

地球观测知识枢纽：内涵、关键技术与展望

赵利民^{1,2,3}, 苗晨⁴, 邢进⁵, 李国庆^{1,6}, 侯宇葵⁷, 刘闯⁸,
李家国^{1,2,3}, 陈兴峰^{1,2,3}, 刘军^{1,2}, 杨健^{1,2,3}, 周翔^{1,2}, 顾行发^{1,2,3}

1. 中国科学院空天信息创新研究院, 北京 100094;
2. 遥感卫星应用国家工程研究中心, 北京 100094;
3. 国家航天局航天遥感论证中心, 北京 100094;
4. 国家遥感中心, 北京 100036;
5. 国防科工局重大专项工程中心, 北京 100101;
6. 国家对地观测科学数据中心, 北京 100094;
7. 中国空间技术研究院钱学森空间技术实验室, 北京 100094;
8. 中国科学院地理科学与资源研究所, 北京 100101

摘要: 智能化地共享和复用地球观测应用实践所形成的知识, 有序促进应用要素的高通量流动, 是激活和释放地球观测系统效能的关键。地球观测知识枢纽EOKH (Earth Observation Knowledge Hub) 将分散的应用要素有机耦合, 形成专题知识集合的协同设计与吞吐能力, 是地球观测成果治理与智能服务的研究前沿。针对地球观测组织GEO (Group on Earth Observations) 发展GEO知识枢纽GKH (GEO Knowledge Hub) 得到的启示, 结合地球观测知识本体研究最新进展, 系统解析了EOKH的内涵与特征, 梳理了发展EOKH的技术需求与挑战, 分析存在的关键问题及其应对策略。在此基础上, 提出了EOKH系统架构原型, 结合地球观测知识包转移复用和知识推理两个EOKH典型应用场景进行了原理实现与概念验证。最后, 就如何促进地球观测知识“人—机—物”要素的协同共生, 提出了中国发展EOKH的相关认识与思考。探索“人在回路”的高通量知识协同生产与转移技术, 以及“知识驱动”的成果复用技术, 是提升地球观测隐性知识可解释性、促进EOKH服务质量与发展力的必要举措。

关键词: 遥感, 知识枢纽, 地球观测, 本体, 知识包, 知识复用, 人—机—物, 人在回路

中图分类号: P2

引用格式: 赵利民, 苗晨, 邢进, 李国庆, 侯宇葵, 刘闯, 李家国, 陈兴峰, 刘军, 杨健, 周翔, 顾行发. 2024. 地球观测知识枢纽: 内涵、关键技术与展望. 遥感学报, 28(6): 1375–1394

Zhao L M, Miao C, Xing J, Li G Q, Hou Y K, Liu C, Li J G, Chen X F, Liu J, Yang J, Zhou X and Gu X F. 2024. Earth observation knowledge hub: Implications, key technologies and perspectives. National Remote Sensing Bulletin, 28(6): 1375–1394 [DOI:10.11834/jrs.20222302]

1 引言

在“高分辨率对地观测系统重大专项”和《国家民用空间基础设施中长期发展规划（2015年—2025年）》等国家重大科技任务的推进下, 中国地球观测应用已融入国民经济和社会发展多领域（国家航天局, 2019）。长期应用实践积累了大量观测数据、软件代码、发表物等成文形式的“显

性知识”, 也产生诸多技能、思想、经验等认知形式的“隐性知识”（附表A1）。“显性知识”可通过开放共享等手段网络化传播, 而“隐性知识”则需在教育、实践等方面付出较大时间和资源成本, 传播难度大（Bathelt等, 2004）。在地球观测初期阶段, 应用效益的增长主要取决于以观测数据为核心的“显性知识”质量及其开放程度（Wulder和Coops, 2014; Zhu等, 2019; 李国庆等, 2016);

收稿日期: 2022-06-20; 预印本: 2022-11-14

基金项目: 国家重点研发计划(编号:2019YFE026600); 航天五院CAST创新基金项目(编号:ZC-1050-2022-11-005)

第一作者简介: 赵利民, 研究方向为光学遥感卫星应用工程。E-mail: zhaolm@aircas.ac.cn

通信作者简介: 顾行发, 研究方向为遥感定标与真实性检验。E-mail: guxf@aircas.ac.cn

随着地球观测体系的发展,应用技术的复杂度、集成度提升,应用要素开始向局部集中,催生了数字鸿沟,降低数据红利(Esch等,2018; Klein等,2017),也令思想、经验等“隐性知识”的影响凸显(Nativi等,2015;何国金等,2018; Bloom等,2020)。新知识的涌现还容易造成用户理解新观点所需的认知空缺,产生技术上的“马太效应”(Chu和Evans,2021)。知识的流通障碍为综合性地球观测应用的推广带来全方位挑战。

知识经济的竞争优势主要源于知识、特别是隐性知识转移的流量(Howells,2002; Bathelt等,2004)。有观点认为,信息技术领域的重大变迁存在“20年周期律”(梅宏等,2022),着重整体任务吞吐率而非单项技术复杂度的“李特尔定律”(Little's Law)有望成为未来信息系统的重要指导原理(Little,2011)。当前,全球地球观测体系正由数据共享向开放科学(Directorate-General for Research and Innovation (European Commission),2018)和开放知识(GEO Secretariat,2021a)迈进,“人在回路”(Humans in the loop)的利用知识,将人类智能与机器智能形成迭代系统成为数据科学与智能服务的研究前沿(Esch等,2020;程学旗等,2020,2022)。知识枢纽作为知识生产与共享网络中的中转与代理节点,能够将分散的知识库耦合在一起,通过人机协同形成强大的知识共享能力(Evers等,2010)。基于知识枢纽的地球观测成果高通量转移复用,为协同群体智能、消弭数字鸿沟和促进地球观测的信息化带来新思路。

在地球观测领域,中国逐渐形成了基础设施强、应用场景多、创新要素多、从业人员多和市场规模大等特色,具备群体协同促进知识共享的内生动力和优势条件。作为地球观测组织GEO(Group on Earth Observations)的4个联合主席国之一,中国也有职责和义务协调GEO及区域地球观测体系的协同发展,为消弭数字鸿沟做出大国贡献。设计并发展地球观测知识枢纽EOKH(Earth Observation Knowledge Hub),对国内国际相互促进加速应用要素流转效率,进一步激活和释放地球观测系统效能具有重要意义。

2 地球观测知识枢纽内涵与发展启示

2.1 内涵与特征

EOKH是知识共享网络中高通量地吞吐可信地

球观测知识的交汇节点,为应用系统发挥推动赋能(Enabler)作用。其核心任务是及时发现和转移有价值、有潜力的研究,促进应用要素高通量流动。主要功能包括:规模获取和甄别可重现的地球观测知识,测试地球观测知识的可复制性,以及将地球观测知识快速转移到新的应用场景,实现知识在不同人或机构间的复用。EOKH的治理对象是地球观测“知识包”,其含义为针对地表特定科学问题或应用需求的观测手段、研究数据、算法模型、技术流程等应用要素及其概念、关系、规则和过程的集合,是显性知识和隐性知识的综合体。EOKH的用户重点面向开发地球观测应用专业技术人员和培训人员,旨在汇聚应用资源,提升开发效率,降低应用门槛。

EOKH具有连通性、高通量和轻量化3方面典型特征。首先EOKH是本地化的、领域性的,意味着它不是一个通用技术平台,而是由多个不同领域知识枢纽节点构成的知识共享网络。每个节点拥有1个或多个本地化知识库,节点内部以及节点之间具有高度的信息连通性和网络化水平,将分散的知识库动态耦合,形成针对问题或主题的全景应用要素连结能力。其次,基于EOKH的知识共享是高通量的,它从本体的角度识别与表述知识资源、刻画需求,自动化地对应用要素进行组合和筛选,实现“知识包”的整体吞吐,形成高效的资源转移能力。最后,由于需支撑知识的测试与扩展复用,EOKH具备一定的应用设计与开发能力,但从存储计算角度看,EOKH是轻量化的:其主要操作地球观测知识本体,较少涉及数据实体的处理,因此并不适用于构建业务化地球观测应用产品生产平台。

2.2 国内外发展现状

GEO最早于2019年提出设计GEO知识枢纽GKH(GEO Knowledge Hub)(GEO Secretariat,2019),定位为一种管理编码化地球观测知识的中央数字档案馆,系统结构如图1所示。通过开发GKH,GEO试图协调全球综合地球观测系统GEOSS(Global Earth Observation System of Systems)汇集的海量地球观测成果,打通从地球观测能力到决策分析的无障碍信息通道(GEO Secretariat,2020a,2020b)。GKH的主要功能是有序策展特定地球观测应用关联要素,提供构建应用所需的可获取知识。由于

知识要素能够借助语义机制实现有机动态关联，使得GKH理论上具备在开放环境下协同设计、生产和转移知识的能力。为消除技术锁定，GKH特别强调在开放的科学环境中管理地球观测应用资源，以及在云平台上检索和分析地球观测大数据。目前，GKH已成为GEOSS基础设施开发的基本任务

之一，也是GEO设施发展工作组2020年—2022年的重点任务。借助欧洲核子中心CERN（法语：Conseil Européen pour la Recherche Nucléaire）开发的开源数据管理框架InvenioRDM，GKH已初步实现知识检索，并开发了若干专题知识包（gkhub.earthobservations.org [2022-05-01]）。

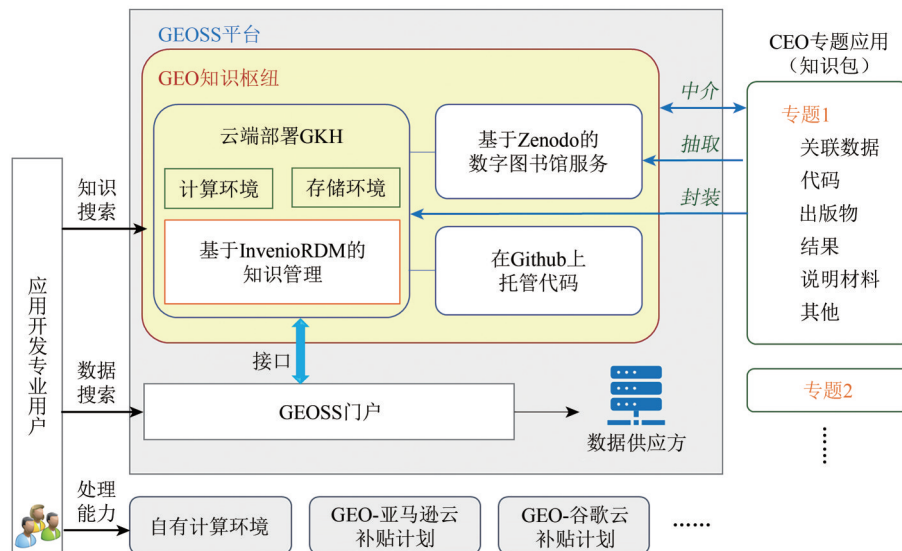


图1 GEO知识枢纽系统结构

Fig. 1 System structure GEO knowledge hub

类似地，为促进地球观测知识共享，在“地平线2020”科研规划（Horizon 2020）的支持下，欧盟研建了服务人类健康与生态系统研究的知识枢纽平台GOS4M-KH（GOS4M Knowledge Hub）（De Simone等，2021），主要以生物地球化学模型驱动的交互式地球观测应用为特色；美国国家航空航天局NASA（National Aeronautics and Space Administration）正建设NEX（NASA Earth Exchange）平台（www.nasa.gov/nex[2022-05-01]），拟通过同步提供数据、工具软件和超算资源，实现科学合作与知识共享；德国航空航天中心提出了U-TEP平台（Urban Thematic Exploitation Platform）（Esch等，2020），用以共享和复用地球观测数据应用于城市专题研究的解决方案；欧洲航天局ESA（European Space Agency）联合NASA推出MAAP平台（Multi-Mission Algorithm and Analysis Platform）（Albinet等，2019），为ESA和NASA当前及未来卫星计划提供开放协作的地球观测数据、产品、算法、工具等服务；世界气象组织开发的OSCAR（Observing Systems Capability Analysis and Review

