

招 标 文 件

项目名称：馆藏文献元数据及全文 OCR 数据制作

项目编号：310000000250327197494-00227538

预算编号：0025-00000245、0025-00000249、
0025-00000246、0025-00000248、
0025-00000250、0025-00000247

内部编号：STC25A114

采购人：上海图书馆（上海科学技术情报研究所）

2025年03月27日
采购代理机构：上海中招招标有限公司

二〇二五年三月
2025年03月27日

第一章 投标邀请书

投 标 邀 请 书

上海中招招标有限公司（以下简称“采购代理机构”）受上海图书馆（上海科学技术情报研究所）（以下简称采购人）委托，就馆藏文献元数据及全文 OCR 数据制作项目进行国内公开招标。现邀请有意向的投标人参加投标。

一、项目概况：

1. 项目编号：310000000250327197494-00227538
2. 内部编号：STC25A114
3. 预算编号：0025-00000245、0025-00000249、0025-00000246、0025-00000248、0025-00000250、0025-00000247
4. 项目名称：馆藏文献元数据及全文 OCR 数据制作项目
5. 项目数据量：中文老旧报纸全文 OCR 元数据制作 154,600 版，科技现刊篇名元数据制作 900,000 篇，老旧期刊篇名元数据制作（中）100,000 篇，会议论文篇名元数据制作 40,000 篇，学术图片库元数据制作 40,000 篇，老旧期刊篇名元数据制作（英）35,000 篇。
6. 预算金额：人民币 443.408 万元，其中中文老旧报纸全文 OCR 元数据制作 298.378 万元，科技现刊篇名元数据制作 104.4 万元，老旧期刊篇名元数据制作（中）15.6 万元，会议论文篇名元数据制作 4.24 万元，学术图片库元数据制作 9.24 万元，老旧期刊篇名元数据制作（英）11.55 万元。超过预算的投标将被拒绝。
7. 项目完成时间：2025 年 10 月 31 日

二、供应商资格要求：

1. 满足《中华人民共和国政府采购法》第二十二条规定：
 - 1) 具有独立承担民事责任的能力；
 - 2) 具有良好的商业信誉和健全的财务会计制度；
 - 3) 具有履行合同所必需的设备和专业技术能力；
 - 4) 有依法缴纳税收和社会保障资金的良好记录；
 - 5) 参加政府采购活动前三年内，在经营活动中没有重大违法记录；
 - 6) 法律、行政法规规定的其他条件。

2. 落实政府采购政策需满足的资格要求：本次招标执行支持中小微企业、强制/优先节能产品、鼓励环保产品、促进残疾人就业、支持监狱和戒毒企业、扶持不发达地区和少数民族地区等相关政策。

3. 本项目的特定资格要求：

1) 根据《上海市政府采购供应商登记及诚信管理办法》已登记入库的供应商：本次招标为网上投标，投标人必须获得上海市电子签名认证证书（CA 认证证书）及成功注册“上海政府采购网”（<http://www.zfcg.sh.gov.cn>）；

2) 供应商应在投标文件中提供无不良信用记录的相关证明（相应网页截图）。不良信用记录指：供应商在“信用中国”网站（www.creditchina.gov.cn）被列入失信被执行人、重大税收违法案件当事人名单或在中国政府采购网（www.ccgp.gov.cn）被列入政府采购严重违法失信行为记录名单；

3) 本项目面向中小企业；

4) 本项目不允许采购进口产品；

5) 是否接受联合体投标：**不允许**

三、招标文件的获取：

凡愿参加本项目的供应商可于本公告发布之日 **2025-03-27** 起至 **2025-04-03** 截止，每天 **00:00:00~12:00:00**，下午 **12:00:00~23:59:59**（北京时间）登陆“上海政府采购网”（<http://www.zfcg.sh.gov.cn>）在网上招标系统中报名。

本项目采用电子化采购方式，供应商可在网上免费下载电子招标文件。

四、投标文件上传截止时间：**2025-04-17 09:30:00**（北京时间）。

供应商应在截止时间之前将投标文件的商务部分、技术部分、保证金凭证上传至“上海政府采购网”（<http://www.zfcg.sh.gov.cn>）。逾期上传的投标文件将被视为无效。

五、开标时间和开标地点

投标地点：电子投标文件：<http://www.zfcg.sh.gov.cn>。

开标地点：本项目将在上海政府采购网/采购云平台（www.zfcg.sh.gov.cn）

以线上远程形式开标，不再进行现场开标。投标人应根据《上海市政府采购云平台供应商-项目采购操作手册》的相关操作要求，在开标时间登录上海政府采购网/采购云平台（www.zfcg.sh.gov.cn）参加开标。

纸质投标文件：投标人可在投标截止时间前将纸质文件提前寄至采购代理机构。邮寄地址：上海市静安区恒丰路 600 号机电大厦 5 楼 C 座，蒋颖收，13818207652。快递外包装上请清晰标注项目名称及投标人名称。请投标人保证纸质文件在投标截止时间之前寄到，以快递签收时间为准。纸质投标文件须与上传的电子投标文件内容一致，纸质文件仅作备查使用。

六、联系方式

采 购 人：上海图书馆(上海科学技术情报研究所)

地 址：上海市淮海中路 1555 号

联 系 人：韩春磊、张宏玲

电 话：021-64468868

采购代理机构：上海中招招标有限公司

地 址：上海市静安区恒丰路 600 号机电大厦 5 楼 C、A

邮 编：200070

联 系 人：蒋颖

邮 箱：jiangyinging@163.com

电 话：021-66271152

第二章 投标须知

投标人须知前附表

项号	条款号	编 列 内 容
1	1.1	项目名称：馆藏文献元数据及全文OCR数据制作项目 预算编号：0025-00000245、0025-00000249、0025-00000246、0025-00000248、0025-00000250、0025-00000247 内部编号：STC25A114
2	1.2	项目数据量：中文老旧报纸全文OCR元数据制作154,600版，科技现刊篇名元数据制作900,000篇，老旧期刊篇名元数据制作（中）100,000篇，会议论文篇名元数据制作40,000篇，学术图片库元数据制作40,000篇，老旧期刊篇名元数据制作（英）35,000篇。
3	10.1	投标保证金金额：80000元。 投标保证金形式：银行转账、支票、汇票、本票、保函等非现金形式。 以银行转账形式支付的投标保证金，应在投标截止时间前以转账方式从投标人基本账户转出，并一次性汇入采购代理机构账户（以保证金实际到账为准），保证金有效期应与投标有效期一致。逾期不交者，将被视为自动放弃取得参加本项目投标的法定资格。 开户银行及帐号： 户 名：上海中招招标有限公司 开户银行：中国民生银行上海分行虹桥支行 帐 号：0208014210004789 投标人应在上传电子投标文件后在上海政府采购网进行投标保证金的缴纳登记，并上传缴纳截图凭证，未进行缴纳登记的投标无效。
4	11.1 17.1	开标时间和投标截止时间：2025-04-17 09:30:00（北京时间） 开标地点：本项目将在上海政府采购网/采购云平台（www.zfcg.sh.gov.cn）以线上远程形式开标，不再进行现场开标。投标人应根据《上海市政府采购云平台供应商-项目采购操作手册》的相关操作要求，在开标时间登录上海政府采购网/采购云平台（www.zfcg.sh.gov.cn）参加开标。 纸质投标文件：投标人可在投标截止时间前将纸质文件提前寄至采购代理机构。邮寄地址：上海市静安区恒丰路600号机电大厦5楼C座，蒋颖收，13818207652。快递外包装上请清晰标注项目名称及投标人名称。请投标人保证纸质文件在投标截止时间之前寄到，以快递签收时间为准。纸质投标文件须与上传的电子投标文件内容一致，纸质文件仅作备查使用。

5	11.1	投标有效期：开标后90天
6		行业说明：按照《中小企业划型标准规定》（工信部联企业【2011】300号，本项目采购的（标的）属于：软件和信息技术服务业。
政策功能		（1）残疾人福利性单位：根据财政部、民政部、中国残疾人联合会《三部门联合布关于促进残疾人就业政府采购政策的通知》（财库〔2017〕141号）的规定，在政府采购活动中，残疾人福利性单位视同小型、微型企业，享受预留份额、评审中价格扣除等促进中小企业发展的政府采购政策。向残疾人福利性单位采购的金额，计入面向中小企业采购的统计数据。残疾人福利性单位属于小型、微型企业的，不重复享受政策。
		（2）中小企业：按照国家财政部、工信部《关于印发〈政府采购促进中小企业发展管理办法〉的通知》（财库【2020】46号）中规定的划分标准确定的中小型企业，并提供《中小企业声明函》。
政策功能		a) 参加政府采购活动的中小企业应当提供按照《关于印发〈政府采购促进中小企业发展管理办法〉的通知》（财库【2020】46号）规定的《中小企业声明函》。若提供虚假材料谋取中标、成交的，按照《政府采购法》有关条款处理，并记入供应商诚信档案。 b) 对于非专门面向中小企业的项目，对小型和微型企业产品的价格给予 10%的扣除，用扣除后的价格参与评审。对于专门面向中小企业的项目，无评审价格的优惠扣除。 c) 若小（微）企业与其他规模企业组成联合体，联合协议中约定，小型、微型企业的协议合同金额占到联合体协议合同总金额 30%以上的，可给予联合体 2%的价格扣除。联合体各方均为小型、微型企业的，联合体视同为小型、微型企业享受扶持政策。 d) 履约保证金、付款期限、付款方式等方面给予中小企业适当支持。 e) 提供本企业制造的货物、承担的工程或者服务，或者提供其他中小企业制造的货物。本项所称货物不包括使用大型企业注册商标的货物。小型、微型企业提供中型企业制造的货物的，视同为中型企业。 f) 政府采购监督检查和投诉处理中对中小企业的认定，由企业所在地的县级以上中小企业主管部门负责。 g) 事业单位、团体组织等非企业性质的政府采购供应商，不属于中小企业划型标准确定的中小企业，不得按《关于印发中小企业划型标准规定的通知》规定声明为中小微企业，也不适用《政府采购促进中小企业发展管理办法》。

(3) 依据《财政部司法部关于政府采购支持监狱企业发展有关问题的通知》(财库〔2014〕68号)规定,监狱企业视同小型、微型企业,享受预留份额、评审中价格扣除等政府采购促进中小企业发展的政府采购政策。

中标服务费:

中标服务费:中标人在收到采购代理机构发出的中标通知书后5个日历日之内向采购代理机构支付中标服务费,中标服务费以中标金额作为收费的计算基数,按以下表格按差额累进制收取:

中标金额	中标服务费
服务费不足8000元	8000元
100万以下	1.5%
100万-500万	0.8%

如中标金额为150万元,中标服务费=100万元*1.5%+50万元*0.8%=1.9万元

中标服务费银行及账号:

开户银行及帐号:

户名:上海中招招标有限公司

开户银行:中国民生银行上海分行虹桥支行

帐号:0208014210004789

电子投标特别提醒

注册登记与 安全认证

(1) 为确保电子采购平台数据的合法、有效和安全,各参与主体均应在电子采购平台上注册登记并获得账号和密码。采购人、供应商、采购代理机构还应根据《上海市数字证书使用管理办法》等规定向本市依法设立的电子认证服务机构申请用于身份认证和电子签名的数字证书,并严格按照规定使用电子签名和电子印章。

(2) 使用电子采购平台采购的项目恕不接受电子采购平台以外其他形式的投标。

采购文件澄清、补充 与修改

采购人和采购代理机构可以依法对采购文件进行澄清、补充与修改,澄清、补充与修改的文件应在电子采购平台上公告,并通过电子采购平台发送至已下载采购文件的供应商工作区,或者通过电子邮件发送给已下载采购文件的供应商。

投标文件的编制、加密和上传	(1) 供应商下载采购文件后，应使用电子采购平台提供的客户端投标工具编制投标文件。 (2) 投标供应商应按照采购文件要求提交已完成的投标文件的彩色扫描文件，并在网上投标系统中采用 PDF 格式上传所有资料，文件格式参考采购文件有关格式。 (3) 如因上传、扫描、格式等原因导致评审时受到影响，由投标供应商承担相应责任。采购人认为必要时，可以要求投标供应商提供商务文书和法律文书原件进行核对，投标供应商必须按时提供。否则，视作投标供应商放弃潜在中标资格并对该投标供应商进行调查，有欺诈行为的按有关规定进行处理。 (4) 供应商和电子采购平台应分别对投标文件实施加密。在投标截止前，供应商通过投标工具使用数字证书对投标文件加密后上传至电子采购平台，再经过电子采购平台加密保存。 (5) 由于供应商的原因造成其投标文件未能加密而致投标文件在开标前泄密的，由供应商自行承担责任。
投标保证金	供应商必须在投标截止时间前，在网上投标系统中录入缴纳保证金信息，并把必填项维护完成后，点击“提交”，并提醒采购代理机构及时确认，采购代理机构在网上投标系统进行确认后生效。 需要缴纳保证金的项目若因供应商线上未录入缴纳保证金信息，而导致投标被拒绝由供应商自行承担责任。
网上投标	(1) 登入招投标系统：投标供应商用上海市电子签名认证证书（CA 证书）登陆上海市政府采购中心网上投标系统。 (2) 填写网上投标文件：投标供应商在“网上投标”栏目内选择要参与的投标项目，在投标截止时间前按照网上投标系统和采购文件要求填写网上投标内容。对于有多个包件的招标项目，投标供应商可以选择要参与的包件进行投标。投标供应商用上海市电子签名认证证书对填写内容加密后上传到投标系统。 (3) 正式投标：投标供应商填写好所有投标内容后，须在网上投标截止时间前通过上海市电子签名认证证书在网上投标系统中递交投标文件，并下载投标回执。对于有多个包件的招标项目，需要对每个包件分别进行投标。

投标签收	各供应商在投标（响应）文件加密上传后，须及时联系采购代理机构进行投标签收，投标截止时间之后，采购代理机构将无法投标签收。投标人应及时查看签收情况。未签收的投标（响应）文件视为投标(响应)未完成，投标失败。 对已完成上传投标的项目进行撤销或重新修改，在“投标管理”菜单中点击左侧导航“已完成投标”内，勾选当前项目的所有包且状态为待签收，点击“撤销”按钮，并进行确认即可。
投标截止	(1) 投标截止后电子采购平台不再接受供应商上传投标文件。 (2) 投标截止与开标的时间以电子采购平台显示的时间为准。
开标	(1) 参加开标会议。根据采购文件规定的要求完成网上投标文件提交后，按照采购文件规定的时间和地点，投标供应商的法定代表人或其授权的投标代表携带要求的材料及设备（笔记本电脑、无线上网卡、电子签名认证证书、纸质投标文件等）出席开标会议。 (2) 开标程序在电子采购平台进行，所有上传投标文件的供应商应登录电子采购平台参加开标。 (3) 因投标人自身原因，在电子采购平台开标规定时间内，未在投标网页上完成签到的，视作投标人主动放弃投标。 (4) 开标时参加开标的投标人仅以开标系统显示为准，此时不寻求不考虑其他外部证据，诸如上传遇阻，格式不符，系统故障等原因。 (5) 开标时若发生影响正常开标的系统故障，开标会延后，项目再次开标时间，另行公告或通知。
投标文件解密	(1) 投标截止、电子采购平台显示开标后，由采购代理机构解除电子采购平台对投标文件的加密。投标供应商应在规定时间内使用数字证书对其投标文件解密。 (2) 投标供应商因自身原因未能将其投标文件解密的，视为放弃投标。
开标记录的确认	(1) 投标文件解密后，电子采购平台根据投标文件中标开一览表的内容自动汇总生成开标记录表。 (2) 投标供应商应及时检查开标记录表的数据是否与其投标文件中的投标报价一览表一致，并作出确认。投标供应商因自身原因未作出确认的视为确认开标记录表内容。投标供应商发现开标记录表与其投标文件开标一览表数据不一致的，应及时向采购人（采购代理机构）提出更正，采购人（采购代理机构）应核实开标记录表与投标文件中的开标一览表内容。

其他	(1) 本项目按规定通过上海市政府采购信息管理平台电子招投标系统（以下简称电子招标系统）实施招标，投标供应商须按照该系统的设置要求对本项目进行电子投标。 (2) 本项目招标过程中因以下原因导致的不良后果，本采购代理机构不承担责任： a) 因电子采购平台由市财政局建设、维护和管理，故电子招标系统发生技术故障或遭受网络攻击对项目所产生的影响； b) 本采购代理机构以外的单位或个人在电子招标系统中的不当操作对本项目产生的影响；
电子投标软件平台帮助 电话	95763

投 标 人 须 知

一、总 则

1. 项目概况

- 1.1 项目名称：见“投标人须知前附表”第1项
- 1.2 项目内容：详见第五章招标要求。
- 1.3 服务区域：见“投标人须知前附表”第2项规定。

2. 投标人的资格

- 2.1 单位负责人为同一人或者存在直接控股、管理关系的不同供应商，不得参加同一合同项下的政府采购活动。
- 2.2 为采购项目提供整体设计、规范编制或者项目管理、监理、检测等服务的供应商，不得再参加该采购项目的其他采购活动。
- 2.3 投标人在参加政府采购活动前3年内因违法经营被禁止在一定期限内参加政府采购活动，期限届满的，可以参加政府采购活动。
- 2.4 合格的投标人：见第一章投标邀请书“供应商资格要求”。

3. 投标费用

- 3.1 投标人应承担所有与编写和提交投标文件有关的费用，不论投标的结果如何，采购人在任何情况下均无义务和责任承担这些费用。

二、招标文件

4. 招标文件的构成

- 4.1 招标文件由目录所列内容和根据本须知第5条和第6条所发出的澄清和修改等组成，当所发出的招标文件澄清和修改不一致时，以最后发出的为准。
- 4.2 投标人应认真阅读招标文件中全部内容，如果投标人没有按照招标文件要求提交

全部资料或者投标文件没有对招标文件在各方面都做出实质性响应，则其投标将被拒绝。

5. 招标文件的澄清

5.1 任何要求对招标文件进行澄清的投标人，均应在投标截止时间 15 日以前按投标邀请书的通讯地址以书面形式(书面形式是指文书原件、传真或法律认定的其他形式，下同)通知采购代理机构，采购代理机构对投标截止时间 5 日以前收到的任何澄清要求以书面形式予以答复，同时将书面答复通知每个购买招标文件的投标人，答复中包括所问问题，但不包括问题的来源。

政府采购代理机构：上海中招招标有限公司

地 址：上海市静安区恒丰路 600 号机电大厦 5 楼 C、A

邮 编：200070

电 话：021-66271152

联 系 人：蒋颖

6. 招标文件的修改

6.1 在投标截止时间前任何时候，无论出于何种原因，采购人可主动地或在解答投标人提出的澄清问题时对招标文件进行修改，

6.2 采购人在投标截止时间前对招标文件所作的补充、修改或者更正、澄清一并构成招标文件的组成部分。招标文件的修改将以书面形式或网上公告通知所有购买招标文件的投标人，并对其具有约束力，投标人应立即以书面形式确认已收到修改文件。

6.3 为使投标人编写投标文件时有充分时间对招标文件的修改部分进行研究，采购人将根据相关法规规定延长投标截止时间。

三、投标文件的编制

7. 投标的语言及度量衡单位

7.1 投标人提交的投标文件以及投标人与采购代理机构就有关投标的所有来往函电均应以中文书写。

7.2 投标文件中所使用的度量衡单位，除招标文件中有特殊要求外，一律用公制。

8. 投标文件的构成(招标文件附件中提供格式的，请按照附件的格式填写)

8.1 商务文件，包括报价一览表、投标人资格证明文件、投标人商务偏离表等其他商务部分内容。

8.1.1 报价一览表

投标人只允许有一个报价，采购人不接受有任何选择的报价。

8.1.2 投标人资格证明文件，至少应包括以下内容：

- (1) 投标书；
- (2) 投标保证金；
- (3) 报价明细表；
- (4) 最新年审合格的营业执照（复印件加盖单位公章）；
- (5) 法定代表人身份证明书原件及身份证复印件；
- (6) 授权委托书原件(授权代表投标时须提供)及委托双方身份证复印件；
- (7) 投标保证金凭据复印件；
- (8) 资格声明函；

(9) 投标人没有在“信用中国”网站（www.creditchina.gov.cn）被列入失信被执行人、重大税收违法案件当事人名单或在中国政府采购网（www.ccgp.gov.cn）被列入政府采购严重违法失信行为记录名单的证明；

(10) 财务状况及税收、社会保障资金缴纳情况声明函。

(11) 无重大违法记录承诺书；

(12) 供应商资格要求中提到的但未在上述注明的相关资质文件（复印件加盖公章）。

8.1.3 商务偏离表

8.2 技术文件：

- (1) 技术偏离表；
- (2) 实施方案；
- (3) 完成本项目的人员配置情况；
- (4) 应急方案；
- (5) 售后服务方案及承诺；
- (6) 特色服务方案；
- (7) 招标文件中要求提供及投标人认为需要提供的其他资料。

9. 投标报价

9.1 报价应为一次性人民币含税价(元)，应包括投标人完成全部工作所需的费用。

9.2 投标报价大小写金额不一致的，以大写金额为准；总价金额与按单价汇总金额不一致的，以单价金额计算结果为准；单价金额小数点有明显错位的，应以总价为准，并修改单价；对不同文字文本投标文件的解释发生异议的，以中文文本为准。

9.3 任何有选择的报价将不予接受，本项目只允许有一个报价。

9.4 投标报价表中的所报总价在合同执行过程中是固定不变的，不得以任何理由予以变更。

9.5 投标人不得哄抬报价，也不应低于成本价(或进价)报价，否则一经查实，其投标将可能被拒绝或被宣布为废标。

10. 投标保证金

10.1 投标保证金作为投标文件的组成部分之一，投标人应按招标文件“投标人须知前附表”第 6 项所规定数额、时限提交投标保证金，提交投标保证金的凭据的复印件作为投标文件的组成部分之一装订在投标文件中。

10.2 投标保证金交纳方式为电汇或银行转账，采购人将拒绝其他方式提交的投标保证金。

10.3 未按上述规定足额提交投标保证金的投标，将被视为无效投标。

10.4 未中标人的投标保证金，请在收到《中标结果通知书》后 5 个工作日内，**由投标人及时联系采购代理机构办理无息退还投标保证金手续**。因投标人自身原因导致无法及时退还的，采购人或采购代理机构将不承担资金占用费；如有质疑或投诉，将在质疑和投诉处理完毕后按流程退还。中标人的投标保证金，在中标人签订合同后 5 个工作日内无息退还。

10.5 下列任何情况之一发生时，投标保证金将被没收：

- (1) 投标人在投标书中规定的投标有效期内撤回其投标；
- (2) 拒绝按本须知第 9.3 条修正报价的；
- (3) 中标人未在规定时间内交纳招标文件所规定的费用的；
- (4) 中标人未能在规定期限内签订合同的。

11. 投标有效期

11.1 投标人的投标文件应从“投标人须知前附表”第4项规定的开标之日起，在“投标人须知前附表”第5项所规定的以日历天计算的投标有效期内有效。投标有效期比规定的时间短的应视为非响应标予以拒绝。

11.2 在特殊情况下，在原投标有效期期满之前，采购人可征得投标人同意延长投标有效期，这种要求与答复均应以书面形式如电传、传真等。投标人可以拒绝采购人的这种要求但不被没收投标保证金，同意延长的投标人既不能被要求也不允许修改其投标文件，但要相应延长其投标保证金的有效期。

12. 投标文件的式样和签署

12.1 投标人应准备一份投标文件正本，五份副本。投标文件分为商务部分和技术部分，在签字盖章扫描后上传至“上海政府采购网”(<http://www.zfcg.sh.gov.cn>)。

12.2 投标文件的正本应打印或用不褪色墨水书写，并由投标人法定代表人或其授权人签字或加盖公章。授权代表须将以书面形式出具的“法定代表人授权委托书”附在投标文件中。

12.3 除投标人对错处做必要修改外，投标文件不得行间插字、涂改和增删，如有修改错漏处，必须由同一签署人在修改错漏处签字或盖公章。

四、投标文件的递交

13. 投标文件的递交

13.1 投标人须在上海政府采购网下载、安装“上海市政府采购信息管理平台投标工具”，使用该工具完成投标文件录入、投标、投标文件加密等内容的操作。投标人应在截止时间之前将投标文件的商务部分、技术部分、保证金凭证上传至“上海政府采购网”(<http://www.zfcg.sh.gov.cn>)。

