



河南工业大学
HENAN UNIVERSITY OF TECHNOLOGY

软件工程概论作业

作业 1：如何提升搜索质量

学 号：201816040209

姓 名：程 长

年 级：2018 级

专 业：软件工程

授课教师：王珂

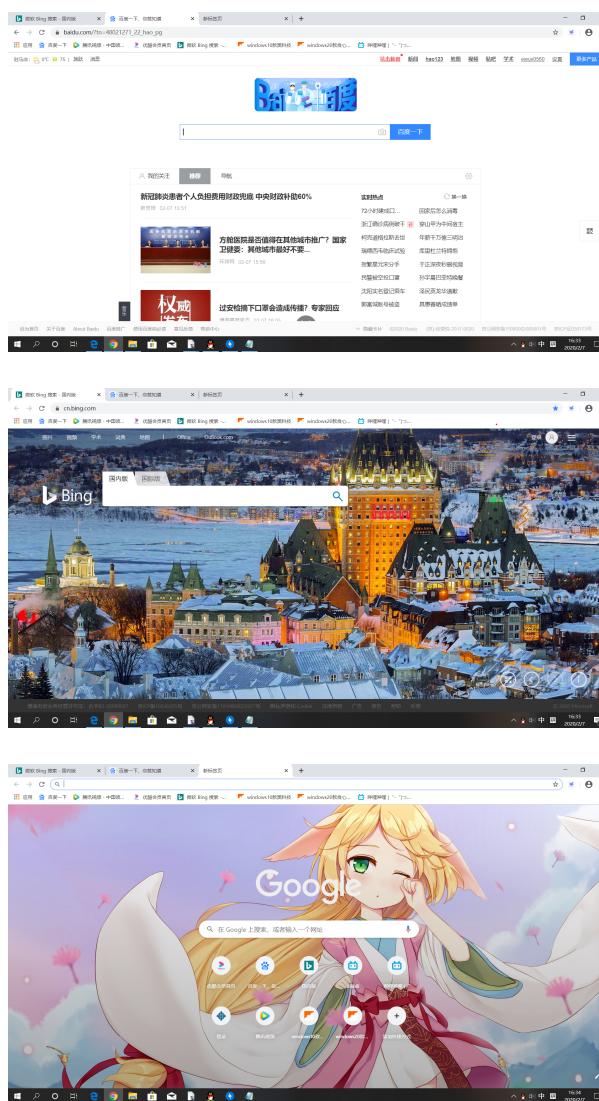
完成日期：2020 年 2 月 5 日

目录

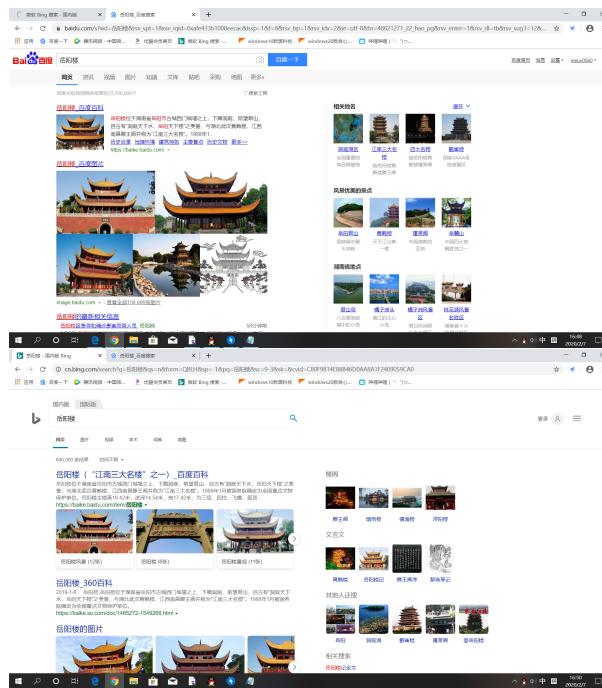
1 问题 1	1
1.1 搜索质量的差异	1
2 问题 2	3
2.1 如何进行高质量的搜索	3

第一章 问题 1

1.1 搜索质量的差异



- 从搜索界面上来看，必应背景图更好看，之前 Goole 和百度都是白色背景图（现在百度和 Goole 支持自定义背景图），必应的背景图会每日更换，有图片索引支持查看图片中的资料，也可以切换回之前的图片，百度主页面搜索框下是新闻，Goole 搜索框下面是常用网页的链接，Bing 搜索框下面则什么都没有，它们都有菜单栏



