

第六章 用户需求书

按《2018年数字图书馆推广工程建设任务(首批)立项通知要求》(国图函[2017]57号)、《关于开展数字图书馆推广工程2018年度数字资源联合建设(首批)工作的通知》(国图函[2018]2号)、海南省财政厅关于提前下达2018年中央补助地方公共文化服务体系建设专项资金(第二批)的通知(琼财教[2017]2173号)要求,结合海南省图书馆2018年自身数字资源建设规划,开展以下数字资源建设的招标投标工作:

项目名称	建设量	提交成果说明
网事典藏	100(个)	<p>▲1. 按推广工程数字资源联合建设网事典藏项目建设要求,建设完成100条网站录入工作,内容包含海南省行政区域的政治、经济、文化发展等信息的网站。并将采集到的网站进行编目、索引,质检合格后并发布。</p> <p>▲2. 按规定日期内完成抓取、录入工作,并提交采集清单、元数据及第三方质检报告。</p>
地方文献 报纸数字 化	10000版	<p>1. 根据《关于开展数字图书馆推广工程2018年度数字资源联合建设工作的通知》内数字资源联合建设地方文献数字化项目建设方案实施。</p> <p>▲2. 需要提交的内容包括:TIFF图像(单页)、</p>

		<p>双层 PDF（单板和合并版）、TXT 文件（单板和合并版）、对应数据库、元数据、说明文件及文献的书目数据（ISO 格式）、XML 文件。数据提交时以上内容各上交一份，具体建设标准见“地方文献数字化具体要求”部分。</p> <p>▲3. 提交“地方文献验收数据提交单”和第三方质检报告（供应商提供承诺函）。</p> <p>▲4. 提供报纸发布平台，并将全部数字化资源录入到该报纸发布平台内，采购方提供平台发布的相关网络和硬件设备。</p>
<p>备注：所有加工数据保存所需的存储介质（移动硬盘）统一由成交供应商提供。</p>		

一. 推广工程数字资源联合建设地方报纸数字化加工规则（2018）

规则适用范围：

本技术规则用于 2018 年数字图书馆推广工程地方报纸数字化项目。本规则针对地方报纸规定了加工编号、对象数据、元数据以及命名规则存储规则等应遵循的标准。

1 元数据加工标准

1.1 规范性引用文件

下列文件对于本规则的应用是必不可少的。

《中国文献编目规则》（第二版）

《GB/T25100-2010 信息与文献 都柏林核心元数据》

《信息资源的内容形式和媒体类型标识》(GB/T3469-2013)

