

doi:10.3799/dqkx.2026.010

基于地学大数据和人工智能的多层次矿产预测数据构建

张振杰

中国地质大学（北京）深时数字地球前沿科学中心，地质过程与矿产资源国家重点实验室，地球科学与资源学院，北京 100083

摘要：随着大数据和人工智能技术的迅猛发展，矿产预测正在经历一场由数据稀疏型向数据密集型的技术变革，成为矿产勘查和找矿突破的重要“科技引擎”。尽管地质信息化的进程中已经积累了大量的地质矿产、地球物理、地球化学和遥感等多源异构工程探测数据，以及丰富的地质报告和文献资源，但如何高效整合并深度挖掘这些数据，以进一步优化矿产预测指标体系、构建高质量矿产预测数据集、提升预测精度，仍是当前研究亟待解决的关键问题。针对这些挑战，本文提出利用大模型和知识图谱技术，整合地球系统、成矿系统、勘查系统与预测评价系统的多层次、多维度知识信息，构建多系统耦合的矿产预测知识图谱，实现矿产预测指标体系的智能化构建。同时，基于大数据和人工智能技术，形成以地学探测数据智能挖掘、科学数据智能抽取和时空分析、以及数据智能反演和模拟为核心的多层次矿产预测数据构建体系。这一方法体系通过推动预测数据与预测指标的深度耦合，有效提升预测结果的准确性和可靠性，为矿产勘查和找矿突破提供强有力的技术支持。

关键词：矿产预测；大数据；人工智能；矿产预测模型；矿产预测数据构建

中图分类号：P628

收稿日期：2025-11-17

Construction of multilayered mineralization prediction data based on big geological data and artificial intelligence

Zhang Zhenjie

基金项目：国家重点研发计划项目（2023YFC2906402），教育部基础学科和交叉学科突破计划（JYB2025XDXM803），国家自然科学基金项目（42472358、42430111 和 42050103），中央高校基本科研业务费（2652023001）

作者简介：张振杰（1988—），男，副教授，博士生导师，主要从事地学大数据和智能矿产预测方面的研究。E-mail: zjzhang@cugb.edu.cn, ORCID: <http://orcid.org/0000-0003-1079-8784>

Frontier Science Center for Deep-time Digital Earth; State Key Laboratory of Geological Process and Mineral Resources; School of Earth Sciences and Resources, China University of Geosciences (Beijing), Beijing 100083, China

Abstract: With the advancement of big data and artificial intelligence technologies, mineralization prediction is undergoing a series of technological innovations, transitioning from data-sparse to data-intensive approaches. This shift is expected to become a new “technology engine” for breakthroughs in mineral exploration and discovery. Despite the substantial accumulation of heterogeneous multisource geological, geophysical, geochemical, and remote sensing data, as well as rich geological reports and literature, effectively integrating and deeply mining these valuable data resources to further optimize mineralization prediction indicator systems and construct high-quality mineralization prediction datasets remains a critical research challenge. To address these challenges, this paper proposes integrating multilayered and multidimensional mineralization prediction knowledge across the Earth system, metallogenic system, exploration system, and prediction–evaluation system through large models and knowledge graph technologies. A multilayered, multi-system-coupled mineralization prediction knowledge graph will be constructed, and intelligent construction of mineralization prediction indicator systems will be achieved through knowledge graph mining. Based on big data and artificial intelligence technologies, a multilayered method system for intelligent construction of mineralization prediction data will be developed. This system will focus on intelligent mining of geological exploration data, automated scientific data extraction and spatiotemporal analysis, and intelligent data inversion and simulation. Such advancements are expected to strengthen the deep coupling between prediction data and indicators, enhancing the accuracy and reliability of prediction results, and providing more robust technical support for mineral exploration and discovery breakthroughs.

Key words: Mineralization prediction; big data; artificial intelligence; mineralization prediction models; mineralization prediction data construction

矿产资源在国家经济、国防以及战略性新兴产业的发展中具有至关重要的作用，其供应安全直接影响国家整体战略。然而，我国在多个关键矿产资源领域存在显著的供需缺口，并对外依存度较高，这在复杂的国际局势下对国家经济和安全构成了严峻挑战(干勇等, 2022)。为应对这些挑战，我国已全面启动新一轮找矿突破战略行动，并实施矿产地战略储备工程，旨在提升战略性矿产资源的长期保障能力。

随着我国大部分地区浅部矿产和易识别矿产的发现概率显著降低,找矿难度和风险日益增加,迫切需要发展新型找矿技术手段以实现新的突破。矿产预测作为矿产勘查的核心环节,通过科学方法和技术手段评估潜在矿产资源,为加速找矿、降低风险、提升找矿效率提供了关键的科学依据。自 20 世纪初以来,全球主要经济体和矿产大国持续开展矿产资源预测与评价工作。例如,自 1879 年《组织法》颁布以来,矿产资源评价一直是美国地质调查局(USGS)的核心任务之一,其实施的矿产资源项目 (Mineral Resources Program, MRP) 旨在通过矿产资源预测评价,为公众提供潜在矿产资源的信息。自 2001 年以来,我国已开展了三轮全国矿产资源潜力评价,系统梳理并评估了全国的成矿地质条件与找矿潜力,为国家资源政策和战略的制定提供了科学支撑。

当前,矿产预测已从早期的定性预测,如“求异”和“相似类比”,发展到基于综合信息挖掘与融合的定量预测阶段(成秋明, 2006)。随着地理信息系统(GIS)和人工智能技术的不断进步,矿产预测领域经历了一系列技术革新,机器学习与深度学习方法逐渐成为主流技术手段(陈建平等, 2008; 周永章等, 2021b; 肖克炎等, 2023; 左仁广等, 2024; 成秋明, 2025)。在机器学习,特别是深度学习技术的推动下,矿产预测正从传统的数据稀疏型模式向数据密集型模式转变。这一转变对数据的全面性和深度提出了更高要求。

地质信息化的发展已经积累了大量的地质矿产、地球物理、地球化学和遥感等探测数据,以及丰富的地质报告和文献等文本数据。但如何高效整合、深度挖掘并结构化这些多源异构数据资源,以优化矿产预测指标体系,并构建高质量的矿产预测数据集,提升预测模型的准确性和泛化能力,仍是当前研究领域面临的重大挑战。为应对这些挑战,本文的研究目标是构建一套基于新一代信息技术的矿产预测新范式。提出了(1)以大模型和知识图谱技术为核心的矿产预测指标体系智能化构建思路,实现对多系统(地球系统、成矿系统等)知识信息的整合与深度关联挖掘(图 1);(2)以地学探测数据智能挖掘、科学数据智能抽取和时空分析、以及数据智能反演和模拟为核心的多层次矿产预测数据构建体系(图 2)。有效实现了矿产预测数据与预测指标的深度耦合,为构建高精度的矿产预测模型提供了关键支撑,并最终提升了预测结果的准确性和可靠性,为矿产勘查和找矿突破提供强有力的技术支持。

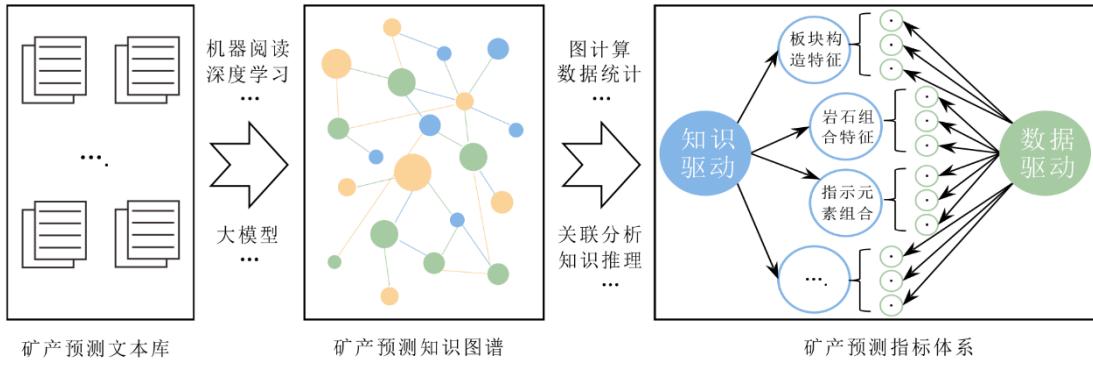


图 1. 基于知识图谱的矿产预测模型自动构建流程

Fig. 1. Automated workflow for mineral prospectivity modeling based on knowledge graphs.

