

“九五”国家科技攻关项目

项目名称：96-F01 全国西文文献信息计算机联合编目系统

专题名称：96-F01-01-01 总体方案研究、规划、设计和标准、规范的制定

分专题名称：96-F01-01-07D 数据处理手册

西文文献数据库数据处理 手 册

上 册

(试用)

《数据处理规范的研究与制定》课题组

1997.3.

前　　言

文献信息资源是人类智慧的结晶，是人类传播科学的媒体，是科技、经济和社会发展的基础条件。为了贯彻执行1995年5月6日《中共中央、国务院关于加速科学技术进步的决定》中对科技信息界提出的要求：“重视科技信息的有效利用和传播，加强科技图书、资料和数据库的建设。”国家科委已决定将“全国西文文献信息计算机联合编目系统”项目(以下简称《项目》)列入“九五”国家科技攻关计划。该《项目》的总目标是：经过“九五”的努力，到本世纪末基本建成以现代化计算机和通信网络为平台的，连接全国各主要文献信息机构的西文文献信息联机编目系统，实现广域网环境下的联机查询、采购协调、联机编目、馆际互借以及电子文献传递等功能，达到科技信息共享和交流的现代化。

标准化是实现资源共享的前提，是保证系统质量和国际接轨的重要技术依据，是保质、保量完成该攻关《项目》的关键。

《西文文献数据库数据处理工作手册》(以下简称《手册》)是该《项目》第一个专题《总体方案研究、规划、设计和标准、规范研究与制订》中的一个分专题。该《手册》是由国家科委信息司、全国情报文献工作标准化技术委员会秘书处邀请国内文献信息处理方面的专家组成专题组来承担的。编制该《手册》目的是为了给各承担西文文献数据库建库任务的数据处理人员提供一份统一的、规范的、操作性、适用性强的工作手册，以保证所建系统能够实现资源共享。

该《手册》是项目组制订的重要标准、规范之一，供各开发任务专题组(或分专题组)研究开发时使用。在执行过程中，如有问题或改进意见和建议，请及时与国家科委信息司信息资源处联系。以使《手册》内容能不断得到充实、提高和完善。

“全国西文文献信息计算机
联合编目系统”第一专题组
一九九七年五月十二日

西文文献数据库数据处理手册

研究和编写人员

顾问：陈松生、周小璞、李廷杰

分专题组长：纪昭民

成员：沈玉兰、宋文、富平、潘岩明、
吴荣荣、吴兵、邱祖斌、段明连、
李秀锦、陈卫

执笔人员：纪昭民、宋文、蒋敏、王德萍、吴荣荣、
吴兵、邱祖斌、潘岩明

特邀专家：袁明敦、张凤楼、熊光莹、蔡蓉华、
朱岩、胡丽娟

编　者　的　话

实现文献信息资源共享是世界文献信息领域的同仁们多年追求的目标，并为此设想过许多办法。国际图书馆协会联合会（IFLA）组织专家制定的国际标准书目著录（ISBD）系列标准为实现全世界范围文献书目信息交流与共享起到了积极的推动作用。国际标准书目著录系列标准规定了文献著录的项目及其顺序、文献著录标识符号等，这就为各种文字、各种类型的文献书目在交流过程中识别及计算机转换创造了一个简单易行而有效的办法。因此，很快为世界各国文献信息界，特别是图书馆界所接受。

目前，由于书刊资料，特别是国外原版书刊资料的价格上涨较快，各文献信息单位用于购买书刊资料的经费都比较紧张，购买书刊资料的数量，尤其是国外原版书刊资料的数量呈逐年下降的趋势，严重地影响了科研、教学、生产等领域的正常使用。因此，国家科委决定有计划、有步骤，分批、分期地组织各文献信息单位共同建设一个全国性的科技文献共享系统，而首先要建立一个最急需的西文文献信息计算机联合编目系统。建立这样一个系统，除必要的设备、网络环境及操作系统外，最重要的就是数据的质量。而要保证许多单位、许多人在不同时间、不同地点处理的数据都能标准、完整、准确、一致，就必须有一个共同遵循的数据处理规范。该规范应包括部分文献编目条例的内容、机读目录数据交换格式的解说、各种相关代码标准等，使之成为一个西文文献信息计算机联合编目人员的工作手册。