1.2 著录对象

以数字化文献为著录对象。一般以具有独立名称、并可独立使用的一个数字资源为著录单位。

报纸的著录单位为一种报纸。

对于连续性资源的著录对象是否发生改变,可依据题名、发行者、版本等的变化情况进行判断,具体规则参照《中国文献编目规则》(第二版)。

1.3 著录信息源

著录信息源是资源本身。资源本身是著录的首选来源。资源本身信息不足,可参考其他信息源。

报纸的著录信息源是本题名下第一期或第一部分,若无法获得本题名下第一期或第一部分,应以可获得的最早发行的一期或部分著录。

1.4 著录用文字

题名、责任者等取自资源本身的信息一般按资源本身的文字著录。

由编目员进行描述或标引的信息均采用简体中文著录。

有规范或规定编码标准的要求按标准著录,如:语种、中图分类。

日期用阿拉伯数字著录。

推荐采用 ISBD 著录标识符。

一个著录项中多个著录内容之间以半角分号分隔。

1.5 元数据著录简表

术语	必备性	著录内容
----	-----	------

加工编号	必备	著录资源的加工编号。
正题名	必备	著录资源的主要名称，根据在资源中出现的形式著录。
其他题名	有则必备	除正题名以外的题名信息。包括其他语种的题名或用于限定、补充、解释正题名的题名信息。
中图分类	必备	描述资源主题内容的中图分类法类号。依照 CLC（《中国图书馆分类法》第五版）进行标引。
主题词或关键词	必备	描述资源主题内容的受控或非受控的词汇。受控主题词依据 CCT（中国分类主题词表）进行标引。
附注	有则必备	<p>关于资源内容的其他说明。凡未在其他字段中著录而又有必要补充说明的内容均可著录于本字段。如：正题名的变化、书目沿革等。可采用自由行文方式著录。</p> <p>版权信息应著录于本字段。</p> <p>应尽可能采用固定导语和规范用语。</p> <p>不同类型信息应重复本字段。</p> <p>例 1：附注：授权范围：互联网使用</p> <p>例 2：正题名：中国晚报</p> <p style="padding-left: 40px;">附注：1990 年 4 月 1 日至 1990 年 4 月 26 日期间与东方晚报联合出版发行</p>
出版者名称	有则必备	使资源可以获得和利用的责任实体，如资源的出版者或颁布者。

		如果连续出版过程中出版者发生变化，若认为这一变化重要，可在附注中注明改变后的出版者。
出版地	有则必备	著录出版者所在的城市名称。如具体出版地不详，可著录所在省名或国名。
开始日期	有则必备	著录报纸最早一期的正式出版日期。著录内容是该报出版情况的客观反映，与编目机构的收藏情况及数字化情况无任何联系。 非公元纪年应如实著录，其后再著录公元纪年，并置于方括号内。 例：开始日期：民国 15 年[1926]
终止日期	有则必备	著录报纸结束出版的日期。著录内容是该报出版情况的客观反映，与编目机构的收藏情况及数字化情况无任何联系。
内容形式	必备	著录资源的内容形式词及内容限定词。取值参见国标《信息资源的内容形式和媒体类型标识(GB/T 3469-2013)》。 例：正题名：中国儿童画报 内容形式：文本；图像（静态）
媒体类型	必备	著录资源的媒体类型。取值参见国标《信息资源的内容形式和媒体类型标识(GB/T 3469-2013)》。 本项目中缺省取值为“电子”。
格式	必备	资源的格式。需用英文大写著录。

		例：TIF;PDF
语种	必备	著录资源的文字语种。采用3位语种代码，参考《新版中国机读目录格式使用手册》（附录A）。
关联	有则必备	描述与本资源相关的其他资源。这些关系包括继承、替代、合并、分自、吸收、其他版本与原版本、其他格式与原格式等。相关资源一般独立存在。可著录资源名称或资源标识符。
时间范围	有则必备	著录资源内容涉及的时间特征，著录内容包括时代、日期、时间段。 例：正题名：人民政协报 时间范围：现代
空间范围	有则必备	著录资源内容涉及的空间特征，著录内容包括地点、地理坐标。
适用对象	必备	资源适宜使用的人群，取值为：普通用户、少年儿童、残障人士。 例：正题名：中国儿童画报 适用对象：少年儿童
出版频率	必备	著录报纸出版的时间间隔。如：日报、周报、每周3期、不定期出版、不详。
馆藏范围	必备	著录馆藏电子报纸的收录范围。 如果某一时间段内的电子资源全部存在，按时间段著录；如果存在缺失情况，则需分段著录，中

		间用半角分号分隔。
ISSN	有则必备	著录在国际标准连续出版物编号中确认资源的唯一标识。
统一刊号	有则必备	著录在国内统一刊号中确认资源的唯一标识。
CDOI	有则必备	著录在 CDOI 中确认资源的唯一标识。
发布地址	有则必备	资源在因特网或局域网的发布地址。
数据提交 单位	必备	著录提交该数据的单位。
所属任务 年份	必备	著录联建工作的任务年度。

2 图像数字化标准和命名规则

2.1 图像扫描加工

(1) 扫描前根据国际色彩协会(International Color Consortium, 简称 ICC) 标准, 做加工设备的基本色彩校正, 及针对各类型文献进行色彩校正;

(2) 逐版扫描;

