

中国矿产资源潜力评价数据模型 研制流程及方法技术

左群超 杨东来 叶天竺

(中国地质调查局发展研究中心, 北京 100037)

提要:本文以中国矿产资源潜力评价工作专项的成果数据建模为实例,提出了复杂问题域数据模型的研制流程和实施策略,系统阐述了研制流程和实施策略的基本内容,简要介绍了设计思想、采用的方法技术以及相关研究成果。通过实践验证了研制流程和实施策略及方法技术的有效性、可行性、实用性和代表性,有一定借鉴意义和推广价值。特别适合以超大型、多专业、工程性以及面向空间数据为主要特征的大型地质项目数据资源建设。

关 键 词:矿产资源潜力评价;数据模型;数据建模;本体;元模型;面向对象技术;模型驱动;问题域

中图分类号:P628+.4 **文献标志码:**A **文章编号:**1000-3657(2012)04-1049-13

1 引 言

1.1 背景

中国矿产资源潜力评价专项是中国国土资源部在矿产资源领域部署的一项基本国情调查工作,总体目标任务是摸清中国矿产资源家底、实现成矿地质理论和技术方法创新、培养一批综合型地质矿产人才。该专项于 2006 年 6 月启动,其成果将为中国制定矿产资源中长期发展规划提供依据。

中国矿产资源潜力评价专项,需要评价的矿产有 25 种(煤炭、铀、铁、铜、铝、铅、锌、锰、镍、钨、锡、钾、金、铬、钼、锑、稀土、银、硼、锂、磷、硫、萤石、菱镁矿、重晶石等),使用的地质基础资料和科研成果资料的时间跨度从新中国成立之初至 2006 年底,覆盖的专业领域包括成矿地质背景研究、成矿规律研究、矿产预测研究以及重力资料应用、磁测资料应用、化探资料应用、遥感资料应用、自然重砂资料应用,还有化工矿产资源潜力评价、铀矿资源潜力评价、煤炭

资源潜力评价等,借助的技术手段是全面全过程应用地理信息系统和计算机技术,采用的评价方法是矿床模型综合地质信息预测方法,矿产资源预测总量描述上要求资源定量、空间定位,参与的部门或行业有中国国土资源部地质勘查司、中国地质调查局及局系统所属地质单位、中国 30 个省(市、自治区)国土资源厅及所属地勘单位、中化地质矿山总局、中国煤炭地质总局、中国核工业地质总局以及相关高等院校和科研院所等^[1]。

1.2 问题

中国矿产资源潜力评价专项具有超大型(参行业多、单位多、人员多、投入经费多、时间周期长)、多专业(覆盖专业广)、工程性(尽管有诸多重大地质问题需要研究和总结,但总体要求与规模方面还是表现为一个工程性项目)等特征,且总体实施采用“全面全过程应用 GIS 信息技术”和“矿床模型综合地质信息预测方法”的技术思路。因此,该专项实施的关键之一是要有一套可供遵循的矿产资源潜力评

收稿日期:2012-04-05;改回日期:2012-06-25

基金项目:中国地质调查局“全国重要矿产资源潜力评价综合信息集成”项目(121201121041)和“国家基础地质数据库整合与集成”项目(1212010815001)共同资助。

作者简介:左群超,男,1964 年生,高级工程师,研究方向为矿产资源潜力评价、地质空间数据整合与集成、计算机应用软件研发或平台设计;E-mail:zuoqc163@163.com。

价数据模型规范。事实上,在中国矿产资源潜力评价专项启动之初,还没有现成的矿产资源潜力评价数据模型等相关规范可用。

本文以中国矿产资源潜力评价专项的成果数据建模为实例,提出了复杂问题域数据模型的研制流程和实施策略,系统阐述了研制流程和实施策略的基本内容,简要介绍了设计思想、采用的方法技术、以及相关研究成果。通过实践验证了研制流程和实施策略及方法技术的有效性、可行性、实用性和代表性,有一定借鉴意义和推广价值。特别适合以超大型、多专业、工程性、面向空间数据为主要特征的大型地质项目数据资源建设。

2 数据模型研制流程

2.1 基本步骤

中国矿产资源潜力评价数据模型研制流程,依序可分五大步骤或阶段:一是制定数据模型建模方案,二是建立专业谱系及特征分类,三是建立图件数据模型,四是建立成果数据库结构逻辑模型,五是建立成果数据库结构物理模型。每个阶段的基本含义及工作内容简述如下。

(1)制定数据模型建模方案。制定数据模型建模技术方案,供数据模型建模人员遵循。数据模型建模技术方案主要内容包括研制数据模型的组织机构与负责人、参与人员、工作目标、工作范围、工作内容、时间进度、采用的方法与技术以及提供若干约定的工作模板或软件工具等。

(2)建立专业谱系及特征分类。建立数据模型所属领域的专业谱系及特征分类。专业谱系及特征分类是关于数据模型所属领域的专业及特征的分类、分级及其编码体系,是数据模型所有组成元素的基本骨架或语义基础,是相关数据或资料分类的基本依据,是专业人员与数据或资料之间联系的桥梁。

(3)建立图件数据模型。建立图件分类、图层分类、图件分类与专业分类之间、图层分类与特征分类之间、特征分类的空间特性、特征分类的属性结构、空间划分方式、空间参考系、比例尺、图式图例、元数据以及相关编码体系等。图件数据模型是数据模型建模阶段性重要成果之一,直接用于指导或规范地质领域各专业人员研究编图与成果表达。

(4)建立成果数据库结构逻辑模型。建立成果数据库的数据库结构逻辑模型。数据库结构逻辑模型

的组成内容、逻辑规则及其约束关系等独立于具体的数据库管理系统实现技术,主要依赖于项目成果数据本身及其组成联系。数据库结构逻辑模型是数据模型建模阶段性成果之一,是实现数据库结构物理模型的依据和基础。

(5)建立成果数据库结构物理模型。建立成果数据库的数据库结构物理模型。数据库结构物理模型是基于已选定硬件环境、计算机操作系统(OS)、数据库管理系统(DBMS)等约束条件下,根据已确定的数据库结构逻辑模型,进行技术实现的结果,使得数据库存储空间占用少、访问效率高和维护代价低,主要包括数据库物理文件、表空间、基本表、字段(数据类型、长度、是否可为空、默认值)、主键、外键、索引、视图、触发器、存储过程等内容。同一数据库结构逻辑模型,依据不同数据库管理系统等技术有不同具体实现,即基于MS SQL、Oracle、或DB2等数据库管理系统分别有不同的具体实现。

中国矿产资源潜力评价数据模型研制流程及每一步取得的阶段性成果如图1所示。

2.2 实施策略

数据模型研制本身是一个繁琐、复杂、耗时的工作,加上地质领域的复杂性、不确定性以及中国矿产资源潜力评价专项的超大型、多专业、工程性、面向地质空间数据或资料等特征因素,更加重了中国矿产资源潜力评价数据模型研制工作的复杂程度,因此,在研制中国矿产资源潜力评价数据模型全过程中,必须采取有效的实施策略。