13.2 投标人应在截止时间之前递交投标文件6份（纸质文件1正5副）并密封（纸质文件与网上上传的电子文件内容一致）。

13.3 逾期上传的投标文件将被视为无效。

13.4 投标文件的密封和标记：

(1) 投标文件需双面打印、胶装成册

(2) 投标人应将响应文件正本、副本、电子版分别密封包装，并在包装上正确标明“正本”、“副本”“电子版”。

(3) 在外层包装上具有下列识别标志：

A、项目名称/项目编号/包号：

B、响应人的名称与地址，并加盖公章

C、 年 月 日 时 分前不能启封

14. 投标截止日期

14.1 投标人上传投标文件的时间不得迟于投标邀请书中规定的投标截止时间。

14.2 推迟投标截止日将在上海政府采购网发布延期公告。

14.3 投标截止时间后送交的投标报价或其他实质性内容修正的函件不作为评标的依据。

15. 迟到的投标文件

15.1 按照第 14 条规定，采购人将拒绝在投标的截止时间后收到的任何投标文件。

15.2 投标文件的递交地点和时间“投标人须知前附表”第 4 项中所规定的地点和时间。

16. 投标文件修改和撤回

16.1 投标人在递交投标文件后，可以修改或撤回其投标，但投标人必须在规定的投标截止期之前将修改或撤回的书面通知递交到采购代理机构。

16.2 在投标截止时间之后，投标人不得对其投标文件做任何修改。

16.3 根据本须知第 10.5 条规定，在投标截止时间至投标有效期内，投标人不得撤回其投标，否则其投标保证金将被没收。

五、开标与评标

17. 开标

17.1 采购代理机构在“投标人须知前附表”中规定的日期、时间和地点组织公开开标。

17.2 有以下情形之一者，将作为无效投标文件处理：

- (1) 投标人没有本项目要求的、相应的经营范围；
- (2) 投标文件未按招标文件要求进行签署、盖章的；
- (3) 在招标文件中，投标人投报两个或多个报价的；
- (4) 报价超过最高投标限价；
- (5) 投标人的报价明显低于通过符合性检查的其他投标人的报价，有可能影响履约的，且未能提供材料证明其报价的合理性；
- (6) 满足招标文件中加*项要求；
- (7) 其他不符合招标文件实质性要求及法律法规规定的其他情形。

17.3 本项目不允许采购进口产品。若无财政针对本项目允许采购进口产品批复，供应商提供的产品若为进口产品，根据财政部关于印发《政府采购进口产品管理办法》的通知（财库〔2007〕119号）精神，投标文件将作为无效处理。所谓进口产品是指通过中国海关验收进入中国境内且产自关境外的产品。

18. 评标

18.1 公开开标后，直至向中标的投标人授予合同时止，凡与审查、澄清、评价和比较投标的有关资料以及授标意见等，均不得向投标人及评标无关的其他人透露。

18.2 开标评标期间，投标人不得与评标委员会成员和采购人及工作人员私下接触，更不得进行旨在影响评标结果的活动，否则一经核实，其投标将被宣布为废标，并将承担相应的法律责任。

18.3 评标委员会在评审过程中发现投标人有行贿、提供虚假材料或者串通等违法行为的，应当及时向财政部门报告。

18.4 评审专家在评审过程中受到非法干预的，应当及时向财政、监察等部门举报。

18.5 采购人将根据本次招标项目的性质和特点依法组建评标委员会；

18.6 评标原则：依照“公平、公正、科学、择优、合法、保密”原则；

18.7 评标办法：“综合评分法”。

18.7.1 “综合评分法”：是指在最大限度地满足招标文件实质性要求前提下，按照招标

文件中规定的各项因素进行综合评审后，以评标总得分最高的投标人作为中标候选人或中标人的评标方法。

若中标人因不可抗力或者自身原因不能履行服务合同的，采购人可以与排位在中标人之后第一位的中标候选人签订服务合同，以此类推。

18.8 如一个分包内包含多种产品的，采购人或采购代理机构将在投标人须知资料表中载明核心产品，多家投标人提供的所有核心产品品牌均相同的，按以下规定处理：

18.8.1 如本项目使用最低评标价法，提供相同品牌产品的不同投标人以其中通过资格审查、符合性审查且报价最低的参加评标；报价相同的，由采购人或者采购人委托评标委员会按照招标文件中评标办法规定的方式确定一个参加评标的投标人；未规定的采取随机抽取方式确定，其他投标无效；

18.8.2 如本项目使用综合评分法，提供相同品牌产品且通过资格审查、符合性审查的不同投标人，按一家投标人计算，评审后得分最高的同品牌投标人获得中标人推荐资格；评审得分相同的，由采购人或者采购人委托评标委员会按照招标文件中评标办法规定的方式确定一个投标人获得中标人推荐资格；未规定的采取随机抽取方式确定，其他同品牌投标人不作为中标候选人。

19. 中标条件

19.1 投标文件能满足招标文件实质性要求且评标总得分最高。

20. 投标文件的澄清

20.1 为有助于对投标文件的审查、评价和比较，评标委员会可分别要求对投标人进行技术质疑、询问、澄清，或对投标人进行实地考察，或要求投标人提供所需的有关资料。有关要求和答复应以书面形式提交，但不得寻求、提供或允许对投标价格或实质性内容做任何更改。

六、中标通知和合同授予

21. 定标

21.1 采购代理机构自评审结束之日起 2 个工作日内将评审报告送交采购人。采购人在收到评审报告之日起 5 个工作日内在评审报告推荐的中标或者成交候选人中按顺序确定中标人。

21.2 接受和拒绝任何或所有投标的权力

采购人保留在授标之前任何时候接受或拒绝任何投标，以及宣布招标程序无效或拒绝所有投标的权力，对受影响的投标人不承担任何责任。

22. 中标通知

22.1 中标人确定后，采购代理机构将在上海政府采购网发布中标公告，并向中标人发出中标通知书。向未中标的其他投标人发出中标结果通知书。

七、合同签订

23. 合同签订方式

23.1 中标人按《中标通知书》指定的时间、地点与采购人签订合同。中标人因不可抗力或者自身原因不能履行服务合同的，采购人可与推荐第二中标候选人签订合同，以此类推，此种情况下，中标人的权利和义务与本须知的所有规定相一致。

23.2 招标文件、中标人的投标文件及其澄清文件、中标通知书，均作为签订合同内容的依据。

23.3 付款方式：详见合同条款。

24. 合同书不可分割之部分

24.1 招标文件；

24.2 中标人的投标文件及澄清文件；

24.3 中标通知书；

24.4 合同书附件。

25. 其他

25.1 其他内容将依照相关法律法规的规定，在合同签订时具体约定。

第三章 上海政府采购网合同范本 包1 合同模板:

[合同中心-合同名称]合同

合同统一编号: [合同中心-合同编码]

合同内部编号:

合同各方:

甲方: [合同中心-采购单位名称]

乙方: [合同中心-供应商名称] (法定代表

人签名: [合同中心-供应商法人姓名], 性

别: [合同中心-供应商法人性别])

地址: [合同中心-采购单位所在地]

地址: [合同中心-供应商所在地]

电话: [合同中心-采购单位联系人电话]

电话: [合同中心-供应商联系人电话]

联系人: [合同中心-采购单位联系人]

联系人: [合同中心-供应商联系人]

根据《中华人民共和国政府采购法》、《中华人民共和国民法典》之规定, 本合同当事人在平等、自愿的基础上, 经协商一致, 同意按下述条款和条件签署本合同:

1. 乙方根据本合同的规定向甲方提供以下服务:

1. 1 乙方所提供的服务其来源应符合国家的有关规定, 服务的内容、要求、服务质量等详见合同附件。

2. 合同价格、服务地点和服务期限

2. 1 合同价格

本合同价格为[合同中心-合同总价]元整 ([合同中心-合同总价大写])。

乙方为履行本合同而发生的所有费用均应包含在合同价中, 甲方不再另行支付其它任何费用。

2. 2 服务地点：甲方指定地点

2. 3 服务期限：~~[合同中心-合同有效期]~~。

3. 质量标准和要求

3. 1 乙方所提供的服务的质量标准按照国家标准、行业标准或制造厂家企业标准确定，上述标准不一致的，以严格的标准为准。没有国家标准、行业标准和企业标准的，按照通常标准或者符合合同目的的特定标准确定。

3. 2 乙方所交付的服务还应符合国家和上海市有关安全、环保、卫生之规定。

4. 权利瑕疵担保

4. 1 乙方保证对其交付的服务享有合法的权利。

4. 2 乙方保证在服务上不存在任何未曾向甲方透露的担保物权，如抵押权、质押权、留置权等。

4. 3 乙方保证其所交付的服务没有侵犯任何第三人的知识产权和商业秘密等权利。

4. 4 如甲方使用该服务构成上述侵权的，则由乙方承担全部责任。

5. 验收

5. 1 服务根据合同的规定完成后，甲方应及时进行根据合同的规定进行服务验收。乙方应当以书面形式向甲方递交验收通知书，甲方在收到验收通知书后的 10 个工作日内，确定具体日期，由双方按照本合同的规定完成服务验收。甲方有权委托第三方检测机构进行验收，对此乙方应当配合。

5. 2 如果属于乙方原因致使系统未能通过验收，乙方应当排除故障，并自行承担相关费用，同时进行试运行，直至服务完全符合验收标准。

5.3 如果属于甲方原因致使系统未能通过验收，甲方应在合理时间内排除故障，再次进行验收。如果属于故障之外的原因，除本合同规定的不可抗力外，甲方不愿或未能在规定的时间内完成验收，则由乙方单方面进行验收，并将验收报告提交甲方，即视为验收通过。

5.4 甲方根据合同的规定对服务验收合格后，甲方收取发票并签署验收意见。

6. 保密

6.1 如果甲方或乙方提供的内容属于保密的，应签订保密协议，甲乙双方均有保密义务。

7. 付款

7.1 本合同以人民币付款（单位：元）。

7.2 本合同款项按照以下方式支付。

7.2.1 付款内容：（分期付款）

7.2.2 付款条件：乙方须适应甲方内部的法律文书审核和财务报批支付等流程规定，并统一提供乙方发票。合同费用将按以下方式分两次支付：第一笔付款—阶段付款(50%)：乙方完成第一阶段工作并验收合格后十个工作日内，支付第一笔合同款；第二笔付款—最终验收付款(50%)：乙方完成所有工作并验收合格后，支付剩余货款。

8. 甲方（甲方）的权利义务

8.1、甲方有权在合同规定的范围内享受，对没有达到合同规定的服务质量或标准的服务事项，甲方有权要求乙方在规定的时间内加急提供服务，直至符合要求为止。

8.2 如果乙方无法完成合同规定的服务内容、或者服务无法达到合同规定的服务质量或标准的，造成的无法正常运行，甲方有权邀请第三方提供服务，其支付的服务费用由乙方承担；如果乙方不支付，甲方有权在支付乙方合同款项时扣除其相等的金额。

8.3 由于乙方服务质量或延误服务的原因，使甲方有关或设备损坏造成经济损失的，甲方有权要求乙方进行经济赔偿。

8.4 甲方在合同规定的服务期限内义务为乙方创造服务工作便利，并提供适合的工作环

境，协助乙方完成服务工作。

8. 5 当或设备发生故障时，甲方应及时告知乙方有关发生故障的相关信息，以便乙方及时分析故障原因，及时采取有效措施排除故障，恢复正常运行。

8. 6 如果甲方因工作需要调整，应有义务并通过有效的方式及时通知乙方涉及合同服务范围调整的，应与乙方协商解决。

9. 乙方的权利与义务

9. 1 乙方根据合同的服务内容和要求及时提供相应的服务，如果甲方在合同服务范围外增加或扩大服务内容的，乙方有权要求甲方支付其相应的费用。

9. 2 乙方为了更好地进行服务，满足甲方对服务质量的要求，有权利要求甲方提供合适的工作环境和便利。在进行故障处理紧急服务时，可以要求甲方进行合作配合。

9. 3 如果由于甲方的责任而造成服务延误或不能达到服务质量的，乙方不承担违约责任。

9. 4 由于因甲方工作人员人为操作失误、或供电等环境不符合合同设备正常工作要求、或其他不可抗力因素造成的设备损毁，乙方不承担赔偿责任。

9. 5 乙方保证在服务中，未经甲方许可不得使用含有可以自动终止或妨碍系统运作的软件和硬件，否则，乙方应承担赔偿责任。

9. 6 乙方在履行服务时，发现存在潜在缺陷或故障时，有义务及时与甲方联系，共同落实防范措施，保证正常运行。

9. 7 如果乙方确实需要第三方合作才能完成合同规定的服务内容和质量的，应事先征得甲方的同意，并由乙方承担第三方提供服务的费用。

9. 8 乙方保证在服务中提供更换的部件是全新的、未使用过的。如果或证实服务是有缺陷的，包括潜在的缺陷或使用不符合要求的材料等，甲方可以根据本合同第 10 条规定以书面形式向乙方提出补救措施或索赔。

10. 补救措施和索赔

10.1 甲方有权根据质量检测部门出具的检验证书向乙方提出索赔。

10.2 在服务期限内，如果乙方对提供服务的缺陷负有责任而甲方提出索赔，乙方应按照甲方同意的下列一种或多种方式解决索赔事宜：

（1）根据服务的质量状况以及甲方所遭受的损失，经过买卖双方商定降低服务的价格。

（2）乙方应在接到甲方通知后七天内，根据合同的规定负责采用符合规定的规格、质量和性能要求的新零件、部件和设备来更换在服务中有缺陷的部分或修补缺陷部分，其费用由乙方负担。

（3）如果在甲方发出索赔通知后十天内乙方未作答复，上述索赔应视为已被乙方接受。如果乙方未能在甲方发出索赔通知后十天内或甲方同意延长的期限内，按照上述规定的任何一种方法采取补救措施，甲方有权从应付的合同款项中扣除索赔金额，如不足以弥补甲方损失的，甲方有权进一步要求乙方赔偿。

11. 履约延误

11.1 乙方应按照合同规定的时间、地点提供服务。

11.2 如乙方无正当理由而拖延服务，甲方有权没收乙方提供的履约保证金，或解除合同并追究乙方的违约责任。

11.3 在履行合同过程中，如果乙方可能遇到妨碍按时提供服务的情况时，应及时以书面形式将拖延的事实、可能拖延的期限和理由通知甲方。甲方在收到乙方通知后，应尽快对情况进行评价，并确定是否同意延期提供服务。

12. 误期赔偿

12.1 除合同第13条规定外，如果乙方没有按照合同规定的时间提供服务，甲方可以应付的合同款项中扣除误期赔偿费而不影响合同项下的其他补救方法，赔偿费按每（天）赔偿延期服务的服务费用的百分之零点五（0.5%）计收，直至提供服务为止。但误期赔偿费的最高限额不超过合同价的百分之五（5%）。（一周按七天计算，不足七天按一周计算。）一旦达到误期赔偿的最高限额，甲方可考虑终止合同。

13. 不可抗力

13.1 如果合同各方因不可抗力而导致合同实施延误或不能履行合同义务的话，不应该承担误期赔偿或不能履行合同义务的责任。

13.2 本条所述的“不可抗力”系指那些双方不可预见、不可避免、不可克服的事件，但不包括双方的违约或疏忽。这些事件包括但不限于：战争、严重火灾、洪水、台风、地震、国家政策的重大变化，以及双方商定的其他事件。

13.3 在不可抗力事件发生后，当事方应尽快以书面形式将不可抗力的情况和原因通知对方。合同各方应尽可能继续履行合同义务，并积极寻求采取合理的措施履行不受不可抗力影响的其他事项。合同各方应通过友好协商在合理的时间内达成进一步履行合同的协议。

14. 履约保证金：无

15. 质量保证金按以下比例收取：

序号	采购内容	预算编号	质量保证金
1	中文老旧报纸全文 OCR 元数据制作	0025-00000245	合同金额的 3%
2	科技现刊篇名元数据制作	0025-00000249	合同金额的 3%
3	老旧期刊篇名元数据制作（中）	0025-00000246	无
4	会议论文篇名元数据制作	0025-00000248	无
5	学术图片库元数据制作	0025-00000250	无
6	老旧期刊篇名元数据制作（英）	0025-00000247	无

16. 争端的解决

16.1 合同各方应通过友好协商，解决在执行本合同过程中所发生的或与本合同有关的一切争端。如从协商开始十天内仍不能解决，可以向同级政府采购监管部门提请调解。

16.2 调解不成则提交上海仲裁委员会根据其仲裁规则和程序进行仲裁。

16.3 如仲裁事项不影响合同其它部分的履行，则在仲裁期间，除正在进行仲裁的部分外，本合同的其它部分应继续执行。

17. 违约终止合同

17.1 在甲方对乙方违约而采取的任何补救措施不受影响的情况下，甲方可在下列情况下向乙方发出书面通知书，提出终止部分或全部合同。

（1）如果乙方未能在合同规定的期限或甲方同意延长的期限内提供部分或全部服务。

（2）如果乙方未能履行合同规定的其它义务。

17.2 如果乙方在履行合同过程中有不正当竞争行为，甲方有权解除合同，并按《中华人民共和国反不正当竞争法》之规定由有关部门追究其法律责任。

18. 破产终止合同

18.1 如果乙方丧失履约能力或被宣告破产，甲方可在任何时候以书面形式通知乙方终止合同而不给乙方补偿。该终止合同将不损害或影响甲方已经采取或将要采取任何行动或补救措施的权利。

19. 合同转让和分包

19.1 除甲方事先书面同意外，乙方不得转让和分包其应履行的合同义务。

20. 合同生效

20.1 本合同在合同各方签字盖章并且甲方收到乙方提供的履约保证金后生效。

20.2 本合同一式叁份，甲乙双方各执一份。一份送同级政府采购监管部门备案。

21. 合同附件

21.1 本合同附件包括： 招标(采购)文件、投标（响应）文件

21.2 本合同附件与合同具有同等效力。

21.3 合同文件应能相互解释，互为说明。若合同文件之间有矛盾，则以最新的文件为准。

22. 合同修改

22. 1 除了双方签署书面修改协议，并成为本合同不可分割的一部分之外，本合同条件不得有任何变化或修改。

23. 备注

[合同中心-补充条款列表]

签约各方：

甲方（盖章）：[合同中心-采购单位名称_1] 乙方（盖章）：[合同中心-供应商名称_1]

日期：[合同中心-签订时间] 日期：[合同中心-签订时间_1]

合同签订点：网上签约

第四章 评标办法

一、评标机构

评标机构为评标委员会，评标委员会由五人或五人以上单数组成。

二、评标原则

评标将遵循下列原则：

1. 坚持公平、公正、科学、择优的原则，本着实事求是的精神，不带有任何主观意愿和偏见，认真负责地做好评标工作，公平、公正地对待每一个投标人。
2. 全面分析，综合评审。
3. 评标只对投标单位的投标文件进行评审，投标文件以外的资料、信息不应作为评标的依据。

三、评标纪律

1. 对评标内容和评标过程要严格保密，不得向投标人或与该过程无关的其它人员泄露；
2. 评标期间的一切资料，包括评标意见、评标记录和评标结论，一律不得向外传和泄露；
3. 任何属于投标文件审查、澄清、评价和比较的资料，不得向投标人或与该过程无关的其它人员泄露；
4. 所有资料(包括招标文件、投标文件、评标表格及各种文字记录)在评标结束后均应分别整理、存档备查，任何人不得复制和保留；
5. 评标结束后，与会人员不得向外界透露评标人员的评标意见，如因此造成的后果由责任者承担；
6. 评标期间，评标人员不得外出，确需外出时应事先请假；
7. 评标期间，所有与会人员均不得私自以任何方式和投标人进行联系，需询问、澄清的问题由评委会统一组织办理；
8. 评标期间，未经允许评标委员会以外的任何单位或部门不得参加评标和采访评标工作。

四、评标目的

评标工作的目的：审查投标人的投标文件与招标文件的所有条款、条件是否相符；是否已提供了必备的资料、文件、证书，做出实质性响应；按评标程序和评标方法推荐出合格的中标候选人。

五、评标程序

评标只对有效投标文件进行评审。评标按：符合性评审、综合评分、评标报告的程序进行。

六、评标办法及打分细则

本次招标所采用的评标方法为“综合评分法”，即：在最大限度地满足招标文件实质性要求前提下，按照招标文件中规定的各项因素进行综合评审后，以评标总得分最高的投标人作为中标候选人或中标人的评标办法。

（一）资格符合性评审

采购人或采购代理机构依据法律法规和招标文件中规定的内容，对投标人的资格进行审查。未通过资格审查的投标人不进入评标；通过资格审查的投标人少于三家的，不进行评标。

评标委员会按照符合性评审内容对投标文件进行符合性检查，只有对符合性评审所列各项作出实质性响应的投标文件才能通过初步评审。如果投标文件实质上没有响应招标文件的要求将予以拒绝，投标人不得通过修正或撤销不合要求的偏离或保留从而使其投标成为实质上响应的投标。对是否实质性响应招标文件的要求有争议的投标，评标委员会将以记名方式表决，被认为响应的得票超过半数的投标人有资格进入下一阶段的评审，否则将被淘汰。

在需要时可要求投标人对投标文件中含义不明确的内容作必要的澄清或说明，澄清或说明应采用书面形式。澄清或说明不得超出投标文件的范围或者改变投标文件的实质性内容。

符合性评审必须满足以下要求，未通过符合性评审的投标人为无效投标，将不进入综合评分：

符合性审查表

检查事项	投标人名称	投标人名称	投标人名称
投标有效期是否满足要求			
是否按招标文件规定的形式和金额交纳投标保证金的；			
是否按照招标文件规定要求签署、盖章的；			
是否没有与其他投标人串通投标，或者与采购人串通投标妨碍其他投标人竞争行为，损害采购人或者其他投标人合法权益。			
是否没有招标文件规定的其他作废标处理或无效投标情形；			
是否没有评标委员会认为投标人的报价明显低于其他通过符合性检查投标人的报价，有可			

能影响履约的，且投标人未按照规定证明其报价合理性的情况；			
是否不存在投标文件含有采购人不能接受的附加条件的；			
是否符合法规和招标文件中规定的其他实质性要求。			
结论			

资格性审查表

磋商文件要求	检查内容	审查结论
(1) 具有独立承担民事责任的能力；	供应商的营业执照复印件。	☐ 通过 ☐ 不通过
(2) 具有良好的商业信誉和健全的财务会计制度；	财务状况及税收、社会保障资金缴纳情况声明函。	☐ 通过 ☐ 不通过
(3) 具有履行合同所必需的设备和专业技术能力；	主要技术人员、售后服务机构人员能力等。	☐ 通过 ☐ 不通过
(4) 有依法缴纳税收和社会保障资金的良好记录；	财务状况及税收、社会保障资金缴纳情况声明函。	☐ 通过 ☐ 不通过
(5) 参加政府采购活动前三年内，在经营活动中没有重大违法记录；	提供无重大违法记录承诺书。	☐ 通过 ☐ 不通过
(6) 供应商应在投标文件中提供无不良信用记录的相关证明（相应网页截图）。不良信用记录指：供应商在“信用中国”网站（www.creditchina.gov.cn）被列入失信被执行人、重大税收违法案件当事人名单或在中国政府采购网（www.ccgp.gov.cn）被列入政府采购严重违法失信行为记录名单；	提供相关证明。	☐ 通过 ☐ 不通过
(7) 供应商为中小企业；	提供企业声明函。	☐ 通过 ☐ 不通过
(8) 本项目不接受联合体投标。		☐ 通过 ☐ 不通过

