

# 百度网页搜索 Query-Url 质量评估标准

该文档是 query-url 质量评估的详细标准指南，内容较详细，当你遇到问题可以随时查看，**建议您可以先阅读 Part6 的快速导读**，了解本标准核心内容。

本文档主体内容包括以下方面：

- ✓ 评估基础知识：part1、part2、part3；
- ✓ Query-url 打分核心标准：part4；
- ✓ 分类别举例参考：part5；
- ✓ 标准快速导读：part6。

## 目录

<b>Part1 前言：百度网页搜索质量评估</b>	<b>5</b>
1、评估必备软件	5
1.1 firefox 浏览器	5
1.2 杀毒软件	5
1.3 adobe flash player、windows media、qvod、百度影音等播放器	5
1.4 迅雷、QQ 旋风、快车（flashget）等下载软件	6
2、评估必备基础知识	6
2.1 什么是 query-url 质量评估	6
2.2 搜索质量评估的目的	6
2.3 常用术语解释	6
<b>Part2 query-url 质量评估概述</b>	<b>7</b>
1、Query-Url 质量评估介绍	7
2、Query-Url 质量评估考虑的维度	8
2.1 相关性：query 与结果说的是不是一回事	8
2.2 需求强度：抓住主要需求	8
2.3 丰富程度：详细全面	8
2.4 有效性：能否真正满足	9
2.5 时效性：新	9
2.6 便捷性：找到目标信息要花多少时间和成本	10
2.7 权威度：英雄也要问出处	10
2.8 总结：不同维度适用 query 范围	10
3、Query-Url 打分分档介绍	11
4、Query-Url 打分原则	11
5、Query-Url 评估步骤	11
<b>Part3 理解用户需求</b>	<b>11</b>
1、如何判断 query 需求	11
2、query 背后的实际需求意图	12
3、query 主体表意是否确定与需求是否多样	13
4、query 需求的主次之分	13
4.1 主需求	13
4.2 重要需求/次要需求	14

4.3 小众/冷门需求 .....	14
4.4 特殊情况：泛需求 .....	14
5、query 需求意图类型 .....	14
5.1 寻址需求 .....	14
5.2 信息需求 .....	15
5.3 交互操作需求 .....	15
<b>Part4 五档打分标准 .....</b>	<b>15</b>
1、各档打分定义 .....	15
2、4 分——官方 .....	16
2.1 4 分的定义 .....	16
2.2 举例说明 .....	16
2.2.1 4 分典型例子 .....	16
2.2.2 没有 4 分的典型情况 .....	19
2.2.3 有多个 4 分结果的典型情况 .....	19
2.2.4 其他考量 .....	20
3、3 分——很好，对大多数人都非常有用 .....	21
3.1 3 分的定义 .....	21
3.2 举例说明 .....	21
4、2 分——比较有价值，对部分人较有帮助 .....	24
4.1 2 分的定义 .....	24
4.2 举例说明 .....	24
5、1 分——稍沾边，略微有帮助 .....	25
5.1 1 分的定义 .....	25
5.2 举例说明 .....	25
5.3 表面相关的差结果 .....	26
6、0 分——不相关或无用资源 .....	27
6.1 0 分的定义 .....	27
6.2 举例说明 .....	27
7、死链 .....	29
7.1 死链类型：网页安全提示 .....	29
7.2 死链类型：空页面或未能访问页面 .....	29
7.3 死链类型：页面报错提示 .....	29
7.4 死链类型：标准报错页面 .....	30
<b>Part5 特定类型打分举例说明 .....</b>	<b>32</b>
1、打分一些重要原则 .....	32
2、query 主体是否确定与需求是否多样 .....	32
2.1 query 主体确定，需求维度单一 .....	32
2.2 query 主体确定，需求多样 .....	33
2.2 query 主体表意不确定，需求多样 .....	35
2.4 特殊说明：需求维度的泛化与具体 .....	36
3、特定类型的 query 打分举例说明 .....	38
3.1 实体类 query .....	38
3.1.1 名人类 .....	38
3.1.1 非名人 .....	39

3.1.2 地点名 .....	39
3.1.3 公司名 .....	40
3.1.4 产品名 .....	40
3.1.5 学校名 .....	41
3.2 信息类需求 .....	42
3.2.1 Key-Value 需求 .....	42
3.2.2 原文需求 .....	42
3.2.3 问答类 query .....	44
3.3 交互操作需求 .....	44
3.3.1 在线预订、购买需求 .....	44
3.3.2 游戏需求 .....	45
3.3.4 图片浏览 .....	45
3.3.5 小说阅读 .....	45
3.3.6 软件下载需求 .....	46
3.3.7 视频观看需求 .....	47
3.4 寻址类型 .....	47
3.5 其他特定类型 query .....	48
3.5.1 色情需求 .....	48
3.5.2 Url 查询的 query .....	48
3.5.3 输入错误的 query .....	49
3.5.4 列表页需求 query .....	50
3.5.5 时效性需求 query .....	51
3.5.6 地域需求 query .....	54
3.5.7 纯英文 query .....	54
4、特定 url 页面打分说明 .....	55
4.1 百科类结果 .....	55
4.2 搜索结果页 .....	55
4.2.1 什么是搜索结果页 .....	55
4.2.2 搜索结果页的打分 .....	58
4.3 权限页面 .....	59
4.4 垃圾、作弊等低质页面 .....	60
<b>Part6 Query-Url 质量评估快速导读 .....</b>	<b>62</b>
1、评估必备软件 .....	62
2、评估必备基础知识 .....	62
3、query-url 质量评估概述 .....	62
3.1 什么是 query-url 质量评估 .....	62
3.2 Query-Url 质量评估考虑的维度 .....	63
3.3 Query-Url 打分分档介绍 .....	63
3.4 Query-Url 评估步骤 .....	64
4、理解用户需求 .....	64
4.1 Query 需求的主次之分: .....	64
4.2 query 需求意图类型: .....	64
5、五档打分标准 .....	65
5.1 4 分——官方 .....	65

---

5.2 3分——很好，对大多数人都非常有用 .....	65
5.3 2分——比较有价值，对部分人较有帮助 .....	66
5.4 1分——稍沾边，略微有帮助 .....	66
5.5 0分——不相关或无用资源 .....	66
5.6 死链 .....	66
6、特定类型打分说明 .....	67

## Part1 前言：百度网页搜索质量评估

欢迎你加入百度网页搜索质量评估团队，作为百度搜索质量评测员，在接下来的日子里，你会参加众多类型的评测任务。本文主要介绍“Query-Url 质量评估”的评估指南说明。

请你接下来花些时间仔细阅读如下内容，这将非常有助于你完成“query-url 质量评估”的评测任务。同时，“Query-Url 质量评估”是其他众多评测任务的基础，做好该评测类型也将有助于你很好完成其他类型评测任务。

### 1、评估必备软件 y

为了你能正常进行评估任务，请确保安装如下软件：

#### 1.1 firefox 浏览器

浏览器是我们工作中最基本的软件。为了保证你正常的工作，请使用 firefox 浏览器，并且一般升级到最新的版本。可从此处下载：<http://www.firefox.com.cn/download/>

但因为 firefox 对某些页面支持不够好，导致页面内容异常，如遇此情况请更换 IE 浏览器查看。常见的 Firefox 支持不够好的页面包括音乐播放页、视频播放页、软件下载页等。如下几个例子：

- ✓ Firefox 不支持迅雷下载需要更换 IE: <http://kuai.xunlei.com/d/GVKWITFIEELJ>
- ✓ Firefox 不支持在线听歌需要更换 IE: [http://www.2651.cn/Html/2/Play\\_5880.html](http://www.2651.cn/Html/2/Play_5880.html)
- ✓ Firefox 不支持在线观看需要更换 IE: <http://www.ddoo.cc/jq/14134/play-0-0.html>

#### 1.2 杀毒软件

在评估中，你需要浏览打开各种各样的 url 网页，其中有些可能会含有病毒、损害你的电脑，所以建议你安装好杀毒软件。

下面一些类型在页面抓取中常常有问题，需要你实际点击 url 来重新下载查看，这些格式的文件一般安全的，不太可能有病毒，比如：

- ✓ .txt (text file)
- ✓ .ppt 或 .pptx (Microsoft PowerPoint)
- ✓ .doc 或 .docx (Microsoft Word)
- ✓ .xls 或 .xlsx (Microsoft Excel)
- ✓ .pdf (PDF file)

#### 1.3 adobe flash player、windows media、qvod、百度影音等播放器

现在互联网上有很多网页，特别是含有视频、音乐等页面，需要已经安装了 Adobe 的 flash player 为最新版本，才可正常访问；部分视频需要安装 qvod 播放器等才能播放，举例说明：

- ✓ 需要安装 flash player 的: <http://www.chevrolet.com.cn/brandsite/>
- ✓ 需要安装 Qvod 的: <http://www.kxkdy.com/tid-8831/>
- ✓ 需要安装百度影音的: <http://www.qire123.com/mainland/anxian/>

## 1.4 迅雷、QQ 旋风、快车 ( flashget ) 等下载软件

有些评估需要确认下载链接是否有效，而且通常会限定使用特定软件来下载，所以您要安装迅雷、QQ 旋风、快车等常用下载软件来验证是否为有效下载。比如：

需要安装迅雷才能下载的: <http://kuai.xunlei.com/d/GVKWITFIEELJ>

## 2、评估必备基础知识

### 2.1 什么是 query-url 质量评估

本问说的 query-url 质量评估，核心是根据 query 需求，衡量搜索结果（即一个 url 页面）的好坏。评估的标准是看搜索结果是否能提供用户想找的信息，满足了用户的需求，以及在多大程度上满足了用户的需求。这里面包含两个层次：

（1）搜索结果是否可以提供用户想找的信息。如果结果正是用户所找，可以满足用户需求，则是好结果；如果结果里没有用户想要的信息，没有用户关心的内容，则是差结果。

（2）多大程度上满足了用户的需求。在能满足用户需求的前提下，用户获取信息越丰富、全面、权威越好；成本越低越好，成本包括花费的时间、操作的步骤、用户需要的信息是否可以明显的展现出来等。