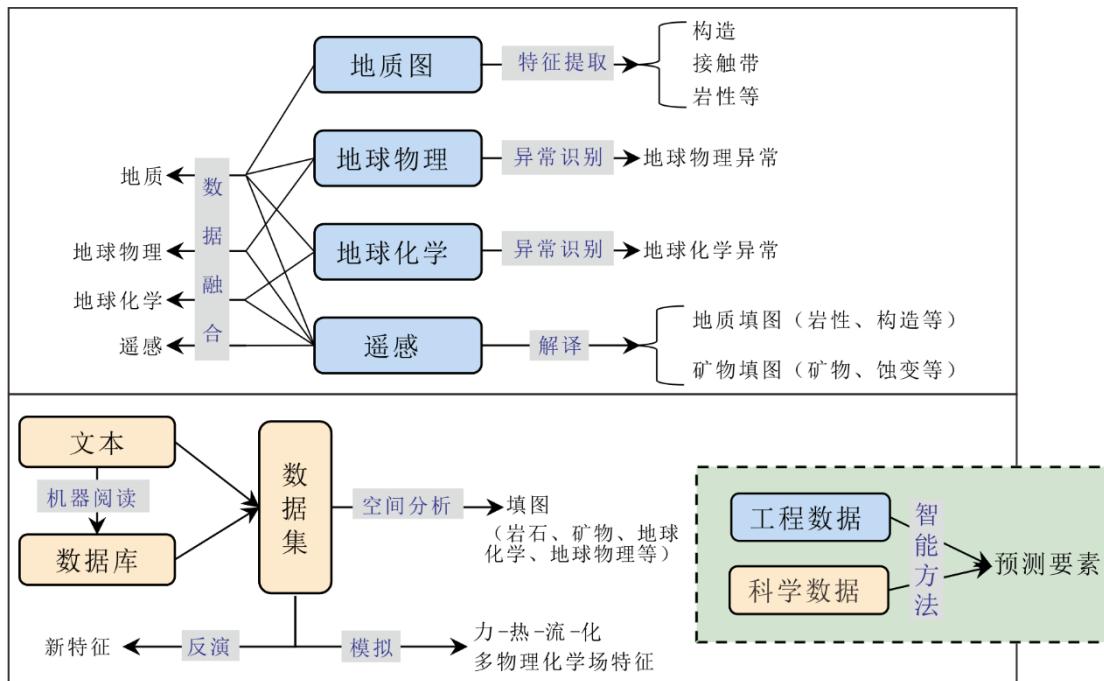


图 2. 基于人工智能和大数据的多层次矿产预测数据智能构建方法

Fig. 2. Multi-level intelligent data construction method for mineral prospectivity prediction based on artificial intelligence and big data.

1 地质大数据与矿产预测

地质大数据是矿产预测的“物质基础”，在矿产预测过程中发挥着关键作用。地质大数据涵盖了大量的地质矿产、地球物理、地球化学和遥感信息，包含从地壳到地幔的深部数据，为理解地球内部结构、地质过程及资源分布提供了全面视角。与通用领域的大数据相比，地质大数据具有海量、多源、多时相、异构和多尺度的特点，且表现出显著的空间属性，如强时空关联和物理关联（赵鹏大，2019）。通过深入挖掘和分析这些数据，大数据技术为地质科学研究和工程实践提供了前所未有的精确度和洞察力（周永章等，2021b）。地质大数据不仅为地质灾害预测与环境保护提供支持，也推动了矿产资源的综合评估与勘查优化。

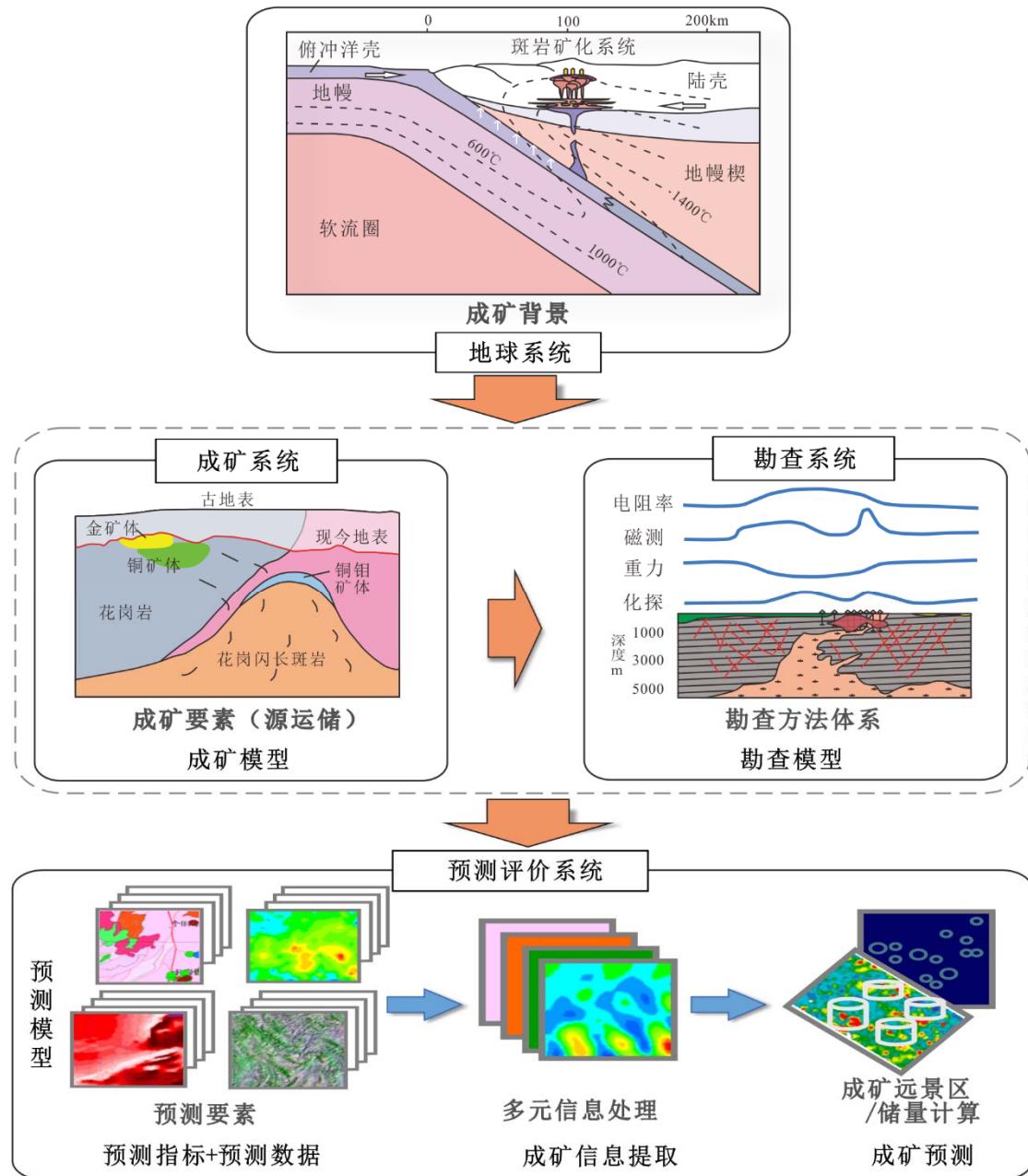


图 3. “地球系统-成矿系统-勘查系统-预测评价系统”耦合的矿产预测流程

Fig. 3. Coupled metallogenetic prediction workflow integrating the “Earth system–metallogenic system–exploration system–prediction and evaluation system”.