Tool）（space.oscar.wmo.int[2022-05-01]）用于发现和分析地球观测需求、变量与遥感器之间的关联，为地球观测应用与系统的协调发展提供定量判据。中国地球大数据科学工程数据共享服务系统（data.casearth.cn[2022-05-01]）致力于汇集、管理和共享地球大数据，形成地球大数据学科驱动平台；国家对地观测科学数据中心（www.noda.org.cn[2022-05-01]）通过科学数据、科技项目的汇交，实现了地球观测科学数据及信息产品的发现与出版；全球变化科学研究数据出版系统（《全球变化数据学报》（中英文）编辑部，2017）也实现科学研究数据的元数据、实体数据、数据论文关联一体出版；开源空间信息基金会中国中心（osgeo.cn[2022-05-01]）主要关注以开源地理信息与遥感软件为主的地理空间信息科学技术、数据和知识的共享。这些地球观测知识共享平台的共性特征是重点面向科学研究用户，强调静态知识要素开放获取；对于知识集合的动态组织管理、系统性转移和交互式复用较少涉及，缺乏知识生产者与用户的互动，人的判断决策等隐性作用在

知识共享链路中难以体现, 尚未能体现知识共享的信息连通、高通量转移和“人在回路”等设计思想。

2.3 问题、需求与挑战

(1) 存在问题。目前公开报道的地球观测知识枢纽以GKH为代表。试运行一年以来, GKH遇到3方面问题 (GEO Secretariat, 2021b): 首先, 开发团队需要付出大量额外的努力用以定制和维护GKH运行所依赖的开源工具InvenioRDM, GKH的全面发展也受到软件版本替换的影响; 其次, 从GEO活动和工作计划中获取GKH所需的知识文档需要知识提供者更强的互动和积极性; 另外, GEO秘书处官员不得不掌握多个领域的专业知识 (如数据与知识基础设施, 云计算, 分析就绪数据, 数据标准等), 才有可能指导成员和相关工作从最新技术进步中获益。除了上述情况, GKH过于追求“零下载”而将数据资产全部托付非受控第三方进行云端操作, 或将产生3方面的知识伦理问题: 1) 所有权和安全性, 主要表现为科学数据和软件“上云容易下云难”, 知识及用户活动存在被云平台未经许可地利用的可能; 2) 开放性和透明性, 主要表现为云服务商变更难, 云平台所提供的人工智能服务的可解释性与准确性不透明, 云端存储和利用知识的社会成本不清 (如云计算所产生的碳排放量); 3) 数据殖民主义, 主要表现为具有优势地位云服务平台或将形成新的技术和资源垄断, 影响知识提供者对知识的控制力, 进而影响知识创造和交互的活跃性。

(2) 中国发展EOKH的需求。中国发展EOKH, 既要发挥技术成果资源的发现、测试和复用功能, 还需无缝对接空间基础设施和地面应用系统, 推动地球观测应用资源数字化。为此, 要求EOKH能够关联数据获取、数据处理和综合分析等多种生产资料 and 工具, 全景呈现和复用针对需求或问题的科技资源, 如图2所示。为实现各类应用资源的智能连接, 需构建一套结构化的、动态关联的地球观测知识管理模型, 支撑基于语义网络的需求自动理解和知识关联推理; 为实现知识的测试复用, 需结合具体案例或场景, 构建应用技术流程的形式化规则模板, 支撑人机交互生成新的、可溯源和扩展的知识包。

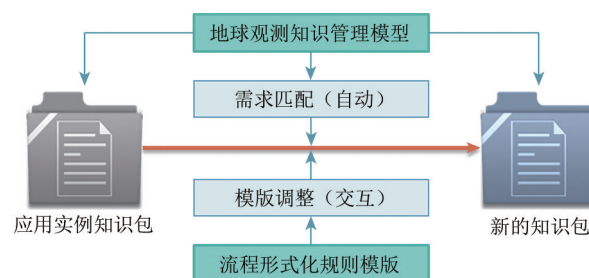


图2 EOKH基础性需求示意

Fig. 2 Demonstration of principal requirements for EOKH

为提升知识服务的可持续性, 结合GKH发展启示, EOKH还需满足以下几方面扩展性需求:

1) 智能化与可解释性结合, 提高知识价值禀赋。中国地球观测正朝天地一体化方向发展, 观测手段、研究数据、算法软件、技术流程都将成为构成高品质应用实体和EOKH知识管理的本体。由此导致知识要素异构多源, 知识关联网络复杂庞大, 迫切需要尽可能自动化地提取和转移知识资源, 提升知识利用效率。此外, 还需将人的智能与决策参与到地球观测知识管理活动中, 提升知识转移过程的透明度和知识复用的可追溯性, 以及对高阶复杂问题的可解释性, 并保证对抗模式下处理模型与方法的鲁棒性和免疫性 (Wing, 2021; Kleinberg 等, 2018)。

2) 开放性与有序性相结合, 提供好的应用体验。为支持协同解决地球观测重大科学问题与应用需求, EOKH需具备在开放协同的环境下桥接和互操作泛在应用成果的能力, 这就要求EOKH能够对天地一体化异构地球观测知识资源的合理有效抽象。为应对开放性带来的知识伦理问题, EOKH知识的转移复用需融合FAIR原则和CARE原则 (Carroll 等, 2021), 即一方面重视知识本体的可查找性 (Findable)、可访问性 (Accessible)、互操作性 (Interoperable) 和可复用性 (Reusable); 另一方面重视知识生产者的集体利益 (Collective Benefit)、控制知识的权力 (Authority to Control)、集体自决的责任 (Responsibility), 以及管理和使用知识时支持知识生产者权利和福祉的道德义务。

3) 基础性与领域性相结合, 支持多种任务类型。地球观测领域普遍存在“一星多用”和“多星一用”现象, 以及一个模型服务多个场景、一个场景使用多个模型的情形。为此, EOKH既需凝练应用共性以支持低代码甚至无代码化应用开发,

也需维护各种领域特性技术成果以适应动态多变的应用，支持专业到行业的嵌合。因此，EOKH不是“大一统”的通用标准或平台，而是一种整体共识与领域共识相结合的智能化应用服务理念，存在方式应是多样性的、可根据领域或场景定制的，实现特定的行业解答。

4) 轻量化与可信度相结合，提供高知识服务效率。EOKH知识管理的重点是对本体而非实体的操作，要求支持灵活多样的资源虚拟化与异构性桥接能力，支持跨层协同模式下的应用敏捷开发与运行支撑（徐志伟 等，2022），应对新要素、新知识涌现的常态。对EOKH知识库中的应用来说，可靠性和准确性是关键禀赋，需要人机结合的手段，对知识本体和实体的质量进行双重监控（陈兴峰 等，2021；杨倩倩 等，2022）。

(3) 挑战与机遇。利用计算机手段管理和复用复杂地球观测知识之所以存在困难，地球观测知识要素在“人—机—物”三元数据空间的复杂关联：“人”指数据处理分析、技术流程设计、判断决策等知识生产和利用的主体，也是知识转移的客体，构成社会空间；“机”指反映地表现象规律的地球观测数据与信息产品，以及承载显性知

识的技术资料，构成信息空间；“物”指感知、采集和处理地球观测数据的软硬件设备装置，构成物理空间。如何将异质异构“人—机—物”对象抽象成计算机可存储、运算的符号集，同时将人的决策融入知识的产生、管理和转移过程，是智能化共享和利用地球观测知识的核心挑战。

利用逻辑和工具回答因果问题是当前数据科学和人工智能变革的方向（Gill，2020）。其中，知识和数据双驱动的群体智能试图在可解释性与执行效率方面取得平衡，是智能技术的前沿（蒲志强 等，2022）。知识驱动（邢立宁和陈英武，2011）以特定应用场景的知识集合为输入，强调群体参与问题求解，“人在回路”地操作已有算法模型和经验。对比模型驱动、数据驱动和知识驱动3种研究模式（表1），不难发现，知识驱动继承了应用实践所形成的高价值知识，也融合了数据驱动的执行效率，同时在可解释性方面具有天然优势，有望改变“瀑布式”地球观测应用开发和“黑盒式”空间信息提取方式，这些潜质有望改进模型驱动和数据驱动等应用模式，尤其适合综合性地球观测应用实践成果的定量集成和转移复用。

表1 模型驱动、数据驱动、知识驱动3种地球观测应用模式特征对比

Table 1 Comparison of the characteristics of three earth observation application approaches: Model-driven, data-driven, and knowledge-driven

对比维度	模型驱动	数据驱动	知识驱动
技术特征	以观测数据为起点，“瀑布式”构建应用，过程透明，机理明确	以观测数据为起点，“黑盒式”挖掘参量关系，效率有意，精度卓越	以技术流程为起点，“人在回路式”操作已有知识，集中转移成果
	观测数据与目标参量因果关系明确	无需精确建模	多学科知识有序集成，群体智慧
优势	过程性数据完备，有利于推理溯源	算法通用性强，开源工具和样本多，技术门槛相对较低	过程可解释性、算法稳定性、最优性、收敛性强，执行效率高
	可解释性强，利用形成认知	可持续学习和进化	技术门槛低，方便知识共享
	受制于模型性能和模型使用者技能	忽视理论规律，可解释性差	知识获取代价高
劣势与典型困境	技术门槛高，推广难度大	受制于样本质量和算力	受制于知识可信度
	高耦合参量病态反演问题	恶意样本干扰问题	利益相关方共享知识的能动性

群体协同、知识驱动的地球观测成果复用技术理念如图3所示。其出发点是对知识包所涉研究数据、方法模型等显性知识在概念层面抽象和编码，建立知识本体模型，有机关联应用要素，实现知识的推理；在此基础上，对技术流程等隐性知识进行形式化表达，抽取变量操作、过程激发

的共性形式结构，结合本体模型实现知识要素的互操作。概念型知识的模型化，以及过程型知识的形式化，将是支撑EOKH系统性吞吐知识包的技术基础；知识驱动的林OKH也有益于构建智能协同地发现、转移和复用地球观测应用成果的交互式操作环境。

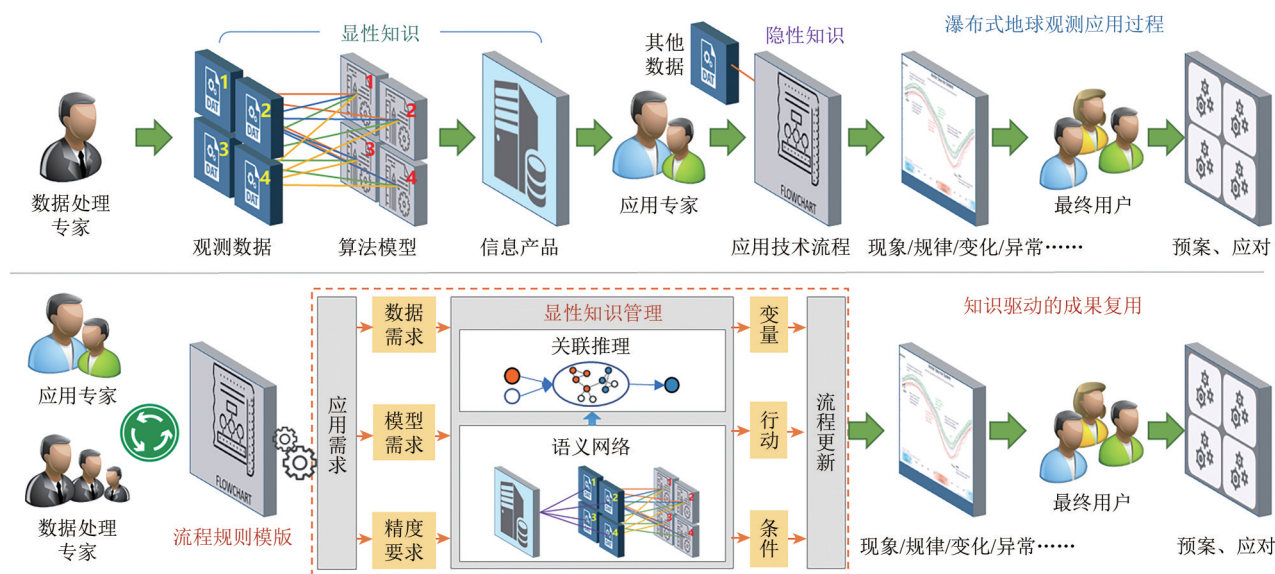


图3 知识驱动的地球观测成果复用概念示意

Fig. 3 Conceptual illustration of knowledge-driven approach for EO application reuse