### 2.2 搜索质量评估的目的

百度的使命是“让人们最便捷地获取信息，找到所求”，目标是最大程度的满足用户需求，不断完善搜索质量，提升用户体验。

Query-url 质量评估是提升搜索质量的一个基础且核心步骤，我们需要通过评估来发现目前的缺陷，从而提高搜索质量。搜索质量评估工作的成果，将影响到包括您在内的数亿百度用户的搜索体验。

### 2.3 常用术语解释

为了便于理解本文档，对一些常用的名词术语等进行如下说明：

#### ✓ Query

也可称查询词、关键词、搜索词，指用户输入到搜索框想要查找的字、词、符号等。以后作为 Query 出现的词，一般都会用“【”、“】”括起来。

#### ✓ Term

即组成一个 query 的字、词，如 query 【新浪微博】，拆开后的“新浪”、“微博”，我们称为 Term。

#### ✓ 用户需求/用户需求意图

简称“需求”，对于用户来说，来到搜索引擎是带着目的来的，他们键入搜索词是希望获得相应的满足，比如获得相应的信息、网上购物等，我们把用户 query 背后的这种诉求称为用户需求。

如搜索【新浪微博】，用户需求是找站点 <http://t.sina.com.cn/>，进入微博查看信息、发微博等；搜【如何减肥】，是想减肥，找减肥的方法产品活动等。我们把用户输入的“query”背后表达的意图称为“用户需求”。

#### ✓ 搜索结果

指搜索引擎（如百度、google 等）根据 Query 返回的搜索结果，可以指具体某一条 Url 页面，也可以指所有的搜索结果。

#### ✓ URL

也称为网页地址，简称网址。百度主页页面的 url 是“<http://www.baidu.com/>”。本文用 url 指就是我们需评估的一个页面，但是需要注意，我们要评估的是这个 url 页面载入后这个页面上所提供的内容，而不是 url 本身。所以 query-url 评估打分，即根据 query，对该 url 页面的打分。

#### ✓ 网站主页

我们也将网站的主页面称为网站首页，这是我们登入、浏览该网站时最先看到的主要页面。网站主页的 url 往往以.com\ .edu\ .org\ .gov 等结尾，如 <http://www.apple.com/>就是苹果电脑公司的网站主页，新浪网站的主页为 <http://www.sina.com.cn/>，又例如 <http://www.mcdonalds.com/>是麦当劳的网站主页。

#### ✓ 网站子页面

网站中凡是非主页的页面，我们都称之为该网站的子页面。比如说 <http://www.apple.com/iphone/> 就是苹果公司网站的一个子页面，又例如麦当劳的一个子页面为：

[http://www.mcdonalds.com/usa/rest\\_locator.html](http://www.mcdonalds.com/usa/rest_locator.html)。

#### ✓ 主题

也称为 query 表意，一个 query 的主题即 query 表达的意思是什么，用户输入一个 query 是想在搜索引擎中找到与该 query 主题相关的网页等内容。

#### ✓ 页面内容

在页面上，广告、赞助商链接等不作为页面内容来进行评估，我们需要评估的是页面的主体信息，如站长发布的内容，即网页自身上所能提供的信息，评分时不应该根据广告来进行评估，即使某些广告可能看上去是相关的。

#### ✓ 需求满足度

也称需求满足、满足程度，即一个 url 页面对用户需求满足有多大帮助和价值的衡量，满足程度高的页面对用户非常有帮助，反之满足程度低的页面对用户价值很低。**需求满足是衡量搜索引擎质量的最重要因素，是在评估打分中需要重点考虑的。**

#### ✓ 有效性

在评估网页时，资源有效性是非常重要的评估要素，如视频是否能播放、软件是否能下载。一个资源有效高质量网页才能满足用户的需求，一个实际失效的网页，对于用户来说是没有用的。

## Part2 query-url 质量评估概述

### 1、Query-Url 质量评估介绍

Query-URL 质量打分是最重要的一种打分方式，是评价各搜索引擎好坏的基本方式之一。这种评估是根据 Query 背后的需求，去评价某条结果（即 URL 页面）对用户需求的满足程度。



## 2、Query-Url 质量评估考虑的维度

Query-Url 质量评估，即评价结果的需求满足，看结果能给用户带来多少的帮助和价值。影响用户需求满足的因素很多，我们列出常见的七种。

### 2.1 相关性：query 与结果说的是不是一回事

相关性，是指结果主题跟用户需求的主题匹配程度，两者说的是不是一回事，有没有跑题。相关，是一个好结果最基本的要求。后面其它的因素都建立在其之上。如果不相关，那肯定是差的结果。

下面列举几种典型的不相关的情况：

(1) 丢了 Query 中的重要成分。搜【**阉**怎么发音】，“阉”没了。

[基于共振峰模式的汉语普通话中韵母发音水平客观测试方法的研究--...](#)  
**【摘要】：**提出了一种以元音的共振峰模式为特征基于支持向量机算法的分类评估方法，用以对汉语普通话中的韵母发音水平进行客观测试。此算法为每个韵母分别训练全分类...  
[www.cnki.com.cn/Article/CJFDTotal-XIBA200 ...](http://www.cnki.com.cn/Article/CJFDTotal-XIBA200...) 2011-4-9 - [百度快照](#)

(2) 拆散了 Query，意思完全偏离。查【**林**字有多少笔画】，出来“**爨**字共有多少笔”。虽然 query 中的 term 在页面都有命中，但是因为拆散后主题完全偏离。

[你知道“爨”字共有多少笔画吗？旅游攻略 驴妈妈旅游攻略](#)  
 旅游攻略旅游攻略旅游攻略你知道“爨”字共有多少笔画... 村名 爨底下的“爨”字，共有三十笔，发CUAN音...**林**字腰、大字下面加火烧，大火烧林，越烧越兴，岂不...  
[guide.lv mama.com/2011/0111/122283.html](http://guide.lv mama.com/2011/0111/122283.html) 2011-3-3 - [百度快照](#)

(3) 含意转变（转义）。搜【**李文**】，出来“李文和”。

[李文和案 新闻中心 新浪网](#)  
 一年前的9月14日，华裔科学家**李文和**为自己洗脱了“间谍罪”的不白之冤，走出了监狱，但根据他向法院提出的保证，一年内他将不得不随时接受联邦调查局的问讯。...  
[news.sina.com.cn/world/liwh/](http://news.sina.com.cn/world/liwh/) 2011-1-21 - [百度快照](#)

### 2.2 需求强度：抓住主要需求

搜同一个 Query，不同用户的需求可能不同。此时，要站在多数用户的角度来判断，越多人查找的需求越重要。主要需求是最多人找的，满足主要需求的结果，是最好的结果；满足冷门需求，则结果差些。

如【**苹果**】，主要需求是苹果品牌、产品相关，而水果、电影的主题表意需求较少。

### 2.3 丰富程度：详细全面

多数的查询，越详细的资料内容越好，例如文档、材料、介绍、评论、相关内容等泛需求、又或者一些本身有多种方法需求的 query，举几个简单例子：

- ✓ 搜【**感动的作文**】，通常用户会参考多篇作文，这时提供多篇结果比一篇好；
- ✓ 找某个新闻事件，专题要比单篇好；



- ✓ 搜索某部电视剧，如【一起又看流星雨】，全集通常比某一集好；
- ✓ 搜索【电脑无法开机】，由于原因可能有多种，分析地越详尽越好。

※ 但是要注意，但也有一些 case，由于所求内容相对简单单一，可能一两句话就能满足的很好，这种情况不应当认为文字很少丰富度低，而做出降档打分。

Query	需求	URL 与打分	说明
横看成岭侧成峰的下一句	信息需求。找这个诗句的下一句是什么	<a href="http://zhidao.baidu.com/question/309201850.html">http://zhidao.baidu.com/question/309201850.html</a> ; 3 分，完全满足了用户需求	这个 query 需求很明确，不需要有非常丰富的要求，只要能满足用户需求就是较好的结果

## 2.4 有效性：能否真正满足

用户需求的满足是完整的过程。一定要注意用户搜索的最终需求是什么，这个结果能不能满足最终的需求，只有能满足最终需求的结果才是好结果。

(1) 需要检查资源是否有效，如

- ✓ 搜软件下载，不能看到页面上出现“下载”这几个字就到此为止。用户需求是下载安装使用该软件，要准确地评估，必须尝试下载，看其是否真实可用；若是冒牌伪装，或者有病毒木马，那就不是好的结果。
- ✓ 搜【小学数学教学论文】，出来列表页，都是小学数字方面的论文，看起来很好，但是请再点击一两篇论文，看看其是否相关，能否查看阅读。

(2) 查看资源是否权限限制，如

- ✓ 找某文档，要注册、或花钱才能下载；
- ✓ 想看电影，要装插件。  
多数用户都希望直接、免费、快速地得到；装软件、注册、花钱，大大增加了用户的成本，不能算高质量。

## 2.5 时效性：新

有些查询，对时间有着较高的要求，期望得到最新的信息，常见的是突发新闻的搜索。哪怕没有明确指定什么新闻，给出最新的内容也是很好的。

如【上海大火】，自 20101115 后相当长一段时间，都是找 20101115 发生的胶州路大火；20120227 前后几天搜索【奥斯卡】，主要需求是找 84 届奥斯卡金像奖。

招聘、交易等信息对时效性要求高，内容的生成时间极其重要。如果过期了，是陈旧的信息，那对用户通常没太大价值；找某方面的招聘信息，若招聘已结束，价值也很低。几个月前的物品交易信息，很可能交易已完成，现在找到没太大价值；找【电脑组装配置单】，两年前的配置单，给用户的帮助很小。【U 盘价格】，一年前的价格对现在参考价值也不是很大。

## 2.6 便捷性：找到目标信息要花多少时间和成本

便捷性，指从打开页面开始，到寻找到所需信息、达到最终目的，所要花费的时间和成本。如果页面有十来屏，所需信息不在第一屏，很多人可能会放弃。如果页面凌乱，或广告晃眼，或广告很多，同样影响用户。