贯穿上述数据模型研制全流程的实施策略,即“一个中心”和“九项基本原则”。“一个中心”,是一切以围绕服务于项目总体目标实现为中心。“九项基本原则”,是统一规划部署、统一组织管理、多领域专家参与、分阶段分层次、可记录迭代、可分割隔离、审核评价、标准化、软件化等多项原则。这些原则的基本涵义描述如下。

(1)统一规划部署原则。一体化考虑用于地质编图的图件数据模型、项目成果数据库结构逻辑模型、项目成果数据库结构物理模型、项目成果数据库建设技术要求以及项目成果数据库基本应用等内容的研制工作。

(2)统一组织管理原则。全流程应有项目总技术负责人的强有力支持、应有一个小组专门负责数据模型研制与维护管理工作。

(3)多领域专家参与原则。应有地质领域专家和
信息领域专家协作参与、不同阶段不同领域专家重
点负责相关内容。

(4)分阶段分层次原则。数据模型研制分阶段、
数据模型成果分层次。专业谱系及特征分类支撑数
据模型研制，图件数据模型直接用于项目地质人员

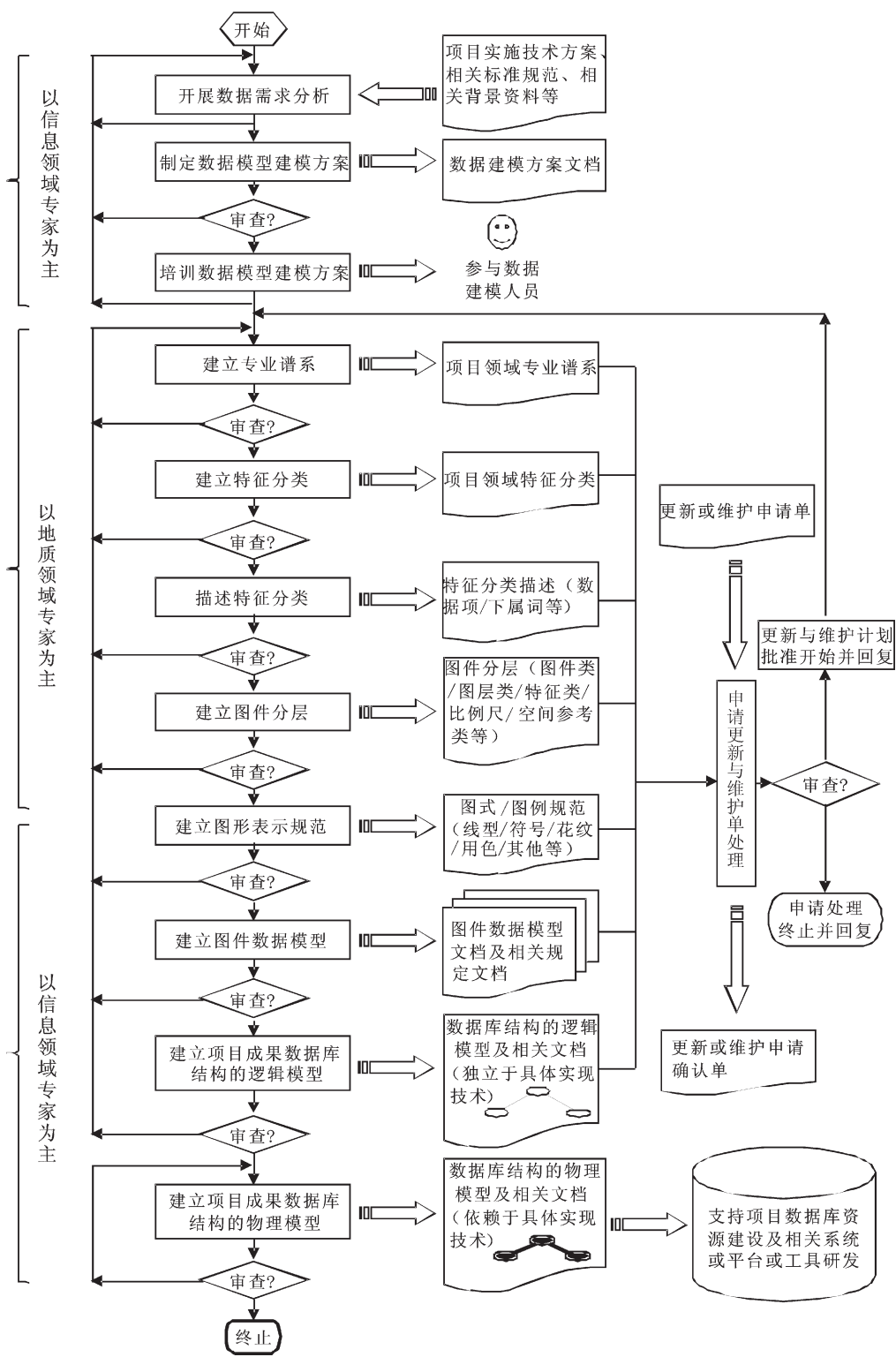


图 1 数据模型研制流程
Fig.1 Data model development process

开展编图与研究,成果数据库结构逻辑模型直接用于成果数据库结构物理模型具体实现,成果数据库结构物理模型直接支撑项目成果数据库及管理系统建设。

(5)可记录可迭代原则。数据模型研制过程的每一步,应有可记录的文档成果,用于迭代与回溯,支持数据模型的管理、更新与维护。

(6)可分割隔离原则。数据模型成果是可分割可隔离的,用以控制或处理数据模型研制复杂度,如:稳定的专业谱系及特征分类应不影响图件数据模型,稳定的图件数据模型应不影响项目成果数据库结构逻辑模型,稳定的项目成果数据库结构逻辑模型应不影响项目成果数据库结构物理模型。

(7)审核评价原则。针对数据模型研制过程的每一阶段性成果和总体成果本身,应按一定审核程序、机制和评价指标(基本评价指标主要包括完整性、一致性、明确性、稳定性、简洁性、有效性),组织专家评审审核数据模型成果本身;对更新或需求变化采取

严格的审核程序,保持数据模型成果本身的相对稳定性,以利于大规模的、工程性的项目成果数据资源建设工作。

(8)标准化原则。在研制数据模型过程中,有标准的,应最大限度采用现有标准或保持与现有标准兼容;无标准的,应依照标准化方法,建立新规范或完善现有标准。

(9)软件化原则。研制数据模型每一步和数据模型成果本身,尽可能借助软件工具,辅助数据模型研制、维护管理和实际应用,提高工作效率、减轻劳动强度、降低复杂度。

中国矿产资源潜力评价数据模型研制流程、组织方式、人员构成、实施策略、审核机制、评价指标、数据模型阶段性成果、数据模型阶段性成果作用等构成要素之间关系及主要实施策略,如图 2 所示。