3 关键技术问题及对策

为实现在开放协同环境下人在回路地操作和转移知识, EOKH将面临系统度量、知识转移、知识复用以及知识的探索与可视化4方面关键技术问题, 如图4所示。

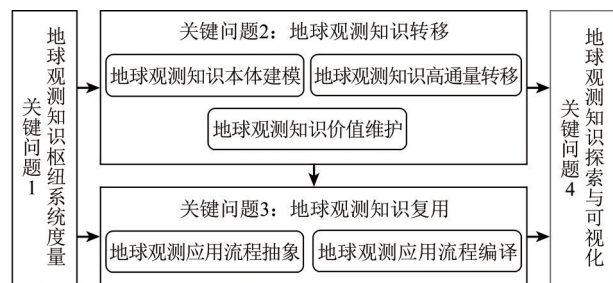


图4 地球观测知识枢纽关键技术问题分布

Fig. 4 Distribution of key technical issues for the EOKH

3.1 知识枢纽系统性能度量

EOKH是知识共享网络中有组织、系统性吞吐地球观测知识集合的新型技术, 有必要设定系统度量依据, 为目标导向、功效测定和迭代优化提供参考基准。主要约定度量单位和度量方法两方面规则。

(1) 度量单位。信息系统的性能一般通过数据或事务的吞吐量(如MFlops, 每秒事务数, 停机时间)来衡量, 对于以高通量吞吐地球观测应用知识为设计目标的EOKH, 可将其性能度量单元设定为反映应用案例要素有序集合的“知识包”。

由于存在包内和包间知识要素通过某种机制动态关联形成新知识包的可能, 将为EOKH知识存量的计量带来混乱。这种情况下, 考虑到EOKH的核心竞争优势源于隐性知识的转移能力, 而隐性知识主要体现为围绕技术流程的要素组织方式, 因此, 将拥有相同流程结构的知识包判定为一个计量单元。

(2) 度量方法。EOKH性能度量判据可约定为单位时间单个用户完成知识转移的任务数, 包含知识生产者行为、知识留存情况、知识包转移通量3类指标。知识生产者行为主要指单位时间内单个用户贡献知识包数量, 以及单个知识包使用频次等。知识留存情况可考虑通过知识包数量、算法脚本和软件模块的数量, 以及与应用相关的研究数据集数量等指标来评估。知识包转移通量可通过计算单位时间内完成保质任务的数量来测定。

3.2 地球观测知识转移

根据知识转移假说(Evers等, 2010), 知识对生产越重要, 知识的复杂度越高, 相关生产的本地化程度越突出, 为开放协同地利用知识带来障碍。其中的技术因素主要涉及地球观测知识本体建模、高通量处理、价值维护等3方面问题。

(1) 地球观测知识本体建模。主要指构建地球观测知识在计算机中的表达形式和运算机制。本体建模既关乎知识的管理, 也关乎知识的抽取和利用, 是影响EOKH效率的关键, 核心在于表

达对象和表达手段两方面因素。

1) 表达对象。地球观测知识存在于“人—机—物”融合的社会空间、信息空间和物理空间，是实体、概念、关系、规则和过程的集合。这些应用要素在技术流程的牵引下有序组织构成一个

“知识包”（图5），人机交互的枢纽是地球观测信息产品。如何以信息产品为核心，将异质、异构、关联和隐性的地球观测知识以统一的符号化手段予以公开描述，是实现知识交互融通、推动隐性知识向显性转化的前提（Bohle, 2014）。

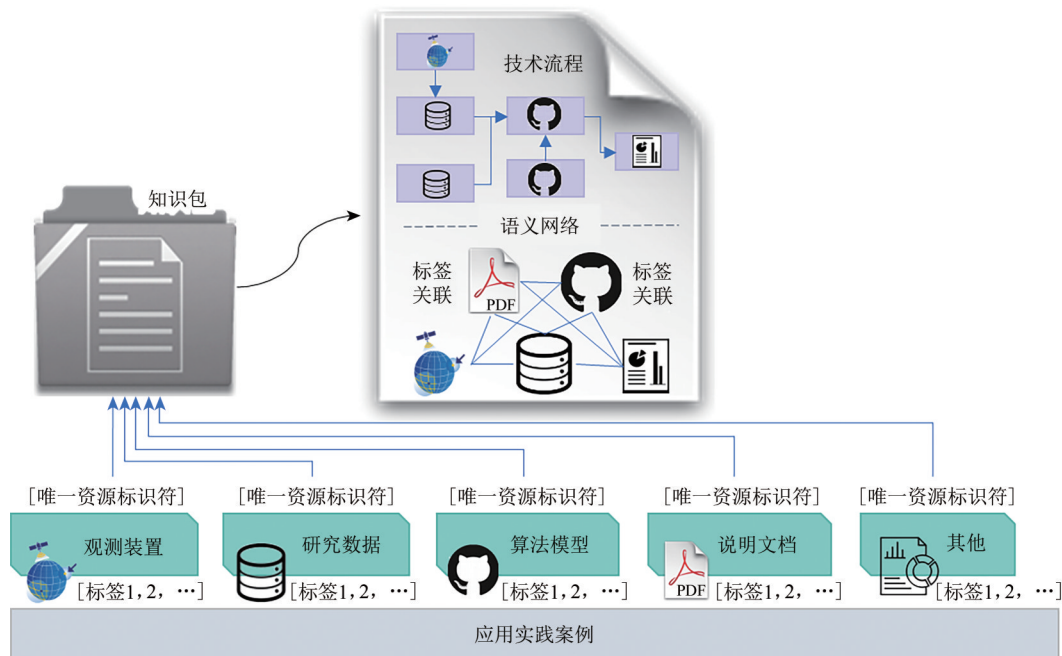


图5 地球观测知识包：目标一致、主题关联的应用要素有序集合

Fig. 5 EO knowledge package: An orderly collection of application elements with consistent objectives and related topics

2) 表达手段。为将地球观测知识人、机、物要素抽象成管理的对象客体 and 计算过程的执行主体，需探索将概念型与过程型相结合的知识表达方法，从而在智能化运行效率和可理解性方面取得平衡。本体研究（Ontology）致力于克服语义障碍所造成的互操作性问题，以结构化的“术语集”抽象表达不同类型知识实体及其关系，有助于形式化地说明地球观测“人—机—物”之间的内在联系，具有较强的可操作性（Abdalla 等，2015；Janowicz 等，2019；Sun 等，2019；Masmoudi 等，2020；Arvor 等，2019；Aldana-Martín 等，2022）。地球观测知识本体用以刻画知识包内各种实体的结构化属性特征，这些特征空间可构成元数据标签体系，通过关联模型与实体的通用资源标识符 URI（Uniform Resource Identifier）建立联系，实现基于语义标签的知识实体定位，例如针对哨兵卫星数据的地球观测本体（melodi.irit.fr/ontologies/eom [2022-05-01]），以及清华大学开发的通用性地球观测本体（github.com/THU-EarthInformation

ScienceLab/Ontology [2022-05-01]）等。本文认为地球观测知识的核心本体有两种：一种是空间信息本体，能够关联不同地球物理参量的时空属性，具有尺度特性；另一种是观测数据本体，能够关联不同观测手段。空间信息本体和观测数据本体通过规则（通常为算法或处理流程）建立联系，形成立体网状关联网络。考虑到地球观测应用基础性和领域性兼备的特性，以及实体及其关系的状态的动态变化现象，很难仅通过标准化的元数据标签体系定位知识实体。需同步开发上层本体和领域本体，建立实体 URI 与元数据动态关联模型，或构建 URI-元数据标准库来标定元数据的一致性。该标准库最好由行业组织维护，使元数据具有良好的扩展性。

(2) 地球观测知识高通量转移。主要指单位时间内吞吐知识包的任务量，与知识转移的数量、种类、系统效率等相关。主要涉及知识的有效获取和互操作等关键技术。

1) 知识抽取。规模获取高质量知识实体是

EOKH运转的前提。地球观测知识混合了概念、过程、控制等多种活动,需研究支持“人一机一物”多源异构知识的高效感知、采集、同构与融合方法,以尽可能自动地摄取知识资源,同时确保知识本体质量和实体可信度。专家系统(Aronson, 2003)从问题和需求出发,通过管理专家知识和经验实现决策辅助,是EOKH知识来源的有益借鉴。然而专家系统主要由专家自上而下地表达、获取和运用知识,知识的规模、质量都有限。知识图谱(Hogan等, 2022; Ji等, 2022)能够把各种结构化、非结构化、半结构化的数据以语义形式连接在一起,借助规则模板或机器学习手段,实现实体、属性、关系的识别和提取,适用于自下而上地大规模挖掘、抽取和补全知识。然而,全自动化地抽取也会令知识的质量和可信度失控,为此,可通过URI-元数据标准库的对标实现知识的质量控制,通过开放评审和定期核验机制确保知识实体可信。对于因大量知识提交导致评审人员不足或评审质量下降问题,可引入“保序机制”(Su, 2021),在激励知识生产者能动性同时增加评议结果的可靠性。需指出的是,知识图谱作为一项尚在发展中的研究,技术尚未见底或定型,成本效益矛盾突出(Paulheim, 2018),在EOKH的技术选型中需综合考虑。

此外还需关注技术流程等隐性知识的获取模式。这类知识的产生和利用受到知识生产者主观行为和外部政策环境等的影响。本文建议,首先由专家人工抽象构建隐性知识管理模型(Pathirage等, 2006),形成机构级的地球观测知识共享策机制,这也将成为EOKH冷启动的数据源;在此基础上,通过应用评价和知识出版机制吸引更多应用开发用户和知识生产者;利用对知识管理形成的经验,迭代优化地球观测知识本体模型,研发部署智能抽取技术和群智手段,形成转移隐性知识的数字化竞争力。

2) 互操作。如何将人的智慧与判断决策融入到知识共享链路,是提升EOKH可信度和竞争力的关键。互操作体现了两个或多个系统或系统组件交换并使用所交换信息的能力(IEEE Standards Association, 2011),能够实现知识或知识集合之间在语义层面的相互通信,确保可以直接使用既有资产,是提升技术流程等隐性地球观测知识的转移效率、优化应用体验的良好途径(Stacey和

Berry, 2018)。比较普适的方式是设计结构化的代码编辑工具编译流程,并通过程序接口调用已有资源。例如,谷歌地球引擎为遥感案例开发和数据、算法等资源调用提供了集成开发环境IDE(Gorelick等, 2017);利用Jupyter Notebook也能实现基于文档形式的地球观测技术流程交互式开发与计算(Rapiński等, 2019; Owusu等, 2022)。对于复杂、专业技术流程开发,还要求知识或知识集合之间在语义层面的相互通信,主要通过制定标准和通用接口实现,如开放地理数据互操作规范OpenGIS(Open Geodata Interoperation Specification)、地理空间数据抽象库GDAL(Geospatial Data Abstraction Library)等;也会通过定制专门的框架或适配系统实现,如观测数据模型(Horsburgh等, 2016)和网格化环境数据协作分析框架(Xu等, 2019)等。地球观测知识的互操作须在借鉴并运用已有规范或模型的基础上,研建采样系统、观测数据、信息产品、处理方法、技术流程等抽象本体间的通信协议。同时针对领域应用和特性工具,开发的泛在应用运行时(runtime),实现EOKH托管资源和非托管资源之间的互操作,方便在泛在环境下加载和管理由各种泛在应用。

(3) 地球观测知识价值维护。为给EOKH知识生产者和用户的互通提供权益保障,促进可持续知识共享生态的形成,维护和保障地球观测知识的价值是必要前提。主要包含两方面内容,一是知识的共享许可类型,二是产权保障手段。

共享许可与知识资源的获取渠道有关。可通过5种手段汇聚成果:1)讨论组或社区级自发征集;2)国家或机构的科技专项要求提交共享的成果;3)在线人工注册;4)在线交互开发;5)在线自动抽取。为此,须支持用户根据需求选择不同共享许可,以在指定范围使用不同资源,如CC-BY(全面许可,只要求引用)、CC-BY-SA(要求公开共享所有衍生作品)、CC-BY-NC(禁止商业化)和CC-BY-SA-NC(结合前两种限制)。

产权保障是为了确保所有知识资源的提供者保留自己生产内容的知识产权,不会因知识转移而发生版权的自动变更。区块链技术通过建立网络环境下信任机制,在保障信息价值转移透明度、安全性和降低交易运营成本方面极具潜力。利用区块链技术维护科学数据等显性地球观测知识的产权是稳健有效的(朱长青等, 2021),但对于技

术流程等由多个显性要素有机组合的联合体，其隐性价值的数字版权维护仍存一定挑战性。当前区块链正经历从1.0到3.0的3个平行发展阶段（清华大学人工智能研究院等，2020），针对每个阶段特征和地球观测知识维护需求，可考虑从以下3方面构建EOKH知识转移过程中的产权保障策略：1）利用块链式数据结构，记录不同主体在不同时间所加工地球观测数据的深度，实现问题解决全过程的可信公布，这也是分布式协同工作的底层保障；2）利用智能合约手段，对地球观测应用知识的复用状态、执行情况进行监视，将知识的调用和转移及时自动化存证，管控数字化知识资产的转移“不作恶”；3）利用超级账本和分片技术，链接、验证地球观测应用实践所形成技术流程或知识要素联合体，实现数据资产、方法过程、工具模块的组合与协调等活动的唯一性资源标识，形成知识互操作的底层通讯协议。