- 从搜索内容上看、对于绝大多数网站，此项指标当属谷歌老大，收录量有时甚至是其他搜索引擎的好几倍，下来排名：谷歌 > 百度 > 必应。例如搜索岳阳楼百度前 2 条是直接与岳阳楼相关的内容，必应的前 4 条都是与岳阳楼直接相关的内容，必应和百度搜索最下面是岳阳楼的相关搜索，百度的右侧显示的是相关地名，景点，地点，以及一个热搜榜单，Bing 右侧显示相关搜索，楼阁，文言文，以及其他人搜索。个人认为必应的质量比百度高，但是百度的收录量很大，有些编程方面的文章百度搜到的更全面
- 如果以搜索结果精准度来说，三家的排位是谷歌、必应、百度。谷歌在搜索结果内容上是最好的，多年搜索域名的开发经验加上强大的技术优势，谷歌在搜索结果上居于全球领先地位。必应的众多优势在于微软拥有自己的内容提供公司。采用的是 Powerset 技术 (Powerset 是微软收购的一家搜索技术公司)。是微软借以取代原有的 Live Search 的新一代搜索引擎，但现在毕竟是刚开始，各个方面不是太完善，个人认为现在跟谷歌和百度比，还是稍差一点。百度是全球最大的中文搜索引擎，在中国居于霸主地位，更贴近中国客户的使用感受，在中文内容（音乐、视频、娱乐、图片等）方面有着先天优势，用起来更符合中国人的使用习惯，但个人认为在搜索内容和结果方面稍逊一筹。

第二章 问题 2

2.1 如何进行高质量的搜索

- 1, 搜索引擎搜索引擎的选择对于搜索结果的影响是最为关键的。综合性的搜索引擎（比如百度谷歌）能解决日常搜索百分之 80 的情况（二八定律），想要更快速的得到结果，你就要使用细分的搜索引擎，比如搜索英文资料时，你应该用”Google No Country Redirect “搜索，而不是”谷歌“搜索，搜索问题解答时，你应该用”百度知道“搜索，而不是”百度“搜索。剩下的那更具搜索价值那百分之 20 内容，你就要使用专业的搜索引擎，它不仅更快，而且也能搜索到综合性引擎搜索不到的内容。比如搜索最新鲜的资讯你需要用”微博“搜索引擎，论文搜索你要用”文献“搜索引擎，图片搜索你需要用专业的图片搜索引擎等。
- 2, 搜索表达式很多用户都还是停留在搜索框中输入一两个关键字，然后点击搜索按钮的阶段，这是非常低效和无用的。运气好你可以在第一页就得到结果，否则你需要不停的翻页来得到结果，学习一些搜索引擎常用的检索表达式是一件低投入高回报，受益终身的事情。而且搜索表达式在绝大多数引擎上都是适用的。检索表达式介绍示例-搜索关键词 A 的同时屏蔽关于关键词 B 的信息 apple -iphone| 同时搜索关键词多个关键词，| 可以用 OR 代替，OR 需要大写 apple|google, apple OR google””要求查询结果要精确匹配，不包括演变形式”Failure is the mother of success”* 只适用于英文，添加一个星号以表示任何未知或不确定的字词。”* is the mother of success”《》只适用于中文，要求查询结果是关于这部作品，而不是普通的词语《Baby》site: 仅从特定网站或网域获得搜索结果 site:qq.com inurl: 查找在 URL 地址里有搜索关键词的页面 inurl:qqintitle: 查找在网页标题里有搜索关键词的页面 intitle:qqfiletype: 查找 pdf,xml,xls,txt,doc,csv 等特定格式的结果 filetype:pdf qq 混合搜索同时使用多个检索表达式，比如：site:qq.com ”firefox” -quantum
- 3, 筛选结果搜索结果筛选的使用可以加速获得搜索结果，不要只是在逛淘宝时才用到搜索筛选，比如搜索视频是按播放次数最多排列，搜索答案时按最多点赞排列，新闻搜索时按最新发布排列，还可以筛选某一个时间阶段的内容。
- 4, 选择结果在得到比较满意的搜索结果时，你可以按住 Ctrl 键点击搜索结果就

会在新的标签页面打开，方便对比结果。如果点击的链接，因为一些原因打不开了，可以利用“网页快照”的功能来重新查看。如果想要知道这个网页的旧版本，可以使用“互联网档案馆”(Internet Archive)来查看。当搜索结果不完整或者不是最开始的资料来源时，可以选择其中的一串文字来重新搜索（10个汉字或者英文单词），在这一串文字的两端最好加上双引号（“”），这样能够更精准的匹配。