

文章编号: 1001-8166(2005)02-0166-07

地学数据产品的开发、发布与共享*

廖顺宝, 孙九林, 李泽辉, 马琳, 彭梅

(中国科学院地理科学与资源研究所, 北京 100101)

摘要:数据是科学研究的基础,数据共享可以最大程度地发挥数据的使用价值,数据是实现数据共享最基本的要素。地学数据具有空间性、综合性、时间性、海量性、多源性等特点。属性数据、遥感数据、矢量数据是加工生产地学数据产品的重要数据源,属性数据空间化是加工、生产地学数据产品的重要手段。地学数据共享发布平台应当具备用户管理、数据目录查询、元数据管理、数据查询与浏览、数据下载等基本功能。推进科学数据共享,必须要有相应的政策措施保证,必须建立公正、合理的数据工作评价体系。

关键词:地学;数据;开发;共享

中图分类号: TP75 **文献标识码:** A

0 引言

数据是科学研究的基础,科研成果的正确性与研究所使用的数据密切相关。科学数据,特别是地学数据的获取往往需要投入大量的人力、物力和财力。因此,地学数据的交换与共享尤为必要^[1]。通过数据共享,不仅能够最大程度地发挥数据的价值,而且还能避免重复投入,为国家节省开支。要实现数据共享,除政策、标准、技术问题外,最核心的是必须收集、加工、整理和开发成系列的数据产品,否则数据共享就成了无米之炊。

数据既是研究的基础,往往又是研究成果的一部分,研究所获得的数据也是数据共享的重要资源。但每个研究项目都有自己的针对性,通过各种研究项目所积累的数据往往系统性不强,难以完全支持其它项目的研究。因此,必须瞄准当前地学研究的前沿领域和问题,有目标、有针对性地开发一批在时间、空间和专题内容上成系列的地学数据产品,才能满足地学工作者的研究需要。

1 地学数据的特点

地学(Geoscience)是复杂科学,它与地球科学(Earth Science)是同义词,包括研究地球的全部学科^[2]。其研究范围涉及大气圈、水圈、岩石圈、生物圈等各个圈层,完全处于开放状态,可控性较差。在一个地区存在的规律在另外一个地区可能就不存在,区域性很强。因此,支持地学研究的科学数据也与支持其它科学研究的数据具有十分不同的特点。概括起来,地学数据主要具有空间性、综合性、时间性、信息海量、多源性等特点,数据正在成为地学研究的重要驱动力。

(1)地学数据的空间性、综合性。地学数据所描述的客观对象是地球系统的一部分,它具有明确的空间坐标(经度、纬度和海拔高度)且有一定的空间范围。以地理学和资源环境科学为例,它们研究的对象主要是地球表层,数据描述的客观对象可以是地表一定大小的规则单元,如:30 m × 30 m、1 km × 1 km等,也可以是不同级别的行政区域,这些对象均

* 收稿日期: 2003-11-10; 修回日期: 2004-06-03

* 基金项目: 国家科学数据共享工程项目“中国地球系统科学数据共享服务网建设”(编号: 2003DEA2C010); 中国科学院知识创新工程项目“中国自然资源数据库”(编号: NF105-SDB-1-18)资助。

作者简介: 廖顺宝(1966-), 男, 四川德阳人, 副研究员, 主要从事遥感与地理信息系统应用以及地学数据产品开发方面的研究。

E-mail: liaosb@igsnrr.ac.cn

具有明确的空间坐标和空间范围。在数据所描述对象的内部,对象特征在空间上的分布是不均匀的,数据往往是对客观对象内要素特征的综合表述,是一种“统计值”或“综合值”。随着对象空间单元的变大(即比例尺变小或空间分辨率降低),数据的综合性、概略性、宏观性更加明显。

(2) 地学数据的时间性。地球系统是一个动态变化的系统。但它的变化速度常常十分缓慢,要科学地揭示和描述其变化的规律,必须要有长时间序列的数据支持,而且往往是时间序列越长,数据越有价值,因为许多变化过程是不可逆的。

(3) 地学数据的海量性。地学数据描述的是地球上一定的空间范围,随着描述单元的变小(比例尺变大或分辨率提高)和时间序列的变长,数据量会急剧上升。现代遥感技术为地学数据的获取提供了强有力的技术手段,数据精度(分辨率)的提高使数据更具开发应用潜力,但同时数据量也呈几何级数增长。从一定意义上说,由于遥感技术的迅猛发展,目前人们能够获得的地学数据源不是太少,而是太多。为了及时处理这些数据,必须借助于现代信息技术,如图像处理技术、地理信息系统技术、全球定位系统以及数据库技术。