(3) 黑白页和灰度页用灰度方式扫描

色彩位深: 8 位

分辨率: 300 dpi

档案典藏级格式: TIFF 不压缩

(4) 彩色页用彩色方式扫描

色彩位深: 24 位

分辨率：300 dpi

档案典藏级格式：TIFF 不压缩

(5) 纠偏处理。对出现偏斜的图像进行纠偏处理，图像歪斜度不可以超过一度，对方向不正确的图像进行旋转还原，以符合阅读习惯；

(6) 去污处理。对图像页面中出现的严重影响图像质量的黑边等进行去污处理。

2.2 图像文件质量要求

(1) 图像文件放大到 1:1 状态，逐版检查。检查文件是否有彩点、彩线、太淡、太浓、黑边、污点、歪斜、模糊（马赛克等）或图像内容不完整等现象。若不符合图像质量要求应进行图像校正或重新扫描；

(2) 图像文件真实反映原件，同种报纸图像尺寸相同，不得有失真现象；按版次顺序由小到大，符合阅读习惯，不能有缺版、错版、数据内容缺失等现象（原件有缺失的除外）；

(3) 图像符合扫描规格要求和技术参数；

(4) 所有文件保存位置以及文件命名正确，可以有效打开和显示；

(5) 图像综合错误率不超过 1%。

2.3 命名规则

注：档案典藏级文件（TIFF 格式）与其相应的发布文件（双层 PDF 格式和 XML 格式）命名规则相同，这里仅以 TIFF 格式文件为例，文件后缀名均为小写。

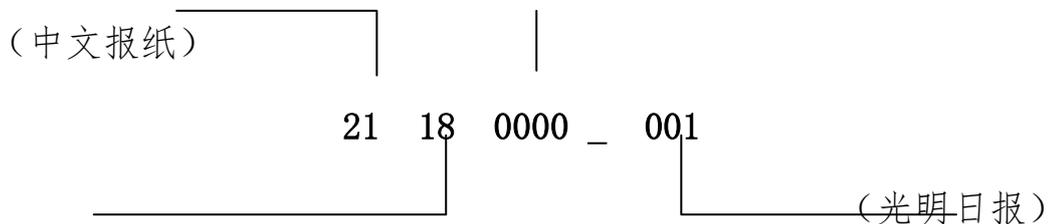
2.3.1 加工编号

加工编号是报纸数字化加工过程中一种报纸的唯一标识号，由

11 位数字和下划线组成。文献基本资料类型（1 位）、文献语种（1 位）、任务年（2 位）、机构代码（4 位）、下划线（1 位）、单位内部报纸种类编号（3 位）。中文报纸的文献基本类型为（2：报纸），文献语种编码为（1：中文）。如：国家图书馆 18 年的光明日报就可以表示为：21180000_001。

例：

文献基本资料类型及语种机构代码（国家图书馆）



任务年（2018）下划线内部报纸
种类编号

图像扫描文件的保存路径为“\加工编号\出版日期（八位数字：年 4 位、月 2 位、日 2 位）”。

如：国图任务年为 2018 年，出版日期为 1950 年 5 月 16 日的光明日报（单位内部编号为 001）的路径应为：

\TIFF\21180000_001\19500516\

2.3.2 文件命名规则

每期报纸按照实际版次命名，用三位数字表示，后缀名小写 tif。如第一版为 001.tif，第 2 版为 002.tif；则若报纸包含号外、增刊、

副刊、特刊等特殊版面，且其不属于报纸原有常规版次内容，为独立的另起版面，其文件的命名规则为：特殊版面种类拼音首字母(大写)+版次(2位，如有实际印刷版次，按实际命名，如没有则按01开始顺序命名)，如号外第一版 H01.tif、增刊第一版 Z01.tif、副刊第一版 F01.tif、特刊第一版 T01.tif；若报纸中存在通版(即打通同一面上两个相邻版而形成的版)情况，应将这两版整体扫描为一个图像文件，并同时将其保存为这两版的图像，如当报纸中的第2版和第3版为通版时，通版图像整体保存为一个文件，并且将这一图像同时命名为 002.tif 和 003.tif。