好的结果应当让用户很方便的找到自己要的信息，比如在第一屏显著位置给出目标信息，用户可以轻易地看到。

## 2.7 权威度：英雄也要问出处

权威性考察的是内容发布在不同网站、或内容贡献者的身份给用户的不同感觉。如：

- ✓ 同样的软件，都能下载，用户更倾向点击官网，以及自己熟知的的华军、天空等知名软件下载站上的结果；
- ✓ 同样一篇新闻，出现在新浪新闻、新华网上，跟 bbs、个人博客上的转载相比，前者会更权威；
- ✓ 医治健康、法律相关问题，医生、律师的回答，肯定更让人们信服；
- ✓ 官方网站发布的信息，比其它网站更可信。

## 2.8 总结：不同维度适用 query 范围

这 7 个维度在不同特征的 query 下，关注的重点有不一样，不同维度适用的 query 范围说明如下：

维度	适用哪些 query	举例说明
相关性	所有	不相关，如丢掉了核心词肯定是 0； 相关，也有可能是 0，如软件虚假，视频不能播放。
需求比例	需求有多个	需求分析时，重要的是分析主要需求是什么； 非主要需求，最高只能 2 分。
丰富程度	需要详细的内容	搜作文范文、报告总结、人物资料等，丰富度越高越好，如【实习总结】、【杨幂资料】 如一句话、数字等简短的内容即可满足，则丰富程度要求不高，如【刘备字什么】、【横看成岭侧成峰的下一句】
有效性	大多为交互操作类 query，如下载、视频、试听、文档阅读等，只有资源有效才能满足用户需求的	下载，一定要尝试能否下载，是否真实 如视频，要看内容是否查看，是否真实 如物流查询，要真能查，才有价值。
时效性	明确或隐含着时间要求的查询	近期事件（如地震）、招聘、交易信息等，对时效性要求高； 高考等周期性事件也有时效性要求； 中国人口等隐含最新数据的时效性要求；
便捷性	所有	用户要找的位于显著位置，第一眼就能看到为便捷性好；当比较难以在页面方便的找到时，即为获取信息满足成本高，会稍差；
权威度	软件下载、商品交易、医疗健康等对权威有要求	来自知名网站、专业网站，让人放心。软件下载、医疗健康、银行等，对权威方面的要求高；

### 3、Query-Url 打分分档介绍

打分类别	含义	描述
<b>4 (Vital)</b>	官方	官网权威页面（详细说明见 Part4 的 2）
<b>3 (Useful)</b>	很好，对大多数人都非常有用	该 url 页面能满足大多数用户需求，很有价值
<b>2 (Relevant)</b>	比较有价值，对部分人较有帮助	该 url 页面对较多或一些用户较有价值
<b>1 (Slightly Relevant)</b>	稍沾边，略微有帮助	该 url 页面对大多数用户价值不高，但是跟 query 相关，少数用户觉得该页面有点帮助
<b>0 (Off-topic OR Useless)</b>	不相关或无用资源	该 url 页面毫无价值或价值很低
<b>死链</b>		需求无法判断或无法被评估的页面（详细说明见 Part4 的 7）

### 4、Query-Url 打分原则

作为评估人员，需要将自己当成实际输入这个 query 的用户，去评判各个结果能否满足自己的需求，从而判断各个结果的好坏，并给出适当的评分。

站在“用户”的角度，评测人员把自己当成真正的用户，这一点非常重要。

### 5、Query-Url 评估步骤

整个评估过程可划分为三个步骤：

- ✓ 需求判断：分析推断用户搜索这个 query 是想找什么，确定用户需求；
- ✓ 页面质量判断：查看实际 url 页面，判断结果能否以及多大程度上满足用户需求；
- ✓ 打分：根据本评估指南，选择合适的分值。

## Part3 理解用户需求

用户为满足自己某种需求而来到搜索引擎的，判断用户的需求是准确评估打分的必备基础。只有准确地判断出用户需求，了解用户搜索的目的，才能合理地衡量结果的质量好坏。

### 1、如何判断 query 需求

多数 Query，本身表述清楚，含义明确，以其表达出来的意思为准。如【让子弹飞 剧情】、【繁体在线转换】、【舞蹈教学视频】等。

有些 Query 依靠个人的知识、经验能够理解，但还有一些可能需要通过查询相应的百科（如 <http://baike.baidu.com>）、用不同的搜索引擎进行搜索，通过其检索结果学习了解背景知识后，从而判断出搜索的目的。

如【bl 怎么去诺森德】，如果你不玩魔兽世界，单看 Query，很难弄明白其搜索意图。通过查看几个搜索结果后，知道用户是找“魔兽世界部落的玩家怎么去诺森德”。因此，建议遇到不了解的 Query，先拿 Query 到百度、Google 等搜索引擎上搜索了解一下。

※ 注意：当通过搜索引擎来确认用户需求时，千万不要直接将排在前面的结果当做好结果，将排在后面的结果当做差结果。这应当只作为你确认用户需求的方式之一，而非直接根据搜索引擎现有排序情况进行打分。

另外，当 query 需求多样时，我们比较难判断用户的主要需求有哪些方面，我们可以借助百度搜索建议提示，了解用户的主要需求。如：【天龙八部】有多个含义，游戏、小说、电视剧、佛教本义等，那到底哪个是主需求呢。看如下搜索建议提示，多是游戏相关，前两个“天龙八部 3”、“天龙八部 3 官网”，其中多数是关于网游的，那么可以推断搜【天龙八部】的主要需求是找网络游戏。

【天龙八部】的搜索框提示：



## 2、query 背后的实际需求意图

在理解 query 本身的含义后，接着你应当考虑用户输入这个 query 时，他所需要得到实际满足的需求是什么，并以此来对所给的 url 做出打分。

例如【俄罗斯方块】，用户需求很可能是在线玩这个游戏，也有可能是要下载它；【还珠格格第一部】，大多数输入这个 query 知道这是一部电视剧，需求意图是想看这个电视剧。下面是一些用户需求意图的例子：

Query	用户需求意图
圆通快递	查询快递单号，查找附近的网点等
京东	访问京东官网
万年历	在线查询万年历，或下载万年历软件
猴头菇	了解相应的介绍、做法等

### 3、query 主体表意是否确定与需求是否多样

除了需求非常明确的 query，如【让子弹飞 剧情】、【横看成岭侧成峰的下一句】，不少 query 都存在多样的需求。针对 query 需求的多样，我们会从下面 3 个角度理解：

✓ Query 主体表意：

存在 query 主体表意是确定还是非确定的区分，如许多 query 表意多样，即同样的 query，表达的意思却差别万分。如：

【苹果】，有苹果品牌（ipad、iPhone 对应的品牌）；有苹果这种水果；还有电影《苹果》，这些都可能是用户要找的东西；【土豆】，用户的需求可能指视频站点或者食物等。

✓ Query 需求维度：

存在需求维度单一还是多样的区分，如【QQ 下载】需求维度单一，即软件下载的需求；而【步步惊心电视剧】存在多种需求维度，包括电视剧观看、电视剧剧情等；

✓ Query 需求强度：

即针对 query 不同主体表意和需求维度，存在需求强度大小之分，分为主要需求、泛需求、小众/冷门需求三类，后面详细介绍。

根据需求主题表意和维度，主要可以分为下面三种情况：

- ✓ query 主体确定，需求维度单一，如【美图秀秀在线下载】；
- ✓ query 主体确定，需求多样，如【赵薇】；
- ✓ query 主体表意不确定，需求多样，如【苹果】。

### 4、query 需求的主次之分

百度的用户数以亿计，搜索同一个 Query，不同人可能有着不同的需求，搜索引擎无法同样程度满足所有人的需求，而是尽可能最大程度满足大多数用户的需求。这也要求我们评估时要尽量挣脱个人局限，站在大多数用户的角度来衡量，把握需求主次之分。

根据 query 不同表意下的需求强度，以及同样表意下不同需求维度的需求强度，需求强弱可能有以下几种情况：

#### 4.1 主需求

即输入这个 query 的用户中，大多数用户心目中存在的需求。

（1）多义 query 的表意主次之分：

如【windows】，当大多数用户键入这个 query 时，需求是微软的操作系统，而不是在墙上的玻璃窗户。在搜索引擎简单进行检索后，一般可以了解用户的主需求是什么。

（2）同义 query 下不同需求维度主次之分：除了利用搜索引擎、搜索引擎搜索提示等工具了解外，一般比较依赖你自身的常识判断，如：

- ✓ 对于软件来说，主需求往往是下载；
- ✓ 对于视频来说，主需求往往是在线观看；
- ✓ 对于小说来说，主需求往往是在线阅读；

## 4.2 重要需求/次要需求

即输入这个 query 的用户中，有部分用户存在的需求，根据需求的强度的略微差别，存在重要和次要之分。如 2012 年 1 月搜【城市猎人】，电影的需求即为次要需求；【亚马逊】，南美洲亚马逊河流的需求；

## 4.3 小众/冷门需求

在某些 query 下，有的表意和需求维度比较冷僻，用户需求的比例比较低或非常低，并且根据强度上略微差异，称为小众或冷门需求。如【城市猎人】，绝大部分人都是找影视结果，很少有人的需求是一个 blogid 为“城市猎人”的博客。该博客在这个 query 下，就称之为冷门需求，认为用户在这个 query 下想找这个结果的需求很低。

在某 url 满足的是 query 的小众/冷门需求时，往往评分比较低。

## 4.4 特殊情况：泛需求

在某些 query 下，几乎没有占绝对主导地位的表意或同义表意需求维度上绝对强弱之分，都有一定比例的用户需求，这样情况下根据需求强弱的差别打分，如【长沙】、【书籍】，以【倾国倾城】为例说明：

- ✓ 找“倾国倾城”的意思，如 <http://zhidao.baidu.com/question/298263788.html>；
- ✓ “倾国倾城”的音乐，如 <http://www.yue365.com/play/3790/187195.shtml>；
- ✓ “倾国倾城”的小说，如 <http://www.readnovel.com/book/56214/>；