3.3 地球观测知识复用

随着地球观测在不同应用领域的泛化，业务专家更需专注本领域业务逻辑的开发维护，对地球观测知识的复用提出新的要求：一是能够低代码甚至无代码化的方式编译和测试现有知识，藉此开发新的应用；二是知识的演替需具备环境感知、自我管理、群体协同的能力，降低特定场景和领域的应用门槛。为此，需探索地球观测应用过程的形式化描述方法，对过程各环节控制要素和执行方向进行抽象，等效表示应用技术流程的静态和动态特性，形成一组形式化、交互式的规则模板，实现隐性知识的符号化表达，并保证过程输入输出特性的完整性；在此基础上，结合知识图谱技术，抽取和推理应用过程各环节适配的知识要素，实现以技术流程为包络的地球观测知识生成、管理和复用。这一知识复用方案主要涉及两方面技术要点：

（1）地球观测应用技术流程抽象与自适应扩展。应用场景的变化会要求既有流程及其关联要素做适应性调整，如何让计算机理解并执行这一需求是地球观测知识复用的首要问题。可通过抽取技术流程中要素联结、转移方式的共性形式结构，将技术流程的逻辑规律和对象抽象为变量、行动、激发条件等形式化约规，同时为人机交互

修订技术流程提供接口。在此基础上，通过流程的嵌套与继承，结合知识实体的唯一资源标识和知识图谱支持下的资源推理推荐技术，实现在已有流程基础上，按需将其他语义关联的科学数据、算法工具等显性知识组合为新的流程，从而派生新的应用。

（2）地球观测应用技术流程计算效率优化。针对存在嵌套知识的复杂地球观测应用流程编译效率问题，需研究在互操作通信协议或运行时基础上，建立融合批处理、流处理、图处理等计算模式的处理框架，例如通过语义驱动自动甄别可并发单元，根据计算资源负载情况同时执行多个上下文无关联的知识，提升复杂隐性地球观测知识利用效率；此外，流程编制过程中强化部分规约，减少过度开发，提升业务体的自适应性。

3.4 地球观测知识探索与可视化

在地球观测应用决策中，不仅要知道某种原因能够造成怎样的结果，还要能够根据结果反推原因，进行归因分析。这为在EOKH中探索和展示知识提出新的要求，即如何用可视化的方式，对问题和需求的产生、解决问题所形成知识的演变，以及预设前提下对问题造成的影响进行交互式图形化表达。为此，可研究通过透明可测的方式对因果假设进行编码，构建探索式规划工具，交互展示技术演进脉络，评估新技术、新应用对体系发展的影响，计算反事实并发现引发效应的原因。

4 原型架构与概念验证

4.1 总体架构

EOKH的架构以管理显性知识和利用隐性知识为主线，以整体转移应用实践形成的案例为目标，借鉴表象状态转移REST（Representational State Transfer）的系统设计理念（Leiserson等，2020），从资源要素角度观察知识枢纽。系统原型结构如图6所示。

根据地球观测知识产生到复用的信息流，上述结构可分为5个层次：（1）实践层，从地球观测应用活动中感知、萃取不同层次、不同复杂度的可信知识，并获得共享许可；（2）实例层，有序

整理不同领域案例所涉知识要素；(3) 本体层，构建概念性知识要素表征模型和过程性知识形式化模型；(4) 管理层，构建地球观测成果知识库，突破知识语义关联网络技术，打造知识资源管理中台；(5) 功能层，策展已有案例及所涉要素，关联推理知识要素，以及应用案例知识包复用。

上述信息流的核心要素是地球观测信息产品：一方面可通过信息产品连接显性知识语义网络和流程规则模板，形成结构一致、要素有序的知识包；另一方面，将应用需求转化为对地球观测信息产品的需求，从而由参考案例派生出新的知识包。

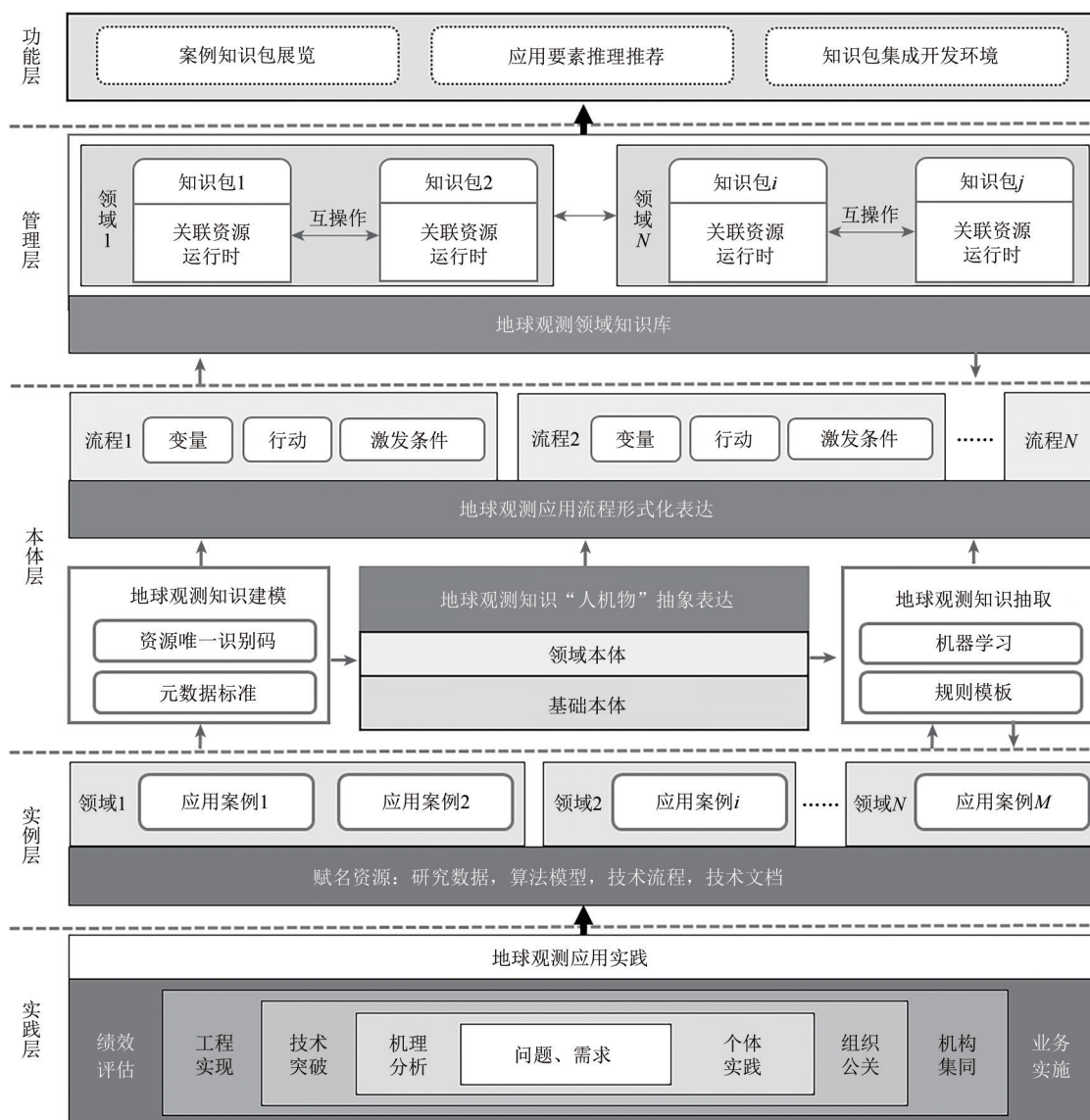


图6 地球观测知识枢纽系统结构

Fig. 6 A system prototype structure of the EOKH

4.2 技术原理实现

(1) 概念型地球观测知识本体建模。本体建模是计算机存储、运算和共享地球观测知识的前提。对于概念型知识，主要采用模型化和规范化手段，使异构异质地球观测知识在转换、传递过程中具备逻辑一致的本体结构，支撑知识包内、

包间语义关联和互操作。其中，模型化主要对知识包所涉载荷、数据、产品、方法等异构要素在概念层面进行抽象，转换成计算机可统一管理分析的数据结构，其核心要义是对异构地球观测知识分类分级以实现同构。概念型地球观测知识的主体是研究数据（即“人—机—物”中的信息空

间，包括输入输出数据、过程性数据以及真实性检验数据等信息要素），观测装置、处理工具等物理空间要素围绕信息产品的生产链路与之形成关联。信息空间要素以观测数据处理深度（顾行发和余涛，2018）为依据进行模型化处理，物理

空间要素以观测装置对电磁波信号的时空谱采样特性为依据进行模型化处理。以信息空间的模型化为例，从数据属性和对象属性两个维度对知识要素进行结构化表征，分类分级体系如图7所示。

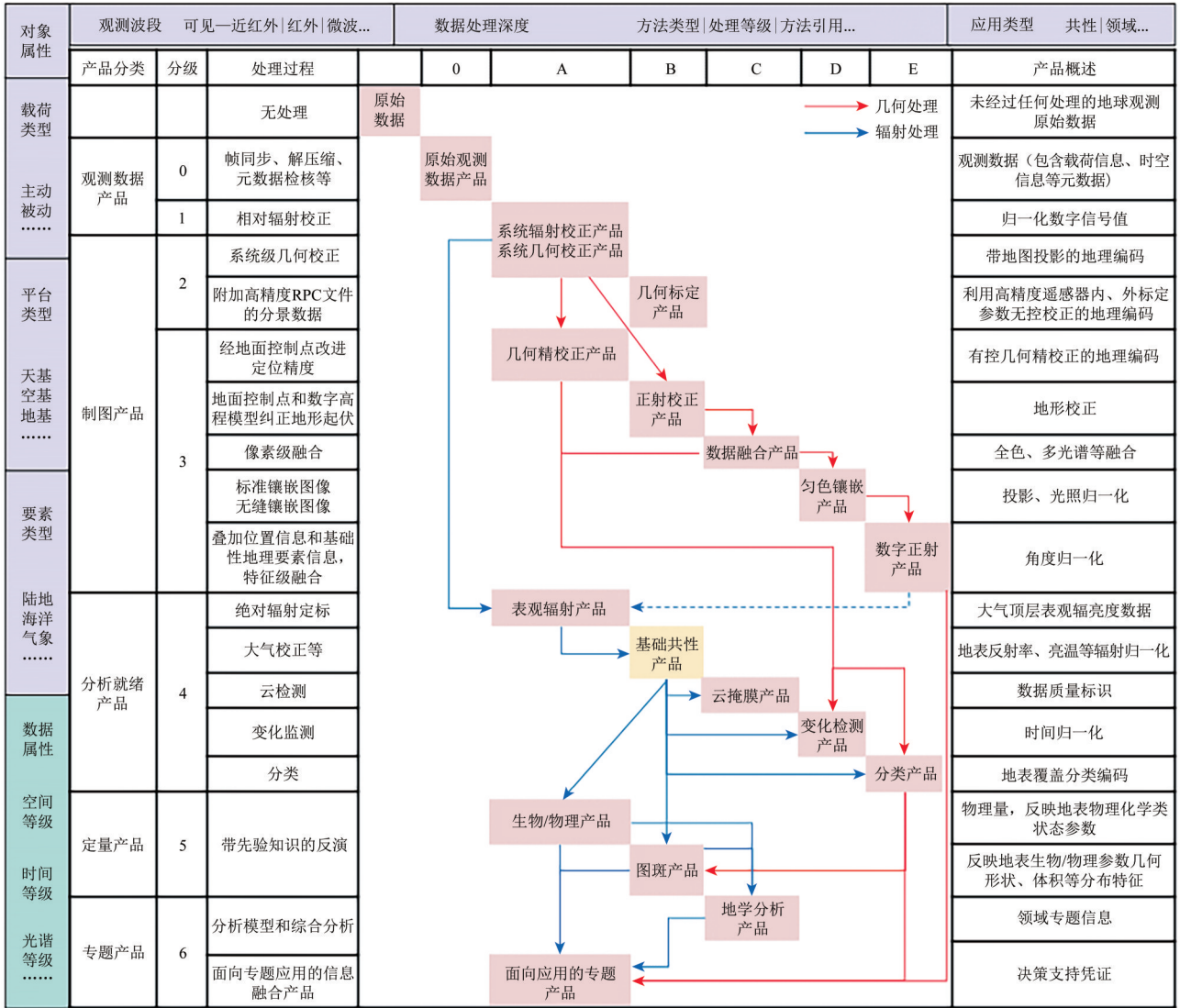


图7 概念型地球观测知识模型化依据:观测数据处理深度

Fig. 7 Basis for structuring conceptual Earth observation knowledge: Depth of observation data processing