如：加工编号为 21180000_001，出版日期是 1953 年 12 月 7 日，报纸常规版面 4 版，有特刊独立版面 2 版，则其存储为：

\TIFF\21180000_001\19531207\001.tif

002.tif

003.tif

004.tif

T01.tif

T02.tif

3 双层 PDF 标准

3.1 双层 PDF 加工

将处理完成后的图像进行报纸篇目文字识别校对。采用图在文上的模式进行双层 PDF 输出。

PDF 文件根据图像尺寸、颜色、数据存储量、按 JPEG2000 有损

压缩，压缩因子适度动态调整，在确保图像清晰的情况下，尽量压缩图像文件所占空间至最小。

3.2 PDF 文件质量要求

- (1) 兼容 adobe reader6.0 及以上版本。
- (2) 在保持图像清晰可读的基础上尽可能减小存储量。
- (3) PDF 文字层所使用的字体以”已嵌入子集”方式嵌入 PDF 文件。
- (4) 双层 PDF 文件的图像层和文字层的文字对位准确，反显区域与文字区域相差 1 毫米以内。
- (5) 双层 PDF 错误率不超过 0.3%。

4XML 文件标准

文件应著录全面的报纸信息，记录标题及篇目位置等信息，标题位置为引题、标题和副题的整体位置，遵守 XML 语法规则，按规范标签进行制作 XML，使用 utf-8 字符集，错误率小于 0.3%。

<报纸元数据>

<加工编号>XXXX</加工编号>

<题名>XXXX</题名>

<出版日期>XXXX</出版日期>

<卷期>XXXX</卷期>

<版次>XXXX</版次>

<整版 PDF 链接>XXXX</整版 PDF 链接>

<备注>

//非必备节点

<题名备注>XXXX</题名备注>

<出版日期备注>XXXX</出版日期备注>

<卷期备注>XXXX</卷期备注>

<版次备注>XXXX</版次备注>

.....

</备注>

</报纸元数据>

<版元数据>

<正文组>

<正文 篇目号="001">

<栏目>XXXX</栏目>

<引题>XXXX</引题>

<标题>XXXX</标题>

<副题>XXXX</副题>

<标题坐标>XXXX</标题坐标>

<篇目坐标>XXXX</篇目坐标>

<内截坐标>XXXX</内截坐标>

<转版信息>

<转版坐标 版次=" XXX" > XXXX </转版坐标>

<转版坐标 版次=" XXX" > XXXX </转版坐标>

.....

</转版信息>

<作者>XXXX</作者>

<附图组>

<附图 附图号="001_01" 版次="XXX">

<附图坐标>XXXX</附图坐标>

</附图>

<附图 附图号="001_02" 版次="XXX">

<附图坐标>XXXX</附图坐标>

</附图>

.....

</附图组>

</正文>

<正文 篇目号="002">

<栏目>XXXX</栏目>

.....

</正文>

</正文组>

<图片组>

<图片 篇目号="003">

<图片标题>XXXX</图片标题>

<标题坐标>XXXX</标题坐标>

<篇目坐标>XXXX</篇目坐标>

<图片作者>XXXX</图片作者>

</图片>

</图片组>

</版元数据>

5 数据库标准

5.1 数据库命名

每种报纸单独制作数据库，其命名与报纸加工编号相同，如：

21180000_001.mdb。

5.2 数据库制作

5.2.1 报纸基本信息表 (book)

序号	中文名称	字段名称	字符类型	必备性	对应书目数据 MARC 内容
1	001	record_id	char	有则必备	MARC 数据中的 001 字段
2	加工编号	book_id	char	必备	加工编号
3	题名	book_name	char	必备	MARC 数据中的 200\$a
4	出版地点	pub_place	char	有则必备	MARC 数据中的 210\$a
5	出版者	pub_house	char	有则必备	MARC 数据中的 210\$c
6	出版时间	pub_date	char	有则必备	MARC 数据中的 210\$d
7	出版周期	pub_T	char	有则必备	MARC 数据中的 326\$a \$b (著录形式为 a(b); 326 字段若重复, 用空格隔 开)

8	起止卷期	start_end _day	备注	有则必备	MARC 数据中的 207\$a (用 空格分隔\$a 重复字段)
9	更名信息	name_chan ge	char	有则必备	MARC 数据中的 311\$a
10	附注信息 (号外、 增刊等)	general_n otes	char	有则必备	MARC 数据中的 300\$a
11	特殊附注 (出版规 律)	specific_ notes	备注	有则必备	MARC 数据中的 315\$a
12	总卷期	pub_num	char	有则必备	MARC 数据中的 215\$a
13	主题词	topic_wor d	char	有则必备	MARC 数据中的 610\$a