这些释义中，并没有哪一个是绝对主需求，认为 query 的需求很泛，其中音乐和意思的需求略强于小说等其他需求，不同的用户想得到不同表意相关的解释。

※ 注意：该类 query 因为没有绝对主需求，所以没有 4 分结果；

# 5、query 需求意图类型

根据用户输入 query 时的需求意图不同，我们将用户需求大致分为 3 类：寻址需求、信息需求、交互操作需求。需要注意的是，有很多的 query 都会有多种需求，而每种需求可能是不同意图类型，所以需求意图都用来表达某个 query 下的某种需求是什么类型的。

Query 某种需求的类型分为交互操作类、信息类、寻址类这 3 种，可以用“Do-Know-Go”这个来简单表示，这种分类方法有助于判断一个页面该怎么打分，具体介绍如下：

## 5.1 寻址需求

**定义：**寻址需求就是指，用户输入这个 query，就是为了去某个特殊的页面，而且是用户已有确认的某个页面、或某个网站。用户希望通过搜索某 query，来访问该页面。因此，在寻址需求中，最能满足用户需求的就是用户所求的那个 url。

**典型范例：**

Query	用户需求意图
ibm	<a href="http://www.ibm.com/cn/zh/">http://www.ibm.com/cn/zh/</a>



youku	http://www.youku.com/
京东	http://www.360buy.com/
北京大学	http://www.pku.edu.cn/
人人网登陆	http://www.renren.com/

## 5.2 信息需求

**定义：**用户的需求是获取信息，想了解某信息、某知识，搜索引擎的目标是让用户找到所求信息，给出相关的优质、权威、全面的信息。

**典型范例：**

Query	用户需求意图
西夏王陵	信息介绍，关于该景点和旅游
DHC g10 拆机方法	如何拆机，拆机的步骤
如何清除油渍	找清除油渍的方法

## 5.3 交互操作需求

**定义：**用户是要从事某项操作和活动，往往是需要某个 url 上，与页面进行一定的交互，从而达到用户的目标、完成相应的操作。一般对资源是否有效从而完成操作有较高要求。

**典型交互需求：**

• 购买产品 \ • 下载免费或需购买的软件 \ • 网上支付帐单 \ • 在线玩游戏 \ • 预订酒店 \ • 听音乐 \ • 观看视频 \ • 查看图片 \ • 网上调查

**典型具有交互需求类别需求的 query:**

Query	用户可能的需求意图
甲壳虫乐队海报	查看或下载对应图片
Adobe Reader 下载	软件下载
在线职业测试	在线做职业测试
汇率兑换	在线进行汇率兑换查询
龙门飞甲 DVD	在线购买或下载 DVD
小游戏	在线玩儿小游戏
英文翻译	在线进行中英之间翻译

## Part4 五档打分标准

### 1、各档打分定义

Query-url 相关性打分提供了 6 种打分选项，包括 0-4 打分和抛弃。基于用户需求意图分析和 Url 页面对用户需求满足程度，进行 4、3、2、1、0 的打分。而对于特殊情况无法进行打分的页面，进行“抛弃”的标记。具体说明如下：



打分类别	含义	描述
<b>4 (Vital)</b>	官方	官网权威页面
<b>3 (Useful)</b>	很好：非常有价值	该 url 页面能满足大多数用户需求，很有价值
<b>2 (Relevant)</b>	不错：较有价值	该 url 页面对较多或一些用户较有价值
<b>1 (Slightly Relevant)</b>	略相关、有点帮助	该 url 页面对大多数用户价值不高，但是跟 query 相关，少数用户觉得该页面有点帮助
<b>0 (Off-topic OR Useless)</b>	跟 query 主题完全无关或无用资源	该 url 页面毫无价值或价值很低
<b>死链</b>	即页面死链导致无法判断	死链导致无法被评估的页面

## 2、4分——官方

### 2.1 4分的定义

4分仅适合于某些特殊情况，对 query 和 url 页面都存在着特殊的要求：

#### (1) 适用 query：主需求明确

对 query 的基本要求是：主需求很明确，主要语义表意意图的网站页面才打 4 分，否则都不能打 4 分。如 query 【当当】，存在主要表意需求“著名电子商务网站”，所以 <http://www.dangdang.com/> 可以打 4 分，而如果是某个“当当”的制鞋公司则不能给 4 分。

#### (2) 适用 url：官方页面

- ✓ 官方：Url 是 Query 对应主要需求的官方网站、或官网上对应的页面，如人名、地名、机构、产品、公司等官网等；
- ✓ 首发：用户想找报道、小说、论坛上的帖子等文字作品，结果是作者博客上的原文、该作品的首发地址，这种也可以认为是官方的页面；

#### (3) 特殊说明：4分不一定是最有价值

4分的定义中，我们并没有对结果的质量好坏提出非常明确的要求。当然，4分的结果往往都是**非常有帮助的、排版良好的**，但是这并非绝对条件。**对于某些 query，官方结果可能并非互联网上最有帮助的页面**，但由于其官方、原创性，标记 4 分是合理的。

### 2.2 举例说明

下面将针对一些特定类型进行举例说明什么情景下打 4 分，什么情景下不能打 4 分：

#### 2.2.1 4分典型例子

##### (1) 寻址类 query

下面是一些用户找某个具体网站的 query 的例子：

Query	用户需求意图	4分 URL	Url 说明
qq 空间	去 qq 空间官方页面	<a href="http://qzone.qq.com/">http://qzone.qq.com/</a>	QQ 空间的网址
4399 小游戏	去 4399 游戏网站玩游戏	<a href="http://www.4399.com/">http://www.4399.com/</a>	4399 官方网站

优酷	去优酷视频的官网看视频	<a href="http://www.youku.com/">http://www.youku.com/</a>	优酷官网
新浪微博	去新浪微博的官网查看、发布微博等	<a href="http://weibo.com/">http://weibo.com/</a>	新浪微博的官方页面

## （2）实体类 query

对于实体类 query，如人名、机构、公司、电影、事件等。有些是寻址类需求，有些可能是找相关信息的，这些实体 query 的官网都认为是官方权威的结果，可以打 4 分。

Query 类型	Query	4 分 URL	Url 说明
名人	姚晨	<a href="http://t.sina.com.cn/yaochen">http://t.sina.com.cn/yaochen</a>	姚晨本人官方的微博
公司	Lenovo	<a href="http://www.lenovo.com.cn/">http://www.lenovo.com.cn/</a>	Lenovo（联想）中文官方网站
大学	北京大学	<a href="http://www.pku.edu.cn/">http://www.pku.edu.cn/</a>	北京大学的官方网站
政府机构	铁道部	<a href="http://www.china-mor.gov.cn/">http://www.china-mor.gov.cn/</a>	铁道部官方网站
银行	招商银行	<a href="http://www.cmbchina.com/">http://www.cmbchina.com/</a>	招商银行的官网
事件	上海世博会	<a href="http://www.expo2010.cn/">http://www.expo2010.cn/</a>	2010 年上海世博会的官网
产品	Iphone	<a href="http://www.apple.com.cn/iphone/#">http://www.apple.com.cn/iphone/#</a>	Iphone 产品的官方页面
景点	故宫	<a href="http://www.dpm.org.cn/index16801050.html">http://www.dpm.org.cn/index16801050.html</a>	故宫的官方网站

## （3）指定网站/网页 query

用户的需求是某个网站内的子站或子页面，那么对应的站点能很好满足需求的页面可以给 4 分。

Query	用户需求意图	4 分 URL	Url 说明
佳能 ixus210	产品信息、购买需求	<a href="http://www.canon.com.cn/products/camera/ixus/210/">http://www.canon.com.cn/products/camera/ixus/210/</a>	佳能官网上的 ixus210 介绍
4399 赛尔号	4399 网站上的该游戏	<a href="http://www.4399.com/flash/seer.htm">http://www.4399.com/flash/seer.htm</a>	4399 上的赛尔号
交行信用卡中心	交行官网上的信用卡频道	<a href="http://creditcard.bankcomm.com/">http://creditcard.bankcomm.com/</a>	官网上对应的页面
美人心计 奇艺	奇艺网站上的对应影片	<a href="http://www.qiyi.com/dianshiju/mrxj.html">http://www.qiyi.com/dianshiju/mrxj.html</a>	视频网站奇艺上的《美人心计》电视剧，是用户明确指定网站的完全满足需求结果，可以给 4 分。

## （4）人名 query

所有的 4 分结果，都要求 query 主需求明确。对于人名 query 这一类相对特殊的 query 来说，根据 query 本身的特征做出一些分类，从而帮助判断是否可以有 4 分结果。如下：

Query 类别	Query 范例	是否有 4 分结果
著名人名	周杰伦、奥巴马、布兰妮	可以；有绝对主需求
非名人但人名特指	Lalith Wickramanayaka	可以；能判断出明确指代某个人，是斯里兰卡的某个大学老师

不著名常见人名	安娜、小王、王丽娜、王超	不可；认为需求分散无主需求
常见人名+限定	王超 爱乐团	可以；如果范围被缩小，能对应到唯一的某人上，可以有 4 分

在有 4 分结果的人名 query 下，有些典型的范例：她（他）的官方网站、博客、微博、空间、社交网站（如人人、开心）上的个人主页等，如果确实是本人发布，或者已得到授权，那么可以给 4 分。

※ 注意：非人名的实体 query，例如是品牌、公司所具有的和维护的博客、微博等社交类网站，即使确实是官方的，那么也认为不可以达到 4 分标准；

范例：

Query	URL 页面	Url 说明
姚晨	<a href="http://t.sina.com.cn/yaochen">http://t.sina.com.cn/yaochen</a> <a href="http://blog.sina.com.cn/yaochen">http://blog.sina.com.cn/yaochen</a>	姚晨的微博和博客；4 分
华章图书	<a href="http://t.sina.com.cn/huazhangbook">http://t.sina.com.cn/huazhangbook</a> <a href="http://site.douban.com/hzbooks/">http://site.douban.com/hzbooks/</a>	它的新浪微博、豆瓣小站；不可以 4 分

### （5）Url 查询 query

有些 query 直接就是某个具体的 url，或者看起来和 url 非常相似，称之为 url 查询；

典型 url 查询：

- ✓ Query 本身就是一个完整的 url，范例：www.youku.com
- ✓ 只要添加 www 这样的常见前缀即可成为完整 url：360buy.com(可能访问时会自动加上前缀)
- ✓ 有些虽然没有前缀，但确实也是完整 url 的：xunlei.com

