

国家标准《中国图书馆机读规范数据格式》

编制说明

一、工作简况

（一）任务来源

根据国家标准化管理委员会2022年国家标准制修订工作安排，由全国图书馆标准化技术委员会（以下简称“图标委”）归口，《中国图书馆机读规范数据格式》（计划号：20220309-T-357）获批为推荐性国家标准制定计划项目，项目实施周期为22个月。

（二）制定背景

机读目录（Machine-Readable Catalogue，简称为MARC），是随着计算机自动化的发展而研制出来的一种计算机可识别和处理的数据格式。MARC的问世，标志着文献编目工作由传统的手工操作迈向计算机处理的自动化时代。我国对MARC的研究起步于20世纪80年代，早期主要偏重于研制中国机读书目格式。1991年，UNIMARC规范格式第1版出版，我国开始研制中国机读规范格式。《中国机读规范格式》（WH/T 15-2002）在2002年作为文化行业标准开始实施。目前，MARC格式已成为我国图书馆及情报部门信息资源交流和交换必不可少的工具。信息技术的迅速发展和信息环境的不断改变，带动了编目理论和实践的变革。为在万维网环境下高效和经济地共享规范数据，以在数量上和权威程度上满足用户的需求，国际图联UNIMARC永久委员会对UNIMARC规范格式进行了多次修订，并于2009年正式出版了《UNIMARC手册/规范格式（第3版）》（UNIMARC Manual: Authorities Format, 3rd edition），并且在近年来持续更新。此外，很多国家加入虚拟国际规范文档项目（Virtual International Authority File，简称为VIAF），以解决在语义网环境下规范数据的开放共享问题。

为了适应新的发展变化，为我国图书馆规范数据库的建立和规范数据的处理提供依据，实现不同书目机构、图书情报部门、文献管理部门的规范数据在国内、国际范围内的共享和交换，非常有必要制定适合我国国情的图书馆机读规范数据格式国家标准。

(三) 起草过程

本标准的主要起草人及分工如下：

姓名	单位	分工
王洋	国家图书馆	框架设计，7.3 编码信息块，最终统稿
顾犇	国家图书馆	1 范围、2 规范性引用文件
毛雅君	首都图书馆	7.1 记录头标、7.2 标识块
张期民	华东师范大学	7.4 规范检索点块
王璐	澳门科技大学	7.5 附注块、7.6 变异检索点块
喻爽爽	北京大学	7.9 其他语言和（或）文字的规范点块、7.10 信息来源块
喻菲	国家图书馆	7.7 相关检索点块、7.8 分类和实体历史块
王彦侨	国家图书馆	5 格式结构、6 记录头标和数据字段——基本说明、统稿
崔云红	国家图书馆	3 术语和定义、4 字段描述符号和约定
王广平	国家图书馆	7.3 编码信息块（协助编制）
张娟	首都图书馆	7.1 记录头标（协助编制）、7.2 标识块（协助编制）
袁芸	北京国图书店有限责任公司	协助统稿
王沁	北京国图书店有限责任公司	前言、引言、附录 A、附录 B

本标准的起草过程如下：

1. 前期工作（2022.4-2022.8）

2022年5月，在调研分析现有国内外相关标准和实践操作的基础上，确定标准制定的指导原则。2022年8月，召开工作组线上工作会议，明确组内分工和编写进度要求。

2. 形成标准草案初稿（2022.8-2023.8）

（1）2022年8月，依据《UNIMARC手册：规范格式》（第3版）以及国际图联网站2012年前（含）发布的修订内容，并在文化行业标准《中国机读规范格式》（WH/T 15-2002）的基础上编写形成草案第一稿，并提交中期检查自查报告。

（2）2022年11月，工作组成员依据国际图联网站在2013—2022年发布的“UNIMARC规范格式”修订内容，完成标准草案第一稿的更新。之后，继续进行汇总和修改，尤其对有疑问的地方作了重点讨论和研究。