注：上述内容从书目数据 (MARC) 的相应字段中提取。

5.2.2 报纸版面篇目信息表 (catalog)

序号	中文名称	字段名称	符类型	必备性	备注
1	加工编号	book_id	char	必备	加工编号
2	标引序号	serial_nu m	数字	必备	
3	出版日期	date	char	必备	4 位年 2 位月 2 位日
4	版次	space_id	char	必备	“1” 表示第一版 “2” 表

					示第二版“H1”表示号外 第一版
5	栏目	column	char	有则必 备	
6	篇目号	specific_ id	数字	必备	“1”表示第一篇“2”表 示第二篇
7	篇名	specific_ name	备注	必备	对应 XML 文件中的“标 题”
8	标题坐标	specific_ position	备注	必备	引题、标题和副题的整体 坐标
9	篇目坐标	text_posi tion	备注	必备	篇目的整体坐标
1 0	内截坐标	other_pos ition	备注	有则必 备	被完全嵌入到本篇的其 他篇的区域
1 1	作者	author	char	有则必 备	
1 2	图像页码	ppage_num	char	必备	对应图像命名,第一版为 “1”,第二版为“2”, 增刊第一版为“Z01”

注: catalog 表与 XML 文件中的对应内容应一致,用于链接对象数据的是加工编号、出版日期、版次、篇名和图像页码。

5.2.3 报纸结构信息表 (struct)

序号	中文名称	字段名称	字符类型	必备性	备注
1	加工编号	book_id	char	必备	
2	总期数	ppub_num	数字	必备	实际加工期数
3	版面数	space_num	数字	必备	与 PDF 文件数量一致
4	篇目数	specific_num	数字	必备	加工的总篇目数
5	特种刊期数	te_num	数字	必备	加工中的特种期刊数
6	保存级图像数量	file_num	数字	必备	TIFF 图像文件数量
7	保存级数据存储量	tiff_mb	数字	必备	MB (小数点后保留两位)
8	发布级数据 PDF 存储量	pdf_mb	数字	必备	MB (小数点后保留两位)
9	发布级数据 XML 存储量	xml_mb	数字	必备	MB(小数点后保留两位)

10	保存级硬盘编号	hdA_plac e	char	必备	保存级文件所在硬盘编号
11	发布级PDF硬盘编号	hdB_plac e	char	必备	发布级PDF文件所在硬盘编号
12	发布级XML硬盘编号	hdD_plac e	char	必备	发布级XML文件所在硬盘编号

5.3 数据库制作及要求

- (1) 标引数据库以 MDB 数据库方式提交，后缀名 mdb；
- (2) 基本信息表的内容应与 MARC 数据相应内容保持一致；
- (3) 版面篇目信息数据库标引要求真实反映报纸原貌；
- (4) 结构信息表应严格按文献实际情况进行描述；
- (5) 无法录入的生僻字等用“■”表示；
- (6) 版面篇目信息表与 XML 文件的对应元素项的内容应一致；
- (7) 各种著录、说明文件的文字、符号、版式、位置和文件名称准确，其综合错误率不超过 0.3%。

6 说明文件

6.1 数字化成品总清单

说明文件格式：Microsoft Excel

说明文件命名：地方文献数据说明（说明文件的 5 张表需保存在同一个 Excel 文件的不同工作表中，每个工作表按下列表命名）。

6.1.1 提交数据总体说明

名称	说明
项目名称	填写项目名称，如 2018 年地方报纸数字化
承建馆	承建馆名称全名
质检单位	第三方质检单位全名
项目时间	填写项目所属年份
数据加工编号	以每种报纸的起始～终止表示，如 (21180000_001) 19540101~19761212
完成报纸总种数	这批数据报纸的总种数
完成报纸总期数	这批数据报纸的总期数
完成报纸总版数	这批数据报纸的总版数
序号	以 1 开始，顺次编号
存储内容	如：扫描图像，PDF 文件，XML 文件等
数据格式	TIFF，PDF 等
存储介质	硬盘
介质编号	硬盘编号，如 21180000
介质数量	存储硬盘的数量
存储量	以 GB 为单位表示
备注	