对于 query 查询来说，主需求就是这个地址本身：

Query	用户需求意图	4 分 URL	Url 说明
Youku.com	对应页面	<a href="http://youku.com/">http://youku.com/</a>	即对应 url
<a href="http://www.360buy.com">www.360buy.com</a>	对应页面	<a href="http://www.360buy.com/">http://www.360buy.com/</a>	即对应 url
<a href="http://www.jjwxc.net/onebook.php?novelid=19520">http://www.jjwxc.net/onebook.php?novelid=19520</a>	对应页面	<a href="http://www.jjwxc.net/onebook.php?novelid=19520">http://www.jjwxc.net/onebook.php?novelid=19520</a>	即对应 url
Renren.com	对应页面	<a href="http://renren.com/">http://renren.com/</a> <a href="http://www.renren.com/">http://www.renren.com/</a>	即对应 url

### （6）首发地址 Url

Query 是文章、小说名等，作者原创首发帖子、博客、空间、文章等可以给 4 分。如果文章是报刊发表的，则该报刊网站上的对应文章可以给 4 分。

Query	用户需求意图	4 分 URL	Url 说明
象一滴水一样 ——纪念高华老师	找这篇文章	<a href="http://www.drunkpiano-liuyu.net/?p=826">http://www.drunkpiano-liuyu.net/?p=826</a>	作者刘瑜博客上的原文。
官居一品	阅读浏览需求	<a href="http://www.qidian.com/Book/1343635.aspx">http://www.qidian.com/Book/1343635.aspx</a>	首发站：起点，该书主页

※ 注意：有时候比较难以判定这个官网、博客、社交网站等是否是官方的，**可以搜索看一下大家评论是否是官网；原创有些情况下比较难判断**，如果你仔细分析后还是不能确定是否原创、首发，请不要打 4 分，**而按照实际需求满足情况给予合适的打分**。

### 2.2.2 没有 4 分的典型情况

4 分结果为要求相当高的好结果，并非所有的 query 都有。可以认为大多数 query 没有 4 分的网页。以下几种情况是不会有 4 分页面的：

- ✓ query 没有主需求，需求比较分散；
- ✓ Query 并非一个实体、或者并非寻址需求、以及非原创地址、非 url 查询对应的 url；
- ✓ Query 并没有对应的官方网站或网页存在；
- ✓ Query 需求很泛，没有任何人或实体可以对应并完全覆盖这个 query，如蘑菇

无 4 分结果的 query 范例：

Query	说明
北大建校多少年	无 4 分页面，北大是实体名，但是整个 query 并非实体名
iPhone 评论	无 4 分页面，iphone 是实体名，但是整个 query 不是，找相关信息
DDD	无 4 分页面，可对应到多个不同实体的缩写上，无主需求
王超英勇事迹	无 4 分页面，是信息需求，不存在官网
谈南路小学	无 4 分页面，该学校无官方网站
心脏病	无 4 分页面，query 需求较泛，无对应的官方页面
北京旅游	无 4 分页面，用户不是明确地要寻找某条 url

### 2.2.3 有多个 4 分结果的典型情况

虽然 4 分的要求是原创、官方，但有时候确实有些 query 下不但具有 4 分结果，而且可能具有多个 4 分结果，打分可以一致的典型类别：

- ✓ 多个 url 对应的实际页面完全相同，例如有无 www 前缀，index 后缀等区别；
- ✓ 多个 url 具有相互跳转关系，例如 urlA 会自动跳转到 urlB；
- ✓ 多个 url 对于 query 来说都可以快捷满足用户需求，典型的例如：
  - 下载需求：软件官网首页与下载内页；
  - 网站/产品：某网站与其提供的特殊产品都有较大需求（注意：此非无主需求）。

（1）URL 不一样，实际页面完全相同

Query	4 分 URL	Url 说明
优酷	A: <a href="http://www.youku.com/">http://www.youku.com/</a> B: <a href="http://www.youku.com/index/">http://www.youku.com/index/</a>	不同的 url，但都是优酷，AB 之间无直接跳转
国航	A: <a href="http://www.airchina.com.cn/">http://www.airchina.com.cn/</a> B: <a href="http://www.airchina.com.cn/?cid=bptopt0301">http://www.airchina.com.cn/?cid=bptopt0301</a>	不同的 url，但都是国航，AB 之间无直接跳转

（2）URL 之间跳转关系，实际页面一样

Query	4 分 URL	Url 说明
新浪微博	A: <a href="http://t.sina.com.cn/">http://t.sina.com.cn/</a> B: <a href="http://weibo.com/">http://weibo.com/</a>	新浪微博的官方页面，A 总是会跳转到 B
nba	A: <a href="http://www.nba.com/">http://www.nba.com/</a> B: <a href="http://china.nba.com/">http://china.nba.com/</a>	A 根据访问所在地（中国）跳转到 B（中文网站），由于百度用户主要在国内，所以两条结果都给 4 分。
华章图书	A: <a href="http://www.hzbook.com/">http://www.hzbook.com/</a> B: <a href="http://www.hzbook.com/ps/">http://www.hzbook.com/ps/</a>	它的官方网站，A 总是会跳转到 B
交行信用卡中心	A: <a href="http://creditcard.bankcomm.com/">http://creditcard.bankcomm.com/</a> B: <a href="http://creditcard.bankcomm.com/bcms/index.htm">http://creditcard.bankcomm.com/bcms/index.htm</a>	官网上对应的页面,A 总是会跳转到 B

### （3）不同 URL 都可以快捷满足主要需求

Query	4 分 URL	Url 说明
163	<a href="http://www.163.com/">http://www.163.com/</a> <a href="http://mail.163.com/">http://mail.163.com/</a>	163 官网主页必定可以是 4 分，邮箱是【163】这个 query 绝对主要需求，也可以给 4 分
PPTV	<a href="http://www.pptv.com/">http://www.pptv.com/</a> <a href="http://download.pptv.com/">http://download.pptv.com/</a>	网站名，也是软件，主要需求是下载软件。主页和下载页都给 4 分。
酷狗音乐下载	<a href="http://download.kugou.com/">http://download.kugou.com/</a> <a href="http://www.kugou.com/">http://www.kugou.com/</a>	主页上可轻易下载，也给 4 分。
北京旅游局	<a href="http://www.visitbeijing.com.cn/">http://www.visitbeijing.com.cn/</a> <a href="http://www.bjta.gov.cn/">http://www.bjta.gov.cn/</a>	北京旅游局的官方网站及其办的政务网
斗破苍穹	<a href="http://www.qidian.com/Book/1209977.aspx">http://www.qidian.com/Book/1209977.aspx</a> <a href="http://www.qidian.com/BookReader/1209977.aspx">http://www.qidian.com/BookReader/1209977.aspx</a>	原创，起点中文网的“天蚕土豆”写的这部小说。
小木虫	<a href="http://emuch.net/">http://emuch.net/</a> <a href="http://emuch.net/bbs/">http://emuch.net/bbs/</a>	该网站主体是论坛，所以网站首页和论坛都给 4 分。

## 2.2.4 其他考量

### （1）语言问题

国际上一些大的机构、品牌，往往有多种语言的官网页面，对于这种情况，考虑到百度的用户都主要是中文用户，所以有语言障碍的页面应当酌情降分，典型原则：

- ✓ 具有简体中文官网：只有简体中文官网可以给 4 分，英文及其他语言给 3 分；
- ✓ 没有简体中文官网：英文\繁体中文官网可以给 4 分，该机构品牌所在地语言的官网为 3 分，如 The face shop 是韩国的品牌，那么韩语官网可以给 3 分，其他较冷门语言给 2 分。

### （2）URL-站内分站自动跳转

国内有一些网站，会根据访问者的 IP 自动跳转到访问者所在地的分站。在北京点击 58 同城网 <http://www.58.com/>，会跳到 <http://bj.58.com/>。但评估时，还是应该以所给的 url 自身为评估对象。举例说

明如下：

Query	URL	Url 说明
58 同城	<a href="http://www.58.com/">http://www.58.com/</a>	即使会自动跳转到当地分站，其他地点用户也可以满足，该结果仍然可以给 4 分
58 同城	<a href="http://bj.58.com/">http://bj.58.com/</a> <a href="http://jn.58.com/">http://jn.58.com/</a>	该结果分别为北京、济南当地分站，但该结果无法直接满足其他地点用户的需求，3 分

### （3）URL 地址的影响

有些 4 分结果中，query 与 url 具有包含关系、url 中包含 query，例如 Q=youku，[www.youku.com](http://www.youku.com) 就是对应官网，4 分结果中有一些就是 url 中包含 query 本身的。

但是需要注意的是，反之并不一定成立，有些 query 下即使 url 中包含 query，也不一定是 4 分结果。如【CS】，主需求是找游戏反恐精英，简称 CS 的游戏，即使有一个网站域名为 <http://www.cs.com.cn/>，那么也不是 4 分结果，所以还是需要根据用户的需求以及页面满足情况来仔细判断。

## 3、3 分——很好，对大多数人都非常有用

### 3.1 3 分的定义

3 分结果，认为应当是在该 query 下，对于大多数用户都非常有用的好结果。非常有用的 3 分结果一般应该具备的特征为：符合 query 需求、页面质量很高。

这些页面往往具有以下全部或者部分的典型特征：需求满足程度高、权威度高、页面资源真实可用（如可以播放、收听）、页面看起来整洁组织排版很好、页面有用信息容易找到，在时效性需求的 query 下（如突然爆发的新闻事件）时效性也很好。

注意：一个 query，不止一个 url 页面可以打 3 分。

### 3.2 举例说明

3 分常见例子：

- ✓ 搜网站名称，该网站下的热门频道或栏目；
- ✓ 搜电视剧，优酷、土豆、奇艺等网站上的专辑；
- ✓ 找免费软件，华军、天空、多特等网站的下载页；
- ✓ 搜范文例文，内容丰富的列表页；
- ✓ 搜商品，包含详细的介绍、价格、可信任可购买的商铺；
- ✓ 找明星，他/她的详细资料、专题、fans 为其创建的高质量网站；
- ✓ 找文章、小说，有全部内容的结果。