(二) 详细评审

各评委按照评标办法对每个投标人进行独立评分，在计算平均分，按照每个投标人最终平均得分的高低依次排名，推荐出中标候选人。

评分标准

一、详细评分表

序号	评审因素	满分	评分要点及标准说明
1	报价得分	15 分	经评委确认满足招标文件要求的有效最低价为满分 15 分。 满足招标文件要求且报价最低的为评审基准价， 价格得分=（评标基准价/投标价格）×15。
2	综合实力和 信誉（客观 分）	8 分	1、具有高新技术企业或软件 who 认定证书的得 3 分；2、具有 OCR 文字识别或数据处理相关软件著作权证书的，有一项得 1 分，最高得 5 分。以上内容均需提供有效证明。
3	管理制度	6 分	对供应商内部规章制度、机构设置等方面进行综合评审。 明确管理职责，管理机构设置合理，工作流程完整、科学、可行，上述内容提供齐全且完整可行、针对性强得 6 分，内容较简单得 3 分，针对性一般得 1 分。
4	实施方案	20 分	1) 对本项目的认知和理解： 重点难点分析最为全面，针对性强得 4 分，内容较简单得 2 分，针对性一般得 1 分，不提供得 0 分。 2) 针对本项目的服务目标、范围和任务： 最贴合招标需求，针对性强得 4 分，内容较简单得 2 分，针对性一般 1 分，不提供得 0 分。 3) 服务方案： 工作计划详细、完整，针对性强得 4 分，内容较简单得 2 分，针对性一般得 1 分，不提供得 0 分。 4) 服务质量和安全保证措施： 提供承诺保证有足够的人力、物力等资源保证按时按质按量完成服务内容，并提供安全保证具体措施，最详细完整得 4 分，保证措施较好得 2 分，针对性一般得 1 分，不提供得 0 分。 5) 服务人员培训计划及方案：最详细完整得 4 分，内容较简单得 2 分，针对性一般 1 分，不提供得 0 分。
5	应急方案	6 分	根据供应商针对本项目应急方案（含应急预案的详细程度、应急状况分类处置的情况、应急保障内容及承诺等内容）的可行性、合理性，针对性进行综合评审。上述内容提供齐全且完整可行、针对性强得 6 分，内容较简单得 3 分，针对性一般得 1 分。

6	特色服务方案	6 分	提供合理化建议或特色服务，方案完整可行且针对性强的，得 6 分，内容较简单得 3 分，针对性一般得 1 分，不提供不得分。
7	项目负责人	6 分	项目负责人从业经历十年（含）以上的得 3 分，具有与本项目类似业绩且担任项目负责人职务的，得 3 分。
8	项目组人员	8 分	1、项目组人员资质，项目组人员具有工程师职称或中级职称、专业人才证书的得 2 分，行业工作年限 5 年以上的得 2 分，需提供以上人员的相关证书，未提供或不满足以上条件的不得分。2、项目组人员配置，项目组人员是否能匹配本项目招标要求，优 4 分，良 2 分。
9	业绩（客观分）	12 分	类似项目业绩指近 3 年以来（2022 年 1 月至今）承接有效的数据处理项目业绩情况；有 1 个得 2 分，本项最高得分为 12 分，没有有效的类似项目业绩的得 0 分。供应商需提供类似项目的合同扫描件，合同扫描件中需体现合同的签约主体、项目名称及内容、合同金额、交付日期等合同要素的相关内容，否则不算有效的类似项目业绩。
10	售后服务	10 分	根据供应商提供的售后服务（售后服务承诺、增值服务、保密承诺、人员固定性承诺等内容）进行综合评审。上述服务内容完整，承诺最优得 10 分，服务内容较简单得 7 分，针对性一般得 4 分，不提供不得分。
11	文件编制	3 分	文件内容完整、简洁明了、上传清晰、编排有序的，得 3 分；内容缺漏、重复繁琐、文字或图片不清晰或者编排混乱的，酌情扣分。未按相关法律法规要求准确填写中小企业声明函（包括行业定义及企业类型）及残疾人福利性单位声明函的，扣 1 分。

二、推荐

评标委员会应根据各投标人所提交的投标方案的评标得分进行排序，推荐名列前三位的投标人为本项目的中标候选人，并按评标得分的高低标明排名顺序，最终推荐排名第一的为中标人。若中标人最终无法完成该项目，或出现法规规定的中标无效情形时，采购人可与排名下一位的中标候选人另行签订合同，或依法重新招标。

注：当两家或两家以上投标人所提交的投标方案的综合得分相等时，则按经核准的投标报价由低到高排序；当两家或两家以上投标人所提交的投标方案的综合得分及其经核准的投标报价均相等时，则由评标委员会根据主要技术指标的优劣作出排序。

三、评标报告

评标委员会根据评审情况和建议推荐中标人名单提出书面评标报告。

第五章 招标要求

项目 序号	项目名称	采购编号	预算
1	中文老旧报纸全文 OCR 元数据 制作	0025-00000245	298.378 万元
2	科技现刊篇名元数据制作	0025-00000249	104.4 万元
3	老旧期刊篇名元数据制作（中）	0025-00000246	15.6 万元
4	会议论文篇名元数据制作	0025-00000248	4.24 万元
5	学术图片库元数据制作	0025-00000250	9.24 万元
6	老旧期刊篇名元数据制作（英）	0025-00000247	11.55 万元

注：超过以上分项报价的投标将被拒绝。

项目 1：中文老旧报纸全文 OCR 元数据制作

项目数据量：154,600 版

项目完成时间：2025 年 10 月 31 日

项目进度要求：

第一阶段：合同签署后，在双方确认技术标准后，到 2025 年 6 月 30 日前至少完成 77,300 版

第二阶段：至 2025 年 10 月 31 日前，完成 154,600 版

一、 招标样本

- 由招标方提供招标所需样本图像，共计约 30 版

- 投标方依据招标方提供的著录标准及各项规范（见下附件）制作招标样本
- 各投标方在将样品提交前，请先在样本上标识所投项目、投标方等信息，并确认标识是否完好
- 投标方需在投标日当天出示招标样本，并做阐述
- 投标方提交的样本将作为评估中标与否的因素之一
- 最终解释权归招标方所有

二、元数据主要著录内容要求

（一） 中文老旧报纸元数据的著录分三种类型

1. 正文内容元数据：8 项≤著录项≤36 项
2. 广告元数据：7 项≤著录项≤17 项
3. 图片元数据：7 项≤著录项≤19 项

（二） 元数据元素主要著录说明

1. 关于字段非空

- 1) 附件各表中标*的元素均为非空字段，著录项不可缺省；未标*的元素视具体内容具体分析，有内容则必须著录。
- 2) 广告标题、广告发布者和广告产品不能同时为空。

2. 关于图片的著录与划框

图片著录时，原则上以单独图片进行著录，对于形状不规则的图片，在保证主题表达完整的前提下可进行组合图片的划框界定。

3. 关于著录项所用语种

除正文的“类别”、广告的“广告类别”、图片的“图片类型”以及“备注”跟随报纸语种外，其余著录项所用语种均依据制作报纸内容照实进行著录。报纸语种由招标方确定。

4. 关于类别著录

正文的类别、广告的类别以及图片的类型均根据事先约定的分类体系进行类别赋予。

5. 关于正文的翻译标题字段

中文报纸中的全英文文章，正文翻译标题应翻译为相应的中文并著录在该字段。

若原文标题与另一种语种的标题混合，则不要求著录翻译字段，完整照实著录在标题字段即可。

6. 关于不要求著录的字段

- 1) 正文的关键词字段不要求著录。
- 2) 图片类型 2 不要求著录。
- 3) 对于所有不要求著录的字段都应保留。

7. 关于标点、符号著录

著录中使用的所有标点、符号需双方协商确认。

8. 关于排序

报纸中每一条正文、广告和图片元数据都应有对应的排序信息，排序应符合报纸实际文脉和阅读次序，以实现数据检索应用过程中的各种排序功能。

9. 关于元数据、图像及 OCR 文本的唯一标识及对应关系

正文唯一标识、广告唯一标识、图片唯一标识等由著录系统自动生成，且唯一标识一旦生成则不可改变，命名规则由招标方确定。元数据、图像及 OCR 文本之间的对应关系一定要满足数据的检索应用要求。

（三）OCR 文本识别数据加工要求

总体原则：以原件内容的客观含义为准，尽量忠实于原文进行识别。要求识别文本内容与原图篇内容应一一对应。

定期优化更新技术软件，应用最新版本技术软件来进行文字识别。

1. 文字识别

对待加工中文老报纸的篇目正文内容进行 OCR 文字识别。遵照原文的篇目范围、正文框顺序和阅读方式按要求进行识别。

2. 文字识别正确率

- 1) OCR 文字识别每批次正确率要求不低于 50%，整体项目各批次汇总后平均文字识别率不低于 60%。
- 2) 影响文字识别正确率的因素涉及待识别篇目的版式、清晰度和文字大小等。

3. PDF 文件

- 1) 双层 PDF 的上层图像为 JPG 格式，下层为篇目全文文本层，且篇目位置已置标。
- 2) PDF 格式，要求 PDF 编码为 1.5 版本(兼容 Adobe reader6.0)，要求在保持图像清晰可读的基础上尽可能减小存储量。

4. XML 文件

- 1) XML 标签合法性检查要求零错误。
- 2) XML 文件要求遵守 XML 语法规则，使用 utf-8 编码方式。

三、验收质量标准

1. 著录规则质量标准：著录规则错误率 $\leq 0.3\%$
2. 著录文字质量标准：著录文字错误率 $\leq 0.05\%$
3. 正文文字质量标准：零漏著
4. 元数据与图像文件的勾链质量标准：元数据与图像文件的勾链无差错
5. 版面内容元数据质量标准：版面内容元数据无漏著
6. 篇名类型著录质量标准：篇名类型无差错
7. 其他元数据的合法性验收标准
 - 1) 返回图像和 PDF 的版数必须和著录的版数完全相同。
 - 2) 返回图像像素、大小无异常，灰度或色彩无明显漏洞，不存在多幅图像相同的问题。
 - 3) 返回图像元数据著录项中非空元素著录项无缺省。

四、数据规范格式

投标方需根据招标方提供的数据格式进行此项目制作。

投标方须提供的数据有：单版 PDF 文件，图像文件（单版图像、著录图像、著录广告图像）的 TIF 及 JPG 格式各一套，元数据著录信息文件，版面分析信息文件等。具体规范详见附件。

五、质保服务

投标方在总项目完成后提供一年的免费质保服务。双方在合同签订后另行签署质保协议书。

附件一：元数据著录规则

一、前言

中文老旧报纸的内容基本分为两部分：正文内容和广告内容。由于报纸的部分正文内容中会附有图片，为了更方便揭示报纸数字化文本的内容，本规则对报纸的文中图片也要求进行著录，著录图片主要为内容独立或附有文字说明的图片。

中文老旧报纸元数据的著录分三种类型：

- 正文内容元数据
- 广告内容元数据
- 图片元数据

二、术语定义

下列术语和定义适用于本附件。

1. 元数据著录单元

指报纸图像中有相对明显物理边界内的内容完整的正文、广告或图片。图像划框据此边界进行画框。下文中简称“著录单元”。

2. 正文

报纸中除广告外的所有内容。

正文著录单元：报纸正文中有相对明显物理边界、内容相对完整独立的内容。

3. 广告

报纸中企业或个人用于推销商品、服务，或向社会传递信息，影响 Yu Lun 的传播内容。政府或其他公务机构发布的各种公告不作为广告处理。

广告著录单元：报纸广告中有相对明显物理边界、内容相对完整独立的广告内容。

4. 图片

报纸的图片包括独立图片和文中附图。

图片著录单元：报纸正文中有相对明显物理边界的内容独立且有意义的图片或单独附图说的图片。

三、正文内容核心元数据及著录信息源

最终著录规则以实测样本为标准。元数据著录规则应该在正式生产前进行讨论，使其完全符合生产流程，同时内容适合于历史背景。

1. 正文著录项目及著录信息源

序号	正文著录项目	正文著录信息源
1	类别*	正文著录单元所属的主题类目，可按事先约定的分类体系著录

序号	正文著录项目	正文著录信息源
2	栏目	在报纸中有明确标识，将内容相近或有某种联系的内容组合在一起
3	子栏目	子栏目从属于栏目
4	标题 1*	著录单元内容的题名
5	翻译标题 1	其他语种的标题 1
6	标题 2	同标题 1
7	翻译标题 2	其他语种的标题 2
8	标题 3	同标题 1
9	翻译标题 3	其他语种的标题 3
10	责任者	报纸正文的第 1 责任者
11	责任者 2	报纸正文的第 2 责任者
12	责任者 3	报纸正文的第 3 责任者
13	责任者单位 1	第 1 责任者所属的单位
14	责任者单位 2	第 2 责任者所属的单位
15	责任者单位 3	第 3 责任者所属的单位
16	著作方式 1	第 1 责任者对应的著作方式
17	著作方式 2	第 2 责任者对应的著作方式
18	著作方式 3	第 3 责任者对应的著作方式
19	关键词	正文内容的关键词
20	新闻来源	新闻正文中显示该条新闻的采写来源，如“路透社电”、“Reuter”等。但只做在文首或文尾的新闻来源，其余为空
21	新闻发布地	只做在文首或文尾的新闻发布地，其余为空，且按原文著录
22	新闻发布时间	只做新闻正文上方显示的新闻发布时间，其余为空，且按原文著录
23	题中人名	正文标题中的人名
24	连载信息	正文中表示连载情况的信息，常见于文末

序号	正文著录项目	正文著录信息源
25	报纸名称*	正文所属的报纸报名
26	附属报名	正文内容所属报纸正刊的附属报纸的名称，包括增刊、副刊、特刊、纪念刊等的名称
27	年*	正文所属的报纸出版日期中的年份信息（公元纪年，格式为 yyyy）
28	月*	正文所属的报纸出版日期中的月份信息（公元纪年，格式为 mm）
29	日*	正文所属的报纸出版日期中的“日”信息（公元纪年，格式为 dd）
30	卷	正文所属报纸的卷信息
31	期	正文所属报纸的期信息
32	总期	正文所属报纸自创刊以来累计的期信息
33	版别	在报纸每版报眉描述版面名称的部分
34	页码	在报纸每版报眉描述报纸实际版序的部分，如 1，第一版，（一）等
35	版次*	招标方提供的原始图像命名中所指定的版序
36	附图数量	附属于正文的图片数量
37	备注	其他待说明事项
38	备用 1	招标方根据加工过程中的内容情况另行确定
39	备用 2	招标方根据加工过程中的内容情况另行确定
40	备用 3	招标方根据加工过程中的内容情况另行确定
41	文本	正文完整内容经过 OCR 后的文字

2. 著录项定义及注释

1) 类别

定义：正文著录单元所属的主题类目。

注释：属主观性著录项，需要通过对著录单元的内容进行判断后给出相应的类别，具体的类别可按事先设定的分类体系供著录时选择。

2) 栏目

定义：内容相近或有某种联系的几条著录单元编排在一起，有客观的名称且有相对

明显的物理边界，这个名称即为正文的栏目。

注释：属客观性著录项，著录时原则上按照著录单元所属栏目客观著录。

3) 子栏目

定义：属于某一个栏目下细分的从属栏目。

注释：属客观性著录项，按照著录单元所属栏目客观著录。

4) 标题

定义：著录单元内容的题名。

注释：属客观性著录项，著录单元中有明确题名时，直接照实著录原始题名。

5) 翻译标题

定义：对应标题的另一种语种的标题。

注释：翻译标题可按相应的翻译原则进行翻译或照实著录原文所列另一种语种的标题。

6) 责任者

定义：创建著录单元内容的主要责任者。

注释：本元素属客观性著录项，是对数字文献内容负主要责任的个人或团体责任者进行著录，本项最多著录三个责任者。

7) 著作方式

定义：责任者创建著录单元时参与的方式。

注释：本元素属客观性著录项，是对文献内容负主要责任的个人或团体参与创建的方式，本项最多著录三个著作方式，与责任者对应。

8) 责任者单位

定义：个人责任者所属单位。

注释：本元素属客观性著录项。

9) 关键词

定义：著录单元的关键词，可采用描述正文内容的非控关键词。

注释：本元素属主观性著录项，依据著录单元内容给出相应的关键词。

10) 新闻来源

定义：著录单元中显示的该条新闻的采写来源。

注释：本元素属客观性著录项。

11) 新闻发布地

定义：新闻发布的地点。

注释：本元素属客观性著录项，依据著录单元的内容客观著录。

12) 新闻发布时间

定义：新闻发布的时间。

注释：本元素属客观性著录项，依据著录单元的内容客观著录。采用公元纪年著录，格式为 YYYY-MM-DD。

13) 题中人名

定义：著录单元的标题中的人名。

注释：本元素属客观性著录项，著录单元的标题 1、标题 2、标题 3 中的人名。

14) 连载信息

定义：正文中表示连载情况的信息，常见于文末。

注释：本元素属客观性著录项，按原件照实著录，如“未完”、“未完待续”等。

15) 报纸名称

定义：正文内容所属报纸的名称。

注释：客观性著录项。

16) 附属报名

定义：正文内容所属报纸正刊的附属报纸的名称，包括增刊、副刊、特刊、纪念刊等。

注释：客观性著录项。

17) 年

定义：正文所属的报纸出版日期中的年份信息。

注释：本元素应为公元纪年并采用 Arabic 数字著录，非公元纪年须转为公元纪年，格式为 yyyy。

18) 月

定义：正文所属的报纸出版日期中的月份信息。

注释：本元素应为公元纪年并采用 Arabic 数字著录，非公元纪年须转为公元纪年，格式为 mm。

19) 日

定义：正文所属的报纸出版日期中的“日”信息。

注释：本元素应为公元纪年并采用 Arabic 数字著录，非公元纪年须转为公元纪年，

格式为 dd。

20) 卷

定义：正文所属报纸的卷信息。

注释：本元素取自报纸当期报头或主要内容版的报眉，一般应采用 Arabic 数字著录。

21) 期

定义：正文所属报纸的期信息。

注释：本元素取自报纸当期报头或主要内容版的报眉，一般应采用 Arabic 数字著录。

22) 总期

定义：正文所属报纸自创刊以来累计的期信息。

注释：本元素取自报纸当期报头或主要内容版的报眉，一般应采用 Arabic 数字著录。

23) 版别

定义：著录单元所在版的版别。

注释：本元素属客观性著录项，在报纸每版报眉描述版面名称的部分，以中文报纸为例，如新闻、娱乐、经济等等。

24) 页码

定义：著录单元所在版的报眉处描述报纸实际版序的部分。

注释：本元素属客观性著录项，取自著录单元所在版报眉，著录人员只照实著录起始版序，若无实际版序应尽量根据其前后版情况进行判断著录。

25) 版次

定义：招标方提供的原始图像命名中所指定的版序信息，属数字化图像过程中产生的报纸扫描版次信息。

注释：版次信息取自招标方提供的原始图像命名中所指定的版序信息，依据当期报纸的版次顺序著录，版次用 4 位 Arabic 数字著录；出现跨版或转版时，跨版请用“-”连接，转版请用“,”连接。

26) 附图数量

定义：附属于正文的图片数量。

注释：客观著录项。

27) 文本

定义：正文完整内容经过 OCR 后的文字。

注释：属客观著录项，应符合正文内容排版顺序。

四、广告内容核心元数据及著录信息源

1. 广告著录项目及著录信息源

报纸中的广告元数据著录，每个著录单元的著录项包括广告类别、广告栏目、卷期信息、页码、版次等。

序号	广告著录项目	广告著录信息源
1	广告标题	广告中对广告内容进行概括，起引导作用的语句
2	广告语	广告中长期、反复使用的简短口号
3	广告发布者	发布广告的个人或机构
4	广告产品	广告宣传的具体产品名称
5	广告类别*	广告著录单元所属的主题类目，可按事先约定的分类体系著录
6	广告栏目	内容相近或有某种联系的几条广告编排在一起，有客观的栏目名称
7	报纸名称*	同正文
8	附属报名	同正文
9	年*	同正文
10	月*	同正文
11	日*	同正文
12	卷	同正文
13	期	同正文
14	总期	同正文
15	版别	同正文
16	页码	同正文
17	版次*	同正文
18	备注	其他待说明事项
19	备用 1	招标方根据加工过程中的内容情况另行确定

序号	广告著录项目	广告著录信息源
20	备用 2	招标方根据加工过程中的内容情况另行确定
21	备用 3	招标方根据加工过程中的内容情况另行确定

2. 广告著录项定义及注释

广告著录单元的卷期、页码、版次著录项的定义和注释同正文。下面只列出广告类别、广告栏目的定义和注释。

1) 广告标题

定义：广告中对广告内容进行概括，起引导作用的语句。

注释：属客观性著录项。广告标题在具体某一则广告中使用，与广告正文内容同时出现，区别于广告语。

2) 广告语

定义：广告中长期、反复使用的简短口号。

注释：属于客观性著录项。广告语具有稳定性和连贯性，可在多个广告或系列广告中使用，区别于广告标题。

3) 广告发布者

定义：发布广告的主体，包括个人和机构。

注释：属客观性著录项。

4) 广告产品

定义：广告宣传的具体产品名称。

注释：属客观性著录项。

5) 广告类别

定义：广告著录单元所属的主题类目。

注释：属主观性著录项，需要通过对著录单元的广告内容进行判断后给出相应的类别，具体的类别可按事先设定的分类体系供著录时选择。

6) 广告栏目

定义：在报纸中有明确标识，且有相对明显的物理边界，将内容相近或有某种联系的几条著录单元组合在一起。

注释：属客观性著录项，著录时原则上按照著录单元所属栏目进行客观著录。

五、图片内容元数据及著录信息源

1. 图片著录项目及著录信息源

序号	图片著录项目	图片著录信息源
1	图片栏目	有明确标识，将内容相近或有某种联系的图片组合在一起，形成图片栏目
2	图片标题*	图片标题或图片所依附正文标题
3	图片说明	图片的文字说明内容
4	图片责任者 1	完成图片内容创建的第一责任者
5	图片责任者 2	完成图片内容创建的第二责任者
6	图片著作方式 1	第一责任者对应的著作方式，如画、绘、摄等
7	图片著作方式 2	第二责任者对应的著作方式，如画、绘、摄等
8	图片类型 1*	漫画、照片、书法、题字、题词、手稿、木刻、篆刻、石刻、雕塑、曲谱、歌谱、表格、图表、画图、地图、拓本、拓片、其他
9	图片类型 2	仅当图片为完整的封面时著录
10	报纸名称*	同正文
11	附属报名	同正文
12	年*	同正文
13	月*	同正文
14	日*	同正文
15	卷	同正文
16	期	同正文
17	总期	同正文
18	版别	同正文
19	页码	同正文
20	版次*	同正文
21	备注	其他待说明事项
22	备用 1	招标方根据加工过程中的内容情况另行确定
23	备用 2	招标方根据加工过程中的内容情况另行确定
24	备用 3	招标方根据加工过程中的内容情况另行确定

2. 图片著录项定义及注释

1) 图片标题

定义：图片标题或图片所依附正文标题。

注释：属客观性著录项，依据图片标题或图片所依附正文标题客观著录。

2) 图片说明

定义：正文中图片的文字说明内容。

注释：属客观性著录项，依据正文中图片的文字说明内容客观著录。

3) 图片责任者

定义：图片的责任者。

注释：属客观性著录项。

4) 图片类型

定义：图片著录单元所属的类型。

注释：属主观性著录项，需要通过对著录单元的图片类型进行判断后给出选择，具体的类型可按事先设定的分类体系供著录时选择。

六、元数据元素著录说明

1. 关于字段非空

1) 上述各表中标*的元素均为非空字段，著录项不可缺省；未标*的元素视具体内容具体分析，有内容则必须著录。

2) 广告标题、广告发布者和广告产品不能同时为空。

2. 关于图片的著录与划框

图片著录时，原则上以单独图片进行著录，对于形状不规则的图片，在保证主题表达完整的前提下可进行组合图片的划框界定。

3. 关于著录项所用语种

除正文的“类别”、广告的“广告类别”、图片的“图片类型”以及“备注”跟随报纸语种外，其余著录项所用语种均依据制作报纸内容照实进行著录。报纸语种由招标方确定。

4. 关于广告的“广告类别”、图片的“图片类型”和正文的“类别”字段

1) 广告类别字段的分类体系一般可包括：启事公告、个人事务、餐饮娱乐、邮电、交通运输、人才招聘、文化教育、医疗保健、食品饮料、百货零售、金融保险、房地产、拍卖、酒店旅游、能源供应、建筑装饰、钟表珠宝、烟草、工业用品、其他等等。

2) 图片类型的分类体系一般可包括：漫画、照片、书法、题字、题词、手稿、木刻、

篆刻、石刻、雕塑、曲谱、歌谱、表格、图表、画图、地图、拓本、拓片、其他等等。

- 3) 正文的类别字段的分类体系一般可包括：消息、通讯、评论、讲话、编辑部来信、公文、小说、散文、诗歌、剧本、原报导读、其他等等。
- 4) 招标方可在样本制作之前或样本制作过程中对正文的类别、广告类别和图片类型的分类体系提出合理的修改。在投标方进入正式生产后，分类体系的修改需双方另行商量决定。