3 方法技术应用

在一般数据模型建模工作中,可使用的方法技

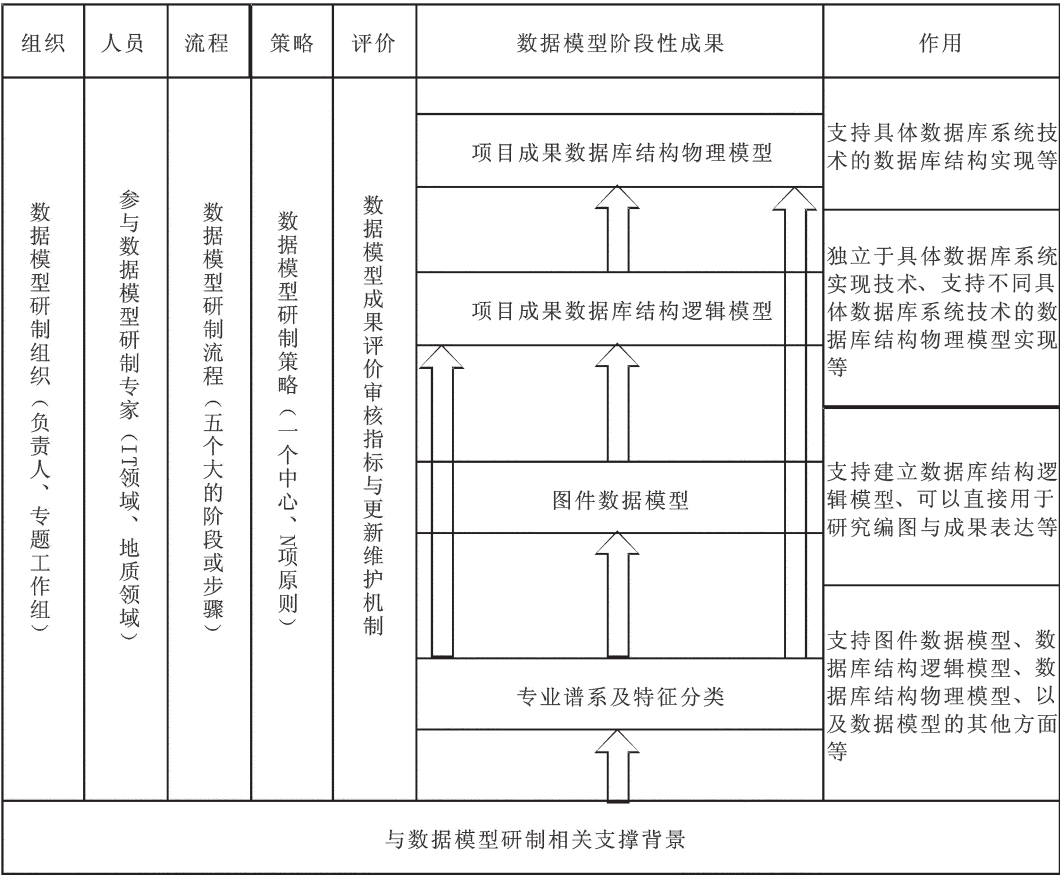


图 2 数据模型研制实施策略

Fig.2 The implementation strategy of the data model development

术很多,如面向对象(Object Oriented Analyze&Design)、实体关系(Entity Relationship Method)、元对象装置(元模型或元-元模型)(Meta Object Facility)、模型驱动(Model Driver Architecture)、信息本体(Information Ontology)、数据字典(Data Dictionary)等方法技术;可采用的表示法或建模语言很多,如文字、表格、图形、ER图(Entity Relationship Diagram)、统一建模语言(UML:Unified Modeling Language)、可扩展标记语言(XML:EXtensible Markup Language)等图形符号或描述语言;可借助的建模工具软件很多,如Microsoft Word、Microsoft Visio、IBM Rational Rose、Sybase PowerDesigner、AllFuusin ERwin Data Modeler等软件工具。在中国矿产资源潜力评价数据模型研制过程中,依据本次数据建模工作的复杂程度、参与建模人员的计算机知识水平,同时考虑数据模型主要面向地质人员等实际情况,笔者采用最常用的方法技术、易交流的表示法和易获取的辅助建模软件工具开展中国矿产资源潜力评价数据模型研制。下面以中国矿产资源潜力评价数据模型研制过程中产生的阶段性重要成果“专业谱系及特征分类”、“图件数据模型”、“成果数据库结构模型”等为例进行说明。

3.1 专业谱系及特征分类

本体是一个源于哲学的概念,意指关于存在及其本质和规律的学说,后来被引入到信息科学并赋予新的含义,因此,在信息科学领域,本体又称“信息本体”。目前信息科学领域已普遍认为:信息本体是一个领域里共享的概念化模型的形式化和显式的规范说明^[2]。依照本体依赖领域程度,本体可分顶级本体(描述最普遍的概念及概念之间的关系,如空间、时间、事件、行为等,与具体应用无关,其他本体都是该类本体的特例)、领域本体(描述特定领域的概念及概念之间的关系)、任务本体(描述特定任务或行为的概念及概念之间的关系)、应用本体(描述依赖于特定领域和任务的概念及概念之间的关系)。依照本体具备推理功能程度,本体可分轻量级本体(不具备逻辑推理功能)、中量级本体(具有简单逻辑推理功能、系统可以识别一阶谓词逻辑的表达式)和重量级本体(具有复杂逻辑推理功能、系统可以识别更加复杂的二阶谓词逻辑的表达式,并为更加复杂的推理功能的实现预留了接口)。

地质领域本体(Geological Domain Ontology),属于一种领域本体(Domain Ontology),是关于地质领域实体概念及关系、领域活动以及该领域所具有的特性和规律的一种形式化描述。建立地质领域本体的过程,是一个浩大、繁琐、耗时、持续的建模过程。理论上,任何具体地质项目的数据模型只是地质领域本体的一个子集或局部,或者只是地质领域本体中与具体地质项目相关部分内容的裁剪、简化、调整,即地质领域本体的一个具体应用而已。

地质领域专业谱系及特征分类是地质领域相关专业及特征的划分、分级、分类及其编码体系,是地质领域本体关于概念化模型的简化与应用。在中国矿产资源潜力评价数据模型中,依据实际需要,专业谱系只划分到五级专业,五级专业划分之下再标识特征分类,将五级专业划分与特征分类按图3所示的代码结构进行统一编码。

基于地质领域本体固有的共享特性和简化处理,专业谱系与特征分类及其编码体系既简单实用又不失共享性,在中国矿产资源潜力评价数据模型中起到了两个重要作用:一是作为数据模型内各组成元素的核心,起骨架支撑作用,二是作为分类标识依据,使计算机识别各类成果数据。