本手册就是根据这种指导思想编写的。手册中涉及到的文献著录规则是根据国际标准书目著录（ISBD）系列标准，并参照《西文文献著

录条例》和AACR2有关条例制定的；手册中所列的有关机读目录通讯格式是经过调研决定采用UNIMARC通讯格式制定的。由于西文文献信息计算机联合编目系统要处理的文献除国外的原版书刊资料外，还要处理我国出版的西文文献资料，所以还要使用CNMARC格式中的一部分字段。本手册中所列的机读目录通讯格式内容的主要部分直接引用了《中国机读目录格式使用手册》(China MARC Format Manual)中的有关部分；手册中所提到的各种代码，除《一般文献型标识代码》采用国际标准书目著录非书资料 (ISBD(NBM)) 附录中的规定外，其它诸如国家和地区名称代码、语种代码等均按照UNIMARC的规定执行。

本手册是按UNIMARC机读目录通讯格式的字段顺序编写的。

考虑到西文文献信息计算机联合编目系统实施的具体情况，本手册在内容划分上没有按通常的文献类型定义，而是根据实际需要划分的。即，除了西文文献数据库数据处理总则外，还分为西文连续出版物、西文会议文献、西文参考工具书、西文非书资料、西文科技报告等几个数据库数据处理细则。在对西文非书资料进行著录时，除“载体形态项”和有关编码不同于一般印刷型的文献外，其它著录规则均与印刷型文献相同。

由于水平所限，手册中一定存在不少问题，请用户在试用过程中及时反馈意见，以便修改补充，使之逐步完善。

《数据处理手册的研究与制定》分专题组

一九九七年

目 录

一 西文文献数据库数据处理总则.....	A
二 西文连续出版物数据库数据处理细则.....	B
三 西文会议文献数据库数据处理细则.....	C
四 西文参考工具书数据库数据处理细则.....	D
五 西文非书资料数据库数据处理细则.....	E
六 西文科技报告数据库数据处理细则.....	F
七 附录.....	G

西文文献数据库数据处理总则

1 通则

1.1 名词术语

1.1.1 文献机读目录数据库 (MARC data base)：将文献目录以代码形式和特定结构记录在计算机存贮装置中，使其成为计算机能够识别和阅读的目录称为文献机读目录。在计算机的存贮设备上合理存放的相互关联的机读目录集合称为文献机读目录数据库。

1.1.2 文献机读目录记录 (MARC record)：是描述某一种或某一件文献实体本身固有的信息和相关信息的集合。

1.1.3 机读目录交换格式 (MARC communication format)：为使计算机可读形式的文献目录能在国家书目机构之间或国内图书情报机构之间进行交换，根据国际标准 ISO 2709 制定的格式。

1.1.4 字节 (Byte)：字节是信息在计算机内存贮的基本单位，由 8 个二进制位构成。

1.1.5 字符集 (character set)：国家标准 GB 1988-80 《信息处理交换用七位编码字符集》（其等效标准为国际标准 ISO646）规定了信息处理与交换用的控制字符、常用标点符号和英文字母的编码及结构，其字符以单字节来表示。

GB 2312-80 规定了信息交换用汉字编码字符集（基本集），其字符以双字节来表示。

1.1.6 填充符 (Fill character)：用于标识没有确切数据值的特定字符位置的字符，而这个字符位置是必用的。

1.1.7 源格式 (Source format)：转换成 UNIMARC 格式之前的机读目录格式。

1.1.8 记录头标 (Record label)：位于记录开始的部分，在该部分给出了本记录的有关参数。

1.1.9 目次区 (Directory)：记录中所有数据字段的标识符、长度和起始位置的标识区。

1.1.10 检索点 (Access point)：用于检索书目记录所使用的名称、主题词、分类号和代码等数据项。

1.1.11 内容标识符 (Content designator)：用于识别数据元素，或提供有关数据元素的附加信息的符号或编码。内容标识符包括字段标识 (Tag)，指示符 (Indicator) 和子字段标识符 (Sub-field identifier)。