6.1.2 保存级对象数据硬盘存储清单

名称	说明
数据内容	存储的数据内容，如 2018 年地方报纸数字化

数据格式	以大写英文字母表示，如 TIFF
序号	以 1 开始，顺次编号
硬盘编号	本硬盘的硬盘编号
数据加工编号 (起始/终止)	以每种报纸的起始～终止表示，如 (21180000_001) 19540101~19761212
种数	本块硬盘存储的报纸总种数
期数	本块硬盘存储的报纸总期数
版数	本块硬盘存储的报纸总版数
存储量	以 GB 为单位表示
备注	需要说明的特殊情况
合计	合计这批数据报纸的种数、期数、版数、存储总量

6.1.3 发布级 PDF 对象数据硬盘存储清单

名称	说明
数据内容	存储的数据内容，如 2018 年地方报纸数字化
数据格式	以大写英文字母表示，如 PDF、XML
序号	以 1 开始，顺次编号
硬盘编号	本硬盘的硬盘编号
数据加工编号 (起始/终止)	以每种报纸的起始～终止表示，如 (21180000_001) 19540101~19761212
种数	本块硬盘存储的报纸总种数
期数	本块硬盘存储的报纸总期数

版数	本块硬盘存储的报纸总版数
存储量	以 GB 为单位表示
备注	需要说明的特殊情况
合计	合计这批数据报纸的种数、期数、版数、存储总量

6.1.4 发布级 XML 对象数据硬盘存储清单

名称	说明
数据内容	存储的数据内容，如 2018 年地方报纸数字化
数据格式	以大写英文字母表示，XML
序号	以 1 开始，顺次编号
硬盘编号	本硬盘的硬盘编号
数据加工编号 (起始/终止)	以每种报纸的起始～终止表示，如 (21180000_001) 19540101~19761212
种数	本块硬盘存储的报纸总种数
期数	本块硬盘存储的报纸总期数
版数	本块硬盘存储的报纸总版数
存储量	以 MB 为单位表示
备注	需要说明的特殊情况
合计	合计这批数据报纸的种数、期数、版数、存储总量

6.1.5 每种文献数据量统计表

名称	说明
----	----

序号	以 1 开始，顺次编号
加工编号	12 位加工编号
题名	报纸名称
总期数	该种报纸加工的总期数
版面数	该种报纸加工的总版面数
TIFF 文件数量	该种报纸的 TIFF 图像数量
篇目数	该种报纸的篇目数
特种刊数量	该种报纸特种刊的数量
备注	
合计	合计这批数据总期数、版面数、TIFF 文件数量、篇目数、特种刊数量

6.2 每个存储介质内的说明文件 (readme.txt)

- 项目名称；如：2018 年地方报纸数字化。
- 文献说明：对象来源（如：馆藏报纸）；载体规格（xx*xx）。
- 技术参数；存储格式；扫描原始设备（设备生产商）；扫描系统软件；扫描分辨率；色彩位深。
- 文件存储量；如：种、期、版、存储容量（GB）。
- 制作单位；如：国家图书馆。
- 完成时间；如：201x 年 xx 月。
- 备注；其他需说明情况。