3.2 图件数据模型

依据面向对象分析方法,提炼出与图件相关的各类基本对象(或对象类):图件、图件空间数据集、图件属性数据集、图层、属性表、图元、特征、属性项、空间划分方式、坐标参照系、比例尺、图式、图例、…

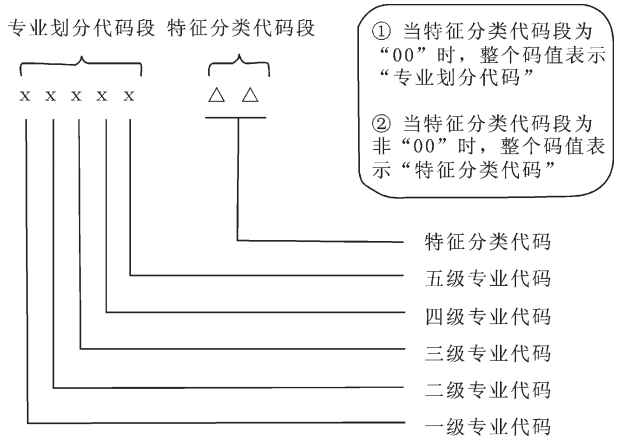


图 3 专业谱系划分及特征分类代码结构
Fig.3 The code structure of professional pedigree division and feature classification

等,以及这些基本对象的若干属性项,如图层对象的基本属性项有图层名称、图层代码、...等,构成了图件数据模型的最基本要素。专业谱系与特征分类、及其编码体系起分类和标识等基础作用。图件由图件空间数据集与图件属性数据集构成,图件空间数据集由若干图层构成、图层由若干图元构成、图元由图形属性与空间属性构成,图件属性数据集类由若干

属性表构成、属性表由若干特征构成、特征由若干专业属性构成,依照特征代码、图元编号建立图元与特征关联。结合面向对象与实体关系等表示法,图件数据模型的概念框架见图 4。

为了专业人员使用方便,基于图件概要描述(图件中文名、图件计算机使用名、图件代码)和图件图层规定(逻辑图层类别、图层中文名、图层是否有属

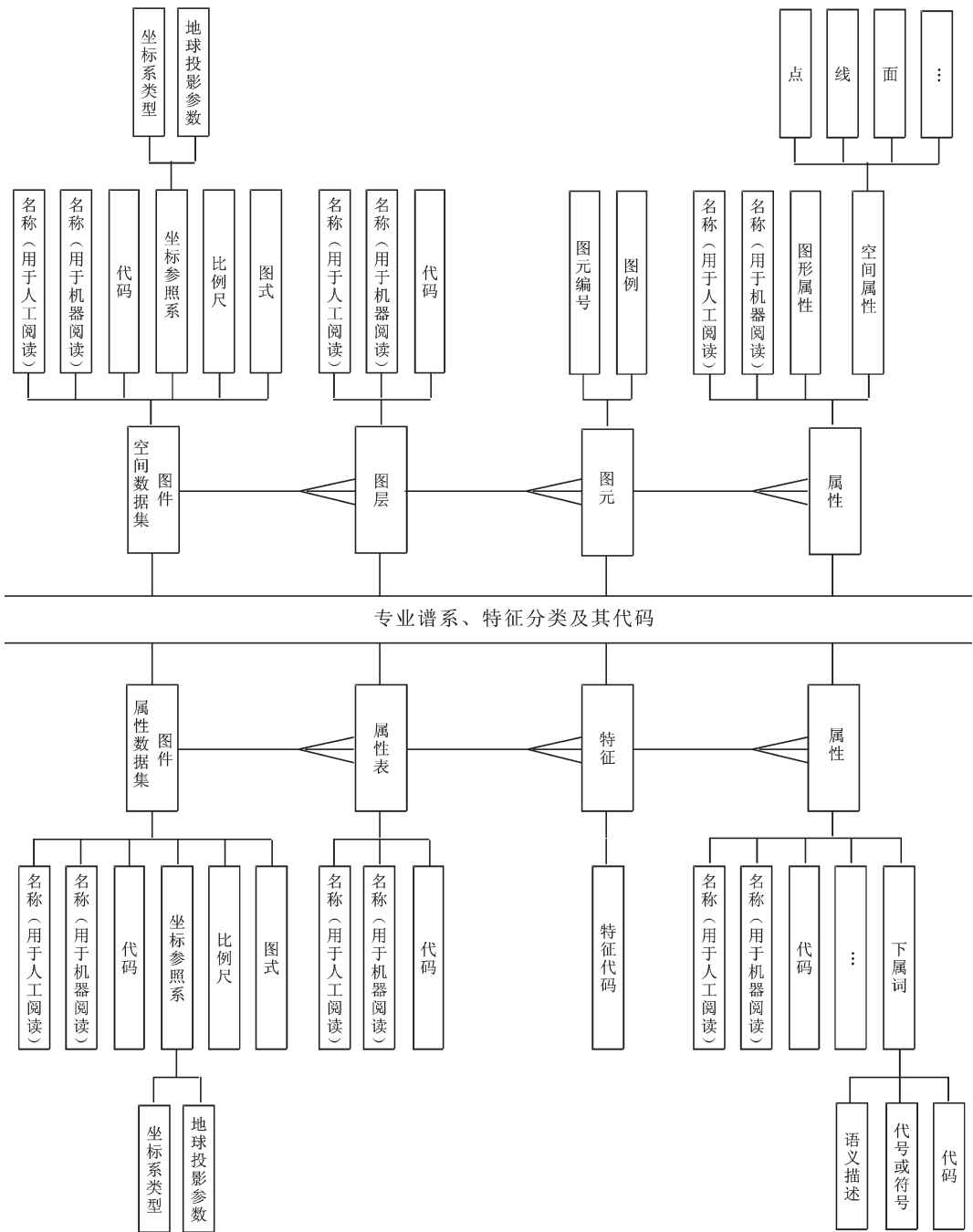


图 4 图件数据模型的概念框架

Fig.4 The concept framework of the map data model

性表、空间几何特性、图层计算机使用名、图层代码、图元及属性数据来源),图件的逻辑模型见表 1。

为了定义或描述特征分类(类)方便,基于特征分类名称、特征分类代码、特征分类概念或描述、特征分类属性结构规定(数据项名称、数据项代码、数据类型、存储长度、小数位数、约束条件、默认值、值域、值单位)以及数据项填写规定、数据项概念或描述、数据项填写示例,特征分类(类)的描述模型见表 2。

3.3 成果数据库结构模型

对复杂问题域 (Problem Domian), 基于 ISO/IEC 的 MOF(Meta Object Facility:元对象装置)思想^[9],从对象级(Object Level)(即对象、数据或信息等)、模型级(Model Level)、元模型级(Meta-model Level),到元-元模型级(Meta-metamodel Level)或更高级,采用逐级抽象方法,建立复杂问题域的层级模型。上元级模型定义或描述次元级模型,次元级模