经典的矿产预测是一个系统化的科学实践过程（图 3），涵盖“地球系统-成矿系统-勘查系统-预测评价系统”之间的紧密耦合（张振杰等，2021）。成功的矿产预测首先依赖于对区域典型“地球系统-成矿系统”的深入分析。通过研究成矿系统的结构特征，明确成矿元素的富集机制，确定矿床形成的地球动力学背景及其时空分布规律，从而构建矿床成因模型。在此基础上，结合成矿地质环境与矿床演化的研究，识别控矿因素和找矿标志，通过“勘查系统”建立矿床勘查模型（即找矿模型），并设计合理的矿产勘查技术和方法体系，为矿产预测提供综合性勘查数据与资料。最终，通过“预测评价系统”建立矿产预测模型，明确综

合勘查数据与控矿因素及找矿标志之间的逻辑关系，并将这些关系转化为数学模型，以进行矿产资源的预测、评价与靶区优选。目前，“预测评价系统”是矿产预测中智能化程度最高的研究领域，其核心在于通过机器学习等先进技术提取并集成地学空间数据的特征信息，以提高预测的精度与科学性。

然而，传统的矿产预测模型通常依赖于专家对一个或多个典型矿床的研究总结。这些模型的智能化和自动化程度较低，通常侧重于矿床及成矿系统的显著因素和成矿标志，而忽略了成矿物质的早期演化、循环过程及远程动态关联的相互制约。这种局限性导致模型的普适性和通用性不足，预测精度受到限制。因此，在地质大数据和地球系统理论的指导下，需要构建“地球系统-成矿系统-勘查系统-预测评价系统”耦合的智能预测模型（包括指标与数据体系），以提高矿产预测的精度和适用性。

2 矿产预测指标体系智能构建

矿产预测指标体系在“预测评价系统”中扮演着至关重要的角色，是预测模型的核心。它是地质矿产学家与信息处理专家之间的桥梁，对预测结果具有决定性影响。其核心任务是通过“地球系统-成矿系统-勘查系统”的耦合研究，构建控制和指示成矿作用的关键指标要素集。这一指标要素集，不仅是预测模型的输入基础，也是指导后续矿产预测数据（地质、地球物理等）收集、处理和特征提取的关键前提。然而，传统指标体系的构建主要依赖于专家的知识积累，这种方式容易受到专家经验和认知的局限，影响指标的全面性和准确性。

知识图谱作为一种创新的知识管理工具，通过三元组（头实体-关系-尾实体）存储大量的事实信息，形成结构化的知识网络（周成虎等，2021）。例如，矿床知识图谱能够将矿床的空间、时间、构造背景、成因类型、元素、矿物、岩石等多维特征以及它们的组合关系进行关联（周永章等，2021a；Zhang等，2025；Zhao等，2025；冯婷婷等，2025）。知识图谱不仅能够提升领域知识的系统性与可解释性，还为人工智能从感知智能向认知智能发展提供了重要基础（诸云强等，2023）。与依赖个体专家的传统方法相比，知识图谱能够以网络图谱形式系统组织领域内的知识，具备语义约束和自动推理补全的能力，从而有效克服专家经验和认知的局限（Ma，2022）。

在矿产预测过程中，面对传统预测指标体系重视直接要素且泛化能力不足等问题，构建耦合“地球系统-成矿系统-勘查系统-预测评价系统”的矿产预测知识图谱显得尤为重要。矿产预测知识图谱的头实体和尾实体包括地质概念和成矿要素，关系则代表这些要素之间的语义逻辑。通过构建这种知识图谱，可以以图谱形式组织和挖掘要素之间的关联（图4），这不仅有助于揭示隐性关系，还为研究复杂的成矿系统及构建成矿全要素的动态关联提供了关键

支持。研究表明,知识图谱的应用能够显著提升预测指标的准确性和科学性,为矿产预测提供更加全面和系统的理论依据(周永章等,2021a;王成彬等,2024;Zhang等,2025)。

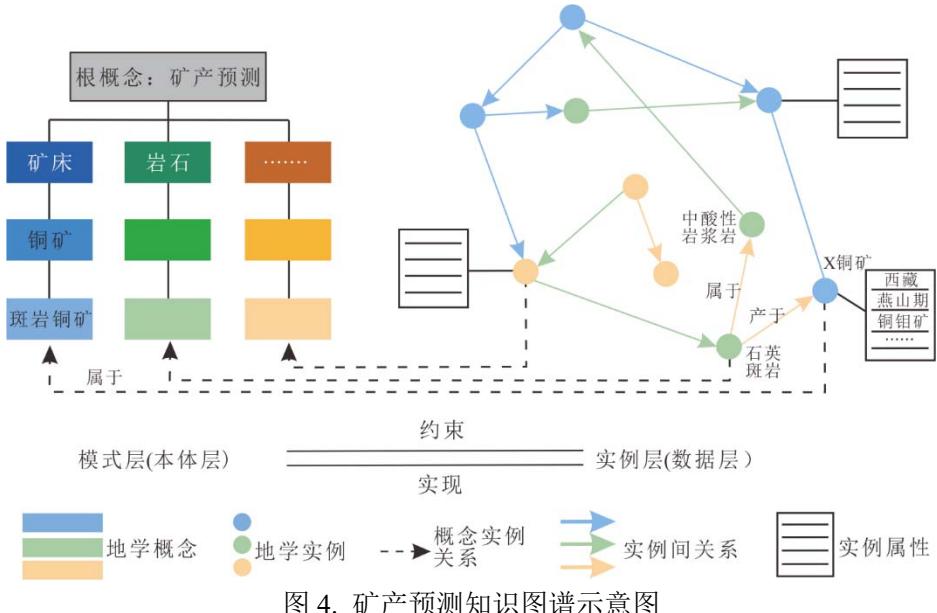


图 4. 矿产预测知识图谱示意图

Fig. 4. Schematic diagram of the knowledge graph for mineral prospectivity prediction.

与通用知识图谱构建方法一致,矿产预测知识图谱的构建方法可分为三种:自顶向下、自底向上和混合模式(诸云强等,2023)。每种方法适用于不同场景和需求。自顶向下模式依赖专家系统,通过人工构建矿产预测领域的实体和关系本体模型(图5),并在此基础上利用人工智能技术实现知识图谱的数据抽取和构建。这种方法虽然依赖专家知识积累,限制了图谱的全面性,但能够确保其结构层次和内容的准确性与系统性。例如,周永章等(2021a)以华南钦杭成矿带斑岩(-夕卡岩)型铜矿为研究对象,利用自顶向下构建了该区域的斑岩成矿系统知识图谱。自底向上模式利用机器学习等自动化技术从大量文本数据中自动抽取实体及关系,进而实现知识图谱的构建(图5)。这种方法能够高效处理大规模数据,但对数据质量和机器学习算法的依赖较大,且需要进一步后处理以确保图谱的准确性和完整性。例如,笔者团队基于18,000余篇关于全球斑岩铜矿的英文论文和摘要,利用Sci-BERT和PFN技术开发了“实体+关系联合抽取”模型,成功建立了一个包含10万余条实体、30万余条关系的斑岩铜矿知识图谱(Zhao等,2025),为斑岩铜矿成矿机理和矿产预测研究提供了基础。混合模式结合了自顶向下和自底向上的优点,通过融合专家知识与自动化技术,不仅提高了构建效率,还增强了图谱的适应性。这种模式尤其适合处理复杂成矿系统中的知识关联,并能够生成更加全面和精确的知识图谱。

近年来,基于大模型的知识图谱构建方法逐渐成为研究热点(Zhang等,2025;师路易

和左仁广, 2025)。这种方法依赖于深度学习技术, 尤其是大规模预训练语言模型(如 BERT、GPT 等), 实现矿产预测领域知识的自动抽取与整合。大模型方法具有高度自动化、强大的数据处理能力、良好的知识更新与扩展能力, 并能够跨领域融合知识, 为未来矿产预测知识图谱的构建提供了新的可能性。例如, 笔者使用大语言模型 (Qwen2-72B) 从地质文献中自动提取稀有金属成矿相关的元素组合, 构建了南岭地区稀有金属矿产预测知识图谱 (Zhang 等, 2025), 突破了人工构建知识图谱效率低和融合难的瓶颈。