7. 存储规则

7.1 存储介质及命名

存储介质为硬盘，硬盘命名由 8 位组成：

文献基本资料类型（1 位）、文献语种（1 位）、加工年（2 位）、机构代码（4 位）。

7.2 数据存储规则

存储内容包括：TIFF 图像、双层 PDF、XML、对应数据库、说明文件、介质说明文件，书目数据文件各一份。

各类型对象数据按加工编号递增顺序依次放置在硬盘内，保存目录及文件名由四级组成：

第一级目录为加工编号的前 8 位，如 21180000，同时将介质说明文件并行保存 (readme.txt)；

第二级目录为 TIFF,PDF,XML 同时将对应数据库、说明文件 (21180000.xls)、书目数据文件 (21180000.iso) 并行保存；

第三级目录为加工编号；

第四级目录为相应格式的对象数据文件；

第五级目录为相应格式的对象数据文件。

|—21180000、readme.txt

| 鏗 |——TIFF

| 鏗 | 鏗 鏗 |—加工编号

| | | |—日期

| | | | |—001.tif

| | | | | 002.tif……

| | | | | 日期……

| 鏗 | 鏗 鏗 | 加工编号

| | | | | 日期

| | | | | 日期……

| 鏗 | 鏗鏗 | ……

| 鏗 | PDF

| | | | | 加工编号

| 鏗 | 鏗鏗 | 加工编号……

| 鏗 | XML

| | | | | 加工编号

| 鏗 | 鏗鏗 | 加工编号……

| | | 21180000.iso

| 鏗 | 21180000_001.mdb

| 鏗 | 21180000.xls

二. 推广工程数字资源联合建设网事典藏网站采集元数据著录规则

(2018)

1. 著录对象

著录对象为存档的网站，包括核心政府机构、事业单位、文化、艺术、科普等网站。以单次存档的网站为一个著录单位。如果一个网站具有多个主页域名，著录时作为一个对象著录。

2. 著录要求

对采集的网站进行编目加工，要求参照著录规则进行编目。元数据以 EXCEL 文件形式提交，元数据著录规则如下：

术语	必备性	著录内容
加工编号	必备	著录元数据的一个明确标识，具体规则见《网站采集加工编号命名规则》。
CDOI	有则必备	著录所采集网站的唯一标识号。
网站名称	必备	赋予资源的名称，一般指网络资源正式公开的名称。
网站其他名称	必备	统一著录为“××网站”，对网站名称进行规范与解释说明，如果是政府网站著录，此处要求按照省、市、区县的顺序著录网站全名。例如：“朝阳区人民政府网”其网站其他名称著录应为“北京市朝阳区人民政府网站”。
摘要	必备	著录网站内容的总结概括性文字。摘要字数要求 200 字以内。建议格式为：摘要内容分两部分，第一部分对网站内容进行整体概括，第二部分把网站包含的栏目名称列出。比如：江西省人民政府门户网站于 1999 年 11 月建成开通，由江西省人民政府办公厅主办，江西省信息中心承办，是互联网上政府信息公开、回应社会关切、提供政务服务的重要载体，是各级政府网站集中展示窗口，是省政府各部门履行职责的重要平

术语	必备性	著录内容
		台，现设有览省情、寻资讯、看政务、找政策、享服务、爱问政和用数据等栏目。
关键词	必备	著录体现网站主要内容的名词或名词短语。如有多个关键词，以英文半角分号间隔。
资源类型	必备	著录所保存资源的类型。统一著录为“网站”。
内容形式	必备	著录内容形式及内容限定。参考国家标准 GB/T 3469—2013《信息资源的内容形式和媒体类型标识》取值。
媒体类型	必备	著录用以承载资源内容的载体类别。参考国家标准 GB/T 3469—2013《信息资源的内容形式和媒体类型标识》取值。网络信息保存资源媒体类型统一著录为“电子”。
语种	必备	著录网站的 3 位语种代码，可参考《新版中国机读目录格式使用手册》。如有多个语种，以英文半角分号间隔。
保存格式	必备	著录所采集的网站资源存档格式。统一著录为“WARC”。
机构名称	有则必备	著录网站的所属机构名称。著录时应以通用性、惯用性为选取原则。如网站中出现多个不同的名称，选择网站最显著位置的名称。

术语	必备性	著录内容
关联	有则必备	著录与当前资源存在某种关系的其他资源。
访问方式	必备	著录资源可以提供服务的范围，取值：互联网访问、推广工程专用网络访问、××图书馆局域网访问等。
采集日期	必备	著录网站采集的日期。如果在审核过程中需重新采集，应对本项内容进行修改。
发布日期	必备	著录存档资源发布的日期。
采集地址	必备	著录网站的原始访问地址。
发布地址	必备	著录存档资源的发布地址。
附注	有则必备	凡未在其他著录项中著录而又有必要进一步补充说明的内容，均可著录于本项。
数据提交单位	必备	著录承建馆的名称。
所属任务年份	必备	著录联建工作的任务年度，2018年度数据则著录2018。