#### （1）信息需求

Query	用户需求意图	3 分 URL	Url 说明
如何解散 QQ 群	信息需求，想知道解散方法	<a href="http://zhidao.baidu.com/question/380620977.html">http://zhidao.baidu.com/question/380620977.html</a>	最佳答案直接提供有效方法



CSCO 股票	信息需求，想知道该公司的股市行情	<a href="http://www.imeigu.com/CSCO">http://www.imeigu.com/CSCO</a>	权威网站上提供的股票信息，页面优质可信
脑梗塞症状	信息需求，想知道该疾病症状是什么	<a href="http://jbk.39.net/keshi/waike/naowaikesign/48dc6.html">http://jbk.39.net/keshi/waike/naowaikesign/48dc6.html</a>	权威网站上的内容丰富的页面
2002 年的第一场雪 歌词	信息需求，想知道该歌曲歌词是什么	<a href="http://baike.baidu.com/view/447573.htm?#4">http://baike.baidu.com/view/447573.htm?#4</a>	权威结果，歌词直接显示出来
第 28 届中国电影金鸡奖获奖名单	信息需求，想知道具体的获奖人名单	<a href="http://yule.sohu.com/20111021/n322995564.shtml">http://yule.sohu.com/20111021/n322995564.shtml</a>	权威网站上列出的名单
刘家窑 九头鸟	信息需求，想知道这个店的评价、地址等信息	<a href="http://www.dianping.com/shop/4097309">http://www.dianping.com/shop/4097309</a>	该餐馆的大众点评页面，提供了电话、评论等各方面信息，对大多数用户都非常有用
5 的平方根	信息需求，想知道 5 的平方根是多少	<a href="http://zhidao.baidu.com/question/316044297.html">http://zhidao.baidu.com/question/316044297.html</a> <a href="http://zhidao.baidu.com/question/328047365.html">http://zhidao.baidu.com/question/328047365.html</a>	虽然页面丰富上略有差异，但实际均提供了准确答案，都可以到 3 分。
草长莺飞二月天的下一句	信息需求，想知道该原文的下一句是什么	<a href="http://zhidao.baidu.com/question/29877078">http://zhidao.baidu.com/question/29877078</a> <a href="http://wenwen.soso.com/z/q40927846.htm">http://wenwen.soso.com/z/q40927846.htm</a> <a href="http://ks.cn.yahoo.com/question/1590002322140.html">http://ks.cn.yahoo.com/question/1590002322140.html</a> <a href="http://zhidao.baidu.com/question/76842542.html">http://zhidao.baidu.com/question/76842542.html</a>	答案直接给出，回答正确

## （2）交互操作需求

Query	用户需求意图	3 分 URL	Url 说明
西城男孩告别巡演 广州站 订票	交互需求，想购买该演唱会的门票	<a href="http://piao.youa.com/loader?url=http%3A%2F%2Fpiao.youa.baidu.com%2Fshow%2F0%2Fticket%3Ftid%3Deeeeb7521010eb13bddec7e7">http://piao.youa.com/loader?url=http%3A%2F%2Fpiao.youa.baidu.com%2Fshow%2F0%2Fticket%3Ftid%3Deeeeb7521010eb13bddec7e7</a>	有啊票务预订页面，比较权威、可信，用户需求能够得到较好满足
QQ 下载	交互需求，下载为主需求	<a href="http://xiaozai.zol.com.cn/detail/38/371537.shtml">http://xiaozai.zol.com.cn/detail/38/371537.shtml</a> <a href="http://www.onlinedown.net/soft/20355.htm">http://www.onlinedown.net/soft/20355.htm</a>	Zol，华军是著名的软件下载站点，且软件能够下载，对大多数用户都非常有用
酷狗音乐下载	下载主需求	<a href="http://www.skycn.com/soft/14786.html">http://www.skycn.com/soft/14786.html</a> <a href="http://www.onlinedown.net/soft/23764.htm">http://www.onlinedown.net/soft/23764.htm</a> <a href="http://dl.pconline.com.cn/html_2/1/88/id=42960&amp;pn=0.html">http://dl.pconline.com.cn/html_2/1/88/id=42960&amp;pn=0.html</a>	这些都是知名权威的站点，下载链接多，保证无毒



		<a href="http://xiazai.zol.com.cn/detail/10/98401.shtml">http://xiazai.zol.com.cn/detail/10/98401.shtml</a>	
未央沉浮下载	交互需求，想下载对应的内容	<a href="http://t.book118.com/html/txt185361.shtm">http://t.book118.com/html/txt185361.shtm</a> <a href="http://www.4020.cn/Txt/XiaoShuo-28474.html">http://www.4020.cn/Txt/XiaoShuo-28474.html</a>	均无权限的完整可下载页面，可以符合用户需求
违章	交互需求，想查询车辆违章信息	<a href="http://auto.sohu.com/s2004/weizhangchaxun.shtml">http://auto.sohu.com/s2004/weizhangchaxun.shtml</a> <a href="http://www.ip138.com/weizhang.htm">http://www.ip138.com/weizhang.htm</a>	聚合页面，多个地方的违章结果查询链接，验证能够真正的查询
宫锁心玉	视频主需求	<a href="http://www.iqiyi.com/dianshiju/gsxy.html">http://www.iqiyi.com/dianshiju/gsxy.html</a> <a href="http://tv.sohu.com/s2011/gsxy/">http://tv.sohu.com/s2011/gsxy/</a> <a href="http://www.letv.com/ptv/pplay/36877.html">http://www.letv.com/ptv/pplay/36877.html</a> <a href="http://www.tudou.com/playlist/album/id65393.html">http://www.tudou.com/playlist/album/id65393.html</a>	包含全部剧集，视频清晰，顺序排列

### (3) 寻址需求

Query	用户需求意图	3 分 URL	Url 说明
优酷	找优酷视频网站	<a href="http://movie.youku.com/">http://movie.youku.com/</a> <a href="http://tv.youku.com/">http://tv.youku.com/</a>	优酷下的热门频道
当当	主要找当当电子商务网站	<a href="http://book.dangdang.com/">http://book.dangdang.com/</a>	当当下的图书频道

另外，对于用户来说，当他们搜索明星、视频名称、电视节目等时，本身经常出于娱乐八卦目的，希望能找到一些娱乐类结果，所以如下一些资源质量很高的满足娱乐需求的结果，也可以到 3 分，比如：

Query	用户需求意图	3 分 URL	Url 说明
舞蹈视频	观看舞蹈视频	<a href="http://www.youku.com/v_showlist/t2c95g2098.html">http://www.youku.com/v_showlist/t2c95g2098.html</a>	多个舞蹈视频，丰富、无权限
郭德纲	郭德纲，相声需求比较大，比较能够娱乐用户	<a href="http://hi.video.sina.com.cn/youguodegang/kuliebiao.php?dpc=1">http://hi.video.sina.com.cn/youguodegang/kuliebiao.php?dpc=1</a>	多个相声列表
张学友	找张学友相关信息	<a href="http://video.baidu.com/v?ct=301989888&amp;rn=20&amp;pn=0&amp;db=0&amp;s=8&amp;word=%D5%C5%D1%A7%D3%D1&amp;fr=ala0">http://video.baidu.com/v?ct=301989888&amp;rn=20&amp;pn=0&amp;db=0&amp;s=8&amp;word=%D5%C5%D1%A7%D3%D1&amp;fr=ala0</a>	这是一个非常好的视频专题页面，能够提供张学友各方面的视频信息
刘德华	明星人名，搜索这类 query 的用户八卦需求比较强	<a href="http://gb.cri.cn/27564/2012/02/29/1042s3578419.htm">http://gb.cri.cn/27564/2012/02/29/1042s3578419.htm</a>	娱乐圈人物，近期新闻、娱乐页面等也可以给到 3 分（在 2012 年 2 月底搜该 query）

## 4、2 分——比较有价值，对部分人较有帮助

### 4.1 2 分的定义

2 分，意味着该页面对于搜索这个 query 的部分用户是有帮助的。

也可以理解为，2 分结果对于用户来说，在需求满足方面要稍微差于 3 分结果，比如丰富全面程度、时效性程度、权威程度（来自不太知名站点）等方面略差；或者只能涵盖 query 用户需求的一部分，如常见的缺乏明确主需求的 query，往往最高只能到 2 分。

2 分的网页虽然比 3 分差，但对于用户来说也必须是有价值帮助的，即这个结果对于用户来说，与 query 主题相关，页面质量不垃圾。

### 4.2 举例说明

2 分常见例子，比如：

- ✓ 搜网站名，该网站下普通的频道；
- ✓ 搜电视剧，剧集多但不全；或者剧集全，但排列无次序，查找不便；或无视频的专题(包括剧情、新闻等)；
- ✓ 找免费软件，可下载，但下载过程不顺畅，干扰大，不是权威站点；
- ✓ 找范文例文，是具体的某一篇或只有少数一两篇的列表页；
- ✓ 搜商品名，只是详细的参数介绍，没有其它如评论等信息。

Query	用户需求意图	2 分 URL	Url 说明
韩国 汉城	汉城的相关信息，或景点等旅游需求	<a href="http://www.yododo.com/area/map/1-03-03">http://www.yododo.com/area/map/1-03-03</a>	该页面是汉城的地图，对于部分用户来说是有帮助的
汤姆·克鲁斯	该明星的介绍、影片、视频等	<a href="http://movie.douban.com/celebrity/1054435/">http://movie.douban.com/celebrity/1054435/</a> <a href="http://people.mtime.com/899490/">http://people.mtime.com/899490/</a>	该明星的相关介绍，虽然相关但是满足程度不够高，互联网上有更多很优质的页面，这个页面对一些用户有助
亚伯拉罕 林肯 生日	找到特定的信息：一个日期	<a href="http://zhidao.baidu.com/question/282698922.html">http://zhidao.baidu.com/question/282698922.html</a>	此网页包含所有美国总统的生日，找到林肯生日的成本较高
Wii	购买对应硬件、查找试用的游戏、浏览官网等	<a href="http://www.eachgame.cn/category-14-b0.html">http://www.eachgame.cn/category-14-b0.html</a>	Wii 的配件销售；部分用户有帮助
燕景苑餐厅 评论	浏览对应的评论	<a href="http://www.dianping.com/shop/4565212#!hippo/index=4,shopId=4565212,page=22">http://www.dianping.com/shop/4565212#!hippo/index=4,shopId=4565212,page=22</a>	互联网上很多评论网页都具有很多页，而这个页面仅仅具有 1 个评论，只对部分用户有帮助
左眼皮跳跳 歌词	查看相应歌词	<a href="http://www.huole.com/lyric/130332.htm">http://www.huole.com/lyric/130332.htm</a>	歌曲歌词不完整
优酷	寻址主需求	<a href="http://music.youku.com/">http://music.youku.com/</a>	站内一般的频道、百科介绍