规范化是在模型化的基础上，设计知识要素编码规则，建立语义层面结构统一的元数据规范和标签体系，规范数据从产生到应用的全过程，实现基于编码的地球观测知识有序组织及检索。具体以信息产品（共性产品、专题产品）为核心，建立多个标签与唯一标识符的动态映射关系，实现基于标签的知识实体的灵活定位，如图8所示。

围绕地球观测知识的产生、处理与应用等信息传递链路，对以研究数据为核心的概念型知识

进行模型化和规范化处理，设计出地球观测知识本体模型如图9所示（仅显示了基础本体的主体结构），把结构各异的地球观测知识转换成基于统一模板的、计算机可读写的数据库模型。最后利用资源描述框架RDF（Resource Description Framework）对标签关系建模，生成知识包与其组成部分（元素）间的关联关系，以及某一组成部分（元素）与其余对象间的关联关系，形成概念型知识语义关联网络，支撑要素推理和溯源。

共性产品（地球物理要素）信息表		专题产品信息表	
主键	地球物理要素唯一标识码	主键	专题产品唯一标识码
外键	观测手段标识码 处理方法标识码	外键	地球物理要素标识码 受益业务标识码 处理方法标识码
属性	要素名称 术语定义 要素类型标识码 计量单位 观测精度 误差单位 平面分辨率等级标识码 垂直分辨率等级标识码 覆盖范围等级标识码 采样周期等级标识码	属性	专题产品名称 术语定义 计量单位 观测精度 误差单位 平面分辨率等级标识码 垂直分辨率等级标识码 覆盖范围等级标识码 采样周期等级标识码

图8 地球观测知识规范化依据:以信息产品属性为索引进行统一编码

Fig. 8 Basis for Earth observation knowledge normalization: Uniform coding with information product attributes

(2) 过程型地球观测知识形式化表达。过程性知识主要通过转换地球观测信息，隐式地表达需求实现或问题求解的过程，存在于“人—机—物”数据空间的社会空间，可理解为技术流程，包括变量、算法、条件等要素。针对地球观测知识复用中技术流程等隐性知识的人机交互开发与管理问题，将地球观测数据处理过程抽象为算子

开发和流程调度两个环节：算子开发约定为变量和变量操作等形式，流程调度约定为激发条件（循环，返回）等形式。其中，变量通过语义标签与地球观测信息产品连接，变量与行动构成一个“输入—操作—输出”三要素组成的有向图。激发条件采用推理方向和调用模式两种形式，从而既可基于已有规则由变量或行动激发新行动并定义执行方向，也可人为修改规则实现过程的控制。通过上述约规，将应用技术流程转变为一组形式化、交互式的可执行规则模板。借助计算机图形技术，将规则模板可视化，如图10所示。图10中矩形代表一个算法节点，矩形上方空心圆是一个输入端口，代表算法一个输入张量；下部蓝色圆圈是一个输出端口，代表算法一个输出张量；箭头代表执行方向。利用多个节点以及连接节点的张量实现人在回路式的知识符号化表达与计算，支撑以应用流程为牵引的地球观测知识包交互式开发、管理和运行。

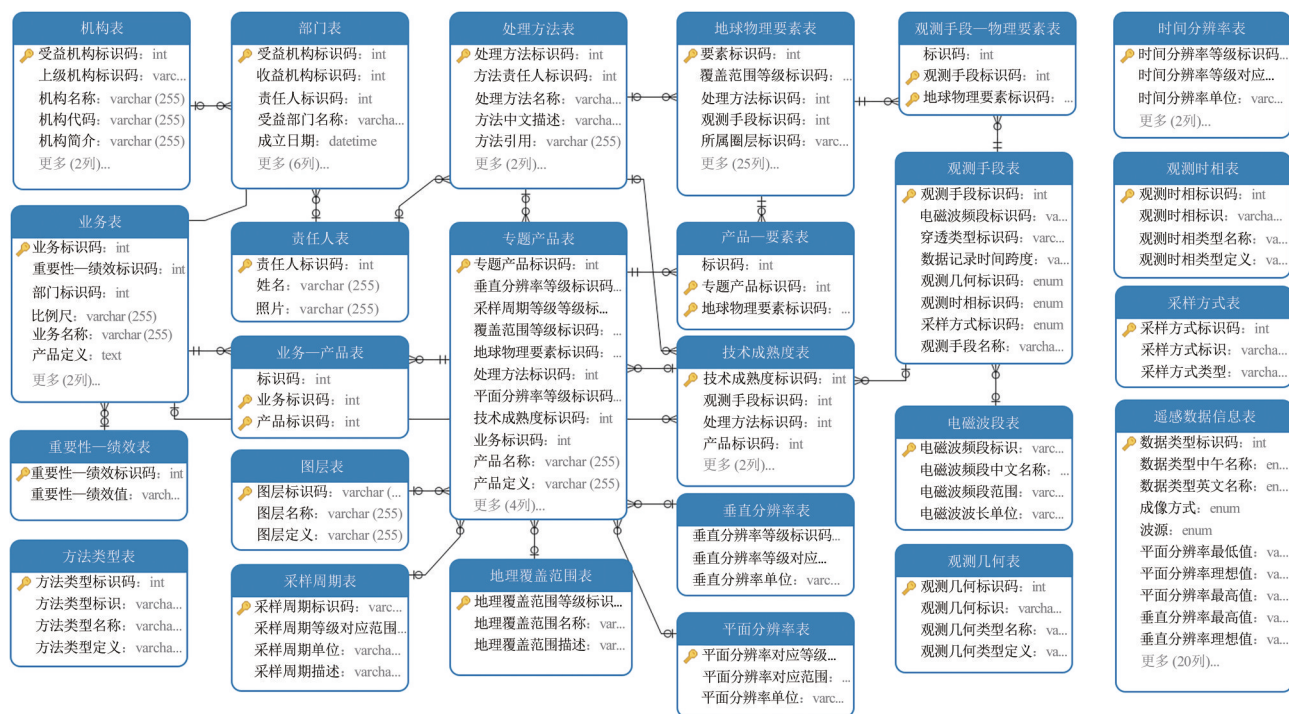


图9 地球观测知识本体模型结构(主体部分)

Fig. 9 An ontology model of conceptual Earth observation knowledge (Main body)

4.3 EOKH的典型应用场景

(1) 应用场景1：知识包转移复用。基于上述地球观测知识关联与技术流程形式化描述框架，对高分辨率光学遥感数据大气校正这一基础实例

进行了复用概念验证，如图11。

图11中A部分为利用葵花-8静止轨道卫星数据反演高时间分辨率气溶胶光学厚度(AOD)的知识包，该知识包已通过可靠性验证(Chen等，

2022)；图 11 中 B1 部分为基于辐射传输模型 6S 反演 GF-2 卫星 PMS 相机地表反射率的知识包，该过程需以 GF-2 同步 AOD 为输入，因此调用知识包 A，在大气校正流程中匹配共性产品，完成知识转移，实现 GF-2 卫星 4 m 多光谱地表反射率的自

动反演，经与地面同步真值对比，精度稳定可靠 (Liu 等, 2022)；图 11 中 B2 为基于知识包 B1 流程，通过匹配 GF-1 PMS 载荷参数和数据时空属性，完成知识包 B1 的整体复用，实现 GF-1 卫星 8 m 多光谱地表反射率的自动反演。

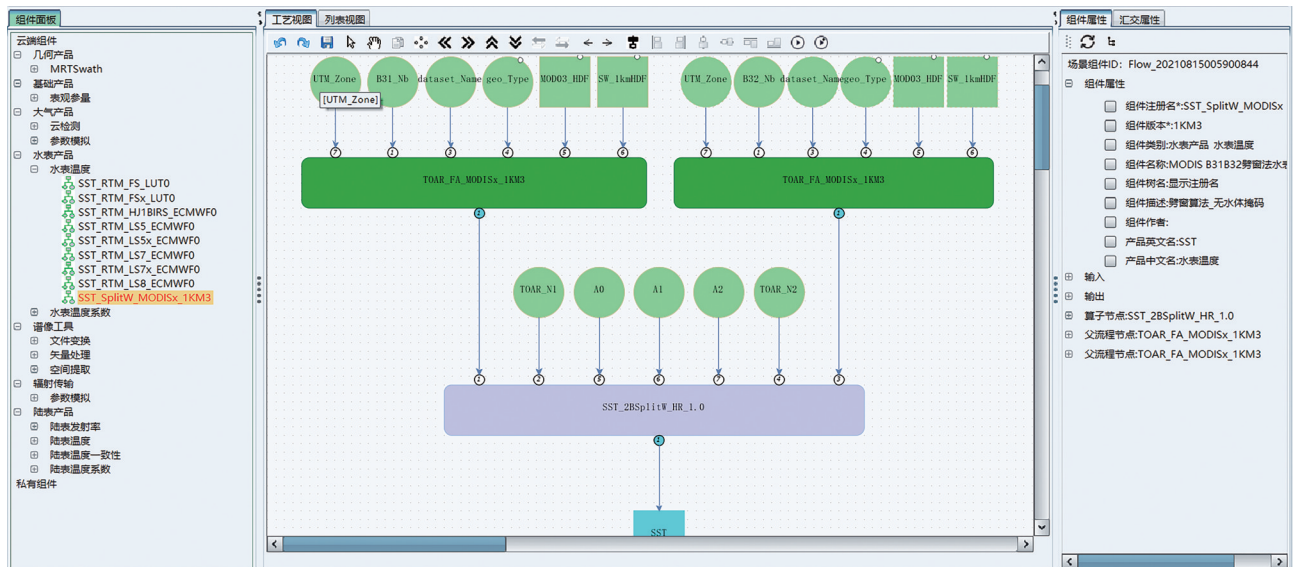


图 10 以流程为牵引的地球观测专题应用场交互式开发、管理和运行

Fig. 10 Interactive development, management, and operation of EO knowledge package driven by application process

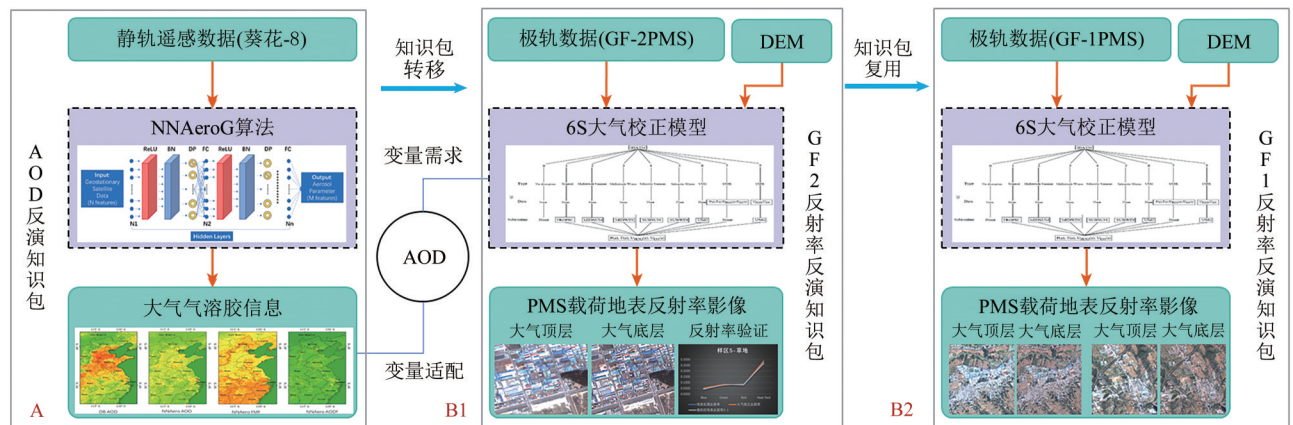


图 11 简单知识包转移、复用测试

Fig. 11 Proof of concept for transferring and reusing knowledge package

上述简单案例测试表明，知识包的转移复用即可极大提升地球观测应用系统开发和复用的便捷性。但对于更为复杂的综合性案例的复用，在领域知识本体建模、复杂流程表征与互操作等方面仍需进一步研究。