（3）2023年8月，形成标准草案第二稿，并再次提交中期检查自查报告。

3. 形成征求意见稿（2023.9-2023.10）

2023年10月，确定标准征求意见稿及相关文件。标准工作组主要成员经过对标准草案第二稿中的文本前后有分歧的样例和文字进行了统一，并且最终在工作组内部取得一致意见，最终形成征求意见稿提交图标委。

二、标准内容及编制情况

（一）编制原则

以现行文化行业标准《中国机读规范格式》（WH/T15-2002）为基础，参照国际图联2009年出版的《UNIMARC手册/规范格式（第3版）》（UNIMARC Manual / Authorities Format, 3rd edition），吸收国际图联网站2022年及以前发布的重要修订内容，结合我国图书馆编目实践和用户需求编制进行编制。

（二）主要内容

本标准规定了计算机可读规范记录的标准结构，包括字段标识符、字段指示符和子字段标识符，以及规范记录的内容标识在磁带、软盘、光盘等载体上的逻辑和物理格式。记录的类型包括规范记录、参照记录和一般说明记录。可记录的规范数据类型包括：个人名称，团体名称，行政管理区或地理名称，商标，印刷者/出版者标志，家族名称，角色，题名，论题主题，出版、演出和出处等的地点和日期，形式、体裁或物理特征等。

正文主要包括7部分：适用范围、规范性引用文件、术语和定义、字段描述和符号约定、格式结构、记录头标和数据字段——基本说明、记录头标和数据字段——详细说明。其中第7章为本标准的核心技术内容，对各种数据字段的名称、定义、出现情况、指示符、子字段及内容等作详细说明，并提供应用示例。

（三）确定主要内容的依据

UNIMARC是国际图书馆通行的数据格式。《中国机读规范格式》（WH/T15-2002）目前仍作为现行的文化行业标准在执行中，是依据UNIMARC规范格式制定的。在实践中，图书馆也是按照该格式编制规范数据。

由于规范记录最终要和书目记录结合在一起使用，因此在格式结构上必须保证兼容，应同样地遵循GB/T 2901。对这两种记录类型中相同的数据元素，要使用相同的内容标识符，在子字段或概念相对应的地方，也尽量采用相同的设计原则。本标准与《中国机读书目格式》（GB/T33286-2016）在格式设计和内容结构上基本保持一致。

三、标准分析综述

（一）试验验证的分析、综述报告

基于已有规范格式标准和工作实践确定采用UNIMARC规范格式。

（二）技术经济论证

国家图书馆编制的名称规范数据超过200万条、主题规范数据超过12万条，CALIS(中国高等教育文献保障系统)编制的名称规范数据超过180万条，且仍在持续增长中。图书馆编制数据依据的都是UNIMARC规范格式。据此格式制定国家标准，可以最大限度地减少因数据格式转换造成的人力、物力、财力等成本。

(三) 预期的经济效益、社会效益和生态效益

新标准实施后可以在国内进一步统一标准，规范相关数据，促进规范数据在我国图书馆业界共享，并提高与相关国际书目机构进行数据交换的效率，提升我国图书馆规范数据格式的影响力。

四、与国际、国外同类标准技术内容的对比情况

1. 与《UNIMARC手册/规范格式(第3版)》相比,根据国内图书馆等信息处理机构的实际使用需求,增加了多项国内使用字段。如:

169 中国民族和历史时代代码

690 中国图书馆分类法(CLC)

692 中国科学院图书馆图书分类法(LCCAS)

696 其他图书馆分类法

2. 与现行行业标准《中国机读规范格式》(WH/T15-2002)技术内容对比,主要变化如下:

(1) 术语的更新调整

“标目”改为“检索点”;“统一标目”改为“规范检索点”;“非规范标目”改为“变异检索点”。

(2) 功能块更名

部分功能块进行了更名。如:

“标目块”改为“规范检索点块”;“单纯参照根查块”改为“变异检索点块”;“相关参照根查块”改为“相关检索点块”;“分类号块”

改为“分类和实体历史块”；“连接标目块”改为“其他语言和（或）文字的规范检索点”。

（3）控制子字段名称和范围变化

扩充了某些控制子字段内容。如：

\$5子字段的名称由“根查控制”改为“关系控制”。由2个字符扩充为5个字符。增加了作品特定关系码，行为者特定关系码，行为者与作品或内容表达之间的关系代码。

\$7子字段的名称由“文字”改为“编目文字和基本检索点文字”。由2个字符扩充为8个字符。增加了编目文字的方向、编目文字的音译法、基本检索点文字、基本检索点的文字方向、基本检索点文字的音译法。

\$8子字段的名称由“编目语种”改为“编目语种和基本检索点语种”。由3个字符扩充为6个字符。增加了基本检索点语种。

（4）新增字段

根据2009年出版的《UNIMARC手册：规范格式》（第3版）、国际图联网站2012—2022年发布的修订以及中文规范工作的实践，本标准新增了90余个字段，如：

003 永久记录标识

010 国际标准名称标识符（ISNI）

017 其他标识符

033 其他系统永久记录标识号

035 其它系统控制号

036 音乐导向

050 国际标准文本代码（ISTC）

051 国际标准音乐作品代码（ISWC）

052 国际标准视听号（ISAN）

- 061 国际标准音像制品编码（ISRC）
- 101 实体语种
- 102 实体国籍
- 104 编码数据字段：实体的主要日期
- 105 编码数据字段：内容表达的表示形式
- 106 编码数据字段：用作主题检索点的实体名称
- 120 编码数据字段：个人名称
- 122 编码数据字段：作品内容的时期
- 123 编码数据字段：行政管辖区或地理名称
- 125 编码数据字段：目标受众
- 127 编码数据字段：时长和采集信息
- 128 编码数据字段：音乐作品的形式以及调或调式
- 169 中国民族和历史时代代码
- 180 编码数据字段：形式、体裁或物理特征
- 216 规范检索点 - 商标
- 217 规范检索点 - 印刷者/出版者标志
- 223 规范检索点 - 角色
- 231 规范检索点 - 题名（作品）
- 232 规范检索点 - 题名（内容表达）
- 241 规范检索点 - 名称/题名（作品）
- 242 规范检索点 - 名称/题名（内容表达）
- 243 规范检索点 - 法律和宗教文本惯用名称/题名
- 280 规范检索点 - 形式、体裁或物理特征
- 333 用户/目标受众附注
- 334 获奖附注

- 340 传记和活动附注
- 341 与印刷/出版者相关的活动附注
- 356 地理附注
- 370 作品附注
- 371 内容表达附注
- 378 作品或内容表达的文摘
- 416 变异检索点 - 商标
- 417 变异检索点 - 印刷者/出版者标志
- 423 变异检索点 - 角色
- 431 变异检索点 - 题名 (作品)
- 432 变异检索点 - 题名 (内容表达)
- 441 变异检索点 - 名称/题名 (作品)
- 442 变异检索点 - 名称/题名 (内容表达)
- 443 变异检索点 - 法律和宗教文本惯用名称/题名
- 480 变异检索点 - 形式、体裁或物理特征
- 501 相关检索点 - 对作品负有责任的个人名称
- 502 相关检索点 - 与内容表达相关的贡献者的个人名称
- 511 相关检索点 - 对作品负有责任的团体名称
- 512 相关检索点 - 与内容表达相关的贡献者的团体名称
- 516 相关检索点 - 商标
- 517 相关检索点 - 印刷者/出版者标志
- 521 相关检索点 - 对作品负有责任的家族名称
- 522 相关检索点 - 与内容表达相关的贡献者的家族名称
- 523 相关检索点 - 角色
- 531 相关检索点 - 题名 (作品)