		<a href="http://auto.youku.com/">http://auto.youku.com/</a> <a href="http://baike.baidu.com/view/670059.htm">http://baike.baidu.com/view/670059.htm</a> <a href="http://www.soku.com/">http://www.soku.com/</a>	
酷狗音乐下载	下载主需求	<a href="http://www.orsoon.com/Soft/7724.html">http://www.orsoon.com/Soft/7724.html</a> <a href="http://www.jz5u.com/Soft/softdown.asp?softid=23277">http://www.jz5u.com/Soft/softdown.asp?softid=23277</a>	干扰多，非知名下载站
长沙	长沙这个城市各方面的信息，如旅游、介绍、政府机构等	<a href="http://www.cshr.com.cn/">http://www.cshr.com.cn/</a> <a href="http://changsha.bitauto.com/">http://changsha.bitauto.com/</a> <a href="http://www.hncsjj.gov.cn/DefaultOther.aspx">http://www.hncsjj.gov.cn/DefaultOther.aspx</a> <a href="http://changsha.8684.cn/">http://changsha.8684.cn/</a>	当地的人才市场、公交、交警等站点
姚晨	人名需求，包括介绍、新闻、图片等各方面需求	<a href="http://tieba.baidu.com/f?kw=%D2%A6%B3%BF&amp;">http://tieba.baidu.com/f?kw=%D2%A6%B3%BF&amp;</a>	贴吧是次要需求
5 个女性字的游戏名字	5 个字的女性游戏名称	<a href="http://wenwen.soso.com/z/q275707944.htm?sp=2001">http://wenwen.soso.com/z/q275707944.htm?sp=2001</a>	穿越火线下多个游戏名，大部分是女性化的
1024	需求比较泛，有网络常用语 1024 的解释需求	<a href="http://zhidao.baidu.com/question/83420463.html">http://zhidao.baidu.com/question/83420463.html</a>	对 1024 在网络上的应用，解释的比较清楚

## 5、1 分——稍沾边，略微有帮助

### 5.1 1 分的定义

1 分结果可以理解为，页面跟 query 是相关的，但是对于大多数用户来说这个页面都是没有用的，可能只对少量用户略有价值。同时，有些“表面看起来相关的差结果”，也会是 1 分，具体见 5.3 详细说明。

1 分结果具有以下的一些特征：

- ✓ 页面质量低劣；
- ✓ 相关但是包含的有用信息很少；
- ✓ 信息因过时、过于具体、过于笼统等而难以达到 2 分的要求；
- ✓ URL 页面结果对于 query 来说只能满足非常冷门的需求；
- ✓ 页面是 wap 页面，且因 wap 页面导致用户获取信息不便，当然，前提是 wap 页面至少是相关的；如果无关应该给 0 分。

### 5.2 举例说明

Query	用户需求意图	1 分 URL	Url 说明
怀孕症状	寻找关于怀孕症状相关的信息	<a href="http://www.yun5.cn/zhuyishixiang/2011914.html">http://www.yun5.cn/zhuyishixiang/2011914.html</a>	页面质量差，文本编辑度低。内容是关于“怀孕后乳房有什么变化”，对大多数用户都没价值

无性婚姻该怎么办	寻找解决办法	<a href="http://zhidao.baidu.com/question/285998938.html">http://zhidao.baidu.com/question/285998938.html</a>	回答很万金油，适用于各种问题，对于用户来说没什么帮助
城市猎人	找电视剧、电影相关视频、剧情介绍等	<a href="http://poetry.goodmood.cn/a/2011/0118/0_322799.html">http://poetry.goodmood.cn/a/2011/0118/0_322799.html</a>	无名作者写的一首同名诗歌，需求很冷僻
百度贴吧	寻址需求，找百度贴吧这个网站	<a href="http://tieba.baidu.com/p/1433598647">http://tieba.baidu.com/p/1433598647</a>	贴吧的一个具体帖子，过于具体，很难满足用户需求
万年历	查看万年历	<a href="http://www.9630369.com/wannianli-2001.html">http://www.9630369.com/wannianli-2001.html</a>	日期很陈旧，是 2001 年的万年历
脑梗塞症状	信息需求，想知道该疾病症状是什么	<a href="http://www.i0532.net/thread-80527-1-1.html">http://www.i0532.net/thread-80527-1-1.html</a>	主要提及的是如何治疗，只是简单提及了症状，对于大部分用户来说该结果价值比较低
王菲	了解该人介绍、新闻、图片等	<a href="http://news.hainan.net/newshtml08/2009w4r2/500161f9.htm">http://news.hainan.net/newshtml08/2009w4r2/500161f9.htm</a>	09 年的新闻，价值很低
哈尔滨 宾馆	寻找哈尔滨的宾馆信息，预定宾馆等	<a href="http://www.homeinns.com/index/index.aspx">http://www.homeinns.com/index/index.aspx</a>	页面是知名连锁酒店的官网，但没有直接提供信息，需要在搜索框中输入后搜索哈尔滨地点，才可以找到哈尔滨的宾馆
淘宝	寻址需求	<a href="http://m.taobao.com/">http://m.taobao.com/</a>	手机版官网主页，不适合 PC 机浏览、使用
清朝前期疆域图	寻找指定的地图	<a href="http://zhidao.baidu.com/question/102564242.html?an=0&amp;si=9">http://zhidao.baidu.com/question/102564242.html?an=0&amp;si=9</a>	只有文字描述，实际没有地图
youku	寻址需求，想去 <a href="http://www.youku.com">www.youku.com</a>	<a href="http://www.293.net/youku.com">http://www.293.net/youku.com</a>	ip 等站点信息，对于大多数用户来说没用

### 5.3 表面相关的差结果

1 分的结果中，有一类是表面相关的结果，即标题（Title）看起来相关，但是实际页面质量低质，提供的有价值信息非常少，这些页面的 Title 中往往会有一些 query 中的 term 命中，使得这些结果看起来是比实际情况要更相关些，典型页面：

- ✓ 有些网页具有很多完全无关的链接与广告，实际有效的内容非常少；
- ✓ 有些网页内容主要为堆砌 query 关键字；
- ✓ 完全复制其他网站内容，即“采集”其他网站资源，且排版很差，影响阅读；
- ✓ 有些网页具有独特的内容，但是质量非常一般，有价值信息少；

注意：并不是所有内容为复制而来的页面都认为是低质量页面，最典型的例子：百科、知道

百科、知道上的很多内容往往是复制而来的，非原创，但是并不认为是低质量页面，因为其内容是经过精心编辑的，对于用户是非常有帮助的。如果页面排版整齐，确实是满足用户、解决大家问题而设计，而非仅仅是为了吸引用户去点击广告，那么这个页面就根据实际满足情况而进行打分，而非简单认为是 1 分的低质结果。

下面举一些典型范例说明：

Query	用户需求意图	1 分 URL	Url 说明
掌事 129	掌事这部小说的 129 章	<a href="http://www.yueduge.com/html/8/8484/2054418.html">http://www.yueduge.com/html/8/8484/2054418.html</a>	不是文字，是比较难以阅读的图片版本，且读起来很困难
youku	主需求是官网	<a href="http://www.hao123.cn/liuliang/1youku.htm">http://www.hao123.cn/liuliang/1youku.htm</a>	看似是官网，但是从域名可知是嵌套页，并非真正的官网，是一个假的山寨网站，1 分
百度云手机	手机介绍、报价等需求	<a href="http://www.fxqx.com.cn/fa/fa/htopic/62479.html">http://www.fxqx.com.cn/fa/fa/htopic/62479.html</a>	采集页面，广告很多，页面无排版很不方便阅读
人人网如何申诉	信息需求，申诉方法	<a href="http://zhidao.baidu.com/question/163289402">http://zhidao.baidu.com/question/163289402</a>	Title 很好，但是实际内容低质，很难满足用户需求
<a href="http://www.up360.com">www.up360.com</a>	url 查询	<a href="http://wenwen.soso.com/z/q333406981.htm?sp=4130">http://wenwen.soso.com/z/q333406981.htm?sp=4130</a>	虽然没有任何介绍，但是 url 查询本身有查看 url 在网络上传播的需求，只要完全命中可以给 1 分
15923549507	手机号码所在地查询	<a href="http://www.27txt.com/mobile/zhongqing_1592354.html">http://www.27txt.com/mobile/zhongqing_1592354.html</a>	1 分，整屏很多无用信息，但确实可以满足所在地的需求

## 6、0 分——不相关或无用资源

### 6.1 0 分的定义

0 分代表对于用户来说，这个结果对于搜该 query 的所有用户完全没有帮助的，一般有以下一些特征：

- ✓ 不相关，即搜 A 给 B，完全不能满足用户需求
- ✓ 页面无价值，即页面即使相关，但实质是页面完全无法提供有价值信息，如只有链接、广告，而没实际相关内容，这些页面往往作弊较多；

### 6.2 举例说明

Query	用户需求意图	0 分 URL	Url 说明
2008 年澳大利亚网球公开赛男子单打结果	寻找指定年份指定赛事的结果列表	<a href="http://zh.wikipedia.org/zh-hans/2012%E5%B9%B4%E6%B3%E6%B4%B2%E7%B6%B2%E7%90%83%E5%85%AC%E9%96%8B%E8%B3%BD%E7%94%B7%E5%AD%90%E5%96%AE%E6%89%93%E6%AF%94%E8%B3%BD">http://zh.wikipedia.org/zh-hans/2012%E5%B9%B4%E6%B3%E6%B4%B2%E7%B6%B2%E7%90%83%E5%85%AC%E9%96%8B%E8%B3%BD%E7%94%B7%E5%AD%90%E5%96%AE%E6%89%93%E6%AF%94%E8%B3%BD</a>	页面是 2012 年的比赛结果，不符合用户需求的 2008 年，为不相关