表 1 图件的逻辑模型
Table 1 The logical model of the maps

| 图件概要描述 | | | | | | | |
|------------------------------|----------------|--|------------------|----------------|---|----------------------|-------------------|
| 图件中文名 | | 图件计算机使用名 | | | 图件代码 | | |
| 例： ××省（市、自治 区）××幅建造构造图 | | 例： MDZJZGZC×××× （按矿产资源潜力评价数据模型《通用代码规定》“图件命名规定”关于图件中文名称取四位汉语拼音首字母的规则填写“××××”） | | | 例： OFBB000C1A×××××××××0 （按矿产资源潜力评价数据模型《通用代码规定》“图件代码规定”的1:25万标准分幅空间范围码填写“×××××××××”） | | |
| 图件图层规定 | | | | | | | |
| 序号 | 逻辑图层类别 | 图层中文名 | 属性表 (有/ 无) | 空间 几何 特性 | 图层计算机 使用名 | 图层代码 | 图元及 属性数 据来源 |
| 1 | 例：建造构造图层 | 例：沉积岩建造 | 有 | 面 | LDZOFBB001 | OFBB001C1A×××××××××0 | XXXX |
| 2 | | 例：火山岩性岩相 | 有 | 面 | LDZOFBB002 | OFBB002C1A×××××××××0 | XXXX |
| 3 | | 例：火山构造 | 有 | 面 | LDZOFBB011 | OFBB011C1A×××××××××0 | XXXX |
| 4 | | 例：侵入岩 | 有 | 面 | LDZOFBB003 | OFBB003C1A×××××××××0 | XXXX |
| 5 | | 例：构造岩浆带 | 有 | 面 | LDZOFBB010 | OFBB010C1A×××××××××0 | XXXX |
| 6 | | 例：变质岩建造 | 有 | 面 | LDZOFBB004 | OFBB004C1A×××××××××0 | XXXX |
| 7 | | 例：大型变形构造 | 有 | 面 | LDZOFBB009 | OFBB009C1A×××××××××0 | XXXX |
| 8 | | 例：地质界线 | 有 | 线 | LDZOFBA002 | OFBA002C1A×××××××××0 | XXXX |
| 9 | | 例：断裂 | 有 | 线 | LDZOFBA003 | OFBA003C1A×××××××××0 | XXXX |
| 10 | | 例：韧性剪切带 | 有 | 面 | LDZOFBA004 | OFBA004C1A×××××××××0 | XXXX |
| 11 | | 例：褶皱 | 有 | 线 | LDZOFBA005 | OFBA005C1A×××××××××0 | XXXX |
| 12 | | 例：同位素年龄 | 有 | 点 | LDZOFBA008 | OFBA008C1A×××××××××0 | XXXX |
| 13 | | 例：产状要素 | 有 | 点 | LDZOFBA016 | OFBA016C1A×××××××××0 | XXXX |
| 14 | 例：重力推断构造图 层 | 重力推断构造图层 | 有 | X | XXXXXXXXXX | XXXXXXXXXX×××××××××0 | XXXX |
| 15 | 例：磁测推断构造图 层 | 磁测推断构造图层 | 有 | X | XXXXXXXXXX | XXXXXXXXXX×××××××××0 | XXXX |
| 16 | 例：地化推断构造图 层 | 地化推断构造图层 | 有 | X | XXXXXXXXXX | XXXXXXXXXX×××××××××0 | XXXX |
| 17 | 例：遥感推断构造图 层 | 遥感推断构造图层 | 有 | X | XXXXXXXXXX | XXXXXXXXXX×××××××××0 | XXXX |
| 18 | 例：地理图层 | 相关地理图层 | X | X | XXXXXXXXXX | XXXXXXXXXX×××××××××0 | XXXX |
| 19 | 例：整饰图层 | 相关整饰图层 | X | X | XXXXXXXXXX | XXXXXXXXXX×××××××××0 | XXXX |
| 20 | 例：非整饰图层 | 相关非整饰图层 | X | X | XXXXXXXXXX | XXXXXXXXXX×××××××××0 | XXXX |

表 2 特征分类描述模型
Table 2 The model for defining or describing the feature classification

| 特征分类名称:XX | | | | 特征分类代码:×××××××× | | | | | 特征分类概念或描述 XX | | | |
|-----------|-------|-----------|------|-----------------|------|----------|-----|-----|--------------|---------|------------|---------|
| 序号 | 数据项名称 | 数据项代码 | 数据类型 | 存储长度 | 小数位数 | 约束条件 | 默认值 | 值域 | 值单位 | 数据项填写规定 | 数据项词语概念或描述 | 数据项填写示例 |
| 1 | 特征代码 | FEATUREID | C | 26 | | NOT NULL | | | | | | |
| 2 | 图元编号 | CHFCAC | C | 6 | ... | NOT NULL | | | | | | |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |

型是上元级模型的实例化。从上至下,模型内容愈来愈丰富具体,模型定义能力愈来愈弱,即模型描述范围愈来愈窄;从下至上,模型形式愈来愈精炼抽象,模型定义能力愈来愈强,即模型描述范围愈来愈广(图 5)。对同一类型问题域,可以建立相同的元模型,使用元数据描述问题域的模型。不同类型的问题域,会有不同的元模型。通过建立元-元模型,使用元-元数据描述元模型,从而实现对不同类型的问题域进行描述。通过 MOF 思想或方法,一方面,有效控制复杂问题域的建模复杂度,另一方面,基于问题域的多层级模型形式描述以及多层级模型之间模型驱动机制,有力支持复杂问题域软件系统设计、多层次集成、模型重用,提高复杂问题域模型快速适用需求变化的能力。

中国矿产资源潜力评价专项所获得的研究成果可分为三个层次:省区级(包括 30 个省(市、自治区)、大区(或地区)级(包括华北、东北、华东、中南、西南、西北)和全国级。省区级成果属于矿产资源潜力评价资料性成果,用于大区范围综合研究;大区级成果,用于全国范围综合研究;全国级成果,代表全国范围、大规模、宏观规律性研究成果,3 个层次研究成果涵盖专业包括成矿地质背景研究、成矿规律研究、矿产预测研究、重力资料应用、磁测资料应用、化探资料应用、遥感资料应用、自然重砂资料应用(图 6)。

中国矿产资源潜力评价专项研究成果数据,从空间上,分为 3 个层次:省区级(省区级包括全省范围、预测工作区、典型矿床研究区三级)、大区(或地区)级和全国级;从专业上,涵盖专业领域:成矿地质背景研究、成矿规律研究、矿产预测研究以及重力资料应用、磁测资料应用、化探资料应用、遥感资料应用、自然重砂资料应用等。通过抽象,使用“成果图