5. 关于正文的翻译标题字段

中文报纸中的全外文文章，正文翻译标题应翻译为相应的中文并著录在该字段。

若原文标题与另一种语种的标题混合，则不要求著录翻译字段，完整照实著录在标题字段即可。

6. 关于不要求著录的字段

- 1) 正文的关键词字段不要求著录。
- 2) 图片类型 2 不要求著录。
- 3) 对于所有不要求著录的字段都应保留。

7. 关于标点、符号著录

著录中使用的所有标点、符号需双方协商确认。

8. 关于排序

报纸中每一条正文、广告和图片元数据都应有对应的排序信息，排序应符合报纸实际文脉和阅读次序，以实现数据检索应用过程中的各种排序功能。

9. 关于元数据、图像及 OCR 文本的唯一标识及对应关系

正文唯一标识、广告唯一标识、图片唯一标识等由著录系统自动生成，且唯一标识一旦生成则不可改变，命名规则由招标方确定。元数据、图像及 OCR 文本之间的对应关系一定要满足数据的检索应用要求。

七、OCR 文本识别数据加工要求

总体原则：以原件内容的客观含义为准，尽量忠实于原文进行识别。要求识别文本内容与原图篇内容应一一对应。

定期优化更新技术软件，应用最新版本技术软件来进行文字识别。

1. 文字识别

对待加工中文老报纸的篇目正文内容进行 OCR 文字识别。遵照原文的篇目范围、正文框顺序和阅读方式按要求进行识别。

2. 文字分类

OCR 得到的字符划分为三类：

- 1) 原图清晰、识别正确的字符集内文字或符号，简称为正常字。
- 2) 将集外字正确认同为字符集内文字或符号，简称为认同字。
- 3) 其他符号。

3. 文字识别正确率

- 1) OCR 文字识别每批次正确率要求不低于 50%，整体项目各批次汇总后平均文字识别率不低于 60%。
- 2) 影响文字识别正确率的因素涉及待识别篇目的版式、清晰度和文字大小等。

4. XML 文件

- 1) XML 标签合法性检查要求零错误。
- 2) XML 文件要求遵守 XML 语法规则，使用 utf-8 编码方式。

5. 其他

图片篇和广告篇目中的文字暂不进行全文文字识别等处理。

6. 文件处理质量要求

- 1) 文件夹和文件命名正确。
- 2) 版面内容全部制作无遗漏。
- 3) 尽量忠于原文，进行文字原貌识别。
- 4) 文字识别正确率不低于 50%，原文文字印刷错误，不记入错误率内。

7. 特殊情况处理

本加工要求针对中文老报纸资源的一般情况制定，因不同种类的报纸在规格、结构、特种刊处理和质量等方面有较大差异，会存在本加工要求未涉及的内容。处理过程中出现任何本加工要求以外特殊情况，最终处理方案应以双方协商结果为准。如有争议，以采购方最终的解释为准。

注：本附件解释权归招标方所有，著录过程中本规则未能详尽的内容由双方协商确定。

附件二：数据质量标准及数据格式规范

一、元数据著录质量标准

由于本合同规定的数据验收方式为分批验收，故本质量标准所指的错误率计算均以批为计量单位。

1. 图像处理质量标准：图像错误率 $\leq 0.05\%$

- 1) 图像错误率=（错误图像版数 $\div$ 该批次图像总版数） $\times 100\%$ （单位：版）
- 2) 每版图像应当同时符合下列质量标准，如有一项不符合，即该版为错误图像
 - （1）图像清晰。
 - （2）图像完整，无缺损。
 - （3）图像无明显污渍。
 - （4）图像无扭曲、倾斜、变形。
- 3) 图像数据出现下列情况之一的，即该批次为验收不合格
 - （1）图像错误率实际值超过规定值。
 - （2）图像出现缺页、多页、错页、重页等情况之一的。

2. 元数据著录质量标准：著录规则错误率 $\leq 0.3\%$

元数据著录规则错误率是元数据著录项错误的数量与需著录项的数据总量比例，计算公式为：元数据著录规则错误率=（元数据著录项错误的数量 $\div$ 需著录项的数据总量） $\times 100\%$ （单位：项）。

著录规则错误是按项为计算单位，凡未能对元数据项进行客观著录的情况均计为著录规则错误。一篇元数据项为多值的，则按值的数量计算项数。

3. 著录文字质量标准：著录文字错误率 $\leq 0.05\%$

著录文字错误率是元数据项内容文字错误的数量（如多、少、错等）与元数据项文字总量的比例，计算公式为：著录文字错误率=（元数据项内容文字错误的数量 $\div$ 元数据项文字总量） $\times 100\%$ （单位：字）。

著录文字错误按文字计算，元数据中文字与原文不符的计为著录文字错误，但不包括对原文的文字错误进行纠正的情况，原文错误纠正记入勘误表。

4. 正文文字质量标准：零漏著

5. 元数据与图像的勾链质量标准：元数据与图像的勾链无差错

6. 版面内容元数据质量标准：版面内容元数据无漏著

7. 篇名类型著录质量标准：篇名类型无差错

8. 其他元数据的合法性验收标准

- 1) 返回图像和 PDF 的版数必须和著录的版数完全相同。
- 2) 返回图像像素、大小无异常,灰度或色彩无明显漏洞,不存在多幅图像相同的问题。
- 3) 返回图像元数据著录项中非空元素著录项无缺省。

非空元素是指《附件二:元数据著录规则》中第三、四、五条第 1 款标*的元素,著录项不可缺省。

二、数据规范格式

根据本项目制作要求,投标方须提供的数据有:单版 PDF 文件,图像文件(单版图像、著录图像、著录广告图像)的 TIF 及 JPG 格式各一套,元数据著录信息文件,版面分析信息文件等,具体规范要求如下:

1. PDF 文件

- 1) 双层 PDF 的上层图像为 JPG 格式,下层为篇目全文文本层,且篇目位置已置标。
- 2) PDF 格式,要求 PDF 编码为 1.5 版本(兼容 Adobe reader6.0),要求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) PDF 像素、大小无异常,灰度或色彩无明显漏洞。
- 6) 在保证图像清晰度的前提下,对保存级别 300DIP/TIF 图像进行压缩。
- 7) 使用 JPG 压缩方式,根据原始图像质量和大小,选取适当的压缩比例(推荐值为 65%)。生产时以样本为准。
- 8) 如有必要,需对招标方提供的源图像进行修正,以保证产出的 PDF 文件画面清晰、无扭曲、倾斜、变形等影响查看的情况发生。
- 9) 除特殊规定外,PDF 内容的完整性和准确性应与招标方提供的源图像保持一致。

2. 成品图像及要求

成品图像包括保存级别和应用级别两套数据。投标方应保持图像为报纸原始扫描色彩。图片和广告,需根据实际情况进行抠图后保存。

- 1) 保存级别图像:包括整版图像、著录图像和著录广告图像
要求:应保持原有的图像样式,保持 300DPI/TIF,使用 LZW 无损压缩格式。
- 2) 应用级别图像:整版图像、著录图像和著录广告图像

要求：在保证图像清晰的前提下，对保存级别 300DIP/TIF 图像进行压缩，根据原始图像的大小和扫描质量，选取恰当的压缩比例（推荐值为 65%），使用 JPG 压缩方式，最终提供整版图像保存为 150DPI、100DPI 两种分辨率的 JPG 格式；著录图像保存为 150DPI 的 JPG 格式，大小视实际抠出大小而定。

抠出的图片、广告等图像务必要保持原始图像的完整性和整洁性，遇到排版不规则、交叉等无法规则抠出等情况，按照尽量保持原始的整体性原则进行处理。

3. 元数据著录信息文件

元数据著录信息文件存储所有元数据信息的可读可写文件，应按照附件一《元数据著录规则》著录单篇元数据、对应的 PDF，图片路径、广告图片路径等。

坐标信息也需要在元数据著录信息文件中标出，包括单篇坐标和抠图图片坐标。

4. 命名规则

文件的命名需与招标方协商后确定，要求规则统一，名称唯一。

建议命名规则：

第一层：交付文件类型，分别为 TIF、XML、PIC、OCR 四个目录。

第二层：批次，统一命名为：shlib-项目简称-资源形态-批次号，其中批次号从 0001 开始标。

注：文件提交的时候必须保证一个批次只有一种报纸，并且还必须要保证批次号是全局唯一号。

第三层：报刊划到号

第四层：报刊代码+日期 8 位

5. 版面分析信息文件

提供该批次图像识别校改后的针对版面 300DPI 图像文件的工程文件（包括版面结构和版面内容），把结构基元（包括行坐标、段落坐标、划框信息等）内容封装在 OCR 内，方便招标方进行验收以及后期再次加工和进一步的信息采集。

6. 其他要求

投标方应根据招标方提供的图像，按照招标文件要求进行样本制作，并提供相应的匹配说明文件，方便招标方进行判断。双方认可的数据样例作为本项目著录及数据规范的参照依据。

项目 2：科技现刊篇名元数据制作

项目数据量：900,000 篇

项目完成时间：2025 年 10 月 31 日

项目进度要求：

第一阶段：合同签署后，双方确认技术标准后，至 2025 年 6 月 30 日前至少完成 450,000 篇

第二阶段：至 2025 年 10 月 31 日前，完成 900,000 篇

一、招标样本

- 由招标方提供招标所需样本图像，共计约 53 页
- 投标方依据招标方提供的著录标准及各项规范（见下附件）制作招标样本
- 各投标方在将样品提交前，请先在样本上标识所投项目、投标方等信息，并确认标识是否完好
- 投标方需在投标日当天出示招标样本，并做阐述
- 投标方提交的样本将作为评估中标与否的因素之一
- 最终解释权归招标方所有

二、元数据主要著录内容要求

1. 正文内容元数据：11 项 ≤ 著录项 ≤ 29 项
2. 元数据制作成品：单篇元数据著录信息文件、单篇 PDF、整本期刊 PDF。
3. 元数据著录要求：依据文献的相关信息和内容客观著录。
4. 著录语种：本项目仅著录中文和英文的正文内容，其它语种内容不予著录。
5. 元数据元素主要著录说明

1) 关于字段非空

“正文著录项目及著录信息源”表中标有“*”的元素为非空字段，著录项不可缺省；未标“*”的元素视具体内容具体分析，有内容则必须著录。

2) 关于不要求著录的字段

- (1) “译名”、“基金项目”和“Web”字段不要求著录。

(2) 对于所有不要求著录的字段都应保留。

3) 关于元数据与图像的对应关系

元数据和 PDF 图像之间保持正确的对应关系, 以满足数据的检索应用要求。

4) 关于元数据的唯一标识

元数据需有全局唯一标识, 且该标识一旦生成则不可改变, 命名规则由招标方确定。

5) 关于著录中的符号使用的说明

著录中使用的标点、符号与招标方协商确认。

三、验收质量标准

1. 著录规则质量标准: 著录规则错误率 $\leq 0.3\%$
2. 著录文字标准: 著录文字错误率 $\leq 0.05\%$
3. 元数据与图像文件的勾链质量标准: 元数据与图像文件的勾链无差错
4. 整册无漏著篇。

四、数据规范格式

投标方需根据招标方提供的数据格式进行此项目制作。

投标方须提供的数据有: 单篇 PDF 文件、整本 PDF 文件、元数据著录信息文件等。具体规范详见附件。

五、质保服务

投标方在总项目完成后提供一年的免费质保服务。双方在合同签订后另行签署质保协议书。

附件一: 元数据著录规则

一、前言

现代期刊(简称现刊)数据库元数据的著录仅针对正文内容元数据。

著录语种: 本项目仅著录中文和英文的正文内容, 其它语种内容不予著录。

二、术语定义

下列术语和定义适用于本附件。

1. 元数据著录单元

指现刊图像中有相对明显物理边界内的内容相对完整的正文。下文中简称“著录单元”。

2. 正文

指现刊中具有收录价值的文章。

三、正文内容核心元数据及著录信息源

1. 正文著录项目及著录信息源

序号	正文著录项目	正文著录信息源
1	分类号	正文上著者赋予的中图分类号
2	主题词/关键词	正文上著者赋予的主题词或关键词
3	题名*	正文的标题
4	译名	正文中文题名对应的英文标题
5	责任者 1	正文的第 1 个人责任者
6	责任者 2	正文的第 2 个人责任者
7	责任者 3	正文的第 3 个人责任者
8	责任者著作方式 1	责任者 1 的著作方式
9	责任者著作方式 2	责任者 2 的著作方式
10	责任者著作方式 3	责任者 3 的著作方式
11	单位 1	责任者 1 所属第一个单位名称
12	单位 2	责任者 2 所属第一个单位名称
13	单位 3	责任者 3 所属第一个单位名称
14	团体责任者 1	正文的第 1 团体责任者
15	团体责任者 2	正文的第 2 团体责任者
16	团体责任者 3	正文的第 3 团体责任者
17	团体责任者著作方式 1	团体责任者 1 的著作方式
18	团体责任者著作方式 2	团体责任者 2 的著作方式
19	团体责任者著作方式 3	团体责任者 3 的著作方式

序号	正文著录项目	正文著录信息源
20	题中人名	正文题名中出现的人名
21	文摘	正文有摘要则照实著录，无摘要则仅著录正文前 200 个汉字/英文单词
22	基金项目 1	正文内容研究所受支持的基金项目
23	基金项目 2	正文内容研究所受支持的基金项目
24	基金项目 3	正文内容研究所受支持的基金项目
25	刊名*	正文所属的期刊名称
26	出版年份*	正文所属期刊的出版年份
27	卷	正文所属期刊的卷号
28	期*	正文所属期刊的期号
29	页码*	正文在所属期刊中的页码
30	Web	正文发表在网络上的网址
31	PDF 的链接*	正文 PDF 的链接
32	索书号*	招标方提供的期刊索书号
33	批次号*	招标方提供每捆期刊的原始批次号，投标方根据要求提交正确的成品批次号
34	上图馆藏信息	招标方提供
35	期刊资源 ID*	招标方提供
36	期刊等级*	招标方提供
37	期刊分类	招标方提供
38	工作组	招标方提供
39	索引库分类号	招标方提供
40	社科、科技标记	招标方提供
41	备注	
42	备用 1	招标方根据加工过程中的内容情况另行确定
43	备用 2	招标方根据加工过程中的内容情况另行确定
44	备用 3	招标方根据加工过程中的内容情况另行确定
45	期刊划到 ID	招标方提供
46	篇 ID*	正文的唯一标识符

注：1) 个人责任者和团体责任者不能同时著录。

2) 34-40 字段无需制作方著录。

2. 著录项定义及注释

1) 分类号

定义：正文中著者所赋予的著录单元的中图分类号。

注释：属于客观性著录项，正文中有则著录，无则不予著录，多个分类号以“|”间隔。

2) 主题词/关键词

定义：正文中著者所赋予著录单元主题内容的主题词/关键词。

注释：属于客观性著录项，正文中有则著录，无则不予著录；多个主题词/关键词以“|”间隔。

3) 题名*

定义：著录单元的著录标题。

注释：属客观性著录项，按照正文照实著录。

4) 译名（该项暂不著录）

定义：著录单元中正文中文题名对应的英文标题。

注释：属客观性著录项，按照正文照实著录。

5) 责任者

定义：创建著录单元内容的主要个人责任者。

注释：本元素属客观性著录项，是对文献内容负主要责任的个人，本项目最多著录三个个人责任者。

6) 责任者著作方式

定义：个人责任者创建著录单元时参与的方式。

注释：属客观性著录项，是对文献内容负主要责任的个人参与创建的方式，本项目最多著录三个著作方式，与个人责任者对应。如著、翻译、评论、口述、笔录等表示责任者著作方式的词等。

7) 单位

定义：著录单元个人责任者所在机构的名称。

注释：本元素属客观性著录项，本项目最多著录三个单位名称，与个人责任者对应，且仅著录个人责任者的第一个单位名称。

8) 团体责任者

定义：创建著录单元内容的主要团体责任者。

注释：本元素属客观性著录项，是对文献内容负主要责任的团体，本项目最多著录三个团体责任者。

9) 团体责任者著作方式

定义：团体责任者创建著录单元时参与的方式。

注释：属客观性著录项，是对文献内容负主要责任的团体参与创建的方式，本项目最多著录三个著作方式，与团体责任者对应。如著、翻译、评论、口述、笔录等表示责任者著作方式的词等。

10) 题中人名

定义：指著录单元题名中出现的人物姓名。

注释：属客观性著录项。

11) 文摘（该项仅针对核心期刊著录）

定义：著录单元内容的概述。

注释：本元素属客观性著录项，原文有摘要则直接照实著录；原文无摘要则取正文前 200 个汉字/英文单词（包括标点符号）作为摘要，并在最后加注省略号（省略号不包括在 200 个汉字/英文单词内）。

12) 基金项目（该项暂不著录）

定义：著录单元内容研究所获得支持的基金项目名称。

注释：本元素属客观性著录项，正文中有则按照招标方提供的规则著录，最多著录三个基金项目名称。

13) 刊名*

定义：著录单元所属期刊的名称。

注释：本元素属客观性著录项。

14) 出版年份*

定义：著录单元所属期刊的出版年份。

注释：采用公元纪年著录，以 Arabic 数字表示。

15) 卷

定义：著录单元所属期刊的实际卷号。

注释：本元素属客观性著录项，以 Arabic 数字表示。

16) 期*

定义：著录单元所属期刊的实际期号。

注释：本元素属客观性著录项，以 Arabic 数字表示。

17) 页码*

定义：著录单元在期刊中所属的实际页码。

注释：本元素属客观性著录项，著录起始及转接页码。

18) Web（该项暂不著录）

定义：著录单元发表在网络上的网址。

注释：本元素属客观性著录项，正文中标注有网址则照实著录。

19) PDF*

定义：著录单元对应的 PDF 的链接。

注释：正文对应的图像 PDF 文件的存储路径，对应单篇 PDF 链接和整本 PDF 链接。

20) 索书号*

定义：著录单元对应的期刊的索书号。

注释：论文所属期刊的索书号，招标方提供。

21) 批次号*

定义：著录单元对应的那捆期刊的批次号。

注释：论文所属期刊归属的批次号，招标方提供。

22) 篇 ID*

定义：正文的唯一标识符

注释：用期刊划到 ID+3 位 Arabic 数字表示（001-999），表示正文的生产顺序，当增篇时，新篇的篇 ID 顺延，删篇时，该篇的篇 ID 不再为其他篇所用。

四、元数据元素著录说明

1. 关于字段非空

“正文著录项目及著录信息源”表中标有“*”的元素为非空字段，著录项不可缺省；未标“*”的元素视具体内容具体分析，有内容则必须著录。

2. 关于不要求著录的字段

- 1) “译名”、“基金项目”和“Web”字段不要求著录；
- 2) 对于所有不要求著录的字段都应保留。

3. 关于元数据与图像的对应关系

元数据和 PDF 图像之间保持正确的对应关系，以满足数据的检索应用要求。

4. 关于元数据的唯一标识

元数据需有全局唯一标识，且该标识一旦生成则不可改变，命名规则由招标方确定。

5. 关于著录中的符号使用的说明

著录中使用的标点、符号与招标方协商确认。

注：本附件解释权归招标方所有，著录过程中本规则未能详尽的内容由双方协商确定。

附件二：数据质量标准及数据格式规范

一、科技现刊篇名元数据制作数据质量标准

1. 元数据著录要求：依据文献的相关信息和内容客观著录。

2. 元数据制作成品

元数据著录信息文件：单篇元数据

PDF 文件：

单篇 PDF：与单篇元数据对应

整本期刊 PDF：与划到对应

3. 元数据著录质量标准

1) 著录规则质量标准：著录规则错误率 $\leq 0.3\%$

(1) 元数据著录规则错误率是规则错误数据量与抽取数据总量的比例，计算公式为：元数据著录规则错误率 = (规则错误数据项 ÷ 该批次著录的项数总和) × 100% (单位：项)。

(2) 著录规则错误按项为计算单位，凡未按照著录规则进行著录的情况均计为著录规则错误。一条元数据中多字段规则错误的，视为多项数据错误。如一篇元数据有 10 个著录项，则计为 10 个项。

2) 著录文字质量标准：著录文字错误率 $\leq 0.05\%$

(1) 元数据著录文字错误率是文字错误数据量与抽取文字数据总量的比例，计算公式为：元数据著录文字错误率 = (文字错误数据项 ÷ 抽取的文字总项) × 100% (单位：项)。

- (2) 著录文字错误按项为计算单位, 凡未按照著录文字进行著录的情况均计为著录文字错误, 但不包括对原文的文字错误进行纠正的情况。一条元数据中多项文字错误的, 视为多项数据错误。
- 3) 元数据与图像的勾链质量标准: 元数据与图像的勾链无差错
- (1) 元数据与图像勾链错误率是元数据与图像勾链错误数据量与抽取数据勾连总量的比例。
- (2) 计算公式为: 元数据与图像勾链错误率 = (元数据与图像勾链错误数据量 ÷ 抽取数据勾连总量) × 100% (单位: 篇)。
- 4) 按附件一所要求的著录单元, 整册需无漏著篇。

二、科技现刊篇名元数据制作数据格式规范

根据本项目制作要求, 投标方须提供的数据有: 单篇 PDF 文件、整本 PDF 文件、元数据著录信息文件, 具体规范要求如下:

1. PDF 文件

每本期刊的 PDF 数据为两套, 一套为单篇 PDF, 一套为整本 PDF。质量要求如下:

- 1) PDF 均为双层 PDF 文件, 底层为文字层, 表层为图像层, 图像层能与文字层精确对应。
- 2) PDF 版本通用性, 需要符合 PDF 国际标准。投标方提交成品时, 需向招标方说明使用的 PDF 版本。
- 3) PDF 像素、大小无异常, 灰度或色彩无明显漏洞。
- 4) 在保证图像清晰的情况下, 对图像层进行压缩, 根据图像质量和大小采用适当的压缩率, 压缩后单篇 PDF 大小控制在 150K 左右 (特殊画幅、质量, 原始图像容量特别大的除外), 生产时以样本为准。
- 5) 如有必要, 需对招标方提供的源图像进行修正, 以保证产出的 PDF 文件画面清晰、无扭曲、倾斜、变形等影响查看的情况发生。
- 6) 除特殊规定外, PDF 内容的完整性和准确性要和招标方提供的源图像保持一致。
- 7) 单篇 PDF 无重复性。当多篇所对应的 PDF 完全相同的情况下, 指向第一篇的 PDF, 不用输出多个重复内容的 PDF。

2. 元数据著录信息文件

元数据著录信息文件是存储所有元数据信息的可读可写文件, 应按照附件一元数据著

录规则著录单篇文章的元数据，及对应的单篇、整本 PDF 路径。

元数据以 MDB 格式提交，包含著录数据各项信息。元数据与 PDF 数据须同步提交。

3. MDB 及 PDF 著录及命名规则

1) MDB 文件以批次号命名，文件存储在批次号下。

2) MDB 文件中相关字段著录及提取规范

(1) m001 字段著录规则为：划到 ID+3 位 Arabic 数字 (001-999)。

(2) reserve2 字段记录单篇 pdf 的存储路径，从成品数据的第二层开始记录。

举例：单篇 PDF 存放路径为：shlib-kjp17-tx-Z2-2017-0002\44317\Z2-2017-0002-11073-5-2017-10\9_028.pdf，则 reserve2 中著录：44317\Z2-2017-0002-11073-5-2017-10\9_028.pdf

(3) 字段内容提取规范

按照招标方提供的批次清单进行数据制作，提取规范如下：

序号	著录字段	清单字段	著录规范
1	m250_a	年	
2	m250_b	卷	
3	m250_c	期	用半角括号，举例：(9A)
4	pinfo_id	资源 ID	
5	hd_id	划到 ID	
6	m463	刊(报)名	
7	m950_a	索书号	1. 索书号中有横杠，则提取索书号中第一个横杠前的字符进行著录 2. 索书号中无横杠有星号，则提取索书号中第一个星号前的字符进行著录 3. 索书号中无横杠无星号，则直接全部提取进行著录
8	m950_b	索书号	1. 索书号中有横杠有星号，则提取索书号中第一个横杠后第一个星号前的字符进行著录 2. 索书号中有横杠无星号，则提取索书号中第一个横杠后的字符进行著录 3. 索书号中无横杠，则著录空