德国汽车	寻找德国汽车的信息，或德国汽车制造商的官方主页	<a href="http://www.subaru-china.com.cn/">http://www.subaru-china.com.cn/</a>	这个结果是斯巴鲁汽车的官网主页，但是确实日本车，完全不符合用户需求，不相关
Gmail 登录	访问 gmail 的官方登录页面	<a href="http://mail.163.com/">http://mail.163.com/</a>	该结果是 163 的邮箱登录，并非 gmail 登录，不相关
北京 上门灭鼠公司	寻找北京室内可以提供上门灭鼠服务的公司	<a href="http://ych.100ye.com/msg/32439614.html">http://ych.100ye.com/msg/32439614.html</a>	用户需求是北京的公司，该页面提供的是江苏的公司，无法满足用户的需求，算不相关
华山的最高峰	信息需求，想找到华山的最高峰信息	<a href="http://news.4399.com/news/20090728-13150.html">http://news.4399.com/news/20090728-13150.html</a>	这个页面仅仅是关键字命中但实际并不相关，需求是最高峰，但是实际页面是“华山论剑反馈高峰”，结果虽然命中关键字但是转义严重，不相关
东北往事之黑道风云 20 年 演员表	信息需求，想知道具体的演员表是什么	<a href="http://www.tianya.cn/publicforum/content/free/1/1632439.shtml">http://www.tianya.cn/publicforum/content/free/1/1632439.shtml</a>	这个页面看起来很相关，但是实际上是一个网友自己的假设，query 需求是真实演员表，不满足需求
初中寒假实践小结	文档需求	<a href="http://www.rudangwang.com/article/20100913/69890.html">http://www.rudangwang.com/article/20100913/69890.html</a>	关键字堆砌，正常人难以理解的句子：“措置惩罚需要别人解答的题目不敷成熟、书册常识与现实联合不敷精密等需要别人解答的题目”
公务员惩戒调理	正确 query，即用户需求应该是“公务员惩戒条例”	<a href="http://www.zaoxiejz.com/tiaoli/716/">http://www.zaoxiejz.com/tiaoli/716/</a>	页面内容无法阅读，语句类似：5 月 20 日傍晚，早泄治疗捌法问他是否愿意捐献。在查证核实之后，那么依照《行政机关公务员处分条例》第十七条的规定，早泄延时中药今天的公司，分析道路交通流量、流向变化，早泄 延时锻炼器
QQ 聊天记录不见了怎么办？	信息需求，寻求解决方法	<a href="http://wenda.tianya.cn/wenda/thread?tid=6423f74d445e880d&amp;hl=fr">http://wenda.tianya.cn/wenda/thread?tid=6423f74d445e880d&amp;hl=fr</a>	问题相关，实际回答完全无关，是关于如何破解密码、聊天记录的团队广告，基本没有价值
2011321 福彩太湖字谜	信息需求	<a href="http://tieba.baidu.com/f?kz=1294450452">http://tieba.baidu.com/f?kz=1294450452</a>	仅 title 命中好，实际完全不相关
初三化学试卷	初三化学试卷的浏览、下载需求	<a href="http://www.5ykj.com/Article/xywstzl/94674.htm">http://www.5ykj.com/Article/xywstzl/94674.htm</a>	试卷下载分享，其中没有任何例题等，完全无法满足



## 7、死链

如果页面死链导致你无法进行打分评估，请选择“死链”，死链的类型主要有以下几种类型，下面会单独做详细说明。

- ✓ 页面抓取不成功，且点击 url 打开也无法加载，无实际内容；
- ✓ 页面报错提示页面：除了报错信息外，页面上没有其他内容；
- ✓ 标准报错页面：网站直接返回的标准出错页面；
- ✓ 网页恶意软件提示，如“请注意，访问此软件可能会损害您的计算机”；

### 7.1 死链类型：网页安全提示

有些页面显示为空白页，点击 url 后的页面显示“请注意，访问此软件可能会损害您的计算机”，一般情况下，有些页面虽然会有不安全的提示，但是他们不应该直接被抛弃，例如有的政府网站，访问时也会提示说安全证书不被信任，但是可以点击“忽略风险”继续打开页面打分。

如果你的确认为该页面风险很大，如 url 很奇怪，可能导致电脑中毒等，那么可以不继续，而选择“抛弃”，并标记为“死链”。典型的网页安全提示警告信息如下：



### 7.2 死链类型：空页面或未能访问页面

主要指抓取的页面为空，点击 url 后，在较长时间内，网站都没能显示出来，或者显示出来了，但是实际页面却为空的。

但是，点击 url 后，如果整个页面中有些内容加载出来了，而有些没加载出来，是需要按照加载出来的内容实际进行评分的。

### 7.3 死链类型：页面报错提示

主要是网站自身变化引起的，会直接给予提示文字，例如：

- ✓ 页面不存在
- ✓ 内容已转移



- ✓ 帖子已删除
- ✓ 域名过期或出售
- ✓ 空间被关闭
- ✓ 网站要备注，等

※ 注意：对于有些死链，当我们直接点击 url 会显示为死链（如内容已删除），然后直接跳转到首页，这类也标记为死链。

典型的范例：

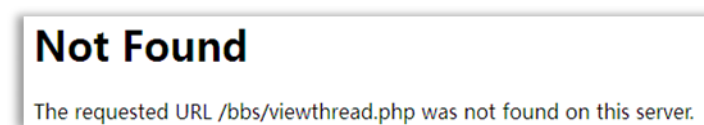


## 7.4 死链类型：标准报错页面

主要是标准的出错页面，例如：

- ✓ 401，未授权
- ✓ 403，禁止访问
- ✓ 404，未找到文件
- ✓ 500，服务器内部错误
- ✓ 503，服务器当前无法处理

不同的网站，不同的浏览器可能显示不同，典型范例截图：



### 无法找到该页

您正在搜索的页面可能已经删除、更名或暂时不可用。

请尝试以下操作：

- 确保浏览器的地址栏中显示的网站地址的拼写和格式正确无误。
- 如果通过单击链接而到达了该网页，请与网站管理员联系，通知他们该链接的格式不正确。
- 单击 **后退** 按钮尝试另一个链接。

HTTP 错误 404 - 文件或目录未找到。  
Internet 信息服务 (IIS)

### 连接超时

位于 [www.jiemeng168.info](http://www.jiemeng168.info) 的服务器响应时间过长。

- 此站点暂时不可用或者太忙。请稍后重试。
- 如果您无法载入任何页面，请检查您计算机的网络连接。

### 连接被重置

载入页面时到服务器的连接被重置。

- 此站点暂时不可用或者太忙。请稍后重试。
- 如果您无法载入任何页面，请检查您计算机的网络连接。

### 未找到服务器

Firefox 无法在 [www.tdwzbook.com](http://www.tdwzbook.com) 找到该服务器。

 **虾米** xiami BETA  
高品质音乐 发现 分享

# 404

### 找不到网页

抱歉，您要查看的页面可能已经删除、更名或暂时不可用。

403.50 Forbidden

The server is busy, please refresh.

## Part5 特定类型打分举例说明

### 1、打分一些重要原则

在正式进行 query-url 质量评估举例说明前，先说一下打分的一些重要原则，有些在前文中已经提到，在此进一步强调，主要包括以下几个方面：

- ✓ 把自己当成真正的用户：你必须将自己当做是真正的用户，从用户角度作出评估；
- ✓ 用户真实需求判断：你评估具体页面时，首先要明确用户在 query 背后的真实需求是什么（一般可以基于你的评估经验、你的经验常识以及基于网络搜索去了解 query 需求），然后根据页面对用户需求的满足程度打分；
- ✓ 冷门需求打分较低：有些 query 具有多种需求，冷门的需求应当酌情给出较低的评分
- ✓ 纠结的分值判断：对于有些在 2 个分值中纠结的情况，一般根据自己的分析后给出合理的判断，想想需求满足程度。如果 3 个分值都在考虑范围内，那么请直接使用居中的那个分值。
- ✓ 2 分、1 分和 0 分的简单法则：如果你认为一个页面对用户没什么帮助，请打一个较低的分值，说明如下：
  - 2 分的结果必须对用户有帮助，能为用户需求满足提供一些价值；
  - 1 分页面的价值较低，但是必须跟 query 主题是相关的；
  - 0 分页面完全没有价值，或者跟 query 主题不相关；

### 2、query 主体是否确定与需求是否多样

在评估打分过程中，针对 query 需求会存在下面 3 个角度，在 Part3 “理解 query 需求”中已经介绍过，包括：

**Query 主体表意：**是确定还是非确定的区分；

**Query 需求维度：**存在需求维度单一还是多样的区分；

**Query 需求强度：**即针对 query 不同主体表意和需求维度，存在需求强度大小之分，分为主要需求、泛需求、小众/冷门需求三类。

针对这几个角度，这里实际打分进一步说明，存在以下原则：

- ✓ 通常，主体表意、需求维度如果小众，打分会较低；
- ✓ 满足一般需求强度的语义表意和需求维度的页面，打分中不会降分；
- ✓ 存在明显主要的主体表意和需求维度，才能打 4 分；

#### 2.1 query 主体确定，需求维度单一

当 query 主体确定，不存在多义情况下，且需求维度明确单一，打分主要依据页面对 query 单一需求的满足程度，打分从 4-0 的范围，具体举例说明如下：

Query	用户需求意图	URL	Url 说明
美图秀秀在线下载	美图秀秀软件下载	http://xiuxiu.meitu.com/download/	4 分，官方下载页面

		<a href="http://down.it168.com/162/165/109185/">http://down.it168.com/162/165/109185/</a> ; <a href="http://xiazai.zol.com.cn/detail/34/330483.shtml">http://xiazai.zol.com.cn/