1.1.12 数据元素 (Data element)：被明确标识的数据最小单元。在一个可变长字段内，数据元素构成子字段，并用子字段标识符标识。在记录头标、目次

和固定长子字段内，由代码构成的数据元素是由其字符所在位置标识的。

1.1.13 数据元素标识符 (Data element identifier)：见子字段标识符。

1.1.14 字段 (Field)：由字段标识符标识的一个被定义的字符串，它包含一个或多个子字段。

1.1.15 字段分隔符 (Field separator)：在每个可变长字段的结尾用以分隔字段的控制字符。它也用于目次区的结尾。该字符是 ISO 2709 中规定的专用字符 IS2 (ISO 646 (1/14))。

1.1.16 固定长子字段 (Fixed length subfield)：长度固定的子字段。它可包含一个或多个数据元。它可用在定长字段里，如 100 字段中的 \$a 子字段。也可用于变长字段，如 101 字段里的 \$a、200 字段里的 \$z 子字段。

1.1.17 层次等级 (Hierarchical level)：表示书目实体（如 200 字段著录的实体）在一个预先规定的层次结构中所处的位置，以及书目实体之间的层次关系，如多卷集与其单册，图书与其章节，丛书与其附属丛书在不同层次上的相互关系。

1.1.18 字段指示符 (Indicator)：与变长字段相关的字符（数字和字母），它提供有关可变长字段内容、记录中不同字段的关系及某些数据处理过程中所需操作的附加信息。

1.1.19 子字段 (subfield)：字段内所定义的数据单元（见数据元素）。

1.1.20 子字段标识符 (Sub-field identifier)：本手册规定为由两个字符组成的代码，用以识别可变长字段中不同子字段。第一个字符为 ISO-2709 中规定的专用符号 IS1 (ISO 646 (1/15))，第二个字符为字母或数字。

1.1.21 字段标识符 (Tag)：用于识别各个字段的三位数字或字符代码。

1.1.22 可变长字段 (Variable field)：长度不定的字段。可包含一个和多个数据元素或子字段。其长度包括文字数据、字段指示符、子字段标识符及字段分隔符。

1.1.23 记录分隔符 (Record separator)：置于每个记录结尾，用来区分记录的控制字符，为 ISO-2709 中规定的专用字符 IS3 [ISO 646 (1/13)]。

1.1.24 必备字段 (子字段) (Mandatory field/Subfield)：在记录中必须出现的字段 (子字段)。

1.1.25 必选字段 (子字段) (Selected field/Subfield)：在文献本身载有某字段 (子字段) 书目信息时，该字段 (子字段) 必须在记录中出现。

1.1.26 可重复性 (Repeatable)：是指同一字段 (子字段) 可以在同一个记录内出现多次。

1.1.27 代码数据 (Code data)：由英文字母和/或数字符号表示的数据，其含义通常与其所在位置有关系。

1.1.28 文字数据 (Textual data)：与编码数据不同，是用于目录款目显示的数据。

1.1.29 文献 (Document)：记录有知识的一切载体。

1.1.30 著录 (Description)：在编制文献目录时，对文献的形式特征和内容特征进行分析、选择和记录的过程。

1.1.31 款目 (Entry)：著录文献的结果，是反映文献的形式特征和内容特征的著录项目的组合。

1.1.32 规定信息源 (Prescribed source of information)：文献的每一个著录项目或单元内容的主要信息来源。

1.1.33 题名页 (Title-page)：一般在文献正文之前，载有反映该文献较完整的题名、责任说明以及出版说明等信息。

1.1.34 代题名页 (Title-page substitute)：当文献没有题名页时，则用其载有通常出现在题名页上的信息的一页作为替代的题名页。

1.1.35 正题名 (Title-proper)：文献的主要题名。正题名包括交替题名，但不包括并列题名和其它题名信息。

1.1.36 共同题名 (Common title)：一组相关文献所共有的正题名，它用于表示这些文献之间的关系，并与分辑题名一起识别某种文献。

1.1.37 从属题名 (Subordinate title)：一题名不足以识别一种文献，需要加上共同题名才能识别该文献的题名。如，分辑题名、补编题名、分丛编题名等。

1.1.38 通用题名 (Conventional title)：缺乏自身特性的题名。如，报告、学报、年度统计等。

1.1.39 交替题名 (Alternative title)：规定信息源中出现的，由两个部分组成的正题名的第二部分。这两个部分之间用“又名”或其它文种的对应词表示。