序号	著录字段	清单字段	著录规范
9	reserve3	索书号	1. 若索书号中有星号，则提取索书号中星号之后的字符 2. 若索书号中无星号，则著录空
10	pkernel	期刊等级	1. 期刊等级=‘核心’，则著录 0 2. 期刊等级=‘次核心’，则著录 1 3. 期刊等级=‘非核心’，则著录 2
11	pclass	期刊分类号	
12	workgroup	批次代码	提取批次代码中的分类组部分，参合同附件中的描述
13	reserve4	索书号	提取索书号中的全部内容

3) PDF 文件要求命名规则统一，名称唯一。

(1) 单篇 PDF 命名为“篇 id_篇起始页码.pdf”。

(2) 整本 PDF 命名为“All.pdf”。

(3) 单篇及整本 PDF 均存储于文件夹的最后一层。

4. 单篇 PDF 排序方式

1) 按文件名左侧序号顺序输出。

2) 左侧序号相同情况下，按右侧序号顺序输出。

5. 其他要求

投标方应根据招标方提供的图像，按照招标文件要求进行样本制作，并提供相应的匹配说明文件，方便招标方进行判断。以上格式规范如需调整，可由双方在样本制作时协商。双方认可的数据样例作为本项目著录及数据规范的参照依据。

三、科技现刊篇名元数据制作数据存储规范

成品数据存储目录由招标方规定，投标方将制作好的数据按实际内容存放在相应的文件夹中。招标方规定的目录架构及命名规则说明如下：

1. 目录架构为三级：批次号\划到 ID\捆号-索书号-年份-期

2. 目录架构命名规则

1) 批次号为：资源来源-项目简称-资源形态-分类组-期刊年份-批次

(1) 资源来源为数字化成果的资源来源，由英文字母(区分大小写)或者数字组成，长度 2-5 个字符，本年项目统一为：shlib。

(2) 项目简称为加工的项目，用 kjp+两位数字（项目年份）表示，如“kjp17”表

示 2017 的本项目。

- (3) 资源形态固定标记为 tx。
 - (4) 分类组为本批次中所有刊所在的期刊分组，由英文字母(区分大小写)或数字组成，如 N、W1、Z6、Z7。
 - (5) 期刊年份为期刊的出版年份，用四位 Arabic 数字表示。
 - (6) 批次用四位 Arabic 数字表示，从 0001 顺延。
- 2) 划到 ID 为每本期刊的唯一标识符，由不定位的 Arabic 数字组成。
- 3) 捆号为：分类组-期刊年份-批次，同批次号中的“分类组-期刊年份-批次”一致。
- 索书号为由不定位 Arabic 数字、符号(-)、字母组成，形如“582”、“258-1”、“1585-A”；年份为四位 Arabic 数字；期由两位 Arabic 数字或文本组成，期为纯数字时，不足 2 位的数字在前面补 0。

举例：2017 年科技现刊篇名项目中的 Z6 的第 1 批次，索书号为 9255 的《光学仪器》2017 年第 1 期的存储路径为：shlib-kjp17-tx-Z6-2017-0001\10210\Z6-2017-0001-9255-2017-01

项目 3：老旧期刊篇名元数据制作（中）

项目数据量：100,000 篇

项目完成时间：2025 年 10 月 31 日

项目进度要求：

第一阶段：合同签署后，在双方确认技术标准后，至 2025 年 6 月 30 日前至少完成 50,000 篇

第二阶段：至 2025 年 10 月 31 日前，完成 100,000 篇

一、招标样本

- 由招标方提供招标所需样本图像，共计约 22 页
- 投标方依据招标方提供的著录标准及各项规范（见下附件）制作招标样本
- 各投标方在将样品提交前，请先在样本上标识所投项目、投标方等信息，并确认标识是否完好
- 投标方需在投标日当天出示招标样本，并做阐述
- 投标方提交的样本将作为评估中标与否的因素之一
- 最终解释权归招标方所有

二、元数据主要著录内容要求

1. 正文内容元数据：6 项 ≤ 著录项 ≤ 29 项
广告元数据：5 项 ≤ 著录项 ≤ 14 项
图片元数据：6 项 ≤ 著录项 ≤ 17 项
2. 元数据制作成品：期刊的整本 PDF、单篇 PDF、图像文件（著录图片、著录广告图片）、元数据著录信息文件。
3. 元数据著录要求：依据文献的相关信息和内容客观著录。
4. 著录语种：期刊的元数据著录项均依据制作期刊内容照实进行著录。
5. 元数据元素主要著录说明
 - 1) 关于元数据元素著录

上述各表中标*的元素均为非空字段，著录项不可缺省；未标*的元素视具体内容具体分析，有内容则必须著录；

2) 关于元数据著录注释中“客观著录”与“照实著录”的说明

客观著录：指以原件内容的客观含义为准，不附加著录人员的主观认识。

照实著录：指按原件实际文字照搬著录。

3) 关于排序

由于老旧期刊的期（或卷）及原始页码不能准确实现检索应用中的期（或卷）排序及每期（或卷）文章的排序，加工时，同种期刊的每一册期刊之间及期刊中每一条正文、广告和图片元数据都需要生成对应的排序信息，以实现数据检索应用过程中的各种排序功能。

4) 关于元数据及图像的唯一标识及对应关系

正文唯一标识、广告唯一标识、图片唯一标识、数字化文本链接标识等由著录系统自动生成，且唯一标识一旦生成则不可改变，命名规则由招标方确定。元数据及图像之间的对应关系一定要满足数据的检索应用要求。

5) 关于类别著录

正文的类别、广告的类别以及图片的类型均根据事先约定的分类体系进行类别赋予。

6) 关于著录中的符号使用的说明

著录中使用的所有标点、符号与招标方协商确认。

三、验收质量标准

1. 著录规则质量标准：著录规则错误率 $\leq 0.3\%$ 。
2. 著录文字标准：著录文字错误率 $\leq 0.05\%$ 。
3. 元数据与图像文件的勾链质量标准：元数据与图像文件的勾链无差错
4. 版面内容质量标准：版面内容无差错
5. 篇名类型著录质量标准：篇名类型无差错
6. 其他元数据的合法性验收标准

1) 返回图像和 PDF 的页数必须和著录的页数完全相同；

2) 返回图像元数据著录项中非空元素著录项无缺省。

非空元素是指《元数据著录规则》中标*的元素，著录项不可缺省。

四、数据规范格式

投标方需根据招标方提供的数据格式进行此项目制作。

投标方须提供的数据有：期刊的整本 PDF、单篇 PDF、图像文件（著录图片、著录广告图片）、元数据著录信息文件等。具体规范详见附件。

五、质保服务

投标方在总项目完成后提供一年的免费质保服务。双方在合同签订后另行签署质保协议书。

附件一：元数据著录规则

一、前言

1949-1977 年期刊篇名数据库元数据的著录分三个部分：

- 正文内容元数据
- 广告内容元数据
- 图片内容元数据

著录语种：期刊的元数据著录项均依据制作期刊内容照实进行著录。

二、术语定义

下列术语和定义适用于本附件。

1. 元数据著录单元

指期刊图像中有相对明显物理边界内的内容完整的正文、广告和图片。下文中简称“著录单元”。

2. 正文

期刊中除广告及独立图片外的所有内容。

3. 广告

期刊中企业或个人用于推销商品、服务，或向社会传递信息，影响 Yu Lun 的传播内容。

4. 图片

包括独立图片和文中附图。独立图片为期刊中独立于文章之外的图片；文中附图为期刊正文中出现的有标题或附有文字说明的图片。

三、正文内容核心元数据及著录信息源

最终著录规格以实测样本为标准。元数据著录规则应该在正式生产前进行讨论，使其完全符合生产流程，同时内容适合于历史背景。

1. 正文内容核心元数据及著录信息源

序号	正文著录项目	正文著录信息源
1	类别	可采用事先约定的分类体系对正文赋予类别
2	栏目	在期刊中有明确标识，将内容相近或有某种联系的内容组合在一起
3	子栏目 1	子栏目 1 从属于栏目
4	子栏目 2	子栏目 2 从属于子栏目 1
5	标题 1*	用来著录的正文标题
6	标题 2	同标题 1
7	标题 3	同标题 1
8	责任者 1	正文的第一责任者
9	责任者 2	正文的第二责任者
10	责任者 3	正文的第三责任者
11	著作方式 1	第 1 责任者对应的著作方式，如著、翻译、评论、口述、笔录等
12	著作方式 2	第 2 责任者对应的著作方式，如著、翻译、评论、口述、笔录等
13	著作方式 3	第 3 责任者对应的著作方式，如著、翻译、评论、口述、笔录等
14	单位 1	正文第一责任者所属工作单位名称
15	单位 2	正文第二责任者所属工作单位名称
16	单位 3	正文第三责任者所属工作单位名称
17	摘要*	正文有摘要则照实著录，无摘要则仅著录正文前 100 个字
18	期刊名称*	正文所属期刊的名称

序号	正文著录项目	正文著录信息源
19	出版年份*	正文所属期刊的出版年份
20	卷	正文所属期刊的卷号
21	期	正文所属期刊的期号
22	号	正文所属期刊的“总期号”、“特刊号”等
23	页码	正文在期刊中的页码
24	起始扫描页码*	该著录单元所在起始页原始图像的命名内容, 为 4 位 Arabic 数字
25	篇排序号*	根据期刊排版等生成该著录单元对应的排序信息, 以实现数据检索应用过程中的各种排序功能
26	连载信息	正文的转接信息, 如“未完”、“未完待续”等
27	中外文对照	正文其他语种和中文同时存在, 并相互对应时, 如著录“中英文对照”
28	图表	著录“附图”、“附表”、“附图表”
29	关键词	表达正文主题内容的词
30	备注	
31	备用 1	招标方根据加工过程中的内容情况另行确定
32	备用 2	招标方根据加工过程中的内容情况另行确定
33	备用 3	招标方根据加工过程中的内容情况另行确定

注：1) 标注*的字段为必著项；

2) 卷、期、号三者不能同时为空；

3) 类别著录项投标方保留空字段，著录内容由招标方完成。

2. 著录项定义及注释

1) 类别

定义：正文著录单元所属的主题类目

注释：属主观性著录项，需要通过正文的内容进行判断后给出相应的类别，具体可选类别事先设定供著录时选择，根据制作内容情况招标方有权对类别进行调整。

2) 栏目

定义：内容相近或有某种联系的几条正文著录单元编排在一起，有客观的名称且有相对明显的物理边界，这个名称即为正文的栏目。

注释：属客观性著录项，按照正文所属栏目客观著录。

3) 子栏目

定义：属于某一个栏目，其下可能再细分子栏目

注释：属客观性著录项，按照正文所属栏目客观著录。

4) 标题

定义：正文著录单元的著录题名

注释：属客观性著录项，正文中有明确题名时，直接著录原始题名。无明确题名时，著录要求见著录细则。

5) 责任者

定义：创建正文著录单元内容的主要责任者

注释：本元素属客观性著录项，是对文献内容负主要责任的个人或团体，本项目最多著录三个责任者。

6) 著作方式

定义：责任者创建正文著录单元时参与的方式

注释：本元素属客观性著录项，是对文献内容负主要责任的个人或团体参与创建的方式，本项目最多著录三个著作方式，与责任者对应。

7) 单位

定义：责任者所属的工作单位名称

注释：本元素属客观性著录项，本项目最多著录三个单位名称，与责任者对应。

8) 摘要

定义：正文著录单元内容的概述

注释：本元素属客观性著录项，原文有摘要则直接照实著录；原文无摘要则取正文前 100 个字（包括标点符号）作为摘要，并在最后加注省略号（省略号不包括在 100 个字内）。

9) 期刊名称

定义：正文著录单元内容所属期刊的名称

注释：客观性著录元素。

10) 出版年份

定义：正文著录单元所属期刊的出版年份

注释：客观性著录元素，采用公元纪年著录，Arabic 数字表示，不能辨别年份的用大写 X 替代相应数字，例如 192X。

11) 卷

定义：正文著录单元所属每期期刊的实际卷号

注释：本元素属客观性著录项，采用 Arabic 数字表示。

12) 期

定义：正文著录单元所属每期期刊的实际期号

注释：本元素属客观性著录项，采用 Arabic 数字表示；若是合期，用 “/” 连接。

13) 号

定义：正文著录单元所属期刊的总期号或特刊号等

注释：本元素属客观性著录项，增刊、特刊、总期号等信息都著录于此字段。

14) 页码

定义：正文著录单元在所属期刊中的实际页码

注释：本元素属客观性著录项，著录起始及转接页码

15) 起始扫描页码

定义：该著录单元所在起始页原始图像的命名内容。

注释：本元素属客观性著录项，为 4 位 Arabic 数字。

16) 篇排序号

定义：根据期刊排版等生成该著录单元对应的排序信息，以实现数据检索应用过程中的各种排序功能

注释：本元素著录 4 位 Arabic 数字。

17) 连载信息

定义：正文著录单元内容的连载信息

注释：本元素属客观性著录项，正文在不同期的期刊中有连载时，根据原件的连载状态照实著录，如“未完”、“未完待续”等。

18) 中外文对照

定义：正文著录单元其他语种和中文同时存在，并相互对应

注释：本元素属主观性著录项，著录如“中英文对照”、“中法文对照”等。

19) 图表

定义：文中附图、附表的状态描述

注释：本元素属主观性著录项，著录如“附图”、“附表”、“附图表”

20) 关键词

定义：用于表达正文著录单元主题内容的词

注释：本元素属客观性著录项，原件中有则照实著录，无则不必著录。

四、广告内容核心元数据及著录信息源

1. 广告著录项目及著录信息源

期刊中的广告元数据著录，每个广告著录单元的著录项包括广告类别、广告栏目、广告标题、广告语、广告产品、广告发布者等。

序号	广告著录项目	广告著录信息源
1	广告类别*	可采用事先约定的分类系统进行类别赋予，如启事公告、书讯、医药、百货零售、金融保险、食品烟酒、其他
2	广告栏目	内容相近或有某种联系的几条广告编排在一起，有客观的栏目名称
3	广告标题	广告中对广告内容进行概括，起引导作用的语句
4	广告语	广告中长期、反复使用的简短口号
5	广告产品	广告宣传的具体产品名称
6	广告发布者	发布广告的机构或个人
7	期刊名称*	广告所属期刊的名称
8	出版年份*	广告所属期刊的出版年份
9	卷	广告所属期刊的卷号
10	期	广告所属期刊的期号
11	号	广告所属期刊的“总期号”、“特刊号”
12	页码	广告所在的页码

序号	广告著录项目	广告著录信息源
13	起始扫描页码*	同正文
14	篇排序号*	根据期刊排版等生成该著录单元对应的排序信息，以实现数据检索应用过程中的各种排序功能
15	备注	
16	备用 1	招标方根据加工过程中的内容情况另行确定
17	备用 2	招标方根据加工过程中的内容情况另行确定
18	备用 3	招标方根据加工过程中的内容情况另行确定

注：1) 标注*的字段为必著项；

2) 卷、期、号三者不能同时为空；

3) 广告标题、广告发布者和广告产品不能同时为空。

2. 广告著录项定义及注释

广告著录单元的期刊名称、出版年份、卷、期、号及页码著录项的定义和注释同正文。下面只列出广告类别、广告栏目、广告标题、广告语、广告产品、广告发布者的定义和注释。

1) 广告类别

定义：广告著录单元所属的主题类目

注释：属主观性著录项，需要通过對著录单元的广告内容进行判断后给出相应的类别，具体可选类别事先设定供著录时选择。

2) 广告栏目

定义：期刊中内容相近或有某种联系的几条广告著录单元编排在一起，有客观的名称且有相对明显的物理边界，这个名称即为广告的栏目。

注释：属客观性著录项，著录时原则上按照广告所属栏目进行客观著录。

3) 广告标题

定义：广告著录单元中对广告内容进行概括，起引导作用的语句。

注释：属客观性著录项。广告标题在具体某一则广告中使用，与广告正文内容同时出现，区别于广告语。

4) 广告语

定义：广告著录单元中长期、反复使用的简短口号。

注释：属于客观性著录项。广告语具有稳定性和连贯性，可在多个广告或系列广告中使用，区别于广告标题。

5) 广告产品

定义：广告宣传的具体产品名称。

注释：属客观性著录项。

6) 广告发布者

定义：广告著录单元的发布者。

注释：属客观性著录项，指发布广告的机构或个人。

五、图片内容元数据及著录信息源

1. 图片著录项目及著录信息源

著录图片包括独立图片和文中附图，文中存在的没有标题或说明的图片不著。图片著录单元对应的图片责任者、图片著作方式、期刊名称、出版年份、卷、期、号及页码的元素著录同正文。

序号	图片著录项目	图片著录信息源
1	图片栏目	有明确标识，将内容相近或有某种联系的图片组合在一起，形成图片栏目
2	图片标题*	独立图片、文中附图的标题
3	图片说明	独立图片、文中附图的具体文字说明内容
4	图片责任者 1	完成图片内容创建的第一责任者
5	图片责任者 2	完成图片内容创建的第二责任者
6	图片著作方式 1	第一责任者对应的著作方式，如画、绘、摄等
7	图片著作方式 2	第二责任者对应的著作方式，如画、绘、摄等
8	期刊名称*	图片所属期刊的名称
9	出版年份*	图片所属期刊的出版年份
10	卷	图片所属期刊的卷号

序号	图片著录项目	图片著录信息源
11	期	图片所属期刊的期号
12	号	图片所属期刊的“总期号”、“特刊号”
13	页码	图片在原件中的页码
14	起始扫描页码*	同正文
15	篇排序号*	根据期刊排版等生成该著录单元对应的排序信息，以实现数据检索应用过程中的各种排序功能
16	图片类型 1*	漫画、照片、书法、题字、题词、手稿、木刻、篆刻、石刻、雕塑、曲谱、歌谱、表格、图表、画图、地图、拓本、拓片、其他
17	图片类型 2	仅当图片为完整的封面时著录
18	备注	
19	备用 1	招标方根据加工过程中的内容情况另行确定
20	备用 2	招标方根据加工过程中的内容情况另行确定
21	备用 3	招标方根据加工过程中的内容情况另行确定

注：1) 标注*的字段为必著项；

2) 卷、期、号三者不能同时为空；

3) “图片类型”由招标方事先设定。

2. 图片著录项定义及注释

1) 图片栏目

定义：期刊中，内容相近或有某种联系的图片著录单元编排在一起，有客观的名称，这个名称即为图片的栏目。

注释：属客观性著录项，著录时按照图片所属栏目客观著录。

2) 图片标题

定义：独立图片、文中附图的著录标题

注释：属客观性著录项，按原件照实著录。

3) 图片说明

定义：相对于著录图片而言，除了图片标题以外的其它说明图片的文字

注释：属客观性著录项，按原件照实著录。

4) 图片类型 1

定义：描述图片著录单元的类型

注释：属主观性著录项，需要通过对著录单元的图片类型进行判断后给出选择，具体可选图片类型事先设定供著录时选择。

5) 图片类型 2

定义：描述图片著录单元的类型

注释：仅当图片为完整的“封面”时著录。该字段著录内容为“封面”。

六、元数据元素著录说明

1. 关于元数据元素著录

上述各表中标*的元素均为非空字段，著录项不可缺省；未标*的元素视具体内容具体分析，有内容则必须著录；

2. 关于元数据著录注释中“客观著录”与“照实著录”的说明

客观著录：指以原件内容的客观含义为准，不附加著录人员的主观认识。

照实著录：指按原件实际文字照搬著录。

3. 关于排序

由于老旧期刊的期（或卷）及原始页码不能准确实现检索应用中的期（或卷）排序及每期（或卷）文章的排序，加工时，同种期刊的每一册期刊之间及期刊中每一条正文、广告和图片元数据都需要生成对应的排序信息，以实现数据检索应用过程中的各种排序功能。

4. 关于元数据及图像的唯一标识及对应关系

正文唯一标识、广告唯一标识、图片唯一标识、数字化文本链接标识等由著录系统自动生成，且唯一标识一旦生成则不可改变，命名规则由招标方确定。元数据及图像之间的对应关系一定要满足数据的检索应用要求。

5. 关于类别著录

正文的类别、广告的类别以及图片的类型均根据事先约定的分类体系进行类别赋予。

6. 关于著录中的符号使用的说明

著录中使用的所有标点、符号与招标方协商确认。

注：本附件解释权归招标方所有，著录过程中本规则未能详尽的内容由双方协商确定。

附件二：数据质量标准及数据格式规范

一、元数据著录质量标准

1. 著录规则质量标准：著录规则错误率 $\leq 0.3\%$

元数据著录规则错误率是元数据著录项错误的数量与需著录项的数据总量比例，计算公式为：元数据著录规则错误率 = (元数据著录项错误的数量 ÷ 需著录项的数据总量) × 100% (单位：项)。

著录规则错误是按项为计算单位，凡未能对元数据项进行客观著录的情况均计为著录规则错误。一篇元数据项为多值的，则按值的数量计算项数。

2. 著录文字质量标准：著录文字错误率 $\leq 0.05\%$

著录文字错误率是元数据项内容文字错误的数量与元数据项文字总量的比例，计算公式为：著录文字错误率 = (元数据项内容文字错误的数量 ÷ 元数据项文字总量) × 100% (单位：字)。

著录文字错误按文字计算，元数据中文字与原文不符的计为著录文字错误，但不包括对原文的文字错误进行纠正的情况。

3. 元数据与图像文件的勾链质量标准：元数据与图像文件的勾链无差错

4. 版面内容质量标准：版面内容无差错

5. 篇名类型著录质量标准：篇名类型无差错

6. 其他元数据的合法性验收标准

1) 返回图像和 PDF 的页数必须和著录的页数完全相同；

2) 返回图像元数据著录项中非空元素著录项无缺省。

非空元素是指附件《元数据著录规则》中标*的元素，著录项不可缺省。

二、数据规范格式

根据本项目制作要求，投标方须提供的数据有：期刊的整本 PDF、单篇 PDF、图像文件（著录图片、著录广告图片）、元数据著录信息文件等，具体规范要求如下：

1. PDF 文件

每本期刊的 PDF 数据为二套。一套为整本 PDF，一套为单篇 PDF，质量要求如下：

1) PDF 均为双层 PDF 文件，表层图像能与 OCR 的文字对应且能满足现代行文检索要求；

2) PDF 版本通用性，需要符合 PDF 国际标准。投标方提交成品时，需向招标方说明使用的 PDF 版本；

- 3) PDF 像素、大小无异常, 灰度或色彩无明显漏洞;
- 4) 在保证图像清晰的情况下, 对图像层进行压缩, 根据图像质量和大小采用适当的压缩率, 压缩后大小控制在 200K 左右 (特殊画幅、质量, 原始图像容量特别大的除外), 生产时以样本为准;
- 5) 如有必要, 需对招标方提供的原图进行修正, 以保证产出的 PDF 文件画面清晰、无扭曲、倾斜、变形等影响查看的情况发生;
- 6) 除特殊规定外, PDF 内容的完整性和准确性要和招标方提供的原图保持一致;
- 7) 完全相同的两篇单篇 PDF 只需保存一篇。

2. 图像文件

单页图像, 在原始图上采取合适的压缩率进行压缩, 分辨率分别为 150DPI 和 100DPI, 保存为 JPG 格式, 大小控制在 100K 左右, 生产以样本为准。

抠图图像为著录过程中从原始图像中抠出的图片、广告图像, 尺寸视实际图像大小而定。抠出的图片、广告等图像务必要保持原始图像的完整性和整洁性, 遇到排版不规则、交叉等无法规则抠出等情况, 按照尽量保持原始的整体性原则进行处理。

图片、广告均需保存级别和应用级别两套数据。保存级别的图像应保持原有的图片格式, 不进行压缩, 分辨率为 300DPI。应用图像应为 JPG 格式, 在原始图上采取合适的压缩率进行压缩, 分辨率为 150DPI, 大小控制在 100K 左右, 实际大小可根据原始抠图的大小而变, 生产以样本为准。

3. 元数据著录信息文件

元数据著录信息文件是存储所有元数据信息的可读可写文件, 应包含按照附件一《元数据著录规则》著录单篇文章的元数据, 以及对应的 PDF, 图片、广告图片路径。具体著录要求由招标方制定。

4. 文件命名及存储规范

文件命名及存储规范需与招标方协商后确定, 要求规则统一, 名称唯一。

5. 其他要求

- 1) 投标方所有著录的内容, 除招标方有规定外, 需向招标方说明著录时使用的字库。
- 2) 投标方根据招标方提供的图像, 按照招标文件要求进行样本制作, 并提供相应的匹配说明文件, 方便招标方进行判断。