(2) 应用场景 2：知识推理溯源。利用概念型地球观测知识本体模型，将地球观测应用需求转化为信息产品的结构化需求，可支撑应用需求的自动推理，也能为地球观测系统的规划提供定量

论证数据。以重点湖库水质水华专题产品的生产需求为例，如图 12 所示，首先确定研究数据到观测装置各环节知识要素的对象属性及数据属性，形成映射关系，为需求的关联、统筹、去冗提供计算变量；其次将可形成能力的观测装置知识结构化表征，通过需求匹配，即可推理出针对需求的最佳观测手段；如果当前观测手段不满足，也可推理出为满足需求所需的新增观测能力，为有效应用和科学论证提供了定量判据。

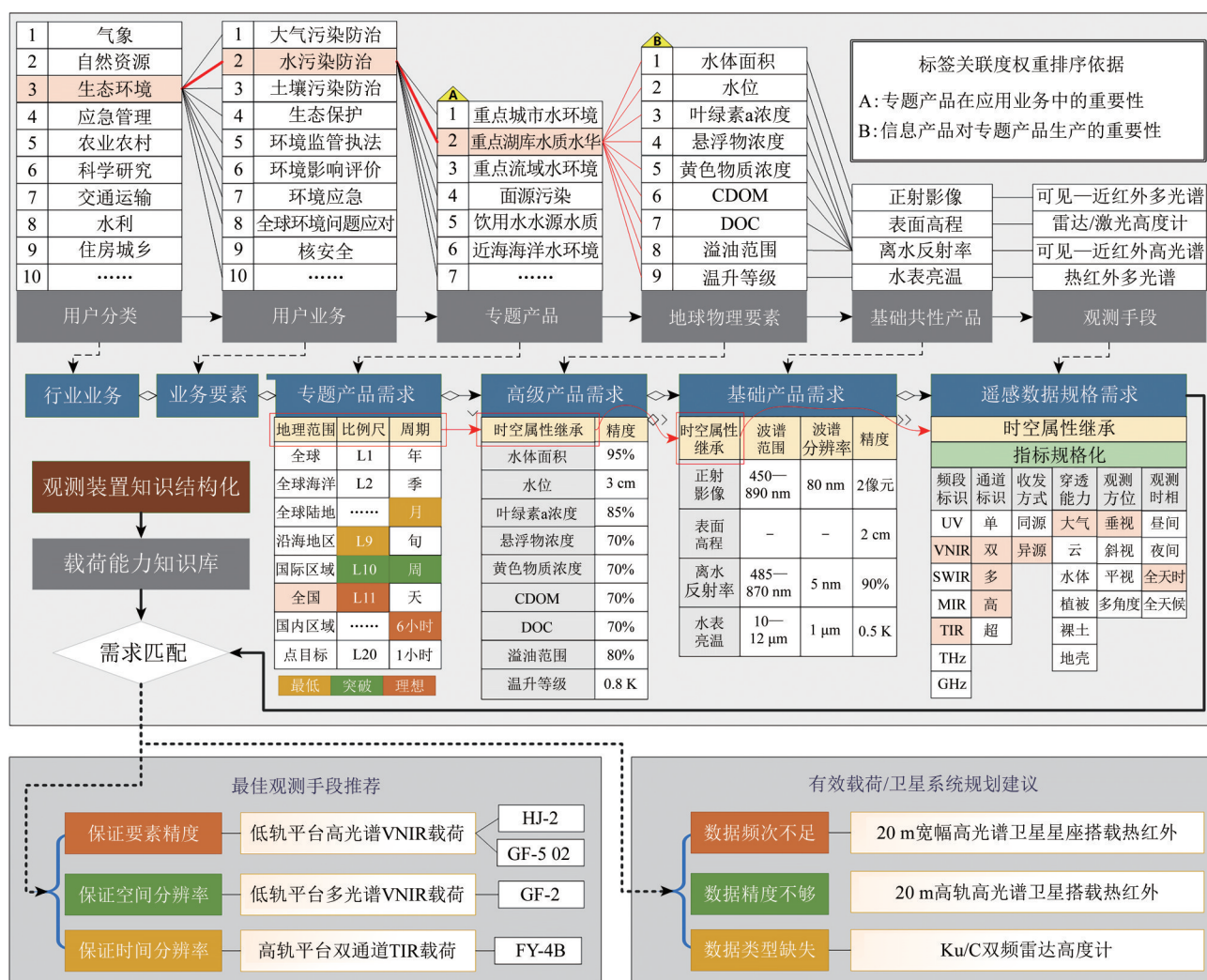


图 12 地球观测应用需求推理溯源

Fig. 12 Reasoning for Earth observation application requirements

上述地球观测知识推理应用案例表明, 本体建模有助于形式化地说明地球观测“人—机—物”之间的内在联系, 增强了高阶复杂问题的可解释性, 也能够支持专业到行业的嵌合, 确保计算机可有效响应应用需求。

5 结 语

人工智能、空天科技是《中华人民共和国国民经济和社会发展第十四个五年规划和2035年远景目标纲要》中强调的战略科技前沿领域, 地球观测应用也是航天强国建设的落脚点。现有地球观测成果共享主要通过文献、数据等形式传递知识, 知识价值的发挥受限于个体的表达与理解能力, 具有较大的过程性损耗。EOKH将知识的共享从点对点的要素级交流, 转变为群体协同式系统级交流, 既降低了地球观测跨行业、社会化综合

应用成本, 又能规避重复性研究投入, 有助于突破技术壁垒和认知障碍, 更全面释放卫星数字经济服务效能。

为有序推进智能化地发现、共享和再利用可信地球观测知识, 促进“人—机—物”要素的协同共生, 构筑可持续的知识共享生态, 对EOKH的发展建议如下:

(1) 强化可信知识的保障能力。“知识包”内容的可信程度是EOKH能够产生实效的前提。为此, 需面向多星多源观测下的地表变量统筹, 设计规格统一、品质可靠的地球观测共性产品生产及真实性检验体系, 实现在计算机环境中有机关联应用需求、信息产品和观测设施, 支持利用“人—机—物”复杂系统协同获取满足需求的可信知识。

(2) 深入探索“人在回路”的地球观测知识

协同生产与转移技术。将概念型地球观测知识本体建模和过程型地球观测知识形式化研究深度融合，强调人、载荷与数据、有效应用需求之间的迭代交互，引入人的智能与决策来提升时空大数据智能处理的可解释性以及高阶复杂问题的解决能力，实现隐性知识和显性知识的同步转移。

(3) 深入探索“知识驱动”的地球观测成果复用技术。融合知识抽取和知识图谱技术进展，增强服务特定需求的可信资源动态关联汇聚与迁移扩展能力，在保障知识贡献者在成果共享过程中权益的同时，逐步降低和消除人为壁垒，推动空间信息应用业务全生命周期的数字化发展和极致资源利用率。

(4) 加强地球观测知识本体研究。通过基础本体建模和领域体建模研究，形成跨领域知识共享协议，支撑地球观测领域的开放科学发展。对应用流程、应用要素在知识包层面进行高级抽象，实现应用知识过程表示和对象表示的结合，支持地球观测应用敏捷开发的实现以确保关键环节技术可控，进而构建地球观测应用智能基础设施(Jordan, 2019)。

(5) 试点代表性行业或领域的地球观测知识枢纽研发。利用牵引性应用场景拉动地球观测应用要素的系统性流动，提升地球观测知识共享服务的成熟度，在相关国际合作计划和任务中发挥消弭数字鸿沟的实质性贡献，推动地球观测应用的数字化发展。

志 谢 论文写作过程中北京百分点科技集团股份有限公司刘译璟博士、清华大学白玉琪教授、中国科学院空天信息创新研究院唐婷研究员参与了讨论，审稿人提出了建设性意见，在此表示衷心的感谢！

参考文献(References)

Abdalla G, Damasceno C D N, Guessi M, Oquendo F and Nakagawa E Y. 2015. A systematic literature review on knowledge representation approaches for systems-of-systems//2015 IX Brazilian Symposium on Components, Architectures and Reuse Software. Belo Horizonte: IEEE: 70-79 [DOI: 10.1109/SBCARS.2015.18]

Albinet C, Whitehurst A S, Jewell L A, Bugbee K, Laur H, Murphy K J, Frommknecht B, Scipal K, Costa G, Jai B, Ramachandran R, Lavallo M and Duncanson L. 2019. A joint ESA-NASA multi-mission

algorithm and analysis platform (MAAP) for biomass, NISAR, and GEDI. *Surveys in Geophysics*, 40(4): 1017-1027 [DOI: 10.1007/s10712-019-09541-z]

Aldana-Martín J F, García-Nieto J, del Mar Roldán-García M and Aldana-Montes J F. 2022. Semantic modelling of Earth Observation remote sensing. *Expert Systems with Applications*, 187: 115838 [DOI: 10.1016/j.eswa.2021.115838]

Aronson J E. 2003. Expert systems//Bidgoli H, ed. *Encyclopedia of Information Systems*. Boston: Academic Press: 277-289 [DOI: 10.1016/B0-12-227240-4/00067-8]

Arvor D, Belgii M, Falomir Z, Mougenot I and Durieux L. 2019. Ontologies to interpret remote sensing images: why do we need them? *Giscience and Remote Sensing*, 56(6): 911-939 [DOI: 10.1080/15481603.2019.1587890]

Bathelt H, Malmberg A and Maskell P. 2004. Clusters and knowledge: local buzz, global pipelines and the process of knowledge creation. *Progress in Human Geography*, 28(1): 31-56 [DOI: 10.1191/0309132504ph469oa]

Bloom N, Jones C I, Van Reenen J and Webb M. 2020. Are ideas getting harder to find? *American Economic Review*, 110(4): 1104-1144 [DOI: 10.1257/aer.20180338]

Bohle S. 2014. Online repository for lab notebooks. *Nature*, 506(7487): 159-159 [DOI: 10.1038/506159e]

Carroll S R, Herczog E, Hudson M, Russell K and Stall S. 2021. Operationalizing the CARE and FAIR Principles for Indigenous data futures. *Scientific Data*, 8(1): 108 [DOI: 10.1038/s41597-021-00892-0]

Chen X F, Zhao L M, Zheng F J, Li J G, Li L, Ding H N, Zhang K N, Liu S M, Li D H and de Leeuw G. 2022. Neural network AEROSol retrieval for geostationary satellite (NNAeroG) based on temporal, spatial and spectral measurements. *Remote Sensing*, 14(4): 980 [DOI: 10.3390/rs14040980]

Chen X F, Zheng F J, Guo D, Wang L L, Zhao L M, Li J G, Li L, Zhang Y H, Zhang K N, Xi M and Li K T. 2021. Review of machine learning methods for aerosol quantitative remote sensing. *National Remote Sensing Bulletin*, 25(11): 2220-2233 (陈兴峰, 郑逢杰, 郭丁, 王莉莉, 赵利民, 李家国, 李雷, 张玉环, 张凯南, 奚萌, 李凯涛. 2021. 气溶胶定量遥感的机器学习方法综述. *遥感学报*, 25(11): 2220-2233) [DOI: 10.11834/jrs.20210451]

Cheng X Q, Liu S H and Zhang R Q. 2022. Thinking on new system for big data technology. *Bulletin of Chinese Academy of Sciences*, 37(1): 60-67 (程学旗, 刘盛华, 张儒清. 2022. 大数据分析处理技术新体系的思考. *中国科学院院刊*, 37(1): 60-67) [DOI: 10.16418/j.issn.1000-3045.20211117005]

Cheng X Q, Mei H, Zhao W, Wah W S B, Shen H W and Li G J. 2020. Data science and computing intelligence: concept, paradigm, and opportunities. *Bulletin of Chinese Academy of Sciences*, 35(12): 1470-1481 (程学旗, 梅宏, 赵伟, 华云生, 沈华伟, 李国杰. 2020. 数据科学与计算智能: 内涵、范式与机遇. *中国科学院院刊*, 35(12): 1470-1481) [DOI: 10.16418/j.issn.1000-3045.20201116005]

China National Space Administration (CNSA). 2019. 2019 National Report of China High-Resolution Earth Observation System