1.1.40 分辑题名 (Section title)：用于识别一组具有共同题名的相关文献中的一个组成部分的专有题名。

1.1.41 一般文献类型标识 (General material designation)：概括地表示文献所属类型的术语。

1.1.42 并列题名 (Parallel title)：文献中对应于正题名的另一个语种的题名。

1.1.43 其它题名信息 (Other title information)：从属于正题名、并列题名；并对它们进行限定、解释、补充等的题名信息。

1.1.44 统一题名 (Uniform title)：为使具有不同正题名的同一著作的不同版本或不同译本在目录中集中在一起，编目机构可选择一个特定的（人们熟悉的）正题名来标识该著作，这个题名就是统一题名。

1.1.45 识别题名 (Key title)：由国际连续出版物数据系统 (ISDS) 国际中心或

国家中心指定的用于识别连续出版物的唯一的题名。

1.1.46 缩略题名 (Abbreviated title)：识别题名的缩略形式。

1.1.47 责任说明 (Statement of responsibility)：对文献的知识内容和艺术内容的创造或完成负有责任或做出贡献的个人、团体及其责任方式。有时责任说明可以是一个没有个人姓名或团体名称的短句。

1.1.48 第一责任说明 (First statement of responsibility)：同时具有几个不同责任说明时，列居首位的责任说明。

1.1.49 其它责任说明 (Other statement of responsibility)：同时具有几个不同责任说明时，除第一责任说明以外的责任说明。

1.1.50 并列责任说明 (statement of parallel responsibility)：文献中对应于责任说明的另一个语种的责任说明。

1.1.51 版本 (Edition)：由同一版型生产，并由同一机构出版的一种文献的全部复本。

1.1.52 版本说明 (Edition statement)：表示一种文献属于某一版本的一个单词、短语。通常是以序数词与“版”字或表示和其它版本不同的词与“版”字相结合的形式出现的。

1.1.53 附加版本说明 (additional edition statement)：在规定信息源中，对版本说明的补充说明。

1.1.54 特定文献类型标识 (Specific material designation)：表示文献属于某种具体文献类型的术语。

1.1.55 附件 (Accompanying material)：文献的主体部分所附带的旨在与主体部分一起使用的资料。

1.1.56 丛编 (Series)：一组相互关联而又各自独立的文献，每种文献除具有各自的正题名外，还有一个适用于整组文献的总题名。

1.1.57 主丛编 (Main series)：包含一个或多个分丛编的丛编。

1.1.58 分丛编 (Sub-series)：作为主丛编的一个组成部分的丛编。

1.1.59 丛编编号 (Numbering within series)：丛编连续出版的各期之识别标识，可以是数字、字母、其它符号或者是它们的组合形式。

1.1.60 多层次著录 (Multi-level description)：文献著录的一种方法，它把所描述的信息分为两个或更多层次。第一个层次包括文献整体的共同书目信息。第二个层次或其余层次包含有关各部分的书目信息。

1.2 著录信息源

1.2.1 著录信息源来自文献本身，如文献本身信息不足时，可参考其它资料。取自文献本身以外的著录信息或编目员自拟的著录信息均需加方括号。

1.2.2 与所著录文献具有永久性联系的信息源优先于仅有临时性联系或易失的

信息源。

1.2.3 各类型文献著录项目的规定信息源请见各类型文献数据处理细则。

1.3 著录用文字

1.3.1 题名与责任说明项（200字段）、版本项（205字段）、特殊细节项（206、207、208、230字段）、出版发行项（210字段）、丛编项（225字段）均按所著录文献本身的文字照录，但大小写、空格等可不照录。无法照录的图形或符号等可改为其它形式的相应内容，并加方括号。

1.3.2 一般文献类型标识（200字段）、载体形态项（215字段）、标准编号与获得方式项（010、011、014、015字段）用英文著录。

1.3.3 附注项（3--附注块）除引用内容（如：题名原文、识别题名等）外均用英文著录。

1.3.4 版次、出版（印制）日期、文献数量、文献开本尺寸、价格等一律用阿拉伯数字著录。

1.3.5 一般情况下，每个著录项目的首词首字母应大写，某些著录单元（如：交替题名、并列题名、分辑题名等）的首词首字母和专有名词的首字母也应大写。其它大写应根据所用文字的规定处理。