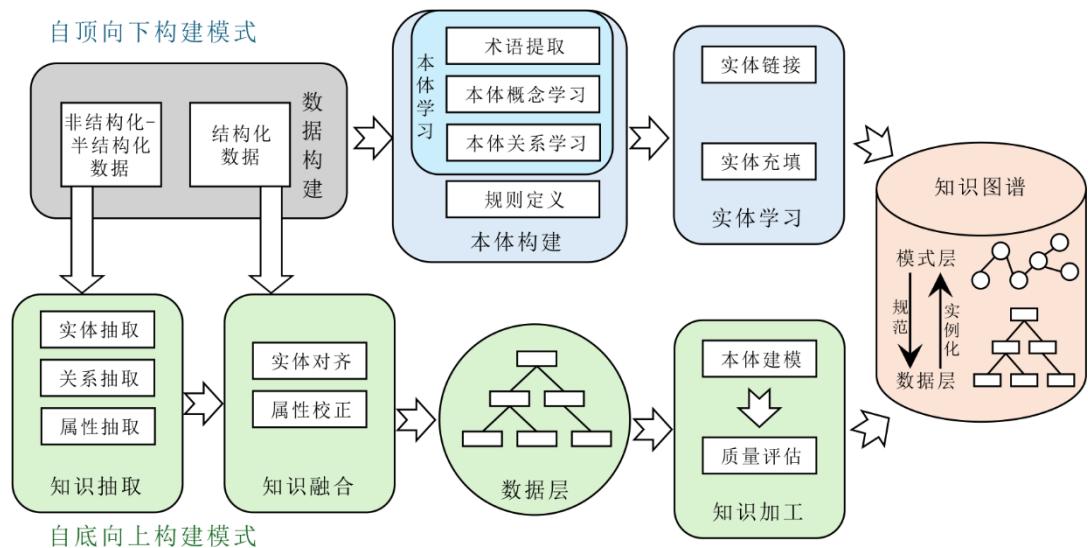


图 5. 矿产预测知识图谱构建流程

Fig. 5. Workflow for constructing the knowledge graph of mineral prospectivity prediction.

已构建的矿产预测知识图谱可以通过图计算、数据统计、关联分析和知识推理等技术进行深入挖掘, 识别图谱中的弱关联和间接关联实体, 揭示其深层次的逻辑关系 (Zhang 等, 2025; Zhao 等, 2025)。这些技术能够帮助揭示复杂成矿系统中的远程关联、隐式依赖关系、模糊关联等, 从而扩展对成矿作用全要素指标集的理解 (图 1)。例如, 一些成矿因素可能通过复杂的间接路径影响矿床的形成, 这些隐性关联通常难以通过直接观察发现, 但可通过图计算和机器学习推理得以揭示。更重要的是, 知识图谱挖掘出的关键控矿指标和关联关系, 能够为后续的矿产预测数据构建提供明确的知识指导和指标筛选依据, 确保所构建的数据集聚焦于关键指标信息, 增强数据构建的科学性和有效性, 从而实现指标体系与数据的深度耦合。

3 矿产预测数据智能构建

矿产预测数据构建是一个系统性过程, 旨在通过工程探测和文本数据处理, 创建全面耦合矿产预测指标的数据体系, 即反映矿床形成机制和找矿标志的数据集合, 并优化预测模型。传统的矿产预测数据构建主要依赖于对地质矿产、地球物理、地球化学和遥感等工程数据的

综合分析,以提取与矿产形成紧密相关的信息。然而,随着大数据和人工智能技术的迅速发展,现代矿产预测数据构建更强调数据的整合与优化。通过多源数据融合与先进数据处理技术的应用,能够显著提高数据的准确性和可靠性,使数据更真实地反映矿产形成的复杂情况。结合前沿数据分析技术,不仅可以挖掘数据中的潜在信息和相关性,还能加深对成矿机制的理解,增强对矿产预测的支持。

在大数据和人工智能背景下,矿产预测数据智能构建主要包括以下途径(图2):(1)工程探测数据挖掘与构建;(2)机器阅读与长尾数据智能提取;(3)人工智能与预测数据反演;(4)人工智能与预测数据模拟。

3.1 工程探测数据挖掘与构建

地质特征、地球物理、地球化学和遥感信息作为区域地质及成矿作用的多维反映,目前广泛应用于矿产预测。这些数据通过地质矿产填图、地球物理和地球化学扫描以及遥感探测等手段获得,被统称为“地物化遥”数据。例如,我国从20世纪70年代起实施的“区域化探全国扫面计划”,通过水系沉积物采样,生成了900幅1:20万(涵盖39种元素)的地球化学测量数据,覆盖近700万平方公里,为矿床发现和勘探提供了重要支持(王学求,2013)。涉及工程探测的预测数据通常可从这些多源异构空间数据中进一步构建,涉及的方法包括空间分析、局部奇异性分析、滤波、机器学习等特征提取和异常识别方法(张振杰等,2021)。

地物化遥数据不仅具有显著的地学空间特征,还展现出多尺度、多源异构的复杂性。在矿产预测中,高效利用多源数据的关键在于数据融合。数据融合可分为三种类型:相同分辨率的不同平台测量数据、不同分辨率的相同平台测量数据以及不同分辨率和不同平台的测量数据。通过多测量平台、多时相和多频谱的数据融合,可以显著提高对地质体、构造和矿体等地质目标的探测和识别能力。例如,遥感图像记录了地质体的光谱特征,通常具有较高分辨率,而地球化学数据揭示了元素的富集或亏损情况,反映地质体的化学属性,但其分辨率较低。为解决这一问题,研究者提出了多种数据融合技术。Ding等(2023)提出基于线性回归模型的融合方法,将高分辨率遥感图像与小比例地球化学数据结合,从而提升地球化学数据的分辨率,适用于缺乏大比例尺地球化学数据的区域勘探。此外,Wang等(2021)通过多源数据融合技术,将低分辨率地球化学数据与高分辨率ASTER遥感图像融合,生成既保留地球化学模式又保持遥感图像纹理结构的综合数据(图6),为岩性单元的识别提供新的数据支持。

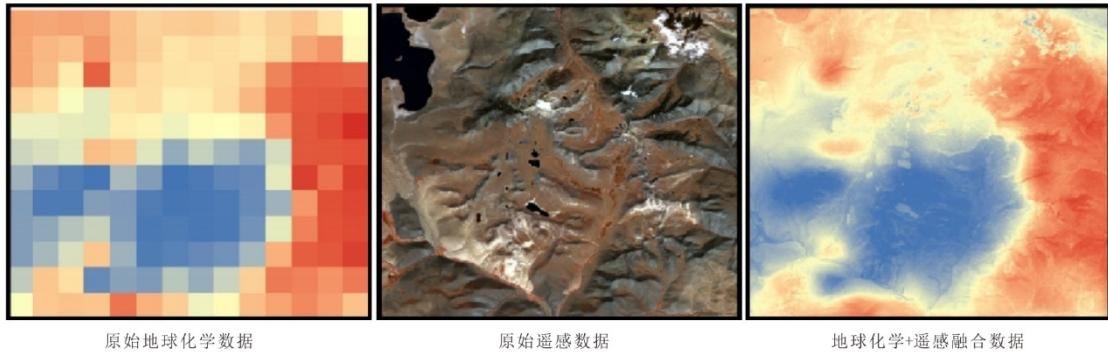


图 6. 小比例地球化学数据与高分辨率遥感图像融合 (据 Wang 等 (2021))

Fig. 6. Integration of low-scale geochemical data and high-resolution remote sensing imagery

随着人工智能技术的发展,地物化遥多源数据在矿物蚀变和岩性填图研究中的应用逐渐成为新兴领域,为矿产预测提供了详实的数据基础。ASTER 和 WorldView 系列遥感数据被广泛应用于泥化、高岭石化和青磐岩化等蚀变矿物的填图研究,尤其在斑岩铜矿和浅成低温热液金矿勘探中展现巨大潜力 (Zhang 和 Zhou, 2017; Dai 等, 2018)。遥感技术的进步,尤其是光谱分辨率的提升,使矿物填图从简单的矿物种类识别发展到矿物亚类、丰度及成分的精细识别。例如,利用高分五号 (GF-5) 高光谱遥感数据,研究人员能够有效区分高铝、中铝和低铝绢云母 (董新丰等, 2020)。此外,深度学习和机器学习技术的应用进一步提升了大规模遥感数据的处理和分类能力。在精细地质体反演和岩性填图方面,除常用遥感数据外,地球化学数据也开始扮演重要角色。例如,笔者团队利用水系沉积物的地球化学数据结合机器学习技术,开展玄武岩地质填图,为第四系覆盖区的隐伏玄武岩填图提供有效技术框架 (Ge 等, 2022)。