- Applications (Business Application Volume). China National Space Administration (国家航天局). 2019. 2019《中国高分卫星应用国家报告》(业务应用卷). 国家航天局)
- Chu J S G and Evans J A. 2021. Slowed canonical progress in large fields of science. *Proceedings of the National Academy of Sciences of the United States of America*, 118(41): e2021636118 [DOI: 10.1073/pnas.2021636118]
- De Simone F, D'Amore F, Bencardino M, Carbone F, Hedgecock I M, Sprovieri F, Cinnirella S and Pirrone N. 2021. The GOS⁴M Knowledge Hub: a web-based effectiveness evaluation platform in support of the Minamata Convention on Mercury. *Environmental Science and Policy*, 124: 235-246 [DOI: 10.1016/j.envsci.2021.06.021]
- Directorate-General for Research and Innovation (European Commission). 2018. OSPP-REC: Open Science Policy Platform Recommendations. Publications Office of the European Union [DOI: 10.2777/958647]
- Editorial Office of Journal of Global Change Data and Discovery. 2017. Guidelines of global change research data publishing & repository. *Journal of Global Change Data and Discovery*, 1(3): 253-261 (《全球变化数据学报》(中英文)编辑部. 2017. 全球变化科学研究数据出版与共享指南. *全球变化数据学报*, 1(3): 253-261) [DOI: 10.3974/geodp.2017.03.01]
- Esch T, Asamer H, Bachofer F, Balhar J, Boettcher M, Boissier E, d'Angelo P, Gevaert C M, Himer A, Jupova K, Kurz F, Kwarteng A Y, Mathot E, Marconcini M, Marin A, Metz-Marconcini A, Pacini F, Paganini M, Permana H, Soukup T, Uereyen S, Small C, Svaton V and Zeidler J N. 2020. Digital world meets urban planet-new prospects for evidence-based urban studies arising from joint exploitation of big earth data, information technology and shared knowledge. *International Journal of Digital Earth*, 13(1): 136-157 [DOI: 10.1080/17538947.2018.1548655]
- Esch T, Üreyen S, Zeidler J, Metz-Marconcini A, Hirner A, Asamer H, Tum M, Böttcher M, Kuchar S, Svaton V and Marconcini M. 2018. Exploiting big earth data from space-first experiences with the timescan processing chain. *Big Earth Data*, 2(1): 36-55 [DOI: 10.1080/20964471.2018.1433790]
- Evers H D, Gerke S and Menkhoff T. 2010. Knowledge clusters and knowledge hubs: designing epistemic landscapes for development. *Journal of Knowledge Management*, 14(5): 678-689 [DOI: 10.1108/13673271011074836]
- GEO Secretariat. 2019. 2020-2022 GEO Work Programme Summary Document (Version 5.1). In, 120-1. Canberra, Group on Earth Observations
- GEO Secretariat. 2020a. Draft implementation plan for the GEO knowledge hub//16th Programme Board Meeting. Geneva: GEO Secretariat
- GEO Secretariat. 2020b. GEO knowledge hub implementation plan//17th Programme Board Meeting. [s.l.]: GEO Secretariat
- GEO Secretariat. 2021a. GEO statement on open knowledge//21st Programme Board Meeting. [s.l.]: GEO Secretariat
- GEO Secretariat. 2021b. Report on the development of the GEO knowledge hub//21st Programme Board Meeting. [s.l.]: GEO Secretariat
- Gill K S. 2020. Pearl, Judea and Mackenzie, Dana: the book of why: the new science of cause and effect (2018). *AI and Society*, 35(3): 767-768 [DOI: 10.1007/s00146-020-00971-7]
- Gorelick N, Hancher M, Dixon M, Ilyushchenko S, Thau D and Moore R. 2017. Google Earth Engine: planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202: 18-27 [DOI: 10.1016/j.rse.2017.06.031]
- Gu X F and Yu T. 2018. Application-Oriented Space Remote Sensing Scientific Argumentation Theory, Method and Technology. Beijing: Science Press (顾行发, 余涛. 2018. 面向应用的航天遥感科学论证理论、方法与技术. 北京: 科学出版社)
- He G J, Wang G Z, Long T F, Peng Y, Jiang W, Yin R Y, Jiao W L and Zhang Z M. 2018. Opening and sharing of big earth observation data: challenges and countermeasures. *Bulletin of Chinese Academy of Sciences*, 33(8): 783-790 (何国金, 王桂周, 龙腾飞, 彭燕, 江威, 尹然宇, 焦伟利, 张兆明. 2018. 对地观测大数据开放共享: 挑战与思考. *中国科学院院刊*, 33(8): 783-790) [DOI: 10.16418/j.issn.1000-3045.2018.08.003]
- Hogan A, Blomqvist E, Cochez M, D'amato C, De Melo G, Gutierrez C, Kirrane S, Gayo J E L, Navigli R, Neumaier S, Ngomo A C N, Polleres A, Rashid S M, Rula A, Schmelzeisen L, Sequeda J, Staab S and Zimmermann A. 2022. Knowledge graphs. *ACM Computing Surveys*, 54(4): 71 [DOI: 10.1145/3447772]
- Horsburgh J S, Aufdenkampe A K, Mayorga E, Lehnert K A, Hsu L, Song L L, Jones A S, Damiano S G, Tarboton D G, Valentine D, Zaslavsky I and Whitenack T. 2016. Observations Data Model 2: a community information model for spatially discrete Earth observations. *Environmental Modelling and Software*, 79: 55-74 [DOI: 10.1016/j.envsoft.2016.01.010]
- Howells J R L. 2002. Tacit knowledge, innovation and economic geography. *Urban Studies*, 39(5/6): 871-884 [DOI: 10.1080/00420980220128354]
- IEEE Standards Association. 2011. IEEE Std 2030-2011 IEEE guide for smart grid interoperability of energy technology and information technology operation with the electric power system (EPS), end-use applications, and loads. New York: IEEE: 1-126 [DOI: 10.1109/IEEESTD.2011.6018239]
- Institute for Artificial Intelligence, Tsinghua University, Beijing Academy of Artificial Intelligence and Tsinghua China Academy of Engineering Knowledge Intelligence Joint Research Center. 2020. Blockchain Development Research Report. Beijing (清华大学人工智能研究院, 北京智源人工智能研究院, 清华—中国工程院知识智能联合研究中心. 2020. 区块链发展研究报告. 北京)static.aminer.cn
- Janowicz K, Haller A, Cox S J D, Le Phuoc D and Lefrancois M. 2019. SOSA: a lightweight ontology for sensors, observations, samples, and actuators. *Journal of Web Semantics*, 56: 1-10 [DOI: 10.1016/j.websem.2018.06.003]
- Ji S X, Pan S R, Cambria E, Marttinen P and Yu P S. 2022. A survey on knowledge graphs: representation, acquisition, and applications. *IEEE Transactions on Neural Networks and Learning Systems*,

- 33(2): 494-514 [DOI: 10.1109/TNNLS.2021.3070843]
- Jordan M I. 2019. Artificial Intelligence—The Revolution Hasn't Happened Yet. *Harvard Data Science Review* [DOI: 10.1162/99608f92.f06c6e61]
- Klein T, Nilsson M, Persson A and Håkansson B. 2017. From open data to open analyses—new opportunities for environmental applications? *Environments*, 4(2): 32 [DOI: 10.3390/environments4020032]
- Kleinberg J, Ludwig J, Mullainathan S and Sunstein C R. 2018. Discrimination in the age of algorithms. *Journal of Legal Analysis*, 10: 113-174 [DOI: 10.1093/jla/laz001]
- Leiserson C E, Thompson N C, Emer J S, Kuszmaul B C, Lampson B W, Sanchez D and Schardl T B. 2020. There's plenty of room at the top: what will drive computer performance after Moore's law? *Science*, 368(6495): eaam9744 [DOI: 10.1126/science.aam9744]
- Li G Q, Zhang H Y, Zhang L C, Wang Y Y and Tian C Z. 2016. Development and trend of earth observation data sharing. *Journal of Remote Sensing (in Chinese)*, 20(5): 979-990 (李国庆, 张红月, 张连翀, 王媛媛, 田传召. 2016. 地球观测数据共享的发展和趋势. *遥感学报*, 20(5): 979-990) [DOI: 10.11834/jrs.20166173]
- Little J D C. 2011. OR FORUM—Little's law as viewed on its 50th anniversary. *Operations Research*, 59(3): 536-549 [DOI: 10.1287/opre.1110.0940]
- Liu S M, Zhang Y L, Zhao L M, Chen X F, Zhou R X, Zheng F J, Li Z L, Li J G, Yang H, Li H F, Yang J, Gao H L and Gu X F. 2022. QUantitative and automatic atmospheric correction (QUAAC): application and validation. *Sensors*, 22(9): 3280 [DOI: 10.3390/s22093280]
- Liu Z M and Wang J. 2020. Human-cyber-physical systems: concepts, challenges, and research opportunities. *Frontiers of Information Technology and Electronic Engineering*, 21(11): 1535-1553 [DOI: 10.1631/FITEE.2000537]
- Masmoudi M, Karray M H, Ben Abdallah Ben Lamine S, Zghal H B and Archimede B. 2020. MEMOn: modular environmental monitoring ontology to link heterogeneous earth observed data. *Environmental Modelling and Software*, 124: 104581 [DOI: 10.1016/j.envsoft.2019.104581]
- Mei H, Cao D G and Xie T. 2022. Ubiquitous operating system: toward the blue ocean of human-cyber-physical ternary ubiquitous computing. *Bulletin of Chinese Academy of Sciences*, 37(1): 30-37 (梅宏, 曹东刚, 谢涛. 2022. 泛在操作系统: 面向人机物融合泛在计算的新蓝海. *中国科学院院刊*, 37(1): 30-37) [DOI: 10.16418/j.issn.1000-3045.20211117009]
- Nativi S, Mazzetti P, Santoro M, Papeschi F, Craglia M and Ochiai O. 2015. Big data challenges in building the global earth observation system of systems. *Environmental Modelling and Software*, 68: 1-26 [DOI: 10.1016/j.envsoft.2015.01.017]
- Owusu C, Snigdha N J, Martin M T and Kalyanapu A J. 2022. PyGEE-SWToolbox: a python jupyter notebook toolbox for interactive surface water mapping and analysis using google earth engine. *Sustainability*, 14(5): 2557 [DOI: 10.3390/su14052557]
- Pathirage C P, Amaratunga R D G and Haigh R P. 2006. Managing construction workers and their tacit knowledge in a knowledge environment: a conceptual framework//6th International Postgraduate Research Conference in the Built and Human Environment. Delft: Delft University of Technology and TNO, Delft University: 733-743
- Paulheim H. 2018. How much is a triple? Estimating the cost of knowledge graph creation//Proceedings of the ISWC 2018 Posters and Demonstrations, Industry and Blue Sky Ideas Tracks co-located with 17th International Semantic Web Conference (ISWC 2018). Monterey: CEUR Workshop Proceedings
- Pu Z Q, Yi J Q, Liu Z, Qiu T H, Sun J L and Li F M. 2022. Knowledge-based and data-driven integrating methodologies for collective intelligence decision making: a survey. *Acta Automatica Sinica*, 48(3): 627-643 (蒲志强, 易建强, 刘振, 丘腾海, 孙金林, 李非墨. 2022. 知识和数据协同驱动的群体智能决策方法研究综述. *自动化学报*, 48(3): 627-643) [DOI: 10.16383/j.aas.c210118]
- Rapiński J, Bednarczyk M and Zinkiewicz D. 2019. JupyTEP IDE as an online tool for earth observation data processing. *Remote Sensing*, 11(17): 1973 [DOI: 10.3390/rs11171973]
- Stacey P and Berry D. 2018. Towards a Digital Earth: using archetypes to enable knowledge interoperability within geo-observational sensor systems design. *Earth Science Informatics*, 11(2): 307-323 [DOI: 10.1007/s12145-018-0340-z]
- Su W J. 2021. You are the best reviewer of your own papers: an owner-assisted scoring mechanism//35th Conference on Neural Information Processing Systems. [s.l.]: [s.n.]: 27929-27939
- Sun K, Zhu Y Q, Pan P, Hou Z W, Wang D X, Li W R and Song J. 2019. Geospatial data ontology: the semantic foundation of geospatial data integration and sharing. *Big Earth Data*, 3(3): 269-296 [DOI: 10.1080/20964471.2019.1661662]
- Wing J M. 2021. Trustworthy AI. *Communications of the ACM*, 64(10): 64-71 [DOI: 10.1145/3448248]
- Wulder M A and Coops N C. 2014. Satellites: make Earth observations open access. *Nature*, 513(7516): 30-31 [DOI: 10.1038/513030a]
- Xing L N and Chen Y W. 2011. Research progress on intelligent optimization guidance approaches using knowledge. *Acta Automatica Sinica*, 37(11): 1285-1289 (邢立宁, 陈英武. 2011. 基于知识的智能优化引导方法研究进展. *自动化学报*, 37(11): 1285-1289) [DOI: 10.3724/SP.J.1004.2011.01285]
- Xu H, Li S, Bai Y Q, Dong W H, Huang W Y, Xu S M, Lin Y L, Wang B, Wu F H, Xin X G, Zhang L, Wang Z Z and Wu T W. 2019. A collaborative analysis framework for distributed gridded environmental data. *Environmental Modelling and Software*, 111: 324-339 [DOI: 10.1016/j.envsoft.2018.09.007]
- Xu Z W, Li G J and Sun N H. 2022. Information superbahn: towards new type of cyberinfrastructure. *Bulletin of Chinese Academy of Sciences*, 37(1): 46-52 (徐志伟, 李国杰, 孙凝晖. 2022. 一种新型信息基础设施: 高通量低熵算力网(信息高铁). *中国科学院院刊*, 37(1): 46-52) [DOI: 10.16418/j.issn.1000-3045.20211117008]
- Yang Q Q, Jin C Y, Li T W, Yuan Q Q, Shen H F and Zhang L P. 2022. Research progress and challenges of data-driven quantitative remote sensing. *National Remote Sensing Bulletin*, 26(2): 268-285 (杨倩倩, 靳才溢, 李同文, 袁强强, 沈焕锋, 张良培. 2022. 数据驱动的定量遥感研究进展与挑战. *遥感学报*, 26(2): 268-285)