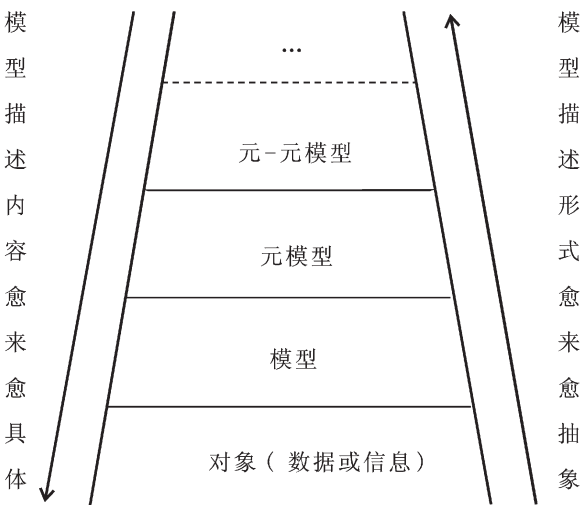


图 5 四层级或多层级元数据模型架构
Fig.5 Four or more layer metadata architecture

件”、“空间实体”、“遥感影像”、“文档资料”、“数据元数据”、“相关基础地质数据”、“系统元数据”等概念描述上述研究成果数据。成果图件,即满足中国矿产资源潜力评价专项相应专业编图技术要求(包括图面要求)和图件数据模型要求的图件;空间实体,即从成果图件或其他相关图件或数据中析取的具有地理空间属性和专业属性的地质要素;遥感影像,即本专项遥感资料应用的遥感影像图件;文档资料,即本专项所有报告、编图说明书、表格或卡片等;数据元数据,即成果图件的元数据以及其他相关成果数据元数据等;相关基础地质数据,即与矿产资源潜力评价工作相关的基础地质数据库,例如:1:5 万/1:20 万/1:25 万/1:50 万地质图库、矿产地数据库、地质工作程度数据库、…等;系统元数据,即指描述成果图件数据库数据字典、空间实体数据库数据字典、遥感影像数据库数据字典、文档资料数据库数据字典、数据元数据数据库数据字典、数据项下属词数据字典、

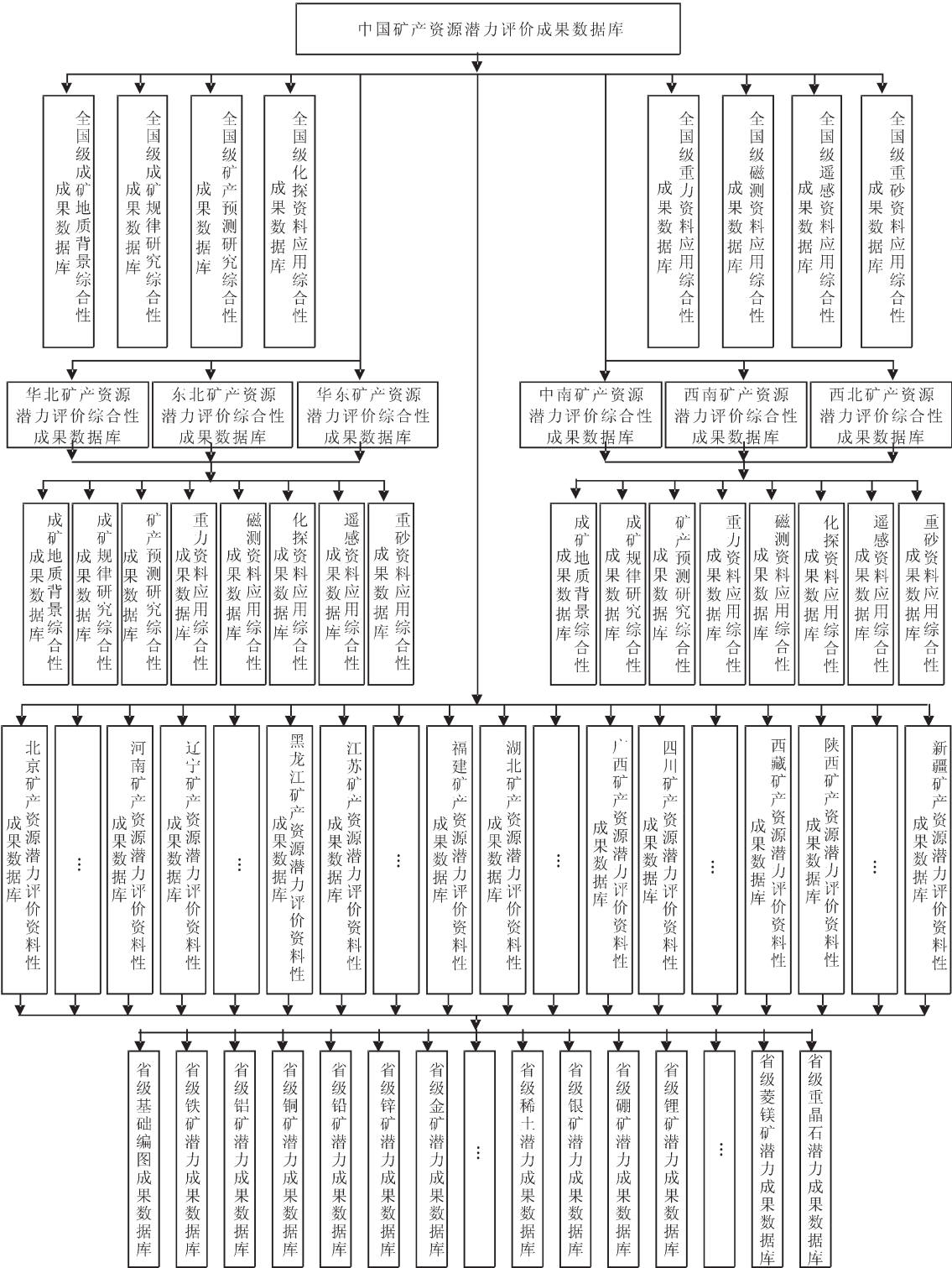


图 6 中国矿产资源潜力评价数据库内容概念框架

Fig.6 The concept framework of China's mineral resources potential evaluation database contents