项目 4：会议论文篇名元数据制作

项目数据量：40,000 篇

项目完成时间：2025 年 10 月 31 日

项目进度要求：

第一阶段：合同签署后，在双方确认技术标准后，至 2025 年 6 月 30 日前至少完成 20,000 篇

第二阶段：至 2025 年 10 月 31 日前，完成 40,000 篇

一、招标样本

- 由招标方提供招标所需样本图像，共计约 27 页
- 投标方依据招标方提供的著录标准及各项规范（见下附件）制作招标样本
- 各投标方在将样品提交前，请先在样本上标识所投项目、投标方等信息，并确认标识是否完好
- 投标方需在投标日当天出示招标样本，并做阐述
- 投标方提交的样本将作为评估中标与否的因素之一
- 最终解释权归招标方所有

二、元数据主要著录内容要求

1. 正文内容元数据：14 项≤著录项≤38 项
2. 元数据制作成品：单篇元数据著录信息文件、单篇 PDF、整本 PDF。
3. 元数据著录要求：依据文献的相关信息和内容客观著录。
4. 著录语种：本项目仅著录中文和英文的正文内容，其它语种内容不予著录。
5. 元数据元素主要著录说明

1) 关于字段非空

“正文著录项目及著录信息源”表中标有“*”的元素为非空字段，著录项不可缺省；未标*的元素视具体内容具体分析，有内容则必须著录。

2) 关于不要求著录的字段

- (1) “分类号”字段不要求著录；

(2) 对于所有不要求著录的字段都应保留。

3) 关于元数据与图像的对应关系

元数据和 PDF 图像之间保持正确的对应关系, 以满足数据的检索应用要求。

4) 关于元数据的篇 id

元数据需有全局唯一标识, 且该标识一旦生成则不可改变, 命名规则由招标方确定。

5) 关于分类号著录

分类均根据事先约定的分类体系进行分类号赋予。

6) 关于著录中的符号使用的说明

著录中使用的所有标点、符号与招标方协商确认。

三、验收质量标准

1. 著录规则质量标准: 著录规则错误率 $\leq 0.3\%$ 。
2. 著录文字标准: 著录文字错误率 $\leq 0.05\%$ 。
3. 元数据与图像文件的勾链质量标准: 元数据与图像文件的勾链无差错。
4. 整本无漏著篇。

四、数据规范格式

投标方需根据招标方提供的数据格式进行此项目制作。

投标方须提供的数据有: 单篇 PDF 文件、整本 PDF 文件、元数据著录信息文件等。具体规范详见附件。

五、质保服务

投标方在总项目完成后提供一年的免费质保服务。双方在合同签订后另行签署质保协议书。

附件一: 元数据著录规则

一、前言

学术会议论文元数据的著录仅针对学术会议论文集正文内容。

著录语种: 本项目仅著录中文和英文的正文内容, 其它语种内容不予著录。

二、术语定义

下列术语和定义适用于本附件。

1. 元数据著录单元

指论文集图像中有相对明显物理边界内的内容相对完整的会议论文正文。下文中简称“著录单元”。

2. 正文

指论文集中符合选收范围的会议论文文章。

三、正文内容核心元数据及著录信息源

本部分内容的著录信息源自与会议及论文集相对应的文章。

1. 正文著录项目及著录信息源

序号	正文著录项目	正文著录信息源
1	分类号 1	采用事先约定的分类体系赋予分类号
2	分类号 2	采用事先约定的分类体系赋予分类号
3	分类号 3	采用事先约定的分类体系赋予分类号
4	题名*	正文的标题
5	题名译名	正文题名对应的外文标题
6	责任者 1	正文的第 1 个人责任者
7	责任者外文名 1	正文的第 1 个人责任者的外文名
8	责任者 2	正文的第 2 个人责任者
9	责任者外文名 2	正文的第 2 个人责任者的外文名
10	责任者 3	正文的第 3 个人责任者
11	责任者外文名 3	正文的第 3 个人责任者的外文名
12	责任者著作方式 1	责任者 1 的著作方式
13	责任者著作方式 2	责任者 2 的著作方式
14	责任者著作方式 3	责任者 3 的著作方式

序号	正文著录项目	正文著录信息源
15	单位 1	责任者 1 所属单位名称
16	单位 2	责任者 2 所属单位名称
17	单位 3	责任者 3 所属单位名称
18	团体责任者 1	正文的第 1 团体责任者
19	团体责任者 2	正文的第 2 团体责任者
20	团体责任者 3	正文的第 3 团体责任者
21	团体责任者著作方式 1	团体责任者 1 的著作方式
22	团体责任者著作方式 2	团体责任者 2 的著作方式
23	团体责任者著作方式 3	团体责任者 3 的著作方式
24	关键词	正文上著者赋予的关键词
25	关键词译文	正文上著者赋予关键词的外文翻译
26	摘要*	正文有摘要则照实著录, 无摘要则仅著录正文前 200 个汉字/外文单词
27	摘要译文	摘要的外文翻译
28	基金项目	正文内容研究所受支持的基金项目
29	页码*	正文在所属论文集集中的页码
30	PDF 链接*	正文 PDF 的链接
31	扫描页码*	正文所属的扫描页码
32	捆号*	每捆论文集的编号
33	论文集名称*	正文所属的论文集名称
34	会议名称*	正文所属的会议名称
35	索取号*	正文所属论文集的索取号
36	索书号	正文所属论文集的索书号

序号	正文著录项目	正文著录信息源
37	论文集 id 号*	正文所属论文集的 id 号
38	会议 id 号*	正文所属会议的 id 号
39	会议年份*	会议召开的年份
40	备注	
41	备用 1	招标方根据加工过程中的内容情况另行确定
42	备用 2	招标方根据加工过程中的内容情况另行确定
43	备用 3	招标方根据加工过程中的内容情况另行确定
44	篇 id*	正文的唯一标识符
45	篇排序号*	论文集正文的排版顺序

注：1) 个人责任者和团体责任者不能同时著录

2) 32-39 字段在篇名数据加工中属于自动生成的著录项

3) 【分类号 1】、【分类号 2】、【分类号 3】著录项投标方保留空字段，著录内容由招标方完成。

2. 著录项定义及注释

1) 分类号(该项暂不著录)

定义：著录单元所属的主题类目

注释：属主观性著录项，依据《中图法》第五版，通过对著录单元的内容进行判断后给出相应的中图法分类号，本项目最多著录三个分类号。

2) 题名*

定义：著录单元的著录标题

注释：属客观性著录项，按照正文照实著录。

3) 题名译名

定义：题名对应的外文标题

注释：属客观性著录项，正文中有则照实著录，没有则无需著录。

4) 责任者

定义：创建著录单元内容的主要个人责任者

注释：本元素属客观性著录项，是对文献内容负主要责任的个人，本项目最多著录三位个人责任者。

5) 责任者外文名

定义：主要个人责任者对应的外文名

注释：属客观性著录项，正文中有则照实著录，没有则无需著录。

6) 责任者著作方式

定义：个人责任者创建著录单元时参与的方式

注释：本元素属客观性著录项，是对文献内容负主要责任的个人参与创建的方式，本项目最多著录三个著作方式，与个人责任者对应。如著、翻译、评论、口述、笔录等表示责任者著作方式的词等。

7) 单位

定义：个人责任者所在机构的名称

注释：本元素属客观性著录项，本项目的单位名称，与个人责任者对应，同一个责任者对应的多个单位，仅著录第一个。

8) 团体责任者

定义：创建著录单元内容的主要团体责任者

注释：本元素属客观性著录项，是对文献内容负主要责任的团体，本项目最多著录三个团体责任者。

9) 团体责任者著作方式

定义：团体责任者创建著录单元时参与的方式

注释：属客观性著录项，是对文献内容负主要责任的团体参与创建的方式，本项目最多著录三个著作方式，与团体责任者对应。如著、翻译、评论、口述、笔录等表示责任者著作方式的词等。

10) 关键词

定义：用于表达著录单元主题内容的词

注释：属客观性著录项，正文中有则著录，无则不予著录；多个关键词以“|”间隔。

11) 关键词译文

定义：关键词对应的外文内容

注释：属客观性著录项，正文中有则著录，无则不予著录；多个关键词翻译以“|”间隔。

12) 摘要*

定义：著录单元内容的概述

注释：本元素属客观性著录项，原文有摘要则直接照实著录；原文无摘要则取正文前 200 个汉字（包括标点符号）/200 个外文单词作为摘要，并在最后加注省略号（省略号不包括在 200 个字内）。

13) 摘要译文

定义：摘要对应的外文内容

注释：本元素属客观性著录项，原文有摘要译文则直接照实著录，原文无摘要译文则不需著录。

14) 基金项目

定义：著录单元内容研究所获得支持的基金项目名称

注释：本元素属客观性著录项，正文中有则按照招标方提供的规则著录；多个基金项目以“|”间隔。

15) 页码*

定义：著录单元所在的实际页码

注释：本元素属客观性著录项，著录起止及转接页码。

16) PDF 链接*

定义：著录单元对应的 PDF 的链接

注释：正文对应的图像 PDF 文件的存储路径，对应单篇 PDF 链接和整本 PDF 链接。

17) 扫描页码*

定义：著录单元所属的扫描页码

注释：将正文所属扫描页码分别列出。

18) 捆号*

定义：著录单元对应的那捆论文集的编号

注释：论文所属论文集归属的捆号，招标方提供。

19) 索取号*

定义：著录单元对应的论文集的流水号

注释：论文所属论文集的索取号，招标方提供。

20) 索书号

定义：著录单元对应的论文集的索书号

注释：论文所属论文集的索书号，招标方提供；目前保留字段即可，无著录内容。

21) 会议年份*

定义：学术会议召开的年份

注释：论文所属论文集对应的学术会议召开的年份，招标方提供。

22) 篇 id*

定义：正文的唯一标识符

注释：篇 id 为单篇元数据在一册论文集内的生产序号，是定义一篇论文的唯一标识，在生产阶段，一旦生成则不可改变。命名规则：会议 id+论文集 id+四位 Arabic 数字（每册从 0001 顺延），会议 id 和论文集 id 均为 7 位，不足部分前面分别加零补足。

注意：返工时，若发生一篇拆分为多篇，则新拆分出的篇 id 在此册内最后一个篇 id 后顺延；若发生删篇，则其他篇的篇 id 不受影响；若发生合篇，则主合篇的篇 id 保留，被合篇消失，其他篇的篇 id 不受影响。

23) 篇排序号*

定义：论文集正文的排版顺序。

注释：本元素属客观性著录项，为该篇在一册论文集中所处的排版顺序，用 Arabic 数字表示，从 1 顺延，增、删篇后，排序号需重新在一册论文集内按篇的排版顺序进行排序。

四、元数据元素著录说明

1. 关于字段非空

“正文著录项目及著录信息源”表中标有“*”的元素为非空字段，著录项不可缺省；未标*的元素视具体内容具体分析，有内容则必须著录。

2. 关于不要求著录的字段

- 1) “分类号”字段不要求著录；
- 2) 对于所有不要求著录的字段都应保留。

3. 关于元数据与图像的对应关系

元数据和 PDF 图像之间保持正确的对应关系，以满足数据的检索应用要求。

4. 关于元数据的篇 id

元数据需有全局唯一标识，且该标识一旦生成则不可改变，命名规则由招标方确定。

5. 关于分类号著录

分类均根据事先约定的分类体系进行分类号赋予。

6. 关于著录中的符号使用的说明

著录中使用的所有标点、符号与招标方协商确认。

注：本附件解释权归招标方所有，著录过程中本规则未能详尽的内容由双方协商确定。

附件二：数据质量标准及数据格式规范

一、 数据质量标准

1. 元数据著录要求

依据文献的相关信息和内容客观著录。

2. 元数据制作内容

本项目制作针对的是学术会议论文的各元素，著录的元数据元素包括：题名、题名译文、责任者、责任者著作方式、团体责任者、团体责任者著作方式、单位、关键词、关键词译文、摘要、摘要译文、基金、页码等；以及单篇元数据的 MDB 文件，对应的单篇及整本论文集 PDF 文件。

3. 元数据著录质量标准

由于本合同规定的数据验收方式为分批验收，故本质量标准所指的错误率计算均以批为计量单位。

1) 著录规则质量标准：著录规则错误率 $\leq 0.3\%$

(1) 元数据著录规则错误率是元数据著录项错误的数量与需著录项的数据总量比例, 计算公式为：元数据著录规则错误率=（元数据著录项错误的数量 $\div$ 需著录项的数据总量） $\times 100\%$ （单位：项）。

(2) 著录规则错误是按项为计算单位，凡未按照著录规则进行著录的情况均计为著录规则错误。一篇元数据有多个著录项，则按著录项的数量计算项数。如一篇元数据有 10 个著录项，则计为 10 个项。

2) 著录文字质量标准：著录文字错误率 $\leq 0.05\%$

(1) 著录文字错误率是元数据项内容文字错误的数量与元数据项文字总量的比例，

计算公式为：著录文字错误率=（元数据项内容文字错误的数量÷元数据项文字总量）×100%（单位：字）。

（2）著录文字错误按文字计算，元数据中文字与原文不符的计为著录文字错误，但不包括对原文的文字错误进行纠正的情况。

3）元数据与图像的勾链质量标准：元数据与图像的勾链无差错

（1）元数据与图像勾链错误率是元数据与图像勾链错误的篇数与元数据总篇数的比例。

（2）计算公式为：元数据与图像勾链错误率=（元数据与图像勾链错误的篇数÷元数据总篇数）×100%（单位：篇）。

4）整本无漏著篇；元数据著录项中“非空元素”著录项均无缺省。

二、 数据格式规范

根据本项目制作要求, 投标方须提供的数据有：单篇 PDF 文件、整本 PDF 文件、元数据著录信息文件。具体规范要求如下：

1. PDF 文件

每本论文集的 PDF 数据为两套，一套为单篇 PDF，一套为整本 PDF。质量要求如下：

- 1）PDF 均为双层 PDF 文件，底层为文字层，表层为图像层，图像层能与文字层精确对应。
- 2）PDF 版本通用性，需要符合 PDF 国际标准。投标方提交成品时，需向招标方说明使用的 PDF 版本。
- 3）PDF 像素、大小无异常，灰度或色彩无明显漏洞。
- 4）在保证图像清晰的情况下，对原始图像进行压缩，根据图像质量和大小采用适当的压缩率，压缩后单页 PDF 大小控制在 150K 左右（特殊画幅、质量，原始图像容量特别大的除外），生产时以样本为准。
- 5）如有必要，需对招标方提供的原图像进行修正，以保证产出的 PDF 文件画面清晰、无扭曲、倾斜、变形等影响查看的情况发生。
- 6）除特殊规定外，PDF 内容的完整性和准确性要和招标方提供的原图像保持一致。
- 7）单篇 PDF 无重复性。当多篇所对应的 PDF 完全相同的情况下，指向第一篇的 PDF，不用输出多个重复内容的 PDF。

2. 元数据著录信息文件

元数据著录信息文件是存储所有元数据信息的可读可写文件，应按照附件二元数据著录规则著录单篇文章的元数据，及对应的单篇、整本 PDF 路径。

元数据以 MDB 格式提交，包含著录数据各项信息。元数据与 PDF 数据须同步提交。

3. 命名规则

文件的命名需与招标方协商后确定，要求具有可读性，规则统一，名称唯一。建议投标方提供的文件及文件夹要求如下：

1) 文件夹层次

- 第一层：批次，统一命名规范为“资源来源-项目简称-资源形态-批次”。其中：资源来源、项目简称、资源形态都由招标方的生产平台提供，为英文字母缩写；批次用 4 位 Arabic 数字表示(0001-9999)，不足 4 位的在前面补“0”，举例：shlib-hylw18-tx-0001；
*样本批次命名从 9999 依次递减表示，样本批举例 shlib-hylw18-tx-9999。
- 第二层：交付的文件类型，分别为 MDB 和 PDF 两个目录；
- 第三层：hylw-捆号-索取号-会议年份，会议年份见招标方提供每册论文集扫描图像文件夹名称的末尾年份，例如：hylw-20160001-0416518-2014
- 第四层：目录底层均按类型存放数据文件

2) 文件命名规则

- 文件建议放于相应的第三层目录下；
- 招标方将提供原始图像的命名规则说明（见附件四），投标方制作时 PDF 的命名需参照招标方提供图像规则进行；
- MDB 命名规则：“hylw-捆号-索取号-会议年份.mdb”
- PDF 命名规则

(1) 单篇 PDF 命名为“篇 id_篇起始扫描页码.pdf”，“篇起始扫描页码”指每篇文章起始的扫描页码编号中的 Arabic 数字；

(2) 整本 PDF 命名为“All.pdf”。

4. 整本 PDF 排序方式

整本 PDF 输出排序方式为：以扫描页码排序。

5. 其他要求

投标方应根据招标方提供的图像，按照招标文件要求进行样本制作，并提供相应的匹配说明文件，方便招标方进行判断。以上格式规范如需调整，可由双方在样本制作时协商。

双方认可的数据样例作为本项目著录及数据规范的参照依据。

项目 5：学术图片库元数据制作

数据量：40,000 篇

项目完成时间：2025 年 10 月 31 日

项目验收进度要求：

第一阶段：合同签署后，在双方确认技术标准后，到 2025 年 6 月 30 日前至少完成 20,000 篇

第二阶段：至 2025 年 10 月 31 日前，完成 40,000 篇

一、招标样本

- 由招标方提供招标所需样本图像光盘，共计约 24 页
- 投标方依据招标方提供的著录标准及各项规范（见下附件）制作招标样本
- 各投标方在将样品提交前，请先在样本上标识所投项目、投标方等信息，并确认标识是否完好
- 投标方需在投标日当天出示招标样本
- 投标方提交的样本将作为评估中标与否的因素之一
- 最终解释权归招标方所有

二、元数据主要著录内容要求

1. 学术图片库元数据的著录

图片元数据：11 项 ≤ 著录项 ≤ 21 项

2. 元数据元素主要著录说明

1) 关于元数据元素著录

各表中标*的元素均为非空字段，著录项不可缺省；未标*的元素视具体内容具体分析，有内容则必须著录。

2) 关于元数据著录注释中“客观著录”与“照实著录”的说明

客观著录：指以原件内容的客观含义为准，不附加著录人员的主观认识。

照实著录：指按原件实际文字照搬著录。

3) 关于元数据及图像的唯一标识及对应关系

图片唯一标识由著录系统自动生成，且唯一标识一旦生成则不可改变，命名规则由招标方确定。元数据及图像的对应关系一定要满足数据的检索应用要求。

4) 关于图片类型著录

图片的类型根据事先约定的分类体系进行类别赋予。

5) 关于著录中的符号使用的说明

著录中使用的所有标点、符号与招标方协商确认。

6) 关于元数据著录与招标方先前元数据的关系

招标方提供先前已有的篇名元数据信息供投标方加工时参考，需明确的是投标方著录时应按本合同的元数据著录规则进行著录。招标方在后续对投标方生产的数据进行校验时按本合同的元数据规则 and 标准进行验收，先前元数据信息仅供参考，但投标方需提供先前篇名与本项目篇名元数据的详细对应关系表供招标方在后续工作中使用，篇名对应表的具体要求由招标方制定。

三、验收质量标准

1. 著录规则质量标准：著录规则错误率 $\leq 0.3\%$
2. 著录文字质量标准：著录文字错误率 $\leq 0.05\%$
3. 元数据与图像的勾链质量标准：元数据与图像的勾链无差错
4. 版面内容质量标准：版面内容无差错
5. 其他元数据与图像的合法性验收标准
 - 1) 图像的分辨率与图像大小需符合招标方技术要求（详见附件二）。
 - 2) 元数据著录项中非空元素著录项无缺省。非空元素是指附件一《元数据著录规则》中标*的元素，著录项不可缺省。

四、数据规范格式

投标方需根据招标方提供的数据格式进行此项目制作。

投标方须提供的数据有：元数据著录信息文件、图像文件（抠图图像、单页图像）、excel文件等，具体规范详见附件。

五、质保服务

投标方在总项目完成后提供一年的免费质保服务。双方在合同签订后另行签署质保协议书。

附件一：元数据著录规则

一、术语定义

下列术语和定义适用于本附件。

1. 元数据著录单元

指期刊图像中有相对明显物理边界内的内容完整的图片。下文中简称“著录单元”。

2. 图片

包括独立图片和文中附图。独立图片为期刊中独立于文章之外图片；文中附图为期刊正文中出现的有意义的图片。

二、图片内容元数据及著录信息源

1. 图片著录项目及著录信息源

序号	图片著录项目	图片著录信息源
1	资源 id*	资源 id
2	划到 id*	划到 id
3	刊名 id*	刊名 id
4	导航 id*	导航 id
5	新篇 id*	新篇 id
6	原篇 id	原篇 id
7	图片栏目	有明确标识，将内容相近或有某种联系的图片组合在一起，形成图片栏目
8	正文标题	文中附图所属正文的标题
9	图片标题*	图片的标题
10	图片说明	图片的具体文字说明内容
11	图片责任者 1	完成图片内容创建的第一责任者
12	图片责任者 2	完成图片内容创建的第二责任者
13	图片著作方式 1	第一责任者对应的著作方式，如画、绘、摄等
14	图片著作方式 2	第二责任者对应的著作方式，如画、绘、摄等
15	期刊名称*	图片所属期刊的名称
16	出版年份*	图片所属期刊的出版年份

序号	图片著录项目	图片著录信息源
17	卷	图片所属期刊的卷号
18	期	图片所属期刊的期号
19	号	图片所属期刊的“总期号”、“特刊号”
20	扫描页码*	按招标方扫描图像的顺序产生的页码
21	图片类型 1*	漫画、照片、书法、题字、题词、手稿、木刻、篆刻、石刻、雕塑、曲谱、歌谱、表格、图表、画图、地图、拓本、拓片、其他
22	图片类型 2	仅当图片类型为完整的封面时著录
23	备注	
24	备用 1	招标方根据加工过程中的内容情况另行确定
25	备用 2	招标方根据加工过程中的内容情况另行确定
26	备用 3	招标方根据加工过程中的内容情况另行确定

注：1) 标注*的字段为必著项。

2) 卷、期、号三者不能同时为空。

3) “图片类型”由招标方事先设定。

4) 图片栏目和图片说明无需著录，仅保留字段。

5) 备注字段仅列出部分，应预留 30 个。

6) 图片类型 2 为“封面”的图片无需著录原篇 id。

2. 图片著录项定义及注释

1) 资源 id

定义：资源 id。

注释：招标方提供。

2) 划到 id

定义：划到 id。

注释：招标方提供。

3) 刊名 id

定义：刊名 id。

注释：招标方提供。

4) 导航 id

定义：导航 id。

注释：招标方提供。

5) 新篇 id

定义：新篇 id。

注释：投标方生成。

6) 原篇 id

定义：原篇 id。

注释：根据招标方提供的信息进行著录。

7) 图片栏目

定义：期刊中，内容相近或有某种联系的图片著录单元编排在一起，有客观的名称，
这个名称即为图片的栏目。

注释：无需著录，保留字段。

8) 正文标题

定义：文中附图所属正文的标题。

注释：参考招标方提供的原篇信息照实著录。

9) 图片标题

定义：图片的标题。

注释：参考招标方提供的原篇信息照实著录。文中附图无标题的著录“无标题”。

10) 图片责任者

定义：图片的主要负责人。

注释：参考招标方提供的原篇信息照实著录。

11) 图片著作方式

定义：图片的著作方式与责任者对应。

注释：参考招标方提供的原篇信息照实著录。

12) 图片说明

定义：相对于著录图片而言，除了图片标题以外的其它说明图片的文字

注释：无需著录，保留字段。

13) 扫描页码

定义：按招标方扫描图像的顺序产生的页码

注释：投标方根据招标方提供文件名提取。

14) 图片类型 1

定义：描述图片著录单元的类型

注释：属主观性著录项，需要通过对所著录单元的图片类型进行判断后给出选择，具体可选图片类型事先设定供著录时选择。

15) 图片类型 2

定义：描述图片著录单元的类型

注释：仅当图片为完整的“封面”时著录。该字段著录内容为“封面”。

三、元数据元素著录说明

1. 关于元数据元素著录

上述各表中标*的元素均为非空字段，著录项不可缺省；未标*的元素视具体内容具体分析，有内容则必须著录。

2. 关于元数据著录注释中“客观著录”与“照实著录”的说明

客观著录：指以原件内容的客观含义为准，不附加著录人员的主观认识。

照实著录：指按原件实际文字照搬著录。

3. 关于元数据及图像的唯一标识及对应关系

图片唯一标识由著录系统自动生成，且唯一标识一旦生成则不可改变，命名规则由招标方确定。元数据及图像的对应关系一定要满足数据的检索应用要求。

4. 关于图片类型著录

图片的类型根据事先约定的分类体系进行类别赋予。

5. 关于著录中的符号使用的说明

著录中使用的所有标点、符号与招标方协商确认。

6. 关于元数据著录与招标方先前元数据的关系

招标方提供先前已有的篇名元数据信息供投标方加工时参考，需明确的是投标方著录时应按本合同的元数据著录规则进行著录。招标方在后续对投标方生产的数据进行校验时按本合同的元数据规则 and 标准进行验收，先前元数据信息仅供参考，但投标方需提供先前篇名与本项目篇名元数据的详细对应关系表供招标方在后续工作中使用，篇名对应表的具体要求由招标方制定。