[DOI: 10.11834/jrs.20211410]

Zhang H Y, Zhou C H, Lv G N, Wu Z F, Lu F, Wang J F, Yue T X, Luo J C, Ge Y and Qin C Z. 2020. The connotation and inheritance of Geo-information Tupu. *Journal of Geo-information Science*, 22(4): 653-661 (张洪岩, 周成虎, 闫国年, 吴志峰, 陆锋, 王劲峰, 岳天祥, 骆剑承, 葛咏, 秦承志. 2020. 试论地学信息图谱思想的内涵与传承. *地球信息科学学报*, 22(4): 653-661) [DOI: 10.12082/dqxxkx.2020.200167]

Zhu C Q, Xu D J, Ren N, Cui H C and Zhao Y Z. 2021. Model and implementation of geographic data transaction certificate and copyright

protection based on blockchain and digital watermarking. *Acta Geodaetica et Cartographica Sinica*, 50(12): 1694-1704 (朱长青, 徐鼎捷, 任娜, 崔瀚川, 赵亚宙. 2021. 区块链与数字水印相结合的地理数据交易存证及版权保护模型. *测绘学报*, 50(12): 1694-1704) [DOI: 10.11947/j.AGCS.2021.20200559]

Zhu Z, Wulder M A, Roy D P, Woodcock C E, Hansen M C, Radeloff V C, Healey S P, Schaaf C, Hostert P, Strobl P, Pekel J F, Lyburner L, Pahlevan N and Scambos T A. 2019. Benefits of the free and open Landsat data policy. *Remote Sensing of Environment*, 224: 382-385 [DOI: 10.1016/j.rse.2019.02.016]

Earth observation knowledge hub: Implications, key technologies and perspectives

ZHAO Limin^{1,2,3}, MIAO Chen⁴, XING Jin⁵, LI Guoqing^{1,6}, HOU Yukui⁷, LIU Chuang⁸, LI Jiaguo^{1,2,3}, CHEN Xingfeng^{1,2,3}, LIU Jun^{1,2}, YANG Jian^{1,2,3}, ZHOU Xiang^{1,2}, GU Xingfa^{1,2,3}

1. *Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China;*

2. *National Engineering Research Center of Satellite Remote Sensing Applications, Beijing 100094, China;*

3. *Demonstration Centre for Spaceborne Remote Sensing National Space Administration, Beijing 100094, China;*

4. *National Remote Sensing Center of China, Beijing 100036, China;*

5. *Earth Observation System and Data Center, China National Space Administration, Beijing 100101, China;*

6. *National Earth Observation Data Center, Beijing 100094, China;*

7. *Qian Xuesen Laboratory, China Academy of Space Technology, Beijing 100094, China;*

8. *Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China*

Abstract: Intelligently sharing and reusing the knowledge developed by application practices are the keys to break through technical barriers and fully activate and release the effectiveness of Earth Observations (EO). The Earth Observation Knowledge Hub (EOKH), which is used to couple decentralized knowledge bases organically, is a research frontier for the governance and intelligent service of global EO applications. In the design of Group on Earth Observations (GEO), the GEO Knowledge Hub (GKH) is intended to provide authoritative, validated, and reproducible content for evidence-based reporting on policy commitments and decision-making. Thus, the GKH offers a platform for users to discover, learn about, and employ methods, analytical tools, and applications; it also provides opportunities for the GEO community to collaborate and provide mutual assistance related to GKH contents. However, important lessons, such as the sensitive issue of intellectual ethics and how to profit GKH from the recent technological advances in information technologies, have been learned during the implementations. In response to the problems and insights encountered by GEO in developing GKH, we systematically analyzed the connotation and characteristics of EOKH and sorted out the fundamental needs and challenges for the development of EOKH in China.

First, we systematically analyzed the connotation and characteristics of EOKH. The study argues that EOKH is the intersection node of high-throughput trusted EO knowledge in knowledge-sharing networks. It has three typical features, i.e., connectivity, high throughput, and lightweight computing. Its core mission is to identify and transfer valuable research in a timely manner and to promote high throughput of application packages. Second, we sorted out the fundamental needs and challenges for the development of EOKH in China. Considering the latest progress in the study of EO ontology, we also analyzed the possible key technical problems and gave strategies to cope with them. On this basis, the system architecture prototype of EOKH, which is drawn on the system design concept of representational state transfer, is proposed, and an ontology model of conceptual EO knowledge and a formalization model of process-oriented EO knowledge are established.

The study argues that EOKH should be in an open collaborative environment where humans are in the loop. The key technologies are system metrics, knowledge transfer, knowledge reuse, and knowledge exploration and visualization. The transfer and reuse of knowledge packages can greatly enhance the ease of development and reuse of EO application practices. Ontology modeling helps formalize the intrinsic connection between the human-cyber-physical systems of EO application and enhances the interpretability of higher-order complex problems. EOKH transforms knowledge sharing from point to point at the element level to group collaborative at the system level, which not only reduces the cost of cross-industry and socially integrated EO applications but also avoids repetitive research inputs, helps break through technical barriers and cognitive obstacles, and releases the effectiveness of satellite digital economy services comprehensively. The study argues that connecting EO knowledge in the human – cyber – physical systems and exploring high-throughput knowledge coproduction and transfer technology in a “humans-in-the-loop” environment are necessary to enhance the interpretability of tacit EO knowledge and promote

the competitiveness and activeness of EOKH.

Key words: remote sensing, knowledge hub, Earth Observations, ontologies, knowledge packages, knowledge recusing, cyber-physical-social systems, humans in the loop

Supported by National Key Research and Development Program of China (No. 2019YFE0126600); CAST Fund of China Academy of Space Technology (No. ZC-1050-2022-11-005)

附表 A1 术语表
Appendix Table A1 Glossary of terms

中文	英文	释义	参考源
地球观测	Earth Observations	通过一系列基于时空坐标的记录行为,将地表现象的特征属性以数据、术语或符号等形式予以表达。该记录行为通常由观测装置、数据传递与处理等活动有序组合的采样过程构成;采样方式则包括远距离遥感、场地测量或非原位采集	Horsburgh 等,2016; 地球观测组织 GEO ¹
知识	Knowledge	通过实践或教育获得的事实、信息或技巧,抑或对某项事物的理论或实践认知。从认知维度可分为事实性知识(Factual Knowledge),概念性知识(Conceptual Knowledge)和过程性知识(Procedural Knowledge)	牛津词典 OED ²
显性知识	Explicit knowledge	以文字、符号、图形等方式表达,以观测数据、软件代码、发表物等成文形式存在的知识,易于传播	术语在线 ³
隐性知识	Tacit knowledge	未以文字、符号、图形等方式表达,以技能、思想、经验等认知形式存在的知识,传播难度大,价值高	术语在线
开放科学	Open science	在所有利益攸关方的积极参与下,让科学数据、信息和研究成果更广泛地可获取并更可靠地加以再利用	联合国教科文组织 ⁴
人在回路	Humans in the loop	人类智能与机器智能形成闭环系统,不断相互作用、相互辅助	程学旗 等,2020
知识枢纽	Knowledge Hub	知识生产与共享网络中的中转与代理节点,能够将分散的知识库耦合在一起,形成强大的知识共享能力,方便用户快速找到知识、学习和利用知识	Evers 等,2010
知识包	Knowledge Package	针对地表特定科学问题或应用需求的观测手段、研究数据、算法模型、技术流程等应用要素及其概念、关系、规则和过程的集合,是显性知识和隐性知识的综合体	
本体	Ontology	利用结构化的“术语集”,实现特定应用领域内对事物概念及概念间关系的规范化描述。本体建模能够克服语义障碍所造成的互操作性问题	术语在线
实体	Entity	客观世界中存在的可互相区分的客观对象或抽象概念,是本体、实例及关系的整合。实例则指事物概念所对应的具体事物个体	术语在线
结构化	Knowledge structuring	一种知识表示形式,按此种形式,由知识要素汇集而成的每个记录的结构都是一致的,并且可以使用关系模型予以有效描述	术语在线
形式化	Knowledge formalization	基于数学原理,以统一表达方式描述和验证过程型知识属性的方法	术语在线
人—机—物	Human-cyber-physical systems	人指人类社会,机指信息空间,物指物理世界,也称为信息—物理—社会系统 CPSS(Cyber-Physical-Social Systems)。人—机—物三元协同与融合是普惠计算(Computing for the masses)的资源依托	Liu 和 Wang, 2020
群体智能	Collective Intelligence	结合了动物的群体行为特性以及人类社会的认知特性的计算智能算法	蒲志强 等,2022
模型驱动	Model driven	通过数学物理模型的分析而发现新知识的研究方法,过程透明,机理明确,有利于对因果关系的理解、设计、构造、部署、操作、维护和修改	术语在线
数据驱动	Data driven	通过对大量实验数据的分析而发现新知识的一种研究方法。一般从初始的数据或观测值出发,运用启发式规则,寻找和建立内部特征之间的关系,从而发现一些定理或定律	术语在线
知识驱动	Knowledge driven	采用已获得知识来指导研究方法的进化,即算法运行之初就已确定知识的基本形式,相关知识按照固定规则随着算法演化而不断调整。早期应用于专家系统	邢立宁和陈英武,2011

续表

中文	英文	释义	参考源
知识模型化 (知识建模)	Knowledge modeling	利用一定的知识表示语言或数据结构,对知识进行形式化和结构化的抽象,从而令知识可由计算机系统所解释和处理	术语在线
互操作性	Interoperability	不同程序、对象或构件之间相互交换信息和协同工作,特别当用户不了解各种功能单元独立特征或对其知之甚少时,实现异构系统之间数据、模型、语义、软件等在线共享和共用的能力	术语在线
知识抽取	Knowledge extraction	通过识别、理解、筛选、格式化,把蕴含于文本中的各个知识点抽取出来,以一定形式存入知识库中的过程。目的是增强信息的可用性和可复用性,这个过程同时又可以看作对现有的非结构化信息的语义标注过程	术语在线
知识图谱	Knowledge Graph	一种图形数据结构,节点为概念或事物,边为概念或事物之间的关系;或指代利用上述数据结构存储的数据内容。该概念有别于地学信息图谱,后者中的“图”侧重于表达地图等地学信息,“谱”侧重于表达信息的时间演进	术语在线 张洪岩 等,2020
知识规范化	Knowledge normalization	将关系模式分解成要求的范式的过程。其基本思想是消除关系模式中存在的合适的数据依赖,从而消除可能存在的数据冗余。一般通过统一编码规则实现	术语在线
保序机制	Isotonic mechanism	在知识评审与发布环节,保证先被送审的知识包先于后送审的知识包被公布	Su,2021
表象状态转移	Representational State Transfer	从资源角度来观察整个网络的一种 Web 软件架构风格。资源由通用资源标志符(Uniform Resource Identifier,URI)确定,客户端应用通过 URI 获取资源的表征,这些表征会使应用程序转变其状态	Leiserson 等,2020

注: ¹GEO at a Glance-Earth Observations (https://www.earthobservations.org/geo_wwd.php); ²OED-Oxford English Dictionary (<https://www.oed.com>); ³全国科学技术名词审定委员会 术语在线 (<https://www.termonline.cn/index>); ⁴<https://www.unesco.org/en/natural-sciences/open-science>