国家在地质方面也自主

研发了大数据平台系统,并在西藏地质调查院进行了试运行,实现了地质大数据的存储、组织管理、快速检索查询与智能分析等功能。

3 地质大数据特点

地质数据独具特性,若忽视或不充分考虑这些特性,在应用大数据时可能遭遇困扰,得出错误结论,甚至造成无谓的浪费与损失。相比于其他数据,地质中的大数据可能会展现出多种的特性,如多元性、混合性、抽样性、时空性、稀缺性、多态性、多源性和分带性等等。

3.1 多元性

地质体是受到多种不同的因素的共同影响而形成的,从而它的外在表征方式也是多样的。总之,地质体可以被看作是多维空间其中的一个点。为了准确地描述某一个地质体,需要找到能精准反映它的空间特征的最小维数,这也体现了地质大数据中多元性的前提条件和必要性所在。如在矿床成因中需要包含源、运、储、变、保五要素才能查清楚矿床的是否存在,需要从这五个方面着手收集并获取所需要的数据,这就是成矿数据的多元性。地质数据中多元性虽然是客观存在,但也不具有确定性,因此在应用地质大数据过程中是必须重视起来的。

3.2 混合性

地球的演变史是一个长达 46 亿年的过程,我们观测到的地质数据往往是在这个演变过程中经过多种作用进行叠加的而形成的结果。如铜矿体中铜品位标准数据,铜矿体中的铜可能会含有地层原本就沉积下来的铜、火山爆发时随岩浆流下来的铜、矿体形成后经过又经过风化而形成的铜或是成矿以后经过变质以后的铜。所以现在的铜矿体中铜品位是地层沉积、岩浆热液、风化、变质等地质作用混合在一起叠加而形成的数据。需要对这些混合的铜品位数据进行分解、分析,分别考察不同成因下铜含量结合其空间分布特征,分别进行研究和评价,才能探究铜矿的成因。这就是地质大数据的混合性。

3.3 抽样性

对各类地质现象、作用和地质体的各种观测、度量、计算和特征描述等,都是在某一需求,基于特定规则,通过点、线、面等不同方式抽取不同的样本进而收集所需的数据。抽样过程需要考虑样本的代表性和可靠性,以确保抽取的样本能够准确地反映整体数据的特征和趋势。同时,抽样也需要考虑数据的分布、密度和相关性等因素,以避免抽样偏差数据分析和处理结果影响。

3.4 时空性

地质数据往往具备显著的时空性特征,这意味着在不同的时间和空间条件下,地质数据可能会呈现出明显的差异。地质体、地质结构、地质资源和地质灾害通常占据庞大的空间范围,并且伴随长时序的发展演化,每一个数据点都

反映了特定时空背景下的某种特性。这种时空性使得地质大数据能够揭示地质现象的动态变化和演化规律,为地质领域的预测和决策提供更加准确的依据。另外,地质大数据中的时空性还蕴含着其他意义,如变异性、序贯性和依赖性等等,这些特性使得地质数据更为复杂且富有研究价值。

3.5 多态性

地质大数据的多态性指的是地质数据的多种形态和表现方式。地质数据可以是来自地质勘探、地震勘探、地质调查、地质灾害监测等多个领域的的数据,包括地质图像、地震波形、地球化学数据、地形地貌数据等多种类型的数据。这些数据的多样性和复杂性使得地质大数据的处理和分析变得更加困难,需要使用先进的数据处理和分析技术来提取有用的信息和知识。如通过对地质现象的直接观测和特征描述来收集定性数据,通过仪器设备的精确测量和不同的测试手段得到定量数据。在这些数据收集过程当中,在进行计数测量时又会产生离散数据,采样测试中会容易获取到连续数据。

为了从这些地质大数据中提炼出有价值的信息和准确的数值,我们需要采用不同的研究方法。同时,不同类型的数据遵循着各自独特的概率分布和数学模型。因此,在利用不同类型的地质数据研究问题时,要选择符合这些地质数据特性以及能够实现研究目标的数学模型是非常重要的。

4 结语

论文旨在深入剖析地质大数据的特点,在运用地质大数据的时候就能够更精准,并且能够有效地解决问题。因此,首先构建一个综合性的知识库,将地质科学中研究方法结果,地质工作中的各种数据,还有就是研究和获取数据过程中所碰到的主要问题进行分类整理。这些问题类型可包括基础地质数据调查、生态环境保护数据调查、勘查数据调查以

及基础地质研究方法等。在知识库的基础上建立数据库、方法库、问题库还有模型库等,以便更具有针对性地收集分析、深度挖掘,从而研究地质大数据,并能解决各种实际的问题。现在,为了推动地质大数据产业的创新发展还需解决很多的挑战,政策、技术、战略等各个方面的问题。专家们认为,在数据安全的前提下,打破行业与行业、部门与部门之间的各种数据壁垒,实现地质大数据的最大程度共享,是关键所在,只有确保不同地质数据的自由流动,才能有效推动地质大数据产业的持续发展。

地质大数据为地质领域的发展提供了有力支持。通过对地质大数据的深入分析和利用,可以为地质勘探、矿产资源开发、地震预测等领域提供更加精准和高效的手段。未来,随着大数据技术的不断发展和完善,地质大数据的利用价值将进一步得到挖掘和发挥,为地质领域的发展注入新的动力。

参考文献

- [1] Porat M U. The Information Economy: Definition and Measurement[J]. Superintendent of Documents, US Government Printing Office, Washington, DC 20402 (Stock No 003-000-00512-7), 1977.
- [2] Baumann P, Mazzetti P, Ungar J. Big Data Analytics for Earth Sciences: the Earth Server Approach[J]. International Journal of Digital Earth, 2015(9):3-29.
- [3] 杨宗喜,唐金荣,周平等.大数据时代下美国地质调查局的科学新观[J].地质通报,2013:1337-1343.
- [4] Earth Cube Guidance for the Community[M].
- [5] Folger P. Geospatial information and geographic information systems (GIS): Current issues and future challenges[M]. 2011.
- [6] 廖谨励,屈红刚,许哲,等.地学大数据技术研究实验平台 GeoBDA[J].地理信息世界,2014,4(8):48-52.