注：本附件解释权归招标方所有，著录过程中本规则未能详尽的内容由双方协商确定。

附件二：数据质量标准及数据格式规范

一、元数据著录质量标准

1. 著录规则质量标准：著录规则错误率 $\leq 0.3\%$

元数据著录规则错误率是元数据著录项错误的数据量与需著录项的数据总量比例，计算公式为：元数据著录规则错误率 = (元数据著录项错误的数据量 ÷ 需著录项的数据总量) × 100% (单位：项)。

著录规则错误是按项为计算单位，凡未能对元数据项进行客观著录的情况均计为著录规则错误。一篇元数据项为多值的，则按值的数量计算项数。

2. 著录文字质量标准：著录文字错误率 $\leq 0.05\%$

著录文字错误率是元数据项内容文字错误的数量与元数据项文字总量的比例，计算公式为：著录文字错误率 = (元数据项内容文字错误的数量 ÷ 元数据项文字总量) × 100% (单位：字)。

著录文字错误按文字计算，元数据中文字与原文不符的计为著录文字错误，但不包括对原文的文字错误进行纠正的情况。

3. 元数据与图像的勾链质量标准：元数据与图像的勾链无差错

4. 版面内容质量标准：版面内容无差错

5. 其他元数据与图像的合法性验收标准

1) 图像的分辨率与图像大小需符合招标方技术要求。

2) 元数据著录项中非空元素著录项无缺省。

非空元素是指附件一《元数据著录规则》中标*的元素，著录项不可缺省。

二、数据规范格式

根据本项目制作要求，投标方须提供的数据有：元数据著录信息文件、图像文件（抠图图像、单页图像）、excel 文件，具体规范要求如下：

1. 图像文件

单页图像，在原始图上采取合适的压缩率进行压缩，分辨率为 150DPI，保存为 JPG 格式，大小控制在 100K 左右，生产以样本为准。

抠图图像为著录过程中从原始图像中抠出的图片，尺寸视实际图像大小而定。抠出的图片务必要保持原始图像的完整性和整洁性，遇到排版不规则、交叉等无法规则抠出等情况，按照尽量保持原始的整体性原则进行处理。

抠图图像均需保存级别和应用级别两套数据。保存级别的图像应保持原有的图片格式，不进行压缩，分辨率为 300DPI。应用图像应为 JPG 格式，在原始图上采取合适的压缩率进行压缩，分辨率为 150DPI，大小控制在 100K 左右，实际大小可根据原始抠图的大小而变，生产以样本为准。

2. 元数据著录信息文件

元数据著录信息文件是存储所有元数据信息的可读可写文件，应包含按照附件一《元数据著录规则》著录单篇文章的元数据，以及对应的单页图像链接、图片抠图链接。具体著录要求由招标方制定。

3. excel 文件

保存为 xlsx 格式。记录招标方提供的原始篇名与本项目所著录篇名的对应关系。具体字段由招标方制定。

4. 文件命名及存储规范

文件的命名及存储规范与招标方协商后确定，要求规则统一，名称唯一。

5. 其他要求

- 1) 投标方所有著录的内容，除招标方有规定外，需向招标方说明著录时使用的字库。
- 2) 投标方根据招标方提供的图像，按照招标文件要求进行样本制作，并提供相应的匹配说明文件，方便招标方进行判断。

项目 6：老旧期刊篇名元数据制作（英）

项目数据量：35,000 篇

项目完成时间：2025 年 10 月 31 日

项目进度要求：

第一阶段：合同签署后，双方确认技术标准后，至 2025 年 6 月 30 日前至少完成 17,500 篇

第二阶段：至 2025 年 10 月 31 日前，完成 35,000 篇

一、招标样本

- 由招标方提供招标所需样本图像，共计约 36 页
- 投标方依据招标方提供的著录标准及各项规范（见下附件）制作招标样本
- 各投标方在将样品提交前，请先在样本上标识所投项目、投标方等信息，并确认标识是否完好
- 投标方需在投标日当天出示招标样本，并做阐述
- 投标方提交的样本将作为评估中标与否的因素之一
- 最终解释权归招标方所有

二、元数据主要著录内容要求

1. 正文内容元数据：6 项 ≤ 著录项 ≤ 29 项
广告元数据：5 项 ≤ 著录项 ≤ 14 项
图片元数据：6 项 ≤ 著录项 ≤ 17 项
2. 元数据制作成品：期刊的整本 PDF、单篇 PDF、图像文件（著录图片、著录广告图片）、元数据著录信息文件。
3. 元数据著录要求：依据文献的相关信息和内容客观著录。
4. 著录语种：期刊的元数据著录项均依据制作期刊内容照实进行著录。
5. 元数据元素主要著录说明
 - 1) 关于元数据元素著录

上述各表中标*的元素均为非空字段，著录项不可缺省；未标*的元素视具体内容具体分析，有内容则必须著录；

2) 关于元数据著录注释中“客观著录”与“照实著录”的说明

客观著录：指以原件内容的客观含义为准，不附加著录人员的主观认识。

照实著录：指按原件实际文字照搬著录。

3) 关于排序

由于老旧期刊的期（或卷）及原始页码不能准确实现检索应用中的期（或卷）排序及每期（或卷）文章的排序，加工时，同种期刊的每一册期刊之间及期刊中每一条正文、广告和图片元数据都需要生成对应的排序信息，以实现数据检索应用过程中的各种排序功能。

4) 关于元数据及图像的唯一标识及对应关系

正文唯一标识、广告唯一标识、图片唯一标识、数字化文本链接标识等由著录系统自动生成，且唯一标识一旦生成则不可改变，命名规则由招标方确定。元数据及图像之间的对应关系一定要满足数据的检索应用要求。

5) 关于类别著录

正文的类别、广告的类别以及图片的类型均根据事先约定的分类体系进行类别赋予。

6) 关于著录中的符号使用的说明

著录中使用的所有标点、符号与招标方协商确认。

三、验收质量标准

1. 著录规则质量标准：著录规则错误率 $\leq 0.3\%$ 。
2. 著录文字标准：著录文字错误率 $\leq 0.05\%$ 。
3. 元数据与图像文件的勾链质量标准：元数据与图像文件的勾链无差错
4. 版面内容质量标准：版面内容无差错
5. 篇名类型著录质量标准：篇名类型无差错
6. 其他元数据的合法性验收标准

1) 返回图像和 PDF 的页数必须和著录的页数完全相同；

2) 返回图像元数据著录项中非空元素著录项无缺省。

非空元素是指《元数据著录规则》中标*的元素，著录项不可缺省。

四、数据规范格式

投标方需根据招标方提供的数据格式进行此项目制作。

投标方须提供的数据有：期刊的整本 PDF 、单篇 PDF、图像文件（著录图片、著录广告图片）、元数据著录信息文件等。具体规范详见下附件。

五、质保服务

投标方在总项目完成后提供一年的免费质保服务。双方在合同签订后另行签署质保协议书。

附件一：近代外文期刊元数据著录规则

一、前言

近代外文期刊的内容基本分为：正文内容、广告内容和图片内容。其中，图片主要为内容独立或附有文字说明的图片。

近代外文期刊元数据的著录分三种类型：

- 正文内容元数据
- 广告内容元数据
- 图片元数据

二、术语定义

下列术语和定义适用于本附件。

1. 元数据著录单元

指期刊图像中有相对明显物理边界内的内容完整的正文、广告或图片。下文中简称“著录单元”。

2. 正文

期刊中除广告外的所有内容。

3. 广告

期刊中企业或个人用于推销商品、服务，或向社会传递信息，影响 YuLun 的传播内容。

4. 图片

期刊的图片包括独立图片和文中附图。独立图片为期刊中独立于文章之外图片；文中

附图为期刊正文中出现的有标题或附有文字说明的图片。

三、正文内容核心元数据及著录信息源

最终著录规则以实测样本为标准。元数据著录规则应该在正式生产前进行讨论，使其完全符合生产流程，同时内容适合于历史背景。

1. 正文著录项目及著录信息源

序号	正文著录项目	正文著录信息源
1	类别	正文著录单元所属的主题类目,可按事先约定的分类体系著录
2	栏目	在期刊中有明确标识,将内容相近或有某种联系的内容组合在一起
3	子栏目 1	子栏目 1 从属于栏目
4	子栏目 2	子栏目 2 从属于子栏目 1
5	标题 1*	著录单元内容的题名
6	翻译标题 1*	其他语种的标题 1
7	标题 2	同标题 1
8	翻译标题 2	其他语种的标题 2
9	标题 3	同标题 1
10	翻译标题 3	其他语种的标题 3
11	责任者 1	期刊正文的第 1 责任者
12	责任者 2	期刊正文的第 2 责任者
13	责任者 3	期刊正文的第 3 责任者
14	责任者单位 1	第 1 责任者所属的单位
15	责任者单位 2	第 2 责任者所属的单位
16	责任者单位 3	第 3 责任者所属的单位
17	著作方式 1	第 1 责任者对应的著作方式
18	著作方式 2	第 2 责任者对应的著作方式
19	著作方式 3	第 3 责任者对应的著作方式
20	语种	正文所采用的语言种类
21	关键词	正文内容的关键词

序号	正文著录项目	正文著录信息源
22	连载信息	正文中表示连载情况的信息，常见于文末
23	期刊名称*	正文所属期刊的名称
24	年*	正文所属的期刊出版日期中的年份信息 (公元纪年，格式为 yyyy)
25	卷	正文所属期刊的卷信息
26	期	正文所属期刊的期信息
27	号	正文所属期刊的“总期号”、“特刊号”等
28	实际页码	正文在期刊中的页码
29	起始扫描页码*	该著录单元所在起始页原始图像的命名内容，为 4 位 Arabic 数字
30	篇排序号*	根据期刊排版等生成该著录单元对应的排序信息，以实现数据检索应用过程中的各种排序功能
31	附图数量	附属于正文的图片数量
32	备注	其他待说明事项
33	备用 1	招标方根据加工过程中的内容情况另行确定
34	备用 2	招标方根据加工过程中的内容情况另行确定
35	备用 3	招标方根据加工过程中的内容情况另行确定

2. 著录项定义及注释

1) 类别

定义：正文著录单元所属的主题类目。

注释：属主观性著录项，需要通过对著录单元的内容进行判断后给出相应的类别，具体的类别可按事先设定的分类体系供著录时选择。

2) 栏目

定义：内容相近或有某种联系的几条著录单元编排在一起，有客观的名称且有相对明显的物理边界，这个名称即为正文的栏目。

注释：属客观性著录项，著录时原则上按照著录单元所属栏目客观著录。

3) 子栏目

定义：属于某一个栏目下细分的从属栏目。

注释：属客观性著录项，按照著录单元所属栏目客观著录。

4) 标题

定义：著录单元内容的题名。

注释：属客观性著录项，著录单元中有明确题名时，直接照实著录原始题名。

5) 翻译标题

定义：对应标题的另一种语种的标题。

注释：翻译标题可按相应的翻译原则进行翻译或照实著录原文所列另一种语种的标题。

6) 责任者

定义：创建著录单元内容的主要责任者。

注释：本元素属客观性著录项，是对数字文献内容负主要责任的个人或团体责任者进行著录，本项最多著录三个责任者。

7) 著作方式

定义：责任者创建著录单元时参与的方式。

注释：本元素属客观性著录项，是对文献内容负主要责任的个人或团体参与创建的方式，本项最多著录三个著作方式，与责任者对应。

8) 责任者单位

定义：个人责任者所属单位。

注释：本元素属客观性著录项。

9) 语种

定义：正文所采用的语言种类。

注释：本元素属客观性著录项，依据正文主体语言种类著录。

10) 关键词

定义：著录单元的关键词，可采用描述正文内容的非控关键词。

注释：本元素属主观性著录项，依据著录单元内容给出相应的关键词。

11) 连载信息

定义：正文中表示连载情况的信息，常见于文末。

注释：本元素属客观性著录项，按原件照实著录，如“未完”、“未完待续”等。

12) 期刊名称

定义：正文内容所属期刊的名称。

注释：客观性著录项。

13) 年

定义：正文所属的期刊出版日期中的年份信息。

注释：本元素应为公元纪年并采用 Arabic 数字著录，非公元纪年须转为公元纪年，格式为 yyyy。

14) 卷

定义：正文所属期刊的卷信息。

注释：本元素取自期刊当期封面，一般应采用 Arabic 数字著录。

15) 期

定义：正文所属期刊的期信息。

注释：本元素取自期刊当期封面，一般应采用 Arabic 数字著录。

16) 号

定义：正文著录单元所属期刊的总期号或特刊号等。

注释：本元素属客观性著录项，增刊、特刊、总期号等信息都著录于此字段。

17) 实际页码

定义：正文著录单元在所属期刊中的实际页码。

注释：本元素属客观性著录项，著录起始及转接页码。

18) 起始扫描页码

定义：该著录单元所在起始页原始图像的命名内容。

注释：本元素属客观性著录项，为 4 位 Arabic 数字。

19) 篇排序号

定义：根据期刊排版等生成该著录单元对应的排序信息，以实现数据检索应用过程中的各种排序功能。

注释：本元素著录 4 位 Arabic 数字。

20) 附图数量

定义：附属于正文的图片数量。

注释：客观著录项。

四、广告内容核心元数据及著录信息源

1. 广告著录项目及著录信息源

期刊中的广告元数据著录，每个著录单元的著录项包括广告类别、广告栏目、卷期信息、实际页码、起始扫描页码等。

序号	广告著录项目	广告著录信息源
1	广告标题	广告中对广告内容进行概括，起引导作用的语句
2	广告语	广告中长期、反复使用的简短口号
3	广告发布者	发布广告的个人或机构
4	广告产品	广告宣传的具体产品名称
5	广告类别*	广告著录单元所属的主题类目，可按事先约定的分类体系著录
6	广告栏目	内容相近或有某种联系的几条广告编排在一起，有客观的栏目名称
7	期刊名称*	同正文
8	年*	同正文
9	卷	同正文
10	期	同正文
11	号	同正文
12	实际页码	同正文
13	起始扫描页码*	同正文
14	篇排序号*	同正文
15	备注	其他待说明事项
16	备用 1	招标方根据加工过程中的内容情况另行确定
17	备用 2	招标方根据加工过程中的内容情况另行确定
18	备用 3	招标方根据加工过程中的内容情况另行确定

2. 广告著录项定义及注释

广告著录单元的卷期、实际页码、起始扫描页码著录项的定义和注释同正文。下面只列出广告类别、广告栏目的定义和注释。

1) 广告标题

定义：广告中对广告内容进行概括，起引导作用的语句。

注释：属客观性著录项。广告标题在具体某一则广告中使用，与广告正文内容同时出现，区别于广告语。

2) 广告语

定义：广告中长期、反复使用的简短口号。

注释：属于客观性著录项。广告语具有稳定性和连贯性，可在多个广告或系列广告中使用，区别于广告标题。

3) 广告发布者

定义：发布广告的主体，包括个人和机构。

注释：属客观性著录项。

4) 广告产品

定义：广告宣传的具体产品名称。

注释：属客观性著录项。

5) 广告类别

定义：广告著录单元所属的主题类目。

注释：属主观性著录项，需要通过对著录单元的广告内容进行判断后给出相应的类别，具体的类别可按事先设定的分类体系供著录时选择。

6) 广告栏目

定义：在期刊中有明确标识，且有相对明显的物理边界，将内容相近或有某种联系的几条著录单元组合在一起。

注释：属客观性著录项，著录时原则上按照著录单元所属栏目进行客观著录。

五、图片内容元数据及著录信息源

1. 图片著录项目及著录信息源

序号	图片著录项目	图片著录信息源
1	图片栏目	有明确标识，将内容相近或有某种联系的图片组合在一起，形成图片栏目
2	图片标题*	图片标题或图片所依附正文标题
3	翻译图片标题*	其他语种的图片标题
4	图片说明	图片的文字说明内容
5	翻译图片说明	其他语种的图片说明
6	图片责任者 1	完成图片内容创建的第一责任者
7	图片责任者 2	完成图片内容创建的第二责任者

序号	图片著录项目	图片著录信息源
8	图片著作方式 1	第一责任者对应的著作方式，如画、绘、摄等
9	图片著作方式 2	第二责任者对应的著作方式，如画、绘、摄等
10	图片类型 1*	漫画、照片、书法、题字、题词、手稿、木刻、篆刻、石刻、雕塑、曲谱、歌谱、表格、图表、画图、地图、拓本、拓片、其他
11	图片类型 2	仅当图片为完整的封面时著录
12	期刊名称*	同正文
13	年*	同正文
14	卷	同正文
15	期	同正文
16	号	同正文
17	实际页码	同正文
18	起始扫描页码*	同正文
19	篇排序号*	同正文
20	备注	其他待说明事项
21	备用 1	招标方根据加工过程中的内容情况另行确定
22	备用 2	招标方根据加工过程中的内容情况另行确定
23	备用 3	招标方根据加工过程中的内容情况另行确定

2. 图片著录项定义及注释

1) 图片栏目

定义：内容相近或有某种联系的图片著录单元编排在一起，有客观的名称，这个名称即为图片的栏目。

注释：属客观性著录项，著录时按照图片所属栏目客观著录。

2) 图片标题

定义：图片标题或图片所依附正文标题。

注释：属客观性著录项，依据图片标题或图片所依附正文标题客观著录。

3) 图片说明

定义：正文中图片的文字说明内容。

注释：属客观性著录项，依据正文中图片的文字说明内容客观著录。

4) 图片责任者

定义：图片的责任者。

注释：属客观性著录项。

5) 图片类型

定义：图片著录单元所属的类型。

注释：属主观性著录项，需要通过对著录单元的图片类型进行判断后给出选择，具体的类型可按事先设定的分类体系供著录时选择。

六、元数据元素著录说明

1. 关于字段非空

1) 上述各表中标*的元素均为非空字段，著录项不可缺省；未标*的元素视具体内容具体分析，有内容则必须著录；

2) 广告标题、广告发布者和广告产品不能同时为空。

2. 关于图片的著录与划框

图片著录时，原则上以单独图片进行著录，对于形状不规则的图片，在保证主题表达完整的前提下可进行组合图片的划框界定。

3. 关于著录项所用语种

除正文的“类别”、广告的“广告类别”、图片的“图片类型”以及“备注”跟随期刊语种外，其余著录项所用语种均依据制作期刊内容照实进行著录。期刊语种由招标方确定。

4. 关于广告的“广告类别”、图片的“图片类型”和正文的“类别”字段

1) 广告类别字段著录统一用英文，分类体系一般可包括：启事公告、个人事务、餐饮娱乐、邮电、交通运输、人才招聘、文化教育、医疗保健、食品饮料、百货零售、金融保险、房地产、拍卖、酒店旅游、能源供应、建筑装饰、钟表珠宝、烟草、工业用品、其他等等。

2) 图片类型字段著录统一用英文，分类体系一般可包括：漫画、照片、书法、题字、题词、手稿、木刻、篆刻、石刻、雕塑、曲谱、歌谱、表格、图表、画图、地图、拓本、拓片、其他等等。

3) 招标方可在样本制作之前或样本制作过程中对正文的类别、广告类别和图片类型的

分类体系提出合理的修改。在投标方进入正式生产后，分类体系的修改需双方另行商量决定。

5. 关于正文与图片的翻译字段

- 1) 正文的翻译标题字段与图片标题、图片说明翻译字段的翻译对象仅为英文，除英语外的其他外文内容无需翻译著录。
- 2) 正文的翻译标题字段与图片标题、图片说明翻译字段著录的是著录单元所对应的中文内容；
- 3) 正文标题 1、图片标题和图片说明有英文著录内容，对应的翻译字段不能为空；
- 4) 正文的标题 2 和标题 3 字段有英文著录内容时，对应的翻译标题 2 和翻译标题 3 字段不能为空；
- 5) 翻译原则如下：

- 忠实原则。翻译应深入研究和探讨不同语种新闻、图片标题和说明的特点，措辞必须准确、通顺。翻译人员应充分考虑题材及风格，翻译时需多兼顾参照当时中文期刊语言表达习惯，如相关事件专有名词、国际惯用表达等；
- 达意原则。准确理解标题或说明的意义，尤其是要透过文字字面理解其深层时代意义和政治背景；
- 可接受性原则。在翻译时应尽可能地呈现原文修辞特点，如双关、押韵等，使译文和原文在修辞上基本吻合；
- 一致性原则。相同人名、地名及事件名称保持一致。

6. 关于不要求著录的字段

- 1) 正文类别不要求著录；
- 2) 对于所有不要求著录的字段都应保留。

7. 关于标点、符号著录

著录中使用的所有标点、符号需双方协商确认。

8. 关于排序

期刊中每一条正文、广告和图片元数据都应有对应的排序信息，排序应符合期刊实际文脉和阅读次序，以实现目录及数据检索应用过程中的各种排序功能。

9. 关于元数据与图像的唯一标识及对应关系

正文唯一标识、广告唯一标识、图片唯一标识由著录系统自动生成，且唯一标识一旦生成不可改变，命名规则由招标方确定。元数据与图像之间的对应关系一定要满足数据的

检索应用要求。

注：本附件解释权归招标方所有，制作过程中本规则未能详尽的内容由双方协商确定。

附件二：近代外文期刊数据质量标准及数据格式规范

一、元数据著录质量标准

由于本合同规定的数据验收方式为分批验收，故本质量标准所指的错误率计算均以批为计量单位。

1. 图像处理质量标准：图像错误率 $\leq 0.05\%$

- 1) 图像错误率=（错误图像版数 $\div$ 该批次图像总版数） $\times 100\%$ （单位：版）
- 2) 每版图像应当同时符合下列质量标准，如有一项不符合，即该版为错误图像：
 - (1) 图像清晰；
 - (2) 图像完整，无缺损；
 - (3) 图像无明显污渍；
 - (4) 图像无扭曲、倾斜、变形。
- 3) 图像数据出现下列情况之一的，即该批次为验收不合格：
 - (1) 图像错误率实际值超过规定值；
 - (2) 图像出现缺页、多页、错页、重页等情况之一的。

2. 元数据著录质量标准：著录规则错误率 $\leq 0.3\%$

元数据著录规则错误率是元数据著录项错误的数量与需著录项的数据总量比例，计算公式为：元数据著录规则错误率=（元数据著录项错误的数量 $\div$ 需著录项的数据总量） $\times 100\%$ （单位：项）。

著录规则错误是按项为计算单位，凡未能对元数据项进行客观著录的情况均计为著录规则错误。一篇元数据项为多值的，则按值的数量计算项数。

3. 著录文字质量标准：著录文字错误率 $\leq 0.05\%$ ；

著录文字错误率是元数据项内容文字错误的数量（如多、少、错等）与元数据项文字总量的比例，计算公式为：著录文字错误率=（元数据项内容文字错误的数量 $\div$ 元数据项文字总量） $\times 100\%$ （单位：字）。

著录文字错误按文字计算，元数据中文字与原文不符的计为著录文字错误，但不包括

对原文的文字错误进行纠正的情况，原文错误纠正记入勘误表。

4. 元数据与图像的勾链质量标准：元数据与图像的勾链无差错

5. 标题及图片中文翻译数据质量标准：标题及图片中文翻译错误率 $\leq 0.8\%$

标题及图片中文翻译错误率是无法达到标题翻译标准的数据与需翻译的数据总量的比例，计算公式为：标题及图片中文翻译错误率=（无法达到标题翻译标准的数据 $\div$ 需翻译的数据总量） $\times 100\%$ （单位：项）。翻译标准参见元数据著录规则中六-5。

