

试题编号：429 试题名称：信息检索

注意：答题一律答在答题纸上，答在草稿纸或试卷上一律无效

一、 名词解释（每小题 3 分，总 30 分）

1. 半停用词
2. 网罗度
3. 自由标引
4. 统计标引法
5. 《汉语主题词表》
6. metadata
7. cross reference
8. KWOC
9. 三次文献
10. 受控语言

二、 问答题（每小题 10 分，总 70 分）

1. 简述概念组配与字面组配的差异。
2. 《中国分类主题词表》具有哪些功能？
3. BA 的索引体系有什么特点？
4. 简述 SCI 的功能。
5. 近期目次的特点与功能。
6. 分析自然语言检索的利弊。
7. 信息检索系统评价中的性能指标有哪些，请分别列出并解释其含义。

三、 论述题（每 1、2 小题 15 分，第 3 小题 20 分，总 50 分）

1. 试分析网络信息检索工具发展的趋势。
2. 分析比较搜索引擎中的命中记录常用的几种排序算法（至少列举三种算法）。
3. 写出小中取大距离分类法的算法，分析其性能，并对其提出一种改进方案，使其满足次序独立性原则。