总的来说,地物化遥多源探测数据与人工智能技术的结合,不仅提升了矿物蚀变和岩性填图的精度,还为矿产预测提供了丰富准确的数据支持,助力提高矿产资源勘探的科学性与效率。

3.2 机器阅读与长尾科学数据智能提取

在地质矿产领域,除了工程探测数据 (如“地物化遥”等),还有大量的科学数据 (如岩石和矿物地球化学、年代学、同位素数据等)。这些数据通常存储在研究报告、书籍和学术论文中。随着地质学科的发展,科学数据的发表和积累显著增加,为矿床形成机制和矿产预测研究提供了新的数据手段。例如,地球深部被认为是大规模成矿作用的“驱动机”、“供应源”和“传输带”。Hou 等 (2023) 通过系统收集西藏冈底斯地区新生代以来的锆石 Hf-O 同位素和全岩 Nd 同位素数据,构建了区域尺度的 Nd-Hf-O 同位素填图 (图 7),揭示了俯冲的印度大陆岩石圈的撕裂作用,成为地幔软流圈上升的通道,从而重新作用于上覆的亚洲

岩石圈并触发熔融，生成含水的花岗岩质富铜熔体。这些岩浆岩“探针”及区域同位素填图不仅是探索深部物质组成与演化过程的主要手段，还提升了对区域成矿规律的认识，为区域矿产预测提供了重要支持。

目前，这些地质科学数据逐渐被系统化存储在专业数据库中，如 EarthChem、PetDB 和 NAVDAT 等岩石地球化学和同位素数据库。这些数据库不仅保存了大量历史数据，还不断更新和扩展，使研究人员能够便捷地获取最新的科学数据，显著降低了数据获取的难度。例如，EarthChem 是一个重要的数据平台，涵盖了超过 100 万件岩石矿物样品和 3000 万条地球化学数据，其数据来源于已发表文献中的高质量条目，并经过同行评议，部分数据由科学家直接上传（张颖慧等，2020）。尽管如此，这些数据库中的数据收录并不完全，尤其是中文文献中的数据往往未被充分覆盖。此外，现有数据库中的数据种类也相对有限，特别是关于矿床和矿石地球化学的专业数据库尚未完全建立。这限制了对矿床特征及成因过程的全面了解，也制约了矿产预测与找矿的突破。

为了有效利用这些海量的地质科学数据，需要将文献中的文本信息抽取并转化为结构化的数据格式。然而，传统的人工抽取方法面临工作量大、效率低的问题。机器阅读技术的出现为解决这一难题提供了有效途径。自动抽取技术能够高效地从大量文本中提取关键信息，并将其转化为结构化的数据格式，具备明确的属性和数值信息。这不仅大大提高了数据处理效率，还确保了数据的准确性和一致性。例如，Martin 等（2022）展示了自动抽取技术在地球化学与地质年代数据获取中的应用潜力。他们利用网络爬虫技术自动搜索、提取并整理 Figshare 数据库中的数据，能够在短时间内收集并下载大量论文附件和数据，为后续研究提供了丰富的数据资源。笔者团队也利用机器阅读技术挖掘文献中的斑岩铜矿时空信息，构建了斑岩铜矿时空数据库，显著提升了对斑岩铜矿时空分布规律及成因机制的研究。

总之，自动化处理技术不仅节省了研究人员的时间和精力，还大幅提高了数据获取的效率与精度，成为未来文本数据智能提取的主要手段。通过机器阅读与长尾数据智能提取技术的应用，科学数据的获取与利用将变得更加便捷，为矿产预测提供更丰富、更精准的数据支持。

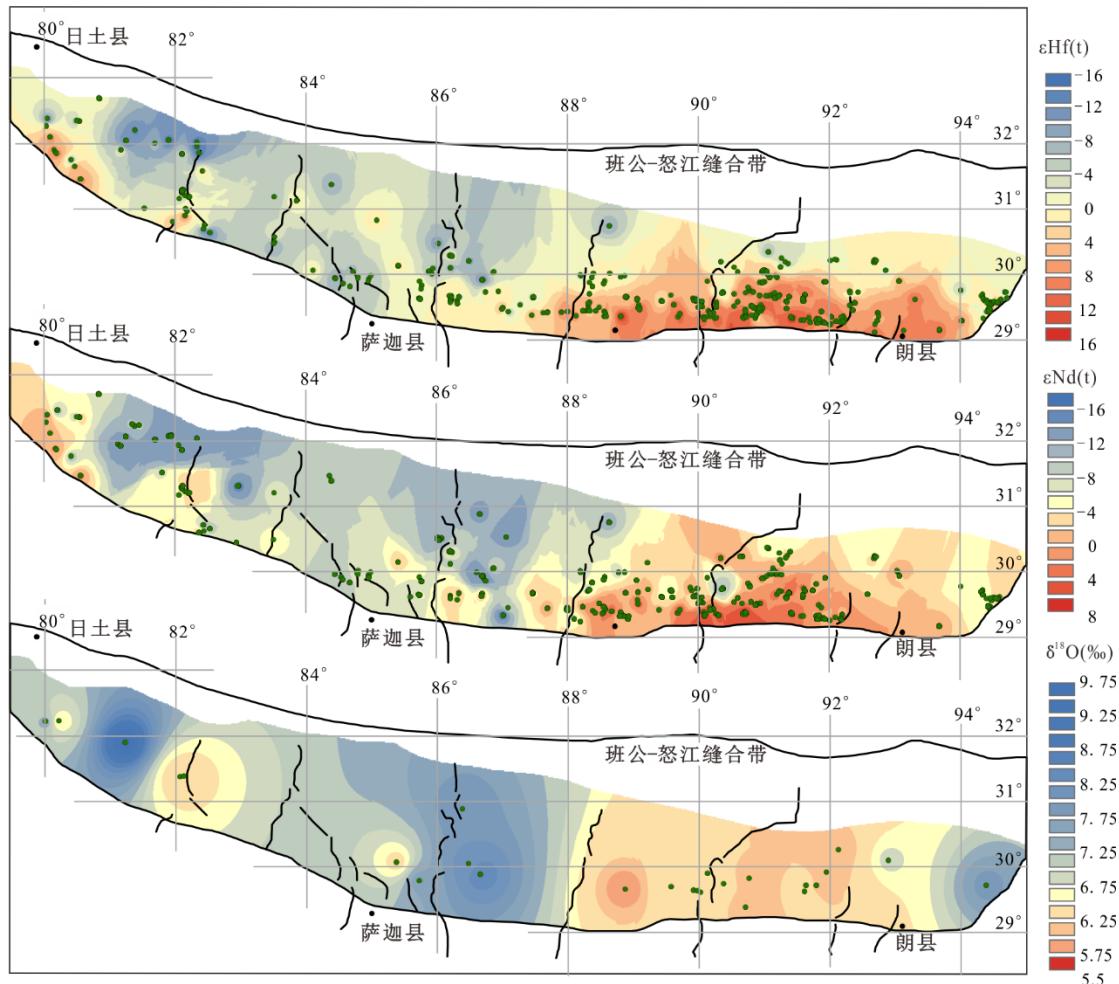


图 7. 西藏冈底斯地区新生代以来火成岩 Nd-Hf-O 同位素填图（据 Hou 等（2023）数据重绘）

Fig. 7. Mapping of Nd–Hf–O isotopes in Cenozoic igneous rocks of the Gangdese region, Tibet

3.3 人工智能与预测数据反演

地质科学数据在矿产预测中具有重要的应用潜力，尤其是岩石和矿物的地球化学及同位素数据。这些数据能够直接反映地质体的化学属性，从而揭示潜在的成矿环境。然而，与成矿相关的远程、隐式依赖因素及间接因素通常难以通过直接的工程数据或文献获得，需要依赖数据反演技术进行深入分析。

例如，最新研究表明，澳大利亚主要的沉积岩容矿贱金属矿床沿着 170 公里岩石圈厚度等值线分布，全球范围内的大型沉积岩容矿贱金属矿床的分布同样与这一等值线高度相关 (Hoggard 等, 2020)。这意味着岩石圈厚度在一定程度上控制了此类矿床的分布和产出。尽管现代岩石圈厚度可以通过地球物理指标分析得出，但对地质历史时期岩石圈厚度的估计仍面临挑战。研究发现，玄武岩的化学成分能够记录岩石圈底部的压力等信息 (Niu, 2021; Zhang 等, 2023a)，因此成为反演岩石圈厚度的理想数据源。基于全球玄武岩地球化学数据，