1.3.6 所著录文献的规定信息源的文字出现错误时，仍需照录，同时将正确的文字著录在其后面方括号内。必要时可在附注项中注明。

1.4 著录详简级次

1.4.1 根据文献收藏单位的类型及规模大小、所著录文献的特点、用户的检索需求、网络化程度等，文献著录可分为简要著录、基本著录、详细著录三个级次。

1.4.2 本手册规定《全国西文文献信息计算机联合编目系统》采取基本级著录。即，除包括所有必备项目外，还在供选择项中规定了“有则必选”字段。

基本级著录需填写的内容：

正题名、一般文献类型标识、并列题名、其它题名信息、第一责任说明、其它责任说明、版本说明、与本版有关的责任说明、特殊细节项、出版地、出版者、出版年、特定文献类型标识与文献数量、图及其它形态、尺寸、附件、丛编正题名、丛编责任说明、丛编国际连续出版物号、丛编编号、分丛编著录事项、附注项、国际标准编号。

1.5 符号约定

本手册采用下列约定符号：

1.5.1 “\$”作为子字段标识的第一个符号（字符集中1/15位）。

1.5.2 “&”表示空格。

1.5.3 “*”作为完整实例中的字段分隔符（字符集中1/14位）。

1.5.4 “#”作为书目记录中的书目记录分隔符（字符集中1/13位）。

1.5.5 “未定义”表示指示符所在位置没有赋值。该指示符位置是用空格“**b**”填充的。

1.5.6 “*NSB*”(08/08位)和“*NSE*”(08/09位)表示不排序开始和不排序结束。即，这两个符号之间的数据不用于排序。

1.5.7 “|”表示在编码数据项中规定不使用的子字段。

1.5.8 “()”和“[]”的前后各需空一格。但当它们前面或后面有用一空格作为结尾或开头的标识符时，则只保留一个空格。而“.”和“,”则只需在其后面空一格。

1.5.9 录入子字段标识符后，应紧接其后录入子字段的数据内容。

1.6 格式结构

UNIMARC格式是ISO2709的一个特定形式。它对每一个用于交换的书目记录规定了必须遵循的标准记录结构。其标准构成为：

1.6.1 记录头标：由24个字符构成。

1.6.2 地址目次区：目次区的每个款目由12个字符构成。其中，三位数字表示字段号，四位数字表示数据字段长度，五位数字表示字段起始字符位置。

1.6.3 数据字段：可变长。每个字段之间由字段分隔符隔开。

记录结构：

记录头标	地址目次区	数据字段	记录分隔符
24	12	m	1

地址目次区结构：

地址目次区款目1 款目2 其它款目

字段号	字段长度	起始字符位置	字段分隔符
-----	------	--------	-------	-------

数据（控制）字段（00-）结构：

数	据	字段分隔符
---	---	-------

数据字段（01-999）结构：

指示符 1	指示符 2	\$a	数据	字段分隔符
-------	-------	-----	----	-------	-------

指示符

子字段标识

其它子字段

1.7 记录头标和字段一览表

记录头标

0 标识块

- 001 记录标识号
- 005 记录处理时间标识
- 009 源记录标识号
- 010 国际标准书号 (ISBN)
- 011 国际标准连续出版物号 (ISSN)
- 014 论文标识号
- 013、015-018 (为其它国际标准号保留)
- 020 国家书目号
- 021 版权登记号
- 022 政府出版物号
- 040 CODEN
- 071 出版者作品号 (录音和音乐)
- 091 统一书刊号
- 092 订购号
- 094 标准号

1 编码信息块

- 100 通用处理数据
- 101 作品语种
- 102 出版或制作国别
- 105 编码数据字段：文字资料、专著
- 106 编码数据字段：文字资料—形态特征
- 110 编码数据字段：连续出版物
- 115 编码数据字段：投影、录像资料和影片
- 116 编码数据字段：书画刻印作品（暂定）
- 117 编码数据字段：三维制品和实物

- 120 编码数据字段：地图与天体图资料—一般性数据
- 121 编码数据字段：地图与天体图资料—形态特征
- 122 编码数据字段：文献内容涵盖期间
- 123 编码数据字段：地图与天体图资料—比例尺与坐标
- 124 编码数据字段：地图与天体图资料—特殊资料标志
- 125 编码数据字段：录音资料与印刷乐谱
- 126 编码数据字段：录音资料—形态特征
- 127 编码数据字段：录音资料与印刷乐谱播放时间
- 128 编码数据字段：音乐演奏与乐谱
- 130 编码数据字段：缩微制品—形态特征
- 131 编码数据字段：地图与天体图资料—大地、坐标网格与垂直测量
- 135 编码数据字段：计算机文件（暂定）
- 191 编码数据字段：拓片
- 192 编码数据字段：民族音乐