6. 版面内容元数据质量标准：版面内容元数据无漏著

7. 篇名类型著录质量标准：篇名类型无差错

8. 其他元数据的合法性验收标准

1) 返回图像和 PDF 的版数必须和著录的版数完全相同；

2) 返回图像像素、大小无异常，灰度或色彩无明显漏洞，不存在多幅图像相同的问题；

3) 返回图像元数据著录项中非空元素著录项无缺省；

非空元素是指《元数据著录规则》中第三、四、五条第 1 款标*的元素，著录项不可缺省。

二、 数据规范格式

根据本项目制作要求，投标方须提供的数据有：整本 PDF、单篇 PDF、图像文件（单版图像、著录图像、著录广告图像）的 TIF 及 JPG 格式各一套、元数据著录信息文件、版面分析信息文件等，具体规范要求如下：

1. PDF 文件

每本期刊的 PDF 数据为两套。一套为整本 PDF，一套为单篇 PDF，质量要求如下：

1) PDF 均为双层 PDF 文件；

2) PDF 版本通用性，需要符合 PDF 国际标准。投标方提交成品时，需向招标方说明使用的 PDF 版本；

3) PDF 像素、大小无异常，灰度或色彩无明显漏洞；

4) 在保证图像清晰度的前提下，对保存级别 300DPI/TIF 图像进行压缩；

5) 使用 JPG 压缩方式，根据原始图像质量和大小，选取适当的压缩比例（推荐值为 65%）。

生产时以样本为准；

6) 如有必要，需对招标方提供的源图像进行修正，以保证产出的 PDF 文件画面清晰、无扭曲、倾斜、变形等影响查看的情况发生；

7) 除特殊规定外, PDF 内容的完整性和准确性应与招标方提供的源图像保持一致。

2. 成品图像及要求

成品图像包括保存级别和应用级别两套数据。投标方应保持图像为报刊原始扫描色彩。图片和广告, 需根据实际情况进行抠图后保存。

1) 保存级别图像: 包括整版图像、著录图像和著录广告图像

要求: 应保持原有的图像样式, 保持 300DPI/TIF, 使用 LZW 无损压缩格式。

2) 应用级别图像: 整版图像、著录图像和著录广告图像

要求: 在保证图像清晰的前提下, 对保存级别 300DPI/TIF 图像进行压缩, 根据原始图像的大小和扫描质量, 选取恰当的压缩比例(推荐值为 65%), 使用 JPG 压缩方式, 最终提供整版图像保存为 150DPI 分辨率的 JPG 格式, 著录图像保存为 150DPI 的 JPG 格式, 大小视实际抠出大小而定。

抠出的图片、广告等图像务必要保持原始图像的完整性和整洁性, 遇到排版不规则、交叉等无法规则抠出等情况, 按照尽量保持原始的整体性原则进行处理。

3. 元数据著录信息文件

元数据著录信息文件存储所有元数据信息的可读可写文件, 应按照元数据著录规则著录单篇元数据、对应的 PDF、图片路径、广告图片路径等。

坐标信息也需要在元数据著录信息文件中标出, 包括单篇坐标和抠图图片坐标。

4. 命名规则

文件的命名需与招标方协商后确定, 要求规则统一, 名称唯一。

建议命名规则:

1) 所有文件的扩展名均为小写

2) PDF

A. 整本 PDF 文件的命名规则: 划到 ID_A11. pdf

B. 单篇 PDF 文件的命名规则: 起始扫描页码(4 位 Arabic 数字)+4 位篇顺序号(从 0001 顺序)+元素类型(正文、广告、图片), 如下:

0003_0004_ZW. pdf (第 4 篇为正文)

0004_0005_TP. pdf (第 5 篇为图片)

0005_0006_GG. pdf (第 6 篇为广告)

注: 此篇顺序号同 xml 中的篇顺序号(下同)。

3) XML

XML 文件的命名规则：起始扫描页码（4 位 Arabic 数字）+4 位篇顺序号（从 0001 顺序）+元素类型（正文、广告、图片），如下：

0003_0004_ZW.xml（第 4 篇为正文）

0004_0005_TP.xml（第 5 篇为图片）

0005_0006_GG.xml（第 6 篇为广告）

4) 图像

单页图像命名同招标方提供的原始图像名

附图图像：附图图像包括正文附图、图片附图、广告附图，命名规则为“起始扫描页码（4 位 Arabic 数字）+4 位篇顺序号（从 0001 顺延）+元素类型（图片、广告）”，如下：

0000_0001_1_TP.jpg（第 1 篇第 1 个图片附图）

0005_0006_1_GG.jpg（第 6 篇第 1 个广告附图）

5. 版面分析信息文件

提供该批次图像识别校改后的针对版面 300DPI 图像文件的工程文件（包括版面结构和版面内容），把结构基元（包括行坐标、段落坐标、划框信息等）内容封装在 OCR 内，方便招标方进行验收以及后期再次加工和进一步的信息采集。

6. 其他要求

投标方应根据招标方提供的图像，按照招标文件要求进行样本制作，并提供相应的匹配说明文件，方便投标方进行判断。双方认可的数据样例作为本项目著录及数据规范的参照依据。

第六章 附 件

附：投标文件封面格式

上海图书馆(上海科学技术情报研究所)馆藏文献
元数据及全文 OCR 数据制作项目

(商务/技术)

投 标 文 件

项目编号：_____

内部编号：STC25A114

预算编号：0025-00000245、0025-00000249、0025-00000246、
0025-00000248、0025-00000250、0025-00000247

投标人全称(加盖公章)：_____

地 址：_____

联系人：_____ 联系电话：_____

日 期：_____

附件一： 开标一览表

项目名称： _____
项目编号： _____
内部编号： _____

上海图书馆(上海科学技术情报研究所)馆藏文献元数据及全文 OCR 数据制作项目包 1

中文老 旧报纸 全文 OCR 元 数据制 作	科技现 刊篇名 元数据 制作	老旧期 刊篇名 元数据 制 作 (中)	会议论 文篇名 元数据 制作	学术图 片库元 数据制 作	老旧期 刊篇名 元数据 制 作 (英)	完成时 间	备注	最终报 价 (总 价、元)

注：

1. 开标时，投标文件中开标一览表(报价表)内容与投标文件中明细表内容不一致的，以开标一览表(报价表)为准。投标文件的大写金额和小写金额不一致的，以大写金额为准；总价金额与按单价汇总金额不一致的，以单价金额计算结果为准；单价金额小数点有明显错位的，应以总价为准，并修改单价；对不同文字文本投标文件的解释发生异议的，以中文文本为准。

法定代表人或授权代表签名：

投标人（盖章）：

日 期：

附件二：资格性自查表

采购文件要求	检查内容	自查结论	证明资料
(1)具有独立承担民事责任的能力；	最新检审合格的营业执照或事业单位法人证书。	☐ 通过 ☐ 不通过	投标文件第 页
(2)具有良好的商业信誉和健全的财务会计制度；	财务状况及税收、社会保障资金缴纳情况声明函；	☐ 通过 ☐ 不通过	投标文件第 页
(3)具有履行合同所必需的设备和专业技术能力；	主要技术人员、售后服务机构人员能力等。	☐ 通过 ☐ 不通过	投标文件第 页
(4)有依法缴纳税收和社会保障资金的良好记录；	财务状况及税收、社会保障资金缴纳情况声明函；	☐ 通过 ☐ 不通过	投标文件第 页
(5)参加政府采购活动前三年内，在经营活动中没有重大违法记录；	提供无重大违法记录承诺书。	☐ 通过 ☐ 不通过	投标文件第 页
(6)投标人应在投标文件中提供无不良信用记录的相关证明（相应网页截图）。不良信用记录指：投标人在“信用中国”网站（www.creditchina.gov.cn）被列入失信被执行人、重大税收违法案件当事人名单或在中国政府采购网（www.ccgp.gov.cn）被列入政府采购严重违法失信行为记录名单；	提供相关证明。	☐ 通过 ☐ 不通过	投标文件第 页
(7)投标人为中小企业；	提供企业声明函。	☐ 通过 ☐ 不通过	投标文件第 页
(8)本项目不接受联合体投标。		☐ 通过 ☐ 不通过	投标文件第 页

法定代表人或授权代表签名：

投标人（盖章）：

日 期：

附件三：投标书

致上海中招招标有限公司：

根据贵方为（招标项目名称）采购的投标邀请（项目编号：_____），签字代表（姓名、职务）经正式授权并代表投标人（投标人名称、地址）提交下述文件：

一、商务文件

二、技术文件

五、我单位(或公司)同意以下事项：

1. 本次招标所投内容为_____，投标总价为：_____（人民币 • 元）；
2. 遵守招标文件的各项条款及一切有关规定；
3. 已详细审查全部招标文件，包括所有补充通知（如果有的话）。我们完全理解并同意放弃对这方面有不明、误解和质疑的权力。；
4. 如我方投标被接受，我方将按招标文件的规定履行合同责任和义务；
5. 同意提供按照贵方可能要求的与其投标有关的一切数据或资料，完全理解贵方不一定接受最低价的投标或收到的任何投标。
6. 本投标书自贵方收到之日生效，有效期至开标后 90 日，在此期间本投标书之规定对我方具有约束力；
7. 如果我方投标被接受，则至合同履行完成和质量保证期满为止，本投标书保持有效。

投标人(盖章)：

地 址：

电 话：

邮 编：

传 真：

法定代表人或授权代表签字：

日 期：

附件四：投标保证书

致上海中招招标有限公司：

本书作为_____ (投标人) 对_____ (项目编号：_____) 的投标邀请提供投标保证金的保证。

由_____ (开户银行名称) 提供的_____ (填投标保证金形式) 总额为_____ (人民币，元)。

一旦发生下述行为，我单位(或公司)将放弃追索保证金的权利：

1. 从开标日起到投标文件有效期满前，投标人撤回投标；
2. 开标、评标到定标期间发生违反招标文件规定的行为；
3. 在收到《中标通知书》后未能按中标通知书规定的时间地点与采购人签订合同；

本保证书自开标日起 90 日有效，除非贵方提前终止或解除本保证书。

投 标 人(盖章)：_____

开户银行：_____

账号：_____

法定代表人或授权代表签字：_____

日期：_____年____月____日

附件五：法定代表人身份证明书

（如为法人直接投标时需提供）

单位名称：_____

单位性质：_____

地 址：_____

成立时间：_____年___月___日

经营期限：_____

姓 名：_____ 性别：_____ 年龄：_____ 职务：_____

系_____（投标人名称）_____的法定代表人。

特此证明。

投标人：_____（盖章）

日 期：_____年___月___日

法定代表人身份证复印件正反面

后附法定代表人《参保人员城镇职工基本养老保险缴费情况》查询表或社保缴纳证明材料。

附件六：法定代表人授权委托书

（如为授权代表投标时需提供）

本授权委托书声明：我_____（姓名）系_____（投标人名称）的法定代表人，现授权委托我单位的_____（姓名）为我公司授权代表，以公司名义参加上海中招招标有限公司代理的_____（项目编号：_____）的投标活动。授权代表在开标、评标、合同谈判过程中所签署的一切文件和处理与之有关的一切事务，我均予以承认。

授权代表(签字)：

性别：

年龄：

部 门：

职务：

授权代表无转委托。特此委托。

投标人(盖章)：

法定代表人(签字或盖章)：

日 期：_____年____月____日

法定代表人身份证复印件正反面

授权代表身份证复印件正反面

后附授权代表《参保人员城镇职工基本养老保险缴费情况》查询表或社保缴纳证明材料。

附件七：投标保证金交纳凭证

上海中招招标有限公司：

（投标人全称）参加贵方组织的、项目编号为_____、包件_____的采购活动。
按招标文件的规定，已通过（电汇、转账）形式交纳人民币（大写）_____元的投标保证金。

投标人名称：_____

投标人开户银行：_____

投标人银行帐号：_____

说明：1. 上述要素供银行转账及银行汇款方式填写，其他形式可不填。其他方式以现场递交为依据。

2. 上述要素的填写必须与银行转账或银行汇款凭证的要素一致，政府采购代理机构依据此凭证信息退还投标保证金。

法定代表人或授权代表签名：_____

投标人（盖章）：_____

日期： 年 月 日

附：

粘贴转账的银行凭证复印件

注：投标人投标响应时，应当按招标文件要求交纳投标保证金。投标保证金可以采用电汇、转账形式交纳。

附件八：业绩清单

(应附合同复印件)

项目名称：

项目编号：

序号	项目名称	用户名称	地点	时间	用户联系方式

(此表可延长)

法定代表人或授权代表签名：

投标人（盖章）：

日 期：

附件九：中小企业声明函

中小企业声明函（服务）

本公司（联合体）郑重声明，根据《政府采购促进中小企业发展管理办法》（财库〔2020〕46号）的规定，本公司（联合体）参加（单位名称）的（项目名称）采购活动，提供的服务全部由符合政策要求的中小企业承接。相关企业（含联合体中的中小企业、签订分包意向协议的中小企业）的具体情况如下：

1. （标的名称），属于（采购文件中明确的所属行业）行业；承建（承接）企业为（企业名称），从业人员 人，营业收入为 万元，资产总额为 万元，属于（中型企业、小型企业、微型企业）；

2. （标的名称），属于（采购文件中明确的所属行业）行业；承建（承接）企业为（企业名称），从业人员 人，营业收入为 万元，资产总额为 万元，属于（中型企业、小型企业、微型企业）；

.....

以上企业，不属于大企业的分支机构，不存在控股股东为大企业的情形，也不存在与大企业的负责人为同一人的情形。

本企业对上述声明内容的真实性负责。如有虚假，将依法承担相应责任。

企业名称（盖章）：

日 期：

说明：（1）本声明函适用于所有在中国境内依法设立的各类所有制和各种组织形式的企业。事业单位、团体组织等非企业性质的政府采购供应商，不属于中小企业划型标准确定的中小企业，不得按《关于印发中小企业划型标准规定的通知》规定声明为中小微企业，也不适用《政府采购促进中小企业发展管理办法》。

（2）从业人员、营业收入、资产总额填报上一年度数据，无上一年度数据的新成立企业可不填报

注：各行业划型标准：

（一）农、林、牧、渔业。营业收入 20000 万元以下的为中小微型企业。其中，营业收入 500 万元及以上的为中型企业，营业收入 50 万元及以上的为小型企业，营业收入 50 万元以下的为微型企业。

（二）工业。从业人员 1000 人以下或营业收入 40000 万元以下的为中小微型企业。其中，从业人员 300 人及以上，且营业收入 2000 万元及以上的为中型企业；从业人员 20 人及以上，且营业收入 300 万元及以上的为小型企业；从业人员 20 人以下或营业收入 300 万元以下的为微型企业。

（三）建筑业。营业收入 80000 万元以下或资产总额 80000 万元以下的为中小微型企业。其中，营业收入 6000 万元及以上，且资产总额 5000 万元及以上的为中型企业；营业收入 300 万元及以上，且资产总额 300 万元及以上的为小型企业；营业收入 300 万元以下或资产总额 300 万元以下的为微型企业。

（四）批发业。从业人员 200 人以下或营业收入 40000 万元以下的为中小微型企业。其中，从业人员 20 人及以上，且营业收入 5000 万元及以上的为中型企业；从业人员 5 人及以上，且营业收入 1000 万元及以上的为小型企业；从业人员 5 人以下或营业收入 1000 万元以下的为微型企业。

（五）零售业。从业人员 300 人以下或营业收入 20000 万元以下的为中小微型企业。其中，从业人员 50 人及以上，且营业收入 500 万元及以上的为中型企业；从业人员 10 人及以上，且营业收入 100 万元及以上的为小型企业；从业人员 10 人以下或营业收入 100 万元以下的为微型企业。

（六）交通运输业。从业人员 1000 人以下或营业收入 30000 万元以下的为中小微型企业。其中，从业人员 300 人及以上，且营业收入 3000 万元及以上的为中型企业；从业人员 20 人及以上，且营业收入 200 万元及以上的为小型企业；从业人员 20 人以下或营业收入 200 万元以下的为微型企业。

（七）仓储业。从业人员 200 人以下或营业收入 30000 万元以下的为中小微型企业。其中，从业人员 100 人及以上，且营业收入 1000 万元及以上的为中型企业；从业人员 20 人及以上，且营业收入 100 万元及以上的为小型企业；从业人员 20 人以下或营业收入 100 万元以下的为微型企业。

（八）邮政业。从业人员 1000 人以下或营业收入 30000 万元以下的为中小微型企业。其中，从业人员 300 人及以上，且营业收入 2000 万元及以上的为中型企业；从业人员 20 人及以上，且营业收入 100 万元及以上的为小型企业；从业人员 20 人以下或营业收入 100 万元以下的为微型企业。

（九）住宿业。从业人员 300 人以下或营业收入 10000 万元以下的为中小微型企业。其中，从业人员 100 人及以上，且营业收入 2000 万元及以上的为中型企业；从业人员 10 人及

以上，且营业收入 100 万元及以上的为小型企业；从业人员 10 人以下或营业收入 100 万元以下的为微型企业。

（十）餐饮业。从业人员 300 人以下或营业收入 10000 万元以下的为中小微型企业。其中，从业人员 100 人及以上，且营业收入 2000 万元及以上的为中型企业；从业人员 10 人及以上，且营业收入 100 万元及以上的为小型企业；从业人员 10 人以下或营业收入 100 万元以下的为微型企业。

（十一）信息传输业。从业人员 2000 人以下或营业收入 100000 万元以下的为中小微型企业。其中，从业人员 100 人及以上，且营业收入 1000 万元及以上的为中型企业；从业人员 10 人及以上，且营业收入 100 万元及以上的为小型企业；从业人员 10 人以下或营业收入 100 万元以下的为微型企业。

（十二）软件和信息技术服务业。从业人员 300 人以下或营业收入 10000 万元以下的为中小微型企业。其中，从业人员 100 人及以上，且营业收入 1000 万元及以上的为中型企业；从业人员 10 人及以上，且营业收入 50 万元及以上的为小型企业；从业人员 10 人以下或营业收入 50 万元以下的为微型企业。

（十三）房地产开发经营。营业收入 200000 万元以下或资产总额 10000 万元以下的为中小微型企业。其中，营业收入 1000 万元及以上，且资产总额 5000 万元及以上的为中型企业；营业收入 100 万元及以上，且资产总额 2000 万元及以上的为小型企业；营业收入 100 万元以下或资产总额 2000 万元以下的为微型企业。

（十四）物业管理。从业人员 1000 人以下或营业收入 5000 万元以下的为中小微型企业。其中，从业人员 300 人及以上，且营业收入 1000 万元及以上的为中型企业；从业人员 100 人及以上，且营业收入 500 万元及以上的为小型企业；从业人员 100 人以下或营业收入 500 万元以下的为微型企业。

（十五）租赁和商务服务业。从业人员 300 人以下或资产总额 120000 万元以下的为中小微型企业。其中，从业人员 100 人及以上，且资产总额 8000 万元及以上的为中型企业；从业人员 10 人及以上，且资产总额 100 万元及以上的为小型企业；从业人员 10 人以下或资产总额 100 万元以下的为微型企业。

（十六）其他未列明行业。从业人员 300 人以下的为中小微型企业。其中，从业人员 100 人及以上的为中型企业；从业人员 10 人及以上的为小型企业；从业人员 10 人以下的为微型企业。

附件十：资格声明函

资格声明函

致：上海中招招标有限公司

关于贵方采购项目名称：_____项目编号：_____包件：_____）投标邀请，本签字人愿意参加投标响应，提供投标文件中规定的货物及服务，并证明提交的下列文件和说明是准确的和真实的。

1. 我方为本次投标所提交的所有证明其合格和资格的文件是真实的和正确的,并愿为其真实性和正确性承担法律责任;

2.我方是依法注册的法人，在法律上、财务上和运作上完全独立于（采购人名称）（采购人）及上海中招招标有限公司（采购代理机构）；

3.我方在参加本次投标前 三 年内，在经营活动中没有重大违法记录；

5. _____

附相关文件如下:

1. 《企业法人营业执照》（或注册登记证书）复印件；
2. 《组织机构代码证》复印件；
3. 《税务登记证》复印件；
4. 行政主管部门或行业协会对所招标货物的经销许可证（若有）复印件。
5. 符合《投标邀请函》第五项“供应商资格条件”要求的相关证明文件。

法定代表人或授权代表签名:

投标人（盖章）：

日期:

附件十一：无重大违法记录承诺书

上海中招招标有限公司：

我公司在参加本次政府采购活动前三年内，在经营活动中没有重大违法记录。

特此声明。

法定代表人或授权代表签名：

投标人（盖章）：

日 期：

附件十二：健全的财务会计制度、依法缴纳税收和社会保障资金的相关证明

财务状况及税收、社会保障资金

缴纳情况声明函

我方（供应商名称）符合《中华人民共和国政府采购法》第二十二条第一款第（二）项、第（四）项规定条件，具体包括：

1. 具有健全的财务会计制度；
2. 有依法缴纳税收和社会保障资金的良好记录。

特此声明。

我方对上述声明的真实性负责。如有虚假，将依法承担相应责任。

供应商名称（公章）

日期：

附件十三:提供投标人没有在“信用中国”网站
(www.creditchina.gov.cn)被列入失信被执行人、
重大税收违法案件当事人名单或在中国政府采购网
(www.ccgp.gov.cn)被列入政府采购严重违法失信行
为记录名单的证明。

附件十四:招标文件中要求提供的其他资质证明
文件。

附件十五：商务条款响应表

项目名称：

项目编号：

包件：

序号	一般商务条款要求	是否响应	偏离说明
1.	完全理解并接受合同条款要求。		
2.	完全理解并接受对合格投标人、合格的货物、工程和服务要求。		
3.	完全理解并接受对投标人的各项须知、规约要求和责任义务。		
4.	投标有效期：投标有效期为自递交投标文件起至确定正式中标人止不少于 90 天。		
5.	报价明细表均涵盖报价要求之一切费用和伴随服务。		
6.	同意接受合同范本所列举的各项条款。		
7.	同意按本项目要求缴付相关款项。		
8.	同意采购方以任何形式对我方投标文件内容的真实性进行审查、验证。		
...	其它商务条款偏离说明		

注：

1、对于上述要求，如投标人完全响应，则请在“是否响应”栏内打“√”，对空白或打“×”视为偏离，请在“偏离说明”栏内扼要说明偏离情况。

2、投标人可根据实际情况填写此表，须包含但不限于此表内容。

法定代表人或授权代表签名：

投标人（盖章）：

日 期：

附件十六：合同偏离表

项目名称：

项目编号：

包件：

序号	招标文件中的合同条款	投标人的响应	是否 偏离	说明

（此表可延长）

法定代表人或授权代表签名：

投标人（盖章）：

日 期：

附件十七：为完成本项目的人员情况表

项目名称：

项目编号：

包件：

序号	姓名	职务	专业	从事本行业的年限	曾从事的类似项目	专长/所获得资质证书
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						

（此表可延长）
后附项目人员《参保人员城镇职工基本养老保险缴费情况》查询表或社保缴纳证明材料。

法定代表人或授权代表签名：
投标人（盖章）：
日 期：

附件十八:技术方案

(格式由供应商自拟,但应包括但不限于如下内容)

供应商应按招标文件要求的内容和顺序,对完成整个项目提出相应的技术方案。对含糊不清或欠具体明确之处,评委会可视为供应商履约能力不足或响应不全处理。组织实施服务方案的内容应包括:

- 1、对项目的理解(项目概述、目标、服务范围、采购人的义务及配合条件);
- 2、针对本项目的服务方案内容:
 - (1)主要需求指标的详细说明,可以是文字资料、电子文本、图纸和数据;
 - (2)服务计划和安排(服务方案、工作方法与流程)
 - (3)项目服务保障(服务质量保证措施、突发事件的处理措施)
 - (4)项目服务人员、设备配置情况;
- 3、供应商认为必要的其它内容。

法定代表人或授权代表签名:

投标人(盖章):

日 期:

附件十九： 保证金退还申请表

公司名称（公章）：

时间：

项目名称：	项目编号：
内部编号：	
包件： /	
交纳保证金时间：	
交纳保证金方式（电汇/转账）：	
交纳保证金金额：小写：	
大写：	
退付保证金金额：小写：	
大写：	
退付保证金账号	
账户：	
账号：	
开户银行：	
联系人：	联系电话：