(4) 地学数据的多源性。地学数据的获取有多种途径,目前最普遍的方法是通过遥感,(航天)遥感具有廉价、周期性和实时性等特点,十分有利于陆地表层的动态监测。此外,野外考察、长期定位观测等也是地学数据获取的重要手段。这些不同的数据获取手段相辅相成,遥感不仅可以减少资源环境调查的工作量而且还可以满足时间序列要求,实地考察和采样又可以提高遥感数据的判读精度。

(5) 数据正在成为深化对地球认识的知识“窗口”。遥感技术的迅猛发展使人们获得高时间分辨率和高空间分辨率的对地观测数据已不再是一种奢望,而已经变成了现实,地学数据被极大地丰富和完善。有了这些高分辨率、时间序列的数据,以前只能是科学家的一些梦想,如今已经通过一些新的研究课题正在成为现实。数据不再仅仅是科学研究的副产品或附属品,它正在成为深化对地球认识的知识“窗口”。

2 地学数据产品的开发

地学数据的品种复杂、种类众多,但从数据产品开发的角度来看,主要是按数据的来源渠道和数据的存储模式进行划分,可分为 3 大类,即:基于属性

数据的地学数据产品、基于遥感信息源的地学数据产品和基于矢量地图的数据产品。

2.1 基于属性数据的数据产品

地学,特别是地理学和资源环境科学中的属性数据主要由统计数据 and 观测数据两部分组成,前者主要描述人文要素,如人口、社会经济统计数据等,也包括一些从专题地图中提取的专题信息,如各种土地资源、土地利用面积等;后者主要描述自然要素,如气象要素、水文要素、生态要素等。

属性数据是非常重要的数据资源。目前我们所获得的各种属性数据都具有综合性(统计数据)和抽样性(观测数据),这也是属性数据进行空间分析时所表现出来的缺陷。综合性主要表现为:以行政区为单元的各类统计数据,如人口、GDP,表示的是该行政区的总数,而不能反映这些要素在行政区内部是如何分布的,另外,由于行政区划的变动使数据的时间序列分析产生困难。抽样性主要表现为:各种定位观测站仅仅布设在一些有代表性的地区,所获得的观测数据只能代表观测站附近很有限的范围,广大无站点地区的数据不能直接获得。属性数据空间化^[3-5]正在成为解决以上问题的有效手段。

属性数据空间化可以解决:反映社会经济要素在统计区域内部的实际分布;插补无观测站地区的观测要素值;使按时间序列进行分析成为可能,时间序列分析是地学、社会经济研究的重要手段,由于行政区划的变动,虽然不同年份行政区名称和代码完全相同,但如果其所辖政区范围发生了变化,则在时间序列上不具有可比性,而属性数据空间化是把属性要素尽可能按照要素的实际分布情况重新分配到规则的网格上去,因此在不同的时间断面上具有可比性。

图 1、2 是属性数据空间化的 2 个具体案例。人口分布与土地利用之间具有十分密切的关系,图 1 是基于土地利用生成的全国 1 km × 1 km 分辨率的人口分布数据,该人口分布图既保持了各市县总人口与统计数据的一致,又反映了各市县内部人口分布的空间变化,信息量明显增加。而传统的统计制图则无法达到这样的效果,因为统计制图只能把同一行政区内的统计要素看成是均匀分布来处理。

图 2 是基于中国 600 多个气象观测站,通过“多元回归+内插残差”^[5]的方法计算获得的约 1 km × 1 km 分辨率的中国 30 年的年平均气温数据,该数据集所包含的信息量已远远超过了原始的属性数据的信息量,因为有气温数据的记录已经从 600 多个

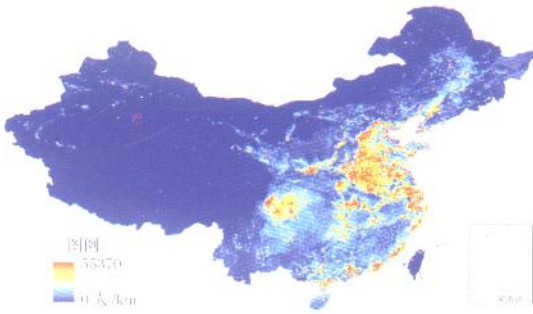


图 1 中国 1 km²分辨率人口分布图 (2000 年)