3. 网站采集加工编号命名规则

网站采集加工编号统一定长为15位。具体组成是：资源类型代码（1位，网站：W）、采集机构代码（4位，数值取自机构代码，2-5字符位）、所属任务年份（4位，6-9字符位）、流水号（6位，10-15字符位）。注意：所属任务年并非采集年，同一任务年的流水号应顺序排列，不同存档资源流水号不可重复。

例如：

网站：W17002018000001

机构代码见《推广工程数字资源联合建设机构代码》。

4. 著录样例

加工编号	W17002018000001
CDOI	108.ndlc.39.4200009031010001/T8F45.1000000004
网站名称	湖北省人力资源和社会保障厅
网站其他名称	湖北省人力资源和社会保障厅网站
摘要	该网站为湖北省人力资源和社会保障厅网站，由湖北省人力资源和社会保障厅办公室、湖北省人力资源和社会保障信息中心主办，主要包括新闻中心、信息公开、服务大厅、互动交流、12333 咨询、政策法规等栏目。
关键词	湖北；人力资源；社会保障
资源类型	网站
内容形式	多种内容形式
媒体类型	电子
语种	chi
保存格式	WARC
机构名称	湖北省人力资源和社会保障厅
关联	湖北省人事厅；湖北省劳动和社会保障厅
访问方式	推广工程专用网络访问

采集日期	2018-03-11
发布日期	2018-06-19
采集地址	http://www.hb.hrss.gov.cn/
发布地址	http://10.118.62.109:8080/wayback/20150414014759/h http://www.hb.hrss.gov.cn/
附注	将原省湖北人事厅、原湖北省劳动和社会保障厅的职责整合，划入省人力资源和社会保障厅。
数据提交单位	湖北省图书馆
所属任务年份	2018

三、参照标准：

本次建设项目供应商应严格参照国家图书馆的下列建设标准：

- 1、专门元数据标准与著录规范——古籍
- 2、专门元数据标准与著录规范——家谱
- 3、专门元数据标准与著录规范——电子图书
- 4、专门元数据标准与著录规范——电子连续性资源
- 5、专门元数据标准与著录规范——学位论文
- 6、专门元数据标准与著录规范——期刊论文
- 7、专门元数据标准与著录规范——网络资源
- 8、专门元数据标准与著录规范——音频
- 9、专门元数据标准与著录规范——视频

- 10、专门元数据标准与著录规范——图像
- 11、国家图书馆数字资源长期保存规范
- 12、国家图书馆管理元数据规范
- 13、本项目在期建设中涉及到的其他相关国家及行业建设标准
- 14、推广工程数字资源联合建设地方文献数字化加工规则(2015)
- 15、推广工程数字资源联合建设唯一标识符注册数据著录规则(2016)
- 16、推广工程数字资源联合建设元数据仓储著录规则(2016)
- 17、推广工程数字资源联合建设政府信息公开信息元数据著录规则(2017)
- 18、推广工程数字资源联合建设专题资源加工及著录规则(2016)
- 19、推广工程数字资源联合建设网事典藏元数据著录规则(2018)
- 20、推广工程数字资源联合建设地方报纸数字化加工规则(2018)

四、其他要求

(1) 供应商在项目完成时提供本次项目建设的详细建设目录表，项目提交数据载体为移动硬盘由成交供应商提供，项目发布载体（服务器、存储等）由采购人提供；

▲ (2) 供应商将相关数据录入到采购方指定的平台内；

▲ (3) 项目完成时间为 2018 年 8 月 31 日前（最终完成时间或以采购方组织的项目终验时间为准），并在规定时间内免费将未通过国家图书馆验收的条目进行修改，直至通过相关验收；

▲ (4) 成交价为最终包干价，此价格包括本项目约定事项所产生

的所有费用（包括但不限于加工、调试、录入、阶段验收、最终验收等项目所产生的一切费用），在项目执行期间为不变价；

▲（5）供应商须承诺：如非采购方原因，未达到采购方对此次项目的各项要求，供应商同意退还采购方已支付的相应款项。