研究人员开发了岩石圈厚度反演的机器学习模型 (Zhang 等, 2023a), 该模型能够在全球范围内对深时岩石圈厚度进行反演, 为矿产预测和靶区优选提供了关键数据支持。

此外, 斑岩铜矿的形成与地壳厚度密切相关。厚地壳有利于岩浆的 MASH 过程 (混合、析出、分离、熔融), 这一过程生成富水、高氧化性的岩浆, 促进铜的富集 (Richards, 2003)。研究发现, 中酸性岩浆岩的痕量元素比值 (如 Sr/Y 、 $(\text{La}/\text{Yb})\text{N}$ 和 Ce/Y) 可用于表征地壳厚度 (Chapman 等, 2015)。进一步研究表明, 利用机器学习方法进行地壳厚度的反演能够获得更加稳健的结果, 为斑岩铜矿预测提供了先进的工具 (Zhang 等, 2023b)。

总之, 随着大数据和人工智能技术的不断进步, 远程和隐式的成矿要素正在变得可量化和可分析。通过对工程数据和科学数据的深度反演, 地质学家能够揭示以往难以直接触及的成矿因素, 为矿产预测开辟新的数据来源, 极大地提升预测的准确性和效率。这种方法不仅拓宽了矿床预测的视角, 还推动了新一代矿产勘查技术的发展。

3.4 人工智能与预测数据模拟

成矿系统是一个涉及动力学形变、流体流动、热传导、物质运移及化学反应等全反馈耦合过程的复杂系统 (袁峰等, 2019)。传统的矿床学研究方法往往难以全面描述这些复杂的反馈机制。为更深入地理解成矿过程, 通过数值模拟对构造特征和成矿环境等条件进行模拟, 能够解析矿床形成中的形变、热力学、流体流动和化学反应等过程 (Hu 等, 2020)。例如, Hu 等 (2020) 的研究通过数值模拟确定了茶亭斑岩型铜金矿成矿系统的关键特征, 将形变、流体流动、热传递、化学反应和溶质运移耦合起来, 估算成矿持续时间在 9,600 到 75,000 年之间, 深度解析了矿床的成因机制。类似的研究如 Zou 等 (2017) 对虎头崖铅锌矿床成矿过程的模拟, 揭示了热液流动是该矿床成矿的关键控制因素。

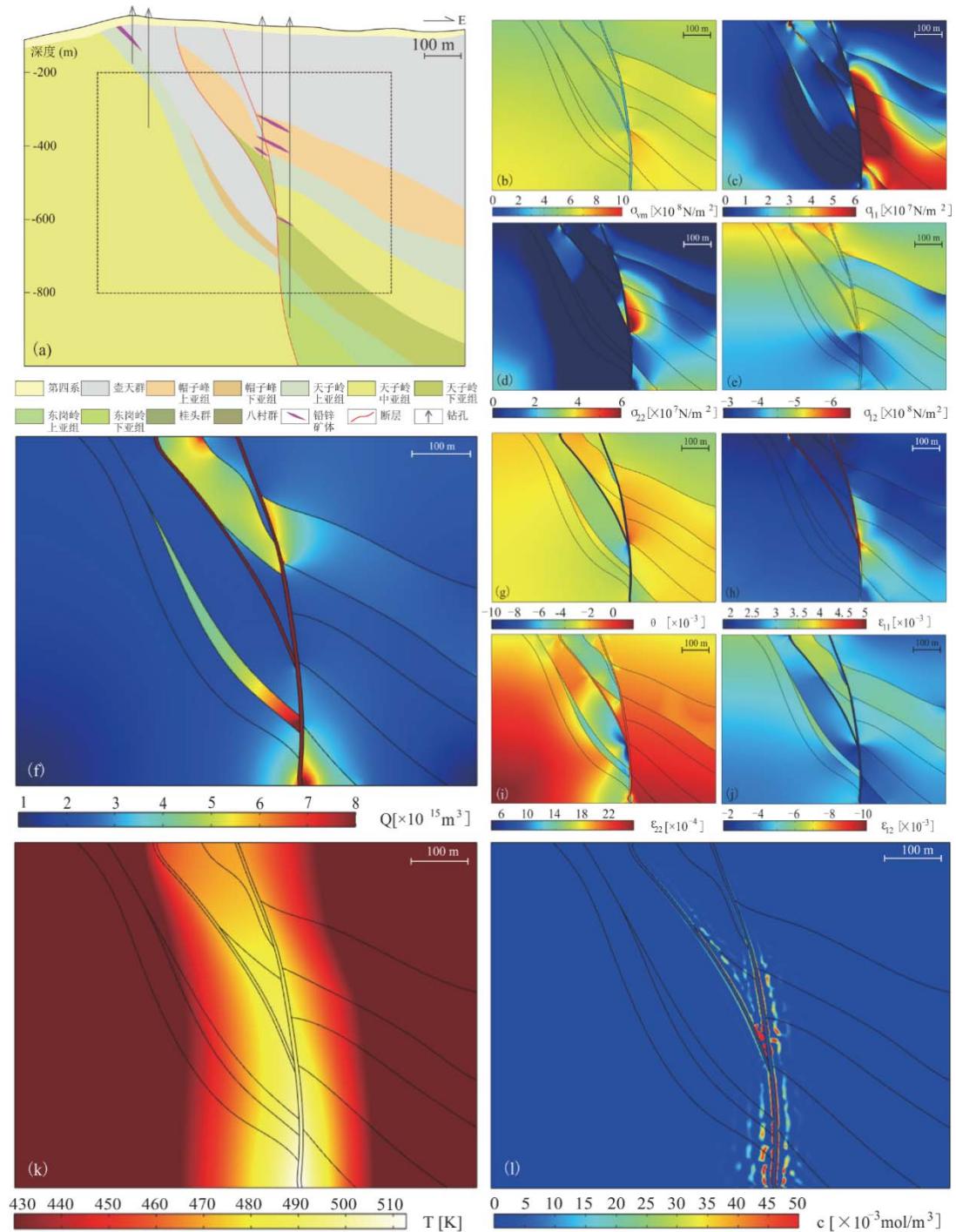


图 8. 粤北凡口铅锌矿成矿力-热-流场模拟 (Xiao 等, 2024)

(a) 钜口铅锌矿勘查剖面图; (b) 冯·米塞斯应力; (c) 第一主应力; (d) 第二主应力; (e) 第三主应力; (f) 流体通量模拟; (g) 体积应变, (h) 第一主应变; (i) 第二主应变; (j) 第三主应变; (k) 热力场模拟; (l) 铅锌矿随机森林预测结果

Fig. 8. Simulation of structural stress, strain, thermal, and fluid fields for the Fankou Pb-Zn deposit, northern Guangdong.

(a) Exploration profile; (b) von Mises stress; (c) first principal stress; (d) second principal stress; (e) third principal stress; (f) fluid flux simulation; (g) volumetric strain; (h) first principal strain; (i) second principal strain; (j) third principal strain; (k) thermal field simulation; (l) random forest prediction results for Pb-Zn mineralization.