相关通用代码规定信息、以及用户信息、权限信息、构件或组件信息、应用系统信息、应用管理与配置信息、...等等(图 7)。

基于上述“成果图件”、“空间实体”、“遥感影像”、“文档资料”、“数据元数据”、“相关基础地质数据”、“系统元数据”等概念抽象,中国矿产资源潜力

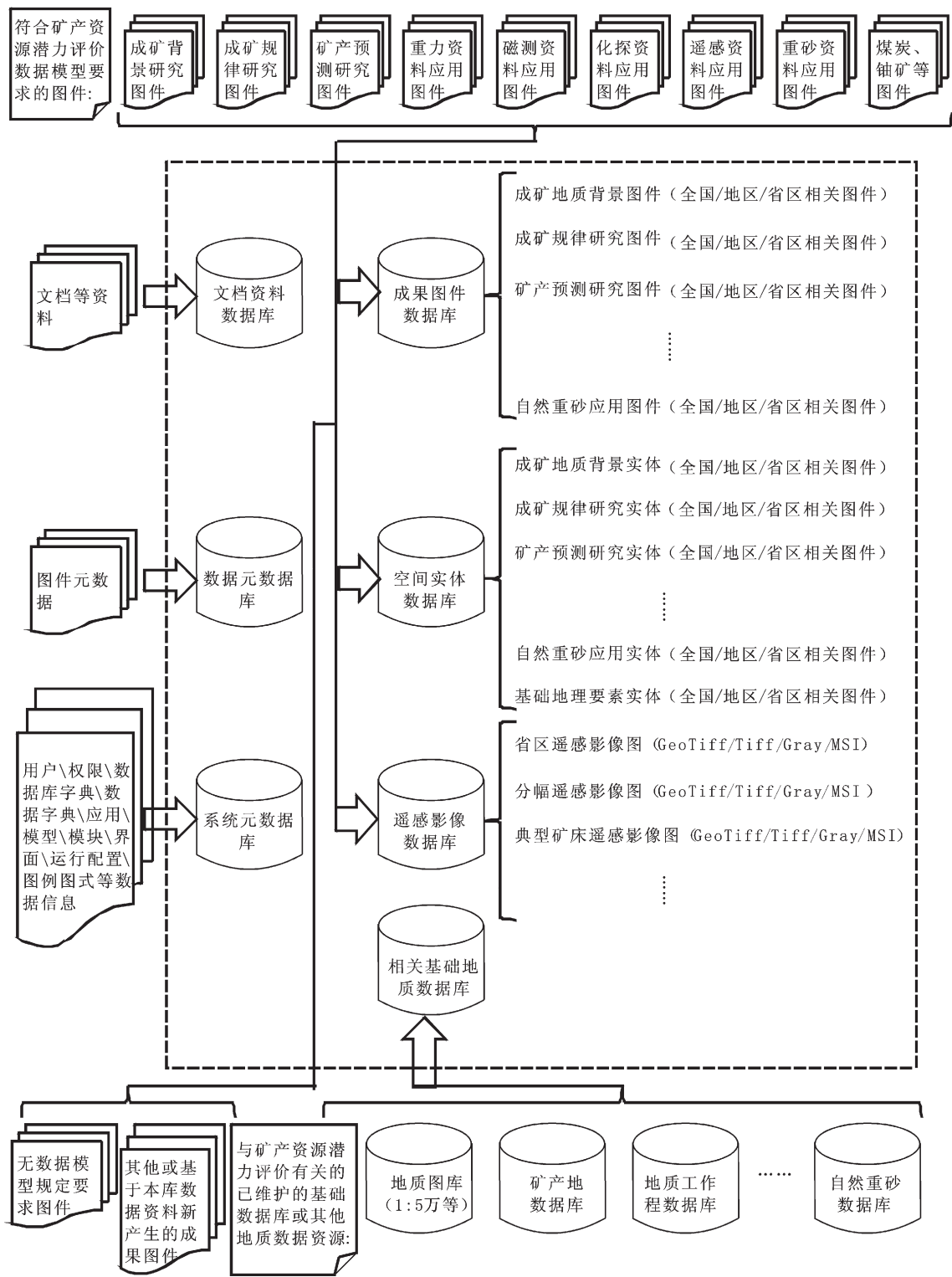


图 7 中国矿产资源潜力评价数据库内容逻辑构成

Fig.7 The logic formation of China's mineral resources potential evaluation database contents

评价成果数据库内容主要由阶段性研究已成形的图件(即已研究成形的地质数据产品:成果图件、遥感影像图件)及图件编图说明书、可持续维护和使用的具有地理空间属性和地质专业属性的地质实体、图件数据元数据、相关专题文档或成果报告、以及相关

基础地质数据、...等等组成,因此,中国矿产资源潜力评价成果数据库,既不同于传统意义上的地质图件数据库,也不同于现代意义上的地理空间数据库,更不同于图件数据库与空间数据库的简单相加。