2 著录信息块

- 200 题名与责任说明
- 205 版本说明
- 206 资料特殊细节项：地图与天体图资料—数学数据
- 207 资料特殊细节项：连续出版物卷期编号
- 208 资料特殊细节项：印刷乐谱的特别说明
- 210 出版发行等
- 211 预定出版日期
- 215 载体形态项
- 225 丛编
- 230 资料特殊细节项：计算机文件特征（暂定）

3 附注块

- 300 一般性附注
- 301 标识号附注
- 302 编码信息附注
- 303 著录信息一般性附注
- 304 题名与责任说明附注
- 305 版本与书目史附注
- 306 出版发行等附注
- 307 载体形态附注
- 308 丛编附注

- 310 装订及获得方式附注
- 311 连接字段附注
- 312 相关题名附注
- 313 主题附注
- 314 知识责任附注
- 315 资料（或出版物类型）特殊细节附注
- 320 书目、索引附注
- 321 被索引、摘要和引用附注
- 322 制作者附注（投影和录像资料及录音）
- 323 演出者附注（投影和录像资料及录音）
- 324 原作版本附注
- 325 复制品附注
- 326 出版周期附注（连续出版物）
- 327 内容附注
- 328 学位论文附注
- 330 提要、文摘或全文
- 332 引文
- 333 使用对象附注
- 336 计算机文件类型附注（暂定）
- 337 计算机文件技术细节附注（暂定）
- 345 采访信息附注

4 款目连接块

- 410 丛编
- 411 附属丛编
- 421 补编、增刊
- 422 正编、正刊
- 423 合订、合刊
- 430 继承
- 431 部分继承
- 434 吸收
- 435 部分吸收
- 436 由……，……和……合并而成
- 440 由……继承
- 441 由……部分继承
- 444 并入

- 445 部分并入
- 446 分成……，……和……
- 447 与……、……合并而成
- 448 改回
- 451 同一载体其它版本
- 452 不同载体版本
- 453 译为
- 454 译自
- 455 复制自
- 456 复制为
- 461 总集
- 462 分集
- 463 单册
- 464 单册分析
- 470 被评论作品
- 488 其它相关作品

5 相关题名块

- 500 统一题名
- 501 作品集 统一题名
- 503 统一惯用标目
- 510 并列正题名
- 512 封面题名
- 513 附加题名页题名
- 514 卷端题名
- 515 逐页题名
- 516 书脊题名
- 517 其它题名
- 520 前题名（连续出版物）
- 530 识别题名（连续出版物）
- 531 缩略题名（连续出版物）
- 532 完整题名（连续出版物）
- 540 编目员补充的附加题名
- 541 编目员补充的翻译题名
- 545 章节题名

6 主题分析块

- 600 个人名称主题
- 601 团体名称主题
- 602 家族名称主题
- 604 名称和题名主题
- 605 题名主题
- 606 专题名称主题
- 607 地名主题
- 610 非控主题词
- 615 主题类目（暂定）
- 620 出版地/制作地检索点
- 626 技术细节检索点
- 660 地区代码
- 661 年代范围代码
- 670 保留词间关系标引法
- 675 国际十进分类法分类号
- 676 杜威十进分类法分类号
- 680 美国国会图书馆分类法分类号
- 686 其它分类法分类号
- 690 中国图书馆图书分类法分类号
- 692 中国科学院图书馆图书分类法分类号

7 知识责任块

- 700 个人名称—主要知识责任
- 701 个人名称—等同知识责任
- 702 个人名称—次要知识责任
- 710 团体名称—主要知识责任
- 711 团体名称—等同知识责任
- 712 团体名称—次要知识责任
- 720 家族名称—主要知识责任
- 721 家族名称—等同知识责任
- 722 家族名称—次要知识责任

8 国际使用块

- 801 记录来源
- 802 ISDS 中心

9 国内使用块

- 905 馆藏信息