**实验一 数字资源信息组织方法与技术实验**

**实验目的**

1. 掌握数字图书馆数字资源的标引原理与方法
2. 掌握数字图书馆数字资源的分类原理与方法

**实验内容**

1.通过网络、文献资料调研基于文本分析（分词、特征词提取）的数字资源标引方法以及基于文本分析的数字资源分类方法，掌握大规模数字资源的标引、分类的相关原理、技术与基本流程。

2.通过网络获取用于文本分析、文本分类相关的工具、软件（或程序代码），并了解其在大规模数字资源信息组织中的功能和作用。

3.利用这些工具和程序代码以及已学过的相关专业知识，完成数字资源标引与分类的实验。

4.设计UI界面，显示数字资源标引与分类的实验过程中所涉及的流程、中间结果以及最终的实验结果。

**实验要求**

1. 在调研基础上，分析描述数字资源标引、分类的基本流程
2. 分析所下载的自动分类相关工具或软件（程序代码）的基本功能，选择其中你认为适合的工具或代码作为本次实验的工具参考。
3. 利用所提供的训练集和实验集，完成大规模数字资源的标引、分类实验。
4. 总结在实验中遇到了哪些问题？如何解决的？本次实验有哪些收获？将上述问题及其心得体会总结成实验报告。实验报告以电子版方式，由学委收齐后，于本课程期末考试之前提交。

**附件**

（1）用于数字资源标引、分类的训练集和测试集