**万方数据全文内容XML结构化加工系统采购需求书**

**一、项目背景**

万方数据是国内较早以信息服务为核心的股份制高新技术企业，经过20年来快速稳定的发展，已经积累了上亿条的全文数据。 目前，国内外学术出版知识服务实践关注行业整体构建，其发展呈现内容加工精细化、用户服务精准化和技术应用深层化的趋势。鉴于此，公司制定了全文内容XML结构化加工战略，于2018年采购了PDF结构化分析软件从事期刊全文内容XML机构化加工。

由于该软件功能简单需要大量人工操作，故生成效率较低，生成成本高，为解决上述主要问题，因此需采购全文内容XML机构化加工系统以解决上述问题。

1. **技术需求**
2. **软件参数**

|  |  |  |
| --- | --- | --- |
| 参数序号 | 要求 | 规格说明 |
| 1 | ★加工资源类型 | （1）支持多种不同类型资源（刊、图书、报纸、方志、新闻信息等）的加工  （2）支持单层PDF、双层PDF等多种格式资源的加工 |
| 2 | ★解析 | （1）支持版面的自动分析，自动提取解析文字、图片、表格等  （2）支持所有文章进行自动切分  （3）对文章的标题、作者、来源、页码、日期、期号等元数据进行自动识别、关联和提取 |
| 3 | ★校对 | （1）支持人工校对，审核界面高亮对照  （2）人工校对过程可加入到机器自学习模型，提升自动分析的准确率 |
| 4 | ★成品数据 | （1）XML文件、mdb数据库文件等  （2）数据准确率不低于3‱（万分之三） |
| 5 | ★加工效率 | （1）平均加工期刊不低于5本/人天 |

注：加★参数为产品必须具备的特性。

**三、商务要求**

**1. 服务要求：**

免费3年技术支持，支持方式如下：

（1）热线支持服务。提供7\*24小时电话支持服务，30分钟响应，提供全年365天×24小时热线电话技术支持服务；如遇严重故障电话远程支持不能及时解决，派工程师到现场处理。

（2）E-Mail支持服务。全年365天24小时E-Mail即时服务，30分钟内响应客户技术需求，2小时内提出解决方案。

（3）现场技术支持服务。接到用户通知后30分钟内提供技术响应，根据客户的要求在2小时内到达现场，提供排错服务，4小时内排除故障。

**2.支付方式：**

第一部分：首付款，在签订合同后的5个工作日内支付合同额的40%；第二部分：在完成系统部署以后，由数据加工人员进行实操后系统验收，并由相关人员签署验收报告，然后支付合同额的50%；第三部分：生产环境稳定运行3个月后支付合同额的10%。

**3.供货周期：**合同签订后20个工作日内工程师到场部署。

**四、报价**

**报价说明：**

1.报价需包含满足项目需求的所有费用，格式可根据实际情况自行调整。

2.以上所有费用都是含税（增值税专用发票）价格。

3.如供应商可提供需求外的其他服务，如提升服务质量的其他产品、收费的技术服务、超出免费服务期的维护服务等，可一并列将其报价列在报价表中，但不计入在此次询价的总价中。

**五、其他要求**

1.投标人需提供企业营业执照复印件、税务登记复印件、法定代表人授权委托书（所有文件如非原件均须加盖投标人公章）。

2.以下资质文件如有，请一并提供（信息安全服务资质、ISO9000,系统集成等认证证书）。

3.提供相关成功案例。

4.投标人拥有体系完善的24x7x365的客户响应中心，提供服务热线，具备完善的疑难问题升级流程及管理流程和能力。

5.请于2021年 11月25日5点前，将加盖公章及骑缝章的报价文件及附加1.2.3条涉及的所有文件，送至我单位，地址：北京市海淀区复兴路15号，联系人：魏铮 010-58882253，逾期报送的报价文件将不予以接受。

北京万方数据股份有限公司

2021年 11 月 19 日。