中国矿产资源潜力评价成果数据库,对成果图

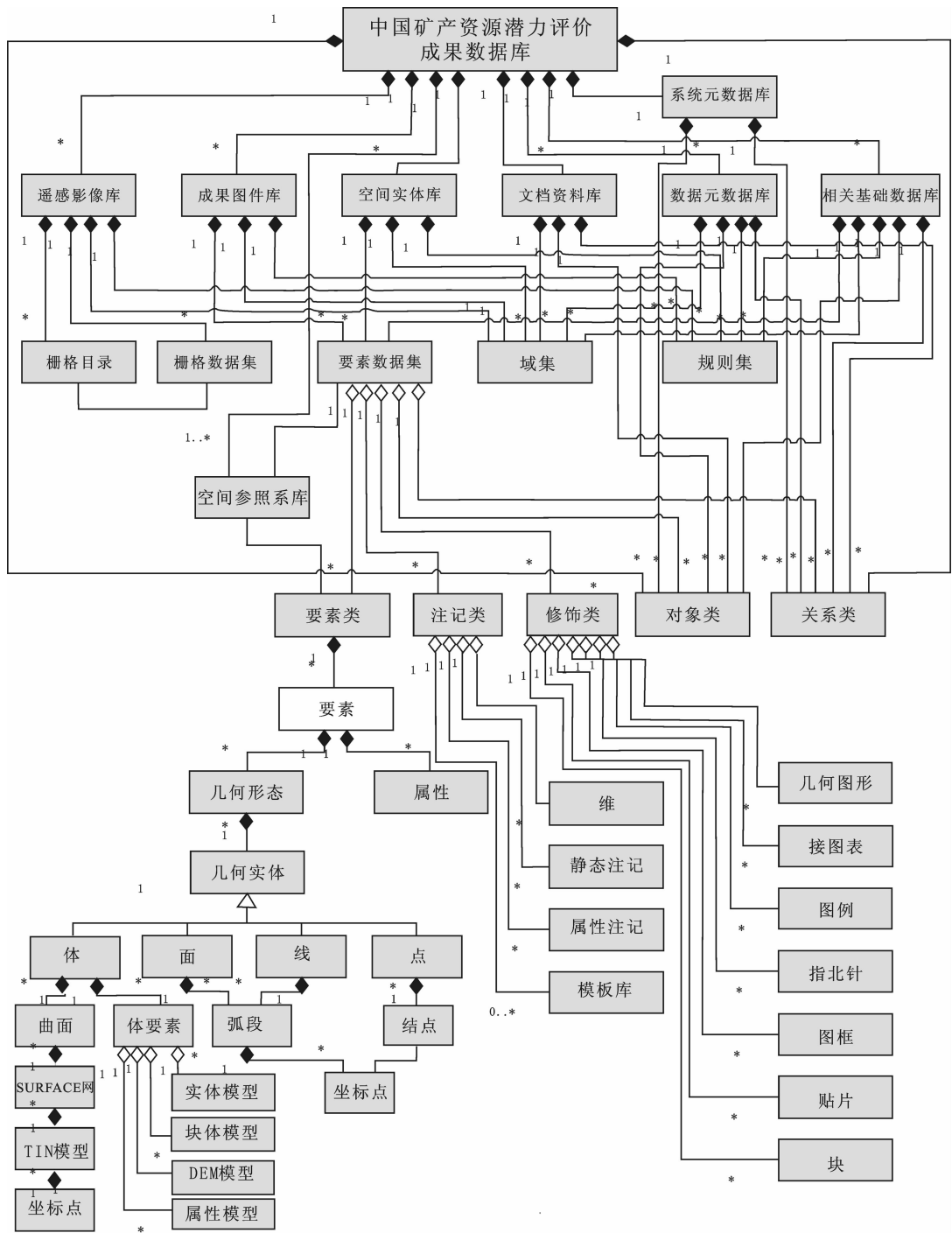


图 8 中国矿产资源潜力评价数据库内容基本结构模型

Fig.8 The basic structure model of China's mineral resources potential evaluation database contents

件管理的支撑,既要原样入库管理,又要原样回放或输出,并实现不同研究阶段或不同研究者的成果图件的版本管理,还要支持基于图件、图层、图元、属性、空间、矿种、专题等条件的检索与裁剪。对空间实体管理的支撑,既要支持不同专题、不同尺度空间实体的管理(包括版本管理),又要支持空间实体与图形表示的动态匹配的显示操作或简单输出,还要支持基于空间、尺度、属性、专题等条件的检索与裁剪。对遥感影像管理的支撑,既要原样入库管理(包括版本管理),又要原样回放或输出,还要支持与空间、矿种等条件的检索与关联。对文档资料、数据元数据管理的支撑,既要入库管理,还要支持与图件、矿种、专题等条件的检索与关联。对相关基础地质数据管理的支撑,主要支持与相关内容的集成处理。简言之,中国矿产资源潜力评价成果数据库,既要支持矿产资源潜力评价阶段性成果的管理与服务,还要支持可持续发展的、动态化的矿产资源潜力评价研究工作。

因此,基于中国矿产资源潜力评价成果数据库内容特点和应用要求的分析与抽象,结合“面向实体的空间数据模型”^[4]研究成果中关于“要素数据集”、“栅格数据集”、“域集”、“规则集”、“要素类”、“对象类”、“关系类”、“注记类”、“修饰类”、“要素”、“几何形态”、“几何实体”、“属性”等概念和模型,设计出中国矿产资源潜力评价成果数据库内容基本结构模型,如图 8 所示。

4 成果应用

基于上述研制流程及方法技术取得的中国矿产资源潜力评价数据模型的图件数据模型成果及研发的配套软件工具(GeoMAG、GeoTOK、GeoPEX、GeoO35 等),已广泛应用到中国矿产资源潜力评价专项的 30 个省(市、自治区)级矿产资源潜力评价工作项目和 6 个地区级矿产资源潜力评价工作项目以及 10 个全国级矿产资源潜力评价专业汇总工作项目,有力而稳步地推进了中国矿产资源潜力评价专项工作,保证了研究成果的规范性。

基于上述成果数据库基本结构模型研发的中国矿产资源潜力评价成果数据集成平台 DipMopa,即将正式投入使用。

另外,图件数据模型及技术方法等部分成果已应用到“全国矿产资源利用现状调查专项”、“全国危

机矿山接替资源找矿专项”工作中。

5 结 语

本文以中国矿产资源潜力评价数据建模为实例,系统地提出了一种研制与建立复杂问题域数据模型的研制流程和实施策略,并以矿产资源潜力评价领域“专业谱系及特征分类”、“图件数据模型”、“成果数据库结构模型”等重要阶段性建模成果为例,简要介绍了相关设计思想、采用的方法技术。通过实践验证了所采用的研制流程和实施策略及方法技术的有效性、可行性、实用性和代表性,有一定借鉴意义和推广价值。

该成果应该说还只是相关研究工作的初步尝试,存在诸多需要深入探讨和研究之处。今后应进一步优化数据模型研制流程,深入开展每一步建模表示的形式化研究、每一步建模结果的语义化研究、每一步建模结果之间转换或驱动机制研究,建立一套集成化的软件工具体系和规范,用于复杂问题域数据模型分析、设计、维护、转换、迁移,一体化、高效、持续地支撑超大型地质项目成果数据库资源建设,持续高效地开展多样化地质数据产品开发和提供地质资料应用服务。

致谢:该研究工作得到严光生、谭永杰、王全明、李裕伟、姜作勤、李景朝、肖庆辉、张克信、张德全、陈辉、王瑞江、肖克炎等研究员的鼓励和指导,特致谢意。

参考文献(References):

- [1] 叶天竺,陈毓川,张洪涛,等.全国矿产资源潜力评价总体实施方案[R].全国矿产资源潜力评价项目办公室,2006.
Ye Tianzhu, Chen Yuchuan, Zhang Hongtao, et al. The Overall Implementation Plan of China's National Mineral Resources Potential Evaluation [R]. The Bureau of China's National Mineral Resources Potential Evaluation Project, 2006(in Chinese).
- [2] Gruber T. Towards principles for the design of ontologies used for knowledge sharing [J]. International Journal of Human-Computer Studies, 1995, 43(5/6):907-928.
- [3] ISO/IEC 19502-2005, Information technology - Meta Object Facility (MOF), First edition 2005-11-01.
- [4] 叶亚琴,左泽均,陈波.面向实体的空间数据模型 [J]. 地球科学——中国地质大学学报, 2006, 31(5):595-599.
Ye Yaqin, Zuo Zejun, Chen Bo. Orient-entity spatial data model [J]. Earth Science —— Journal of China University of Geosciences, 2006, 31(5):595-599(in Chinese with English abstract).

The development process and technique of the mineral resources potential evaluation data model in China

ZUO Qun-chao, YANG Dong-lai, YE Tian-zhu

(Development Research Center of China Geological Survey, Beijing 100037, China)

Abstract: Exemplified by data modeling of China's mineral resources potential evaluation special project, this paper put forward the development process and implementation strategy of the data modeling for the very complex problem domain, systematically illuminated the essential contents of the development process and implementation strategy of the data modeling, briefly introduced the design idea, the methods and techniques, and the related research achievements. Practices have verified the effectiveness, feasibility, practicability and representativeness of the development process, methods and techniques, and the results obtained have certain reference significance and application value. The achievements are especially suitable for large quantities of geological project data resources construction characterized by super-large, multi-professional, engineering and spatial-oriented data.

Key words: mineral resources potential evaluation; data model; data modeling; ontology; meta model; object-oriented technology; model driven; problem domain

About the first author: ZUO Qun-chao, male, born in 1964, senior engineer, engages in mineral resources potential evaluation, geological space data integration, and software development and platform design; E-mail: zuoqc163@163.com.