

中国虚拟天文台

—— 任务、特点、方案

赵永恒 崔辰州

中国科学院国家天文台

摘 要

目前,虚拟天文台已经成为世界天文学界乃至 IT 界的热门话题。美国、欧盟、英国的虚拟天文台计划已经进入实施阶段,更多国家的类似计划正在酝酿之中,国际虚拟天文台联盟初现端倪。如此形势之下,中国为什么有必要建立自己的虚拟天文台?中国虚拟天文台将有何特色,对国际虚拟天文台会有怎样的贡献?中国虚拟天文台将采用何种体系结构,何时能实现?本文就这些问题进行了探讨。

四百年前伽利略首次把望远镜指向天空,结束了人类一直用肉眼进行天文观测的历史。一百五十年前,照相技术和光谱技术开始在天文观测中应用,单纯以人眼作为天文探测器的时代结束,天体物理学诞生并发展成为现代天文学的主流。五十多年前,在第二次世界大战中得到蓬勃发展的无线电技术使得天文学家的视野超出了可见光,射电天文学诞生。此后不久宇航时代到来,空间天文学诞生,人类对宇宙的观测扩展到了伽马射线、X 射线、紫外和红外波段。

从二十世纪九十年代开始,天文学正经历着革命性的变化。这一变化是由前所未有的技术进步推动的,即望远镜的设计和制造、大尺寸探测器阵列的开发、计算能力的指数增长以及互联网络的飞速发展。

望远镜技术的进步使得人类可以建造大型的空间天文台,为伽玛射线、X 射线、光学和红外天文的发展开辟了新的前景,同时也推动了新一代的大口径地面光学望远镜和射电望远镜的建造。现在,天文学家们正在计划建造功能更好口径更大的空间和地面望远镜,并将配备尺寸更大像素更多的探测器。随着众多先进的地面与空间天文设备的投入使用,大规模的观测数据正在产生,例如目前哈勃空间望远镜每天大约产生 50 亿字节的数据,我国正在建造的 LAMOST 望远镜也将产生每天 30 亿字节的数据,美国计划建造的“大口径巡天望远镜”将会达到每天 10 万亿字节的量级!目前,世界天文学数据量增倍的周期已经缩短到一年以内。

使天文学研究发生重大变化的另一个技术进步是快速互联网技术,这使得异地天文数据的交换和处理成为可能,使得观测数据具有巨大的科学产出的潜在意义。

在这样的情况下,美国国家科学院天文学及天体物理学发展规划委员会在题为“新千年的天文学和天体物理学”^[1]的十年发展规划中把建立国家虚拟天文台作为优先推荐项目。

虚拟天文台的概念提出后各国天文学界迅速响应,纷纷提出了各自的虚拟天文台计划。当前国际上已经得到资金支持的主要计划有美国国家虚拟天文台(简称 NVO)^[2]、欧盟天体物理虚拟天文台(简称 AVO)^[3]、英国天文网格(简称 AstroGrid)^[4]。正在积极推动的虚拟天文台计划主要有澳大利亚虚拟天文台、印度虚拟天文台、俄罗斯虚拟天文台、意大利虚拟天文台、加拿大虚拟天文台等。

由欧洲南方天文台(简称 ESO)、欧洲空间局(简称 ESA)、美国宇航局(简称 NASA)和美国国家科学基金会(简称 NSF)共同资助的“通向国际虚拟天文台之路(Toward an International Virtual Observatory)”国际天文研讨会^[5]于 2002 年 6 月在德国召开。会议期间,以 NVO、AVO、AstroGrid 为首提出了成立国际虚拟天文台联盟(简称 IVOA)的倡议,阐述了 IVOA 的历史使命并提出了相应的行动方案。

虚拟天文台工作原理

巡天，就是对整个天区进行观测、普查。如果利用伽马射线巡天、X 射线巡天、紫外巡天、光学巡天、红外巡天和射电巡天所得到的观测数据，用适合的方法对数据进行统一规范的整理、归档，便可以构成一个全波段的数字虚拟天空；而根据用户要求获得某个天区的各类数据，就仿佛是在使用一架虚拟的天文望远镜；如果再根据科学研究的要求开发出功能强大的计算工具、统计分析工具和数据挖掘工具，这就相当于拥有了虚拟的各种研究设施。这样，由数字虚拟天空、虚拟天文望远镜和虚拟研究设施所组成的机构便是一个独一无二的虚拟天文台（如图 1 所示）。