力-热-流-化多物理化学场的耦合模拟不仅帮助研究人员更准确地理解矿床的动态形成机制，还为矿产预测提供了新的数据支持。已有多个不同尺度的应用案例展示了模拟数据在矿产预测中的有效性。例如，斑岩铜矿系统通常沿汇聚板块边缘的弧岩浆带分布。Diaz-Rodriguez 等（2021）通过 GPlate 模拟板块俯冲过程，提取相关动力学参数，并结合机器学习模型，成功将俯冲带的构造演化与斑岩铜矿的成矿过程关联起来，生成了铜矿化预测图。同样，安文通等（2021）通过模拟地南金矿成矿过程中的应变场和温度场，为该区域的三维矿产预测提供了可靠数据支撑。Xiao 等（2024）的研究利用数值模拟方法恢复了粤北凡口铅锌矿的成矿构造应力场、应变场、热力场和流体场，并将这些模拟数据与机器学习工具结合，成功完成了矿体尺度的矿产预测（图 8）。

总体而言，随着地质大数据和人工智能技术的快速发展，对成矿条件和过程的量化模拟已经成为现实。这不仅帮助研究人员深入理解成矿动力学演化过程及其响应，还为解决矿产预测模型中某些难以直接获取的数据提供了新的思路和工具。通过模拟与实际数据的融合，能够进一步提高矿产预测的精度和科学性，推动矿产资源勘探的高效进行。

4 结论与展望

人工智能和大数据技术正在深刻改变地学研究与工程实践的范式，为矿产预测带来了前所未有的机遇。尽管人工智能技术已逐渐融入矿产预测评价系统，并在成矿信息提取和综合方面取得了一定进展，但在预测指标与数据构建方面，人工智能的应用仍需进一步探索与完善。成矿元素与矿产资源的分布受到壳幔动力、深部地质过程、物质交换及物理化学过程的综合影响。这种复杂性要求我们对深部物质交换和物理化学过程有更深入的理解，而传统的专家解析模式往往难以准确描述这些复杂过程，尤其是在远程和隐式因素的把握上，困难尤为突出。

大模型和知识图谱技术能够实现对地球系统、成矿系统、勘查系统及预测评价系统等多层次、多维度信息的整合，从而推动矿产预测指标体系的智能化构建。同时，人工智能和大数据技术也将促进形成一个以地学探测数据挖掘、科学数据智能抽取和时空分析、以及数据智能反演和模拟为核心的多层次矿产预测数据智能构建方法体系（图 2）。然而，某些关键预测指标仍难以转化为空间数据，从而影响矿产预测的准确性。通过构建可解释性的矿产预测人工智能算法，可以实现领域知识在“数据-算法-输出”全链条中的嵌入，从而提升预测结果的可解释性。因此，未来可以在预测评价系统中直接纳入可转化为空间数据的预测指标，并通过算法设计和输出嵌入那些难以构建为空间数据的预测指标。这将不断提高矿产预测指标和数据的准确性和可靠性，为矿产勘查和找矿突破提供更有力的技术支持。

致谢：本文内容在“第八届中国人工智能与大数据地球科学学术研讨会”和“新一轮找矿突破战略行动青年论坛”等会议与同行们进行了有益讨论和交流，博士生简富源对本文部分图件进行了清绘，作者一并表示感谢。

参考文献

- An, W. T., Chen, J. P., Zhu, P. F., 2021. A Two-Way Forecasting Method Based on Numerical Simulation of Mineralization Process for the Prediction of Concealed Ore Deposits. *Earth Science Frontiers*, 28(3): 97-111. (in Chinese with English abstract)
- Chapman, J. B., Ducea, M. N., DeCelles, P. G., et al., 2015. Tracking changes in crustal thickness during orogenic evolution with Sr/Y: An example from the North American Cordillera. *Geology*, 43(10): 919-922. doi: 10.1130/G36996.1
- Chen, J. P., Chen, Y., Wang, Q. M., 2008. Study on Synthetic Informational Mineral Resource Prediction Using GIS: A Case Study in Chifeng Region, Inner Mongolia, China. *Earth Science Frontiers*, 15(4):18-26. (in Chinese with English abstract)
- Cheng, Q. M., 2006. Nonlinear Theory of Mineralization Prediction: Singularity - Generalized Self-Similarity - Fractal Spectrum (3S) Models and Methods. *Earth Science*, 31(3):337 - 348.(in Chinese with English abstract)
- Cheng, Q. M., 2025. A New Paradigm for Mineral Resource Prediction Based on Human Intelligence-Artificial Intelligence Integration. *Earth Science Frontiers*, 32(4): 1-19. (in Chinese with English abstract)
- Dai, J. J., Qu, X. M., Song, Y., 2018. Porphyry copper deposit prognosis in the middle region of the Bangonghu - Nujiang Metallogenic Belt, Tibet, using ASTER remote sensing data. *Resource Geology*, 68 (1) : 65-82. doi: 10.1111/rge.12154
- Diaz-Rodriguez, J., Müller, R. D., Chandra, R., 2021. Predicting the emplacement of Cordilleran porphyry copper systems using a Spatio-temporal machine learning Model. *Ore Geology Reviews*, 137: 104300. doi: 10.1016/j.oregeorev.2021.104300
- Ding, H. F., Jing, L., Xi, M. J., et al., 2023. Research on Scale Improvement of Geochemical Exploration Based on Remote Sensing Image Fusion. *Remote Sensing*, 15 (8) : 1993. doi: 10.3390/rs15081993
- Dong, X. F., Gan, F. P., Li, N., et al., 2020. Fine Mineral Identification of GF-5 Hyperspectral Image. *Journal of Remote Sensing*, 24(4): 454-464. doi: 10.11834/jrs.20209194. (in Chinese)
- Feng, T. T., Cai, S. R., Zhang, Z. J., 2025. Mining Elements of Carbonatite-Type Rare Earth Deposits Based on Knowledge Map. *Earth Science Frontiers*, 32(4): 262-279. (in Chinese with English abstract)
- Gan, Y., Peng, S. P., Mao, J. W., et al., 2022. High-Quality Development Strategy for the Supply Chain of Critical Minerals and Its Material Industry in China. *Engineering Science in China*, 24(3): 1-9. (in Chinese with English abstract)
- Ge, Y. Z., Zhang, Z. J., Cheng, Q. M., et al., 2022. Geological mapping of basalt using stream sediment geochemical data: Case study of covered areas in Jining, Inner Mongolia, China. *Journal of Geochemical Exploration*, 232: 106888. doi: 10.1016/j.gexplo.2021.106888
- Hoggard, M. J., Czarnota, K., Richards, F. D., et al., 2020. Global distribution of sediment-hosted metals controlled by craton edge stability. *Nature Geoscience*, 13 (7) : 504-510. doi: 10.1038/s41561-

- Hou, Z. Q., Wang, R., Zhang, H. J., et al., 2023. Formation of giant copper deposits in Tibet driven by tearing of the subducted Indian plate. *Earth-Science Reviews*, 243: 104482. doi: 10.1016/j.earscirev.2023.104482
- Hu, X. Y., Li, X. H., Yuan, F., et al., 2020. Numerical modeling of ore-forming processes within the Chating Cu-Au porphyry-type deposit, China: Implications for the longevity of hydrothermal systems and potential uses in mineral exploration. *Ore Geology Reviews*, 116: 103230. doi: 10.1016/j.oregeorev.2019.103230
- Ma, X. G., 2022. Knowledge graph construction and application in geosciences: A review. *Computers & Geosciences*, 161: 105082. doi: 10.1016/j.cageo.2022.105082
- Martin, E. L., Barrote, V. R., Cawood, P. A., 2022. A resource for automated search and collation of geochemical datasets from journal supplements. *Scientific Data*, 9: 724. doi: 10.1038/s41597-022-01730-7
- Niu, Y. L., 2021. Lithosphere thickness controls the extent of mantle melting, depth of melt extraction and basalt compositions in all tectonic settings on Earth - A review and new perspectives. *Earth-Science Reviews*, 217: 103614. doi: 10.1016/j.earscirev.2021.103614
- Richards, J. P., 2003. Tectono-Magmatic Precursors for Porphyry Cu-(Mo-Au) Deposit Formation. *Economic Geology*, 98 (8) : 1515-1533. doi: 10.2113/gsecongeo.98.8.1515
- Shi, L. Y., Zuo, R. G., 2025. Mineral Prospecting Large Model. *Earth Science*. doi:10.3799/dqkx.2025.3190. (in Chinese with English abstract)
- Wang, C. B., Wang, M. G., Wang, B, et al., 2024. Knowledge Graph-Infused Quantitative Mineral Resource Forecasting. *Earth Science Frontiers*, 31(4): 26-36. (in Chinese with English abstract)
- Wang, X. Q., 2013. Landmark Events of Exploration Geochemistry in the Past 80 Years. *Geology in China*, 40(1): 322-330. (in Chinese with English abstract)
- Wang, Z. Y., Zuo, R. G., Jing, L. H., 2021. Fusion of Geochemical and Remote-Sensing Data for Lithological Mapping Using Random Forest Metric Learning. *Mathematical Geosciences*, 53 (6) : 1125-1145. doi: 10.1007/s11004-020-09897-8
- Xiao, F, Chen, X. Y., Cheng, Q. M., 2024. Combining numerical modeling and machine learning to predict mineral prospectivity: A case study from the Fankou Pb - Zn deposit, southern China. *Applied Geochemistry*, 160: 105857. doi: 10.1016/j.apgeochem.2023.105857
- Xiao, K. Y., Fan, M. J., Sun, L., et al., 2023. Theoretical Method of Integrated Geological Information Prediction of Metallogenic Series for Mineral Resource Potential Assessment. *Acta Geoscientica Sinica*, 44(5): 769-780. (in Chinese with English abstract)
- Yuan, F., Li, X. H., Hu, X. Y., et al., 2019. A New Approach for Researching Hydrothermal Deposit: Numerical Simulation. *Chinese Journal of Geology*, 54(3): 678-690. doi: 10.12017/dzqx.2019.040. (in Chinese with English abstract)
- Zhang, N. N., Zhou, K. F, 2017. Identification of hydrothermal alteration zones of the Baogutu porphyry copper deposits in northwest China using ASTER data. *Journal of Applied Remote Sensing*, 11 (1) : 015016. doi: 10.1117/1.JRS.11.015016
- Zhang, Y. H., Wang, T., Jiao, S. T., et al., 2020. Review of Igneous Rock Databases and Their Application Prospect. *Geological Journal of China Universities*, 26(1): 11-26. (in Chinese with English abstract)
- Zhang, Z. J., Chen, G. X., Kusky, T., et al., 2023a. Lithospheric thickness records tectonic evolution by controlling metamorphic conditions. *Science Advances*, 9 (50) : eadi2134. doi:

10.1126/sciadv.adi2134

- Zhang, Z. J., Chen, G. X., Yang, J., et al., 2021. Machine Learning for Mineral Prospectivity: A Case Study of Iron-Polymetallic Mineral Prospectivity in Southwestern Fujian. *Earth Science Frontiers*, 28(3): 221-235. (in Chinese with English abstract)
- Zhang, Z. J., Kusky, T., Yang, X. K., et al., 2023b. A paradigm shift in Precambrian research driven by big data. *Precambrian Research*, 399: 107235. doi: 10.1016/j.precamres.2023.107235
- Zhang, Z. J., Yang, Z. X., Jian, F. Y., et al., 2025. Interpretability-Enhanced Mineral Prospectivity Models: A Synergistic Approach Using Large Language Models, Knowledge Graphs, and Machine Learning. *Mathematical Geosciences*: doi: 10.1007/s11004-025-10231-3
- Zhao, M. L., Zhang, Z. J., Yang, J., et al., 2025. Knowledge graph construction and knowledge discovery for porphyry copper deposits. *Ore Geology Reviews*: 106875. doi: 10.1016/j.oregeorev.2025.106875
- Zhao, P. D., 2019. Characteristics and Rational Utilization of Geological Big Data. *Earth Science Frontiers*, 26(4): 1-5. (in Chinese with English abstract)
- Zhou, C. H., Wang, H., Wang, C. S., et al., 2021. Geoscience Knowledge Graph Research in the Big Data Era. *Science China Earth Sciences*, 51(7): 1070-1079. (in Chinese with English abstract)
- Zhou, Y. Z., Zhang, Q. L., Huang, Y. J., et al., 2021. Constructing Knowledge Graph for the Porphyry Copper Deposit in the Qingzhou - Hangzhou Bay Area: Insight into Knowledge Graph Based Mineral Resource Prediction and Evaluation. *Earth Science Frontiers*, 28(3): 67-75. (in Chinese with English abstract)
- Zhou, Y. Z., Zuo, R. G., Liu, G., et al., 2021b. The Great-Leap-Forward Development of Mathematical Geoscience During 2010 - 2019: Big Data and Artificial Intelligence Algorithm Are Changing Mathematical Geoscience. *Bulletin of Mineralogy, Petrology and Geochemistry*, 40(3): 556-573. (in Chinese with English abstract)
- Zhu, Y. Q., Sun, K., Li, W. R., et al., 2023. Comparative Analysis and Enlightenment of Geoscience Knowledge Graphs: A Perspective of Construction Methods and Contents. *Geological Journal of China Universities*, 29(3):382 - 394. doi: 10.16108/j.issn1006-7493.2021111. (in Chinese with English abstract)
- Zou, Y. H., Liu, Y., Dai, T. G., et al., 2017. Finite difference modeling of metallogenic processes in the Hutouya Pb-Zn deposit, Qinghai, China: Implications for hydrothermal mineralization. *Ore Geology Reviews*, 91: 463-476. doi: 10.1016/j.oregeorev.2017.09.008
- Zuo, R. G., Cheng, Q. M., Xu, Y., et al., 2024. Explainable Artificial Intelligence Models for Mineral Prospectivity Mapping. *Science China Earth Sciences*, 67(9): 2864-2875. (in Chinese with English abstract)

中文参考文献

- 安文通, 陈建平, 朱鹏飞, 2021. 基于成矿过程数值模拟的隐伏矿双向预测研究. 地学前缘, 28 (3): 97-111.
- 陈建平, 陈勇, 王全明, 2008. 基于 GIS 的多元信息成矿预测研究——以赤峰地区为例. 地学前缘, 15 (4): 18-26.
- 成秋明, 2006. 非线性成矿预测理论: 多重分形奇异性-广义自相似性-分形谱系模型与方法. 地球科学, 31 (3): 337-348.
- 成秋明, 2025. 面向人类智能与人工智能融合的矿产资源预测新范式. 地学前缘, 32 (4): 1-19.
- 董新丰, 甘甫平, 李娜, 等, 2020. 高分五号高光谱影像矿物精细识别. 遥感学报, 24 (4): 454-

- 冯婷婷, 蔡诗柔, 张振杰, 2025. 基于知识图谱的碳酸岩型稀土矿成矿要素挖掘. 地学前缘, 32 (4): 262-279.
- 干勇, 彭苏萍, 毛景文, 等, 2022. 我国关键矿产及其材料产业供应链高质量发展战略研究. 中国工程科学, 24 (3): 1-9.
- 师路易, 左仁广, 2025. 矿产预测大模型. 地球科学: doi: 10.3799/dqkx.2025.3190.
- 王成彬, 王明果, 王博, 等, 2024. 融合知识图谱的矿产资源定量预测. 地学前缘, 31 (4): 26-36.
- 王学求, 2013. 勘查地球化学 80 年来重大事件回顾. 中国地质, 40 (1): 322-330.
- 肖克炎, 樊铭静, 孙莉, 等, 2023. 矿床成矿系列综合信息预测理论方法及其应用. 地球学报, 44 (5): 769-780.
- 袁峰, 李晓晖, 胡训宇, 李跃, 贾蔡, 2019. 热液矿床成矿作用研究新途径: 数值模拟. 地质科学, 54 (3): 678-690.
- 张颖慧, 王涛, 焦守涛, 等, 2020. 国内外岩浆岩数据库现状与应用前景. 高校地质学报, 26 (1): 11-26.
- 张振杰, 成秋明, 杨玠, 等, 2021. 机器学习与成矿预测: 以闽西南铁多金属矿预测为例. 地学前缘, 28 (3): 221-235.
- 赵鹏大, 2019. 地质大数据特点及其合理开发利用. 地学前缘, 26 (4): 1-5.
- 周成虎, 王华, 王成善, 等, 2021. 大数据时代的地学知识图谱研究. 中国科学: 地球科学, 51 (7): 1070-1079.
- 周永章, 张前龙, 黄永健, 等, 2021a. 钦杭成矿带斑岩铜矿知识图谱构建及应用展望. 地学前缘, 28 (3): 67-75.
- 周永章, 左仁广, 刘刚, 等, 2021b. 数字地球科学跨越发展的十年: 大数据, 人工智能算法正在改变地质学. 矿物岩石地球化学通报, 40 (3): 556-573.
- 诸云强, 孙凯, 李威蓉, 等, 2023. 地球科学知识图谱比较分析与启示: 构建方法与内容视角. 高校地质学报, 29 (3): 382-394.
- 左仁广, 成秋明, 许莹, 等, 2024. 可解释性矿产预测人工智能模型. 中国科学: 地球科学, 54 (9): 2917-2928.