Fig 1 Map of population distribution of China at 1km by 1km in 2000

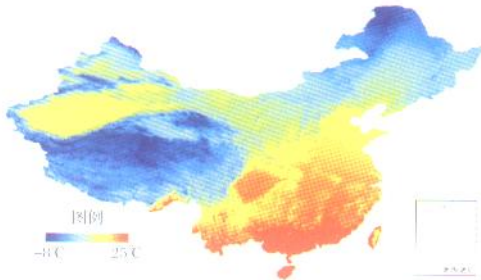


图 2 中国 30 年 (1951—1980) 平均气温分布

Fig 2 Map of 30-year average temperature (from 1951 to 1980) in China

气象站外推到全国。

属性数据可以以 2 种方式提供用户使用:一是直接提供,主要供宏观分析使用、不需要进行准确定位的用户;二是提供空间化的数据产品,主要为进行空间分析的用户提供数据。

2.2 基于遥感的 数据产品

遥感,特别是航天遥感,具有覆盖范围大、周期性和廉价等特点,是获取地表信息的重要技术手段。目前,遥感技术已在测绘、农业、林业、畜牧业、环境监测、地质找矿、考古、旅游、军事等领域得到了广泛的应用,产生了巨大的经济效益和社会效益。遥感技术是调查、监测和分析人类面临的人口、环境、资源和灾害四大难题的最好手段之一^[6]。

虽然通过遥感技术可以获得地面大范围、甚至全球范围内的信息,但传感器收集到的是被探测地物的电磁波信息,而并不是一种反映地面资源环境状况、具有明确地学意义的信息。要从电磁波信息中提取有地学意义的信息,如植被状况、土地利用状况、土壤湿度、地表温度等,必须通过数字图像处理软件和相应的算法才能实现^[7]。一般地,从遥感的原始波段信息获得有地学意义的信息需要经过从图像预处理(辐射纠正、波段配准、几何纠正)到图像分类再到专题图生成的过程。图 3、图 4 分别是基

于 NOAA-AVHRR 的归一化植被指数 (NDVI) 和地温数据。

基于遥感信息源加工生产的数据是地学数据产品的重要组成部分,而用不同遥感平台和传感器获得的信息源生产的数据产品也有不同的特点。

(1) Landsat TM 数据。Landsat TM 数据具有几何分辨率高 (30 m)、重复覆盖地表同一区域的周期长 (16 天) 和价格昂贵等特点,其数据主要用于范围相对较小、非动态监测方面的研究,如土地利用分类。

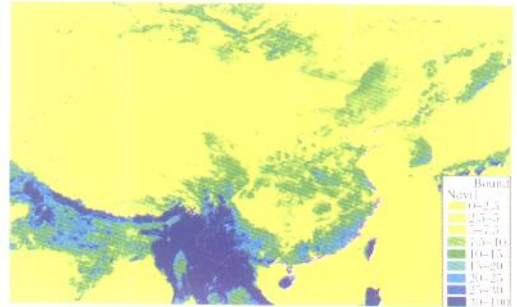


图 3 基于 NOAA-AVHRR 的中国 NDVI 数据 (1998 年 1 月)

Fig 3 Map of NOAA-AVHRR based NDVI in China (January, 1998)

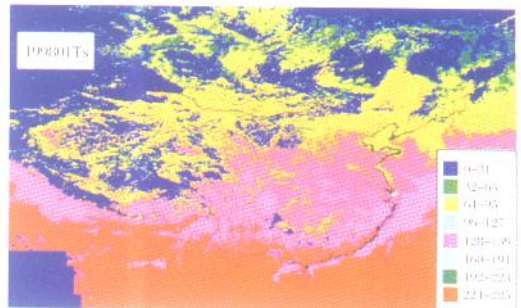


图 4 基于 NOAA-AVHRR 的中国地表温度数据 (1998 年 1 月)

Fig 4 Map of land surface temperature based on NOAA-AVHRR in China (January, 1998)

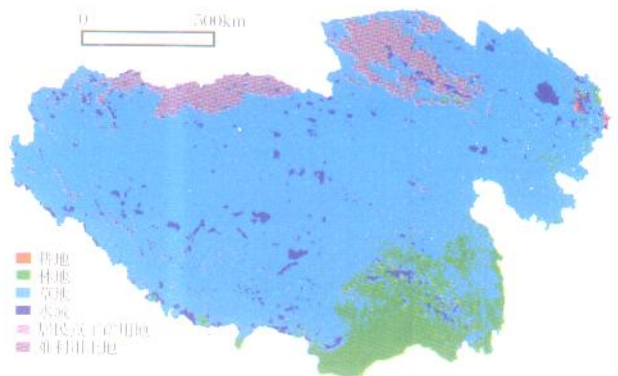


图 5 西藏、青海 2 省区的土地利用图 (90 年代)

Fig 5 Land use map of Tibet and Qinghai (1990s)

(2) NOAA-AVHRR 数据。NOAA-AVHRR 数据具有地面覆盖范围大、价格低廉、时间分辨率高、几何分辨率低(星下点 1.1 km)等特点,适宜于资源环境动态监测,如作物长势监测、森林火灾监测等。

(3) TERRA-MODIS 数据。中分辨率成像光谱仪 TERRA-MODIS 在保留 NOAA-AVHRR 功能的同时,在数据波段数目和数据应用范围、数据分辨率、数据接收和数据格式等方面都作了相当大的改进,这些改进使 MODIS 成为 AVHRR 的换代产品^[8]。其中最明显的改进为:AVHRR 的波段数为 5 个、几何分辨率分别为 1 100 m;MODIS 的波段数为 36 个、几何分辨率分别为 250 m (2 个波段)、500 m (5 个波段)和 1 000 m (29 个波段)。基于 MODIS 可以生产 44 种数据产品,涉及陆地、海洋、大气、云等。

由于具有较高的时间分辨率和廉价获取等特点,NOAA-AVHRR 和 TERRA-MODIS 是当前生产遥感地学数据产品的主要信息源。随着 MODIS 数据处理技术的不断发展和完善,在不久的将来,MODIS 将会成为遥感地学数据产品的主导信息源。

2.3 基于矢量地图的数据产品

地图是表现地学数据的重要手段,是地理学的第二语言。地图通过数字化、编辑、修改、拓扑关系生成、地图数据库的建立形成地图数据产品。按内容划分,地图主要分为普通地图和专题地图,普通地图表达地表的各种基本要素——水系、地貌、土质、植被、居民点、交通网等;专题地图突出反映一种或几种主要要素或现象^[9]。

(1) 普通地图数据产品。普通地图数据有时也称基础地理信息,由政府主管部门负责数据的采集、组织和管理,基础地理信息涉及国界、保密、国家安全等方面的问题,它是其它专题地图的基础底图,因此必须具有权威性和一致性,应由国家主管部门负责对外发布,其它任何部门和个人都无权发布。

(2) 专题地图数据产品。专题地图数据产品是数据生产者在普通地图(如境界)的基础上叠加了自己收集、加工和整理的其它专题要素,因此,它是基础地理信息与专题要素信息的集成。图 5 是西藏、青海 2 省区的土地利用图,它是在 2 省区境界(国界、省区界)的基础上叠加了土地利用信息。

在地图数据产品的生产中,一般的数据生产者只能生产自己专业领域的专题数据产品,而无权、也没有必要去生产由国家主管部门负责生产的基础地理数据产品,但专题数据产品必须与基础地图数据产品保持一致。

2.4 其它地学数据产品

主要指派生地学数据产品,如人口密度、人均 GDP、净初级生产力等。作为数据产品,必须是经过实质性加工、具有智力投入的成果。有的数据虽然表达形式变化了,但由于没有进行实质性加工和智力投入,并不能形成新的数据产品。例如,基于行政区划的统计制图不是一种数据产品,它只是统计数据的另一种表现方式,并没有增加信息量,从用户的角度而言,他们倾向于直接使用属性数据,而不是这种统计制图的结果。

3 地学数据产品的发布

3.1 数据发布平台的基本功能描述

信息技术、网络技术的飞速发展使互联网正在成为数据共享的主要手段。为满足广大地学工作者访问数据的需求,地学数据共享发布平台应当具备用户管理、数据目录查询、元数据管理、数据查询与浏览、数据下载等基本功能。

(1) 用户管理。应当具有普通用户注册及登录、系统管理员为用户设置权限以及限定访问资源的功能。数据共享并不总是完全、无偿和开放的,它往往在不同的层次和级别上进行。根据数据本身的对外开放程度(保密级别)和产生数据产品的投资渠道、方式,某种特定的数据产品具有其特定的共享范围和共享方式,如可以在多大范围内共享?如何共享?是完全免费还是只收取成本费?还是完全实行商业性数据共享?这些功能都需在用户管理模块中完成。科学、合理地对用户和数据进行分类分级,既可以最大程度、最大范围地实现数据共享,又有利于国家安全和知识产权保护。

(2) 数据目录查询。数据目录是数据共享系统的总体概要,是用户了解数据资源的窗口。数据目录的内容信息来自元数据库,但又并不是仅仅简单地从元数据库中提取、组合有关信息。数据目录必须基于用户熟悉、方便的地学数据分类系统,使用户通过数据目录可以很快地定位到自己所需要的数据内容。

地球科学的学科分类系统是地学数据分类的重要参考标准和依据,但又不能完全照搬学科分类标准,因为学科分类的目的和数据分类的目的毕竟是有差别的。例如,地球信息科学是地球科学下一个公认的分支学科,但如果把地球信息科学也作为地学数据分类的一个层次则明显地不合理,因为地球信息科学是有关地学研究技术、方法的学科,并没有

特定的数据对象,因此不存在属于地球信息科学的特定的数据资源。

(3) 元数据管理。元数据是对数据进行描述和说明的数据,即所谓数据的数据。通过元数据查询检索系统,用户可以知道那些数据可以共享、共享的方式和范围、数据的产生背景、数据质量、数据格式、数据量大小等信息。从而决定是否要使用其该数据产品,是否向数据管理员提出使用数据的申请。

除查询检索功能外,元数据管理系统还应具备元数据输入、编辑修改功能。因为,数据提供者随时可能有新的数据集上网共享,需要不断地向元数据库中追加新的记录,同时,已有的元数据信息随着时间的推移也可能做某些修改。

(4) 数据查询与浏览。查询浏览与直接下载是数据共享的 2 种重要方式。政务、管理、商业等信息主要通过查询方式共享,科学数据则主要通过下载方式共享。尽管如此,查询浏览功能也是地学数据共享系统的重要组成部分。主要是因为,用户在下载数据之前通过预览数据可以对数据情况有进一步的了解;或者是用户为了节省数据下载过程中的网络传输时间,并不对整个数据集进行下载,而只是下载一个完整数据集的一部分内容。例如,对数据库中的数据只下载某些记录或者对空间数据集只下载一定空间范围的数据内容,这些都需要通过数据查询浏览功能实现。

对于用数据库管理的数据,一般通过 ODBC、JDBC 等数据库访问工具可以实现数据的查询和浏览;对于用文件方式存储的地理空间数据,则可以通过 WEBGIS 工具实现数据的浏览、查询以及数据子集的提取。

(5) 数据下载。通过网络下载数据有多种方式,FTP、HTTP 是 2 种最常用的下载方式。通过数据查询与浏览,确定所需要的数据集(或子集)后,则可实现数据的下载。作为一个完整的数据发布与共享系统,在进行数据下载的过程中,还应当同时记录与用户、数据有关的信息,如:用户名称、浏览器 IP 地址、下载时间、下载数据集名称、下载数据量等。以便于系统管理员对用户和数据进行监控。

3.2 地学数据发布平台功能结构图

图 6 是具有上述功能的地学数据发布、共享平台的流程结构图。

4 地学数据共享

数据、标准规范、技术平台、共享政策是科学数

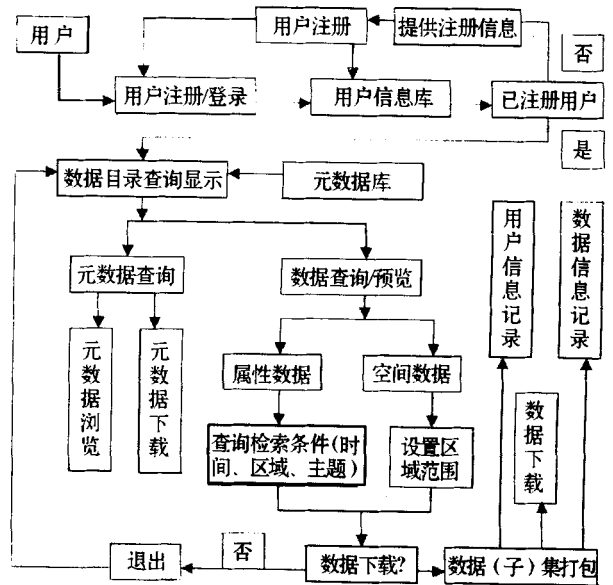


图 6 地学数据发布、共享平台流程结构图

Fig 6 Flow chart of structure and function of platform for geo-science data sharing

据共享系统的 4 大组成子系统,缺一不可。

(1) 数据是共享的基础。科学数据共享所共享的对象是数据,因此,离开了数据,数据共享就无从谈起。由于地学研究的复杂性以及地学数据的多样性,用于共享的地学数据必须是在时间、空间上成序列的数据产品,应当包括能基本上满足地学研究需要的各种类型的数据品种。同时,在数据产品的收集、整理、开发与集成过程中,数据质量控制工作举足轻重。GIS 是地学数据产品生产加工与表达的重要手段,有人认为,建立 GIS 的过程就是和误差作斗争的过程,GIS 产生的误差可以比喻成快速产生的垃圾,如对误差处理不当,GIS 能以相当快的速度产生各种垃圾^[10]。

在高度重视数据质量工作的同时,还必须重视元数据的编写与建库。元数据一方面是数据共享的重要组成部分,它更是数据产品质量的写照。

(2) 标准规范是实现共享的技术保证。同一数据产品可以通过不同的软件来开发和管理,在应用中,不同的用户也可能使用各自熟悉的软件来分析、使用数据。这就涉及到数据格式是否兼容的问题。一种数据产品,即使再有价值、数据质量再高,如果它不能被大多数用户方便地使用,其共享价值也就不存在了。因此,数据产品的标准化、规范化对于数据共享至关重要。与数据共享有关的标准规范很多,主要包括:数据分类标准、元数据标准、数据产品

质量标准、数据产品的格式标准等。

(3) 平台是数据共享的技术支持。共享平台是数据共享系统的界面,数据提供者、数据用户、数据管理者通过共享平台实现数据共享。数据的组织方式应当有利于平台的数据管理,但数据很难按照某种固定的专题或思路进行组织,因为,这样可能无法满足某些用户的需求或喜好,而不能满足另外一些用户的需求或喜好,通过数据组织满足所有用户的需求事实上是不可能的。用户的各种需求可以通过系统平台的功能来实现,例如,按照不同的分类标准对数据进行检索、对查询到的数据通过系统平台进行实时集成等。

(4) 数据共享政策是实现数据共享持续发展的政策保证。数据共享可以最大程度地发挥数据的使用价值,使开发数据资源的投资得到最有效的利用。数据不像物质产品,没有排它性。能否实现数据共享并使之持续发展,国家政策引导是很重要的。

对于国家投资产生的一般数据,应当及时向全社会无偿共享;对于国家投资、单位或个人具有较大智力投入的数据,可以设立一定的保护期,保护期内,由生产单位或个人独占,过了保护期以后,则应当向全社会公开。

必须尊重和保护数据提供者的知识产权。被提供服务者是数据的最终用户,对数据不具备散发权。同时,被提供服务者在其研究成果中要注明所使用数据的来源。

建立公正、合理的数据工作评价体系,进一步推进科学数据及其共享工作。数据是科学研究的基础,数据工作的独特性在于长期地进行数据积累,为科学研究与创新奠定基础,而其本身的研究和创新成分并不多。因此,数据工作有别于一般的科学研究,如果仅仅根据发表论文、获奖及专利来评价数据工作者的工作成绩必然会挫伤广大数据工作者的工作积极性。而必须建立符合数据工作特点的评价机制,如:数据的质量、数量、服务范围、用户满意度等。

参考文献 (References):

[1] Liao Shunbao. Scientific data sharing and its solutions based on in-

ternet[J]. *Science & Technology Review*, 2002, (1): 25-27. [廖顺宝. 科学数据共享及其在互联网上实现的技术途径[J]. 科技导报, 2002, (1): 25-27.]

[2] Sun Shu. Learn new things by reviewing the past, carry on the past and open a way for future-A special issue of review and forecast on subjects related to geoscience (preface) [J]. *Advances in Earth Science*, 2001, 16(5): 597-598. [孙枢. 温故知新继往开来——地球科学有关学科回顾与展望专辑(序)[J]. 地球科学进展, 2001, 16(5): 597-598.]

[3] Liao Shunbao, Sun Jiulin. GIS based spatialisation of population census data in Qinghai-Tibet Plateau[J]. *Acta Geographica Sinica*, 2003, 58(1): 25-33. [廖顺宝, 孙九林. 基于 GIS 的青藏高原人口统计数据空间化[J]. 地理学报, 2003, 58(1): 25-33.]

[4] Liao Shunbao, Li Zehui. Study on spatialisation of population census data based on relationship between population distribution and land use-Taking Tibet as an example[J]. *Journal of Natural Resources*, 2003, (6): 659-665. [廖顺宝, 李泽辉. 基于人口分布与土地利用关系的人口数据空间化研究——以西藏自治区为例[J]. 自然资源学报, 2003, (6): 659-665.]

[5] Liao Shunbao, Li Zehui. A methodology of spatialisation of observed data based on GIS[J]. *Progress in Geography*, 2003, 22(1): 87-93. [廖顺宝, 李泽辉. 基于 GIS 的定位观测数据空间化[J]. 地理科学进展, 2003, 22(1): 87-93.]

[6] Sun Jiabing, Shu Ning, Guan Zequn. Fundamentals, Methods and Application of Remote Sensing [M]. Beijing: Publishing House of Survey and Mapping, 1997. [孙家炳, 舒宁, 关泽群. 遥感原理、方法和应用[M]. 北京: 测绘出版社, 1997.]

[7] Zheng Du. Forming Environment and Development of Qinghai-Tibet Plateau[M]. Shijiazhuang: Hebei Publishing House of Science and Technology, 2003. [郑度. 青藏高原形成环境与发展[M]. 石家庄: 河北科学技术出版社, 2003.]

[8] Liu Chuang, Ge Chenghui. Characteristics and application of remote sensed data from U. S. EOSMODIS[J]. *Remote Sensing Information*, 2000, (3): 45-48. [刘闯, 葛成辉. 美国对地观测系统(EOS)中分辨率成像光谱仪(MODIS)遥感数据的特点与应用[J]. 遥感信息, 2000, (3): 45-48.]

[9] Yin Gongbai, Wang Jiayao, Tian Desen, et al. Tian Desen, An Introduction to Map[M]. Beijing: Publishing House of Survey and Mapping, 1991. [尹贡白, 王家耀, 田德森, 等. 地图概论[M]. 北京: 测绘出版社, 1991.]

[10] Zhao Junxi, Liu Honglin. Usually used theories and methods for study on spatial data quality[J]. *Map*, 2001, (4): 16-18. [赵军喜, 刘宏林. 研究空间数据质量的常用理论和方法[J]. 地图, 2001, (4): 16-18.]

DEVELOPMENT, PUBLISHING AND SHARING OF DATA PRODUCTS FOR GEO-SCIENCES

LIAO Shun-bao, SUN Jiu-lin, LI Ze-hui, MA Lin, PENG Mei

(*Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China*)

Abstract: Abstract: Data is the foundation of scientific research, and in the mean time, it is the result of scientific research. Data sharing can make data more usefully valuable to a great extent. Data sharing is a systematic engineering which includes not only data itself but also software platform for data sharing, technical standards, sharing policy, administrative mechanism and data service. Data is the most basic and important element in data sharing engineering. Geo-science data have attributes of space, time, integration, great capacity and multi-sources. It is costly work to capture geo-science data. Therefore, it is more necessary to share them among different organizations or scientists.

Tabular data, remotely sensed data and vector data are most important and often used types of geo-science data. Because of limitations of various kinds of original data, such as statistical data, remotely sensed spectral data (e.g. TM, NOAA-AVHRR, MODIS), map data and so on, besides them, series of data sets should be developed to meet the need of basic and forward research for geo-science. By means of spatialisation technique, observed data from observation stations can be used to calculate data in areas without observation stations and statistical data from administrative divisions can be re-assigned to regular grid, for example, one kilometer by one kilometer. Through various algorithms, remotely sensed spectral data can be used to produce data sets with obvious geo-science sense, for example, Vegetation Index (VI) and Net Primary Production (NPP).

The software platform is important technological support for geo-science data sharing. It should be easy to use and include the following fundamental functions: (a) users management; (b) data catalogue query; (c) metadata management; (d) data query and view; and (e) data download.

In order to promote scientific data sharing, relevant policies, measures and standards should be drawn up. Reasonable appraising system for people who are engaged in work of data should be set up to encourage they devote themselves to data sharing.

Key words: Geo-sciences; Data; Production; Sharing