- 532 相关检索点 - 题名（内容表达）
- 543 相关检索点 - 法律和宗教文本惯用名称/题名
- 541 相关检索点 - 名称/题名（作品）
- 542 相关检索点 - 名称/题名（内容表达）
- 580 相关检索点 - 形式、体裁或物理特征
- 600 主题检索点 - 个人名称
- 601 主题检索点 - 团体名称
- 602 主题检索点 - 家族名称
- 606 主题检索点 - 论题名称
- 607 主题检索点 - 地理名称
- 608 主题检索点 - 作品的形式或体裁
- 610 主题检索点 - 非控主题词
- 616 主题检索点 - 商标
- 617 主题检索点 - 分级地理名称
- 631 主题检索点 - 题名（作品）
- 632 主题检索点 - 题名（内容表达）
- 640 与实体有关的地点和日期
- 641 主题检索点 - 名称/题名（作品）
- 642 主题检索点 - 名称/题名（内容表达）
- 696 国内其他分类法
- 716 其他语言和（或）文字的规范检索点 - 商标
- 717 其他语言和（或）文字的规范检索点 - 印刷者/出版者标志
- 723 其他语言和（或）文字的规范检索点 - 角色
- 731 其他语言和（或）文字的规范检索点 - 题名（作品）
- 732 其他语言和（或）文字的规范检索点 - 题名（内容表达）

- 741 其他语言和（或）文字的规范检索点 - 名称/题名（作品）
- 742 其他语言和（或）文字的规范检索点 - 名称/题名（内容表达）
- 743 其他语言和（或）文字的规范检索点 - 法律和宗教文本惯用名称/题名
- 780 其他语言和（或）文字的规范检索点 - 形式、体裁或物理特征
- 822 其他系统的相关数据
- 836 被替换检索点信息
- 856 电子资源定位与检索
- 886 无法从源格式转换的数据
- （5）取消字段
- 015 国际标准规范数据号

五、与相关国际标准的关系

本标准规范借鉴了现行的相关标准（含国际标准）研究成果，见规范性引用文件及参考文献。是图书馆领域国际通行标准在我国图书馆事业和图书馆标准化工作中的落实与应用。

六、与有关法律、行政法规及相关标准的关系

本标准符合现行法律、法规和强制性国家标准的相关要求。

本标准在编制过程中坚持维护与现行图书馆有关的各种法律、法规、强制性国家标准一致性和协调性的原则，与现行图书馆相关各种法律法规没有冲突。标准编写过程中的编制依据包括《标准化工作导则 第1部分：标准的结构和编写》（GB/T1.1—2020）及规范性引用文件列出的国家标准。

七、重大分歧意见的处理经过和依据

本标准制定的前期工作成果文化行业标准《中国机读规范格式》（WH/T15-2002）在图书馆业界实际已实施多年，未出现较大分歧。此次制定更具普适性的国家标准中无重大分歧意见。

八、涉及专利的有关说明

无

九、实施标准的要求和措施建议

（一）实施标准的要求

在标准发布后，需要组织对业界相关专职人员进行培训，以增强他们对标准技术内容的认识和理解，同时提高他们在标准使用中的技能水平。

（二）组织措施、技术措施

为更好的宣贯本标准，在标准正式发布后，计划编写标准的应用指南，由归口单位组织开展专项标准宣讲培训工作，同时结合业界重大学术交流活动对本标准加以宣贯。

（三）标准过渡期和实施日期的建议

与之前的文化行业标准相比，本标准作了比较大的修订。为留给业界充分的时间学习，建议标准正式发布一年后实施。

（四）废止现行有关标准的建议

本标准发布实施后，因名称调整，为避免产生使用中的混淆，同时维护国家标准和行业标准的权威性，建议文化行业标准《中国机读规范格式》（WH/T15-2002）予以废止。

十、其他应当说明的事项

无

《中国图书馆机读规范数据格式》标准起草组

2023年10月24日