建设中国虚拟天文台的必要性

目前，国际虚拟天文台（简称 IVO）正处于起步阶段，中国天文界以及作为中国天文界唯一大科学工程的 LAMOST 项目^[6]应该积极参与，从一开始就把自身融入到这股将为天文学带来一场新的革命的浪潮中。

中国虚拟天文台（简称 China-VO^[7]）的建设是非常必要的。

1. 只有加入国际虚拟天文台联盟，才能以平等的身份全方位共享 IVOA 的技术与资源。
2. 建设 China-VO 是实现 LAMOST、BATC 巡天等我国自产数据与 IVO 数据融合的最佳途径。
3. 建设 China-VO 可以培养一批与 IVO 相适应的天文学家和技术人才，为未来中国天文学的发展提供智力支持。
4. 可以利用 IVO 丰富的资源加强教育和科学普及工作，提高我国公众的科学素质。

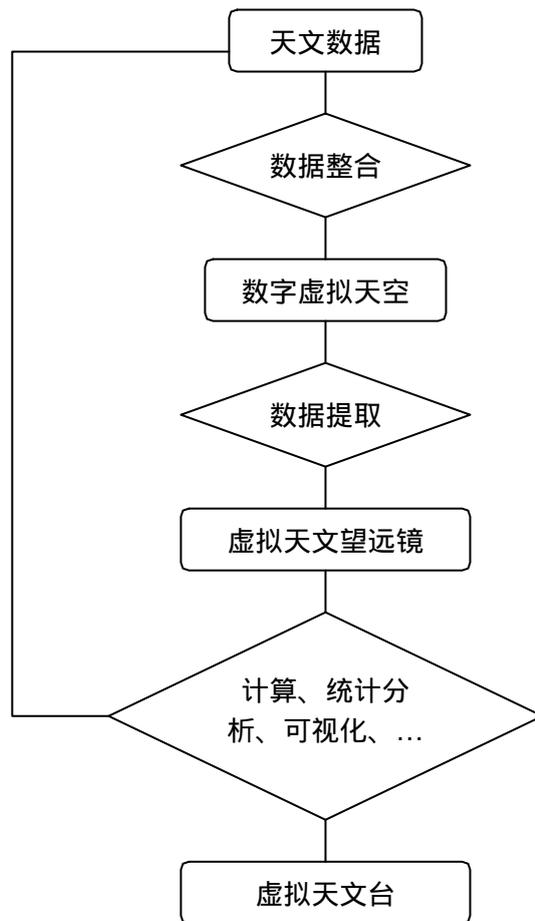


图 1 虚拟天文台工作原理

大样本光谱巡天是 LAMOST 工程的本质特点，把 LAMOST 建成为“VO-oriented LAMOST”不但是 LAMOST 本身的需要，也将是中国对世界天文的重大贡献。它的观测数据将成为 China-VO 的基础，为国际虚拟天文台和天文学的发展做出贡献。

1. LAMOST 的观测目标必将从 IVO 中选取。LAMOST 是对选定目标进行光谱巡天的望远镜，输入星表的选取决定了 LAMOST 的科学产出。IVO 是最富有的天文数据拥有者，

LAMOST 从中提取输入星表是必然的途径。

2. LAMOST 的观测结果，将成为 IVO 的一个重要组成部分。目前正在进行中的大型巡天观测，大部分是成像观测，为了深入研究天体物理过程，大量光谱资料的获得将成为天体物理学发展的瓶颈。可同时获得 4000 个天体光谱的 LAMOST 望远镜建成后将成为世界上威力最大的光谱巡天望远镜，LAMOST 数据中心也将成为世界光学光谱数据中心。

3. LAMOST 望远镜观测获得的海量光谱数据不可能依靠传统的数据处理方法进行处理，必须开发一套自动处理和分析工具，比如光谱处理、谱线提取、天体自动分类、红移测量、统计分析等工具。这些工具对于实现 IVO 的数据挖掘和知识发现功能将起到重要的推动作用。

4. LAMOST 与 IVO 一样都面临 TB、PB 量级数据的归档和管理问题。LAMOST 可以借鉴 IVO 的数据管理经验，同样 LAMOST 的自身经验也可为 IVO 所借鉴。

现有工作基础

天文观测数据的急剧增长和计算机网络技术的突破性进展是国际虚拟天文台计划产生和实现的两大直接动力。同样中国虚拟天文台也必须以这两个基本条件为建立的前提。

经过数十年的奋斗，我国现代天文学取得了长足的发展，一批大型天文观测设备已经投入使用并取得了丰富的观测资料。更重要的是，像 LAMOST、太阳空间望远镜(简称 SST)、五百米口径射电望远镜(简称 FAST)等一批正在研制或预研中的大观测设备在投入使用后将极大丰富我国的天文数据资源，这为我们加入 IVO 奠定了数据资源基础。此外，除了这些自产的天文数据，国家天文台天文数据中心作为世界数据中心唯一的天文学科中心经过多年的努力已经积累了大量的国际天文数据，主要资源如表 1 所示。

数据集	在线情况	光盘数	数据量(GB)
CDS/ADC 在线星表	√	4	2
USNO-A1.0 星表		8	5
USNO-A2.0 星表	√	10	6
Hipparcos/Tycho 星表		10	6
RealSky (第一期巡天图像)	√	8	5
DSS (第一期巡天图像)	√	102	60
DSS-II (第二期巡天图像)		>600	>360
Einstein X 射线卫星数据		9	5
ROSAT X 射线卫星数据		45	27
2MASS 红外星表		>2DVD	>10
ADS 文献数据服务	√		350
(合计)			>836

表 1 国家天文台天文数据中心主要国际天文数据资源

同时，我国的计算机和网络技术也正在飞速发展，取得了许多世界先进水平的研究成果。第二代互联网和网络技术的研究正在与世界其他国家同步进行。国家天文台自行搭建的数据中心数据容量已超过 5TB，网络速度为 100Mbps 并将很快升级为 1000Mbps。国家天文台天文数据中心已经与中国科学院计算研究所就共同建设中国虚拟天文台达成共识。这些都为中国虚拟天文台的建设提供了技术支持。

China-VO 基本体系结构

从目前 IVO 的发展趋势以及 IT 领域的技术动向来看，层次式的体系结构将是未来网络技术发展^[8]、WEB 服务^[9]以及 IVO 体系结构的主流。China-VO 的基本体系框架将如图 2 所示。

从图中可以看出，China-VO 的体系结构大致可以分为三层，从下往上依次是：构造层、汇集层、用户层。

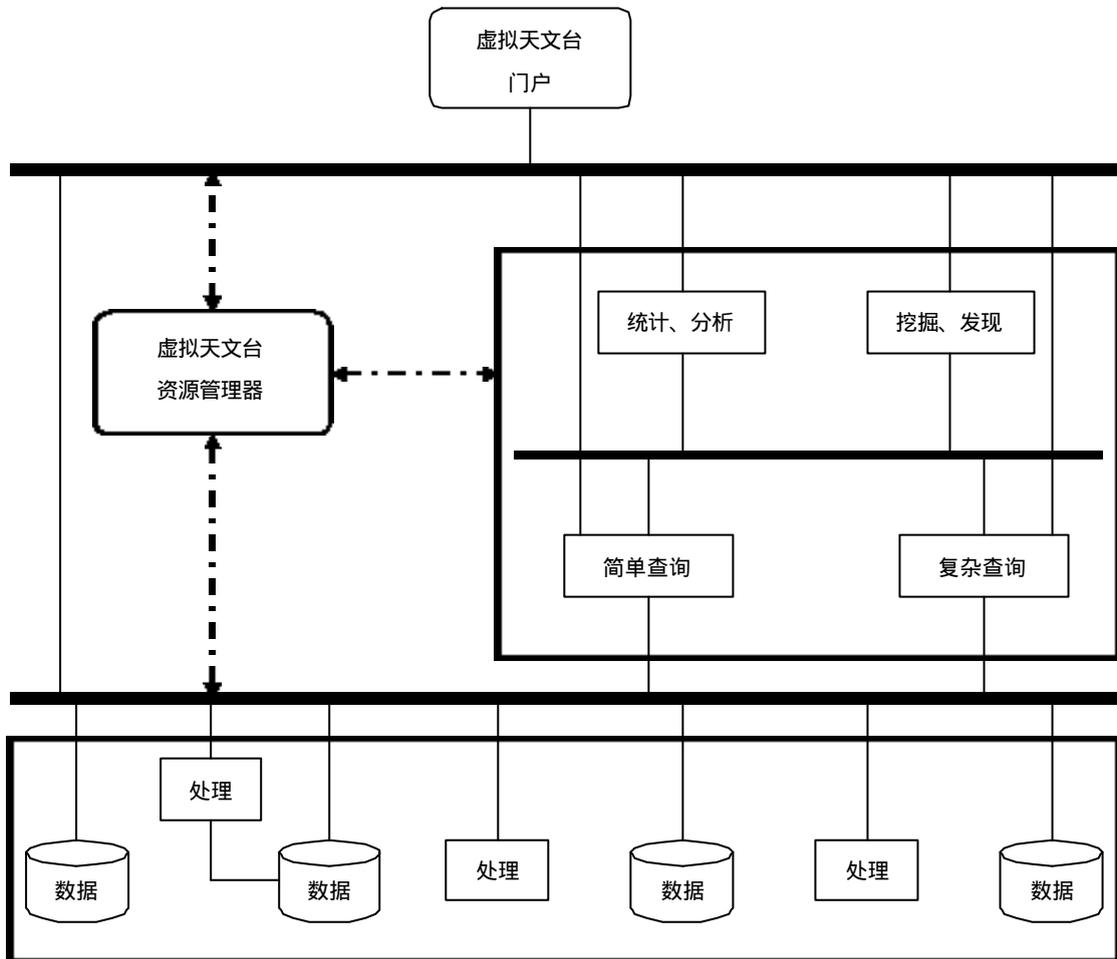


图 2 China-VO 基本体系结构

构造层主要由各种天文数据以及为了实现异构数据间的互操作而提供的各种数据处理工具或称为数据中介。构造层与汇集层利用统一的虚拟天文台数据模型（简称 VODM）进行数据交换。虚拟天文台数据模型将很可能以可扩展标记语言（简称 XML）^[10]为基础并结合天文学的特殊需求对其加以扩展或定制来实现，比如 AVO 与 NVO 联合提出的 VOTable^[11]。

汇集层是虚拟天文台的控制中心和处理中心，主要由虚拟天文台资源管理器（简称 VORM）、数据库检索查询工具、统计分析工具、数据挖掘和知识发现工具构成。其中 VORM 要完成整个系统的资源调度、任务管理、访问控制等职能。IT 领域正在兴起的网格技术，例如 Glubus^[12]将对 VORM 的设计起到重要的参考作用。数据库检索查询工具可对分布于全球各地的天文数据库进行各种简单或复杂查询，实现虚拟天文望远镜的功能。查询结果或者直接提交用户层或者提交给各种统计分析和数据挖掘工具，这些工具将对提取的数据进行更深层的科学处理然后再把结果提交用户层。

用户层，也就是虚拟天文台门户（简称 VOP），是虚拟天文台体系的最高层，直接与虚拟天文台用户接触。用户层的基本职能是用户任务提交和处理结果返回。根据用户类型的不同，比如天文学家、其他学科科研人员、大学生、理科教师、天文爱好者、中小學生、网络冲浪儿等，VOP 将提供相应的任务提交界面和结果显示方式。对于职业天文学家，VOP 将提供最为全面、高级、可自定义的任务提交界面，处理结果的返回方式也可以让天文学家最大限度的进行定制。对于有较高科技知识基础的用户，比如其他学科的科研人员、大学生和理科教师，VOP 将在任务提交、结果返回方式的选择上提供一定的自由度，同时提供与这些用户知识水平相适应的辅助工具和资料文档。具有一般知识水平的普通大众可能会在数量上占据虚拟天文台用户的多数，不过这些用户对虚拟天文台的功能要求相对简单。为了让普通公众也能理解高深的天文学，VOP 将提供各种可视化工具和友好的人机界面，把观测数据和处理结果生动的、以公众乐于接受的形式提供给这些用户。

特色与挑战

相对于欧美这些科技强国而言，China-VO 无论从科技上还是人财物资源上都存在一定差距。China-VO 必须按照“有所为有所不为”的指导思想，挖掘自身潜力，建设出自己的特色。

科学上，China-VO 将采用与 LAMOST 项目紧密结合的方式，充分发挥 LAMOST 光学光谱数据中心的作用，建设“VO-oriented LAMOST”，使光谱数据及其相关处理技术成为 China-VO 的核心与特色。

技术上，China-VO 必须最大程度的集中国内各天文研究机构、IT 研究机构和数学等相关领域的人才资源，共同努力实现目标。

此外，China-VO 必须采用开放的运作方式，与国际上各虚拟天文台计划开展充分的合作，在 IVOA 中发挥积极作用。

IVO 的提出与发展离不开计算机与网络技术的直接推动，但同时也为 IT 技术提出了挑战，比如分布式多数据库联合查询、分布式异构数据的互操作性、大规模多维数据统计分析与可视化、授权认证管理等等。为了实现以“VO-oriented LAMOST”为特色的 China-VO，除了 IVO 共同面临的这些挑战，我们还必须解决下面一些问题。

1. 海量多光纤光谱观测数据的自动处理；
2. 光纤光谱谱线的自动提取；
3. 光谱自动分类；
4. 光谱红移的自动测量；
5. 光纤光谱数据与其它类型天文数据的融合；
6. 光谱数据的可视化。

由于我国是一个非英语国家，按照语言习惯，中国虚拟天文台的用户可分为英文用户和中文用户两大群体。一方面，为了融入 IVO 大家庭，为国际用户提供服务，China-VO 必须把现有的中文文献资料、观测数据转化为英文并提供相应的英文用户界面，即“国际化”。另一方面，为了防止因为语言障碍而把国内广大用户挡在门外的现象出现，China-VO 必须承担起把丰富的英文资源转化为中文的使命并同时提供相应的中文用户界面，即“本地化”。

China-VO 实施方案

根据国家天文台和 LAMOST 工程的具体情况同时参考 IVO 的发展设想，China-VO 制定如下的初步实施方案。

- 2002 年 7 月：完成 China-VO 体系结构草案设计
- 2002 年 10 月：实现大部分现有数据的在线访问；加入国际虚拟天文台联盟
- 2003 年 1 月：完成 China-VO 用户需求在线调查；完成 China-VO 体系结构设计；完成 China-VO “数据中介”设计
- 2003 年 5 月：实现 China-VO 初步在线访问
- 2003 年 8 月：完成 China-VO 0 级演示版
- 2003 年 10 月：完成 VO-oriented LAMOST 数据模型（简称 LmstD）1.0 版
- 2004 年 1 月：完成 China-VO 1.0 版
- 2004 年 10 月：完成 LmstD 2.0 版
- 2005 年 1 月：完成 China-VO 2.0 版

参 考 资 料

1. Astronomy and Astrophysics in the New Millennium (Decadal Survey), National Academy of Science, Astronomy and Astrophysics Survey Committee, <http://www.nap.edu/books/0309070317/html/>
2. National Virtual Observatory, <http://www.us-vo.org>
3. Astrophysical Virtual Observatory, <http://www.eso.org/avo>
4. AstroGrid, <http://www.astrogrid.org>
5. Toward an International Virtual Observatory, Scientific Motivation, Roadmap for Development and Current Status, ESO/ESA/NASA/NSF Astronomy Conference, <http://www.eso.org/gen-fac/meetings/vo2002/>
6. The Large Sky Area Multi-Object Fiber Spectroscopic Telescope , <http://www.lamost.org>
7. Chinese Virtual Observatory, <http://www.china-vo.org>
8. Global Grid Forum, <http://www.gridforum.org/>
9. Web Services, <http://www.w3.org/2002/ws/>
10. Extensible Markup Language, <http://www.w3c.org/XML/>
11. VOTable, <http://cdsweb.u-strasbg.fr/doc/VOTable/> or <http://www.us-vo.org/VOTable/>
12. Globus, <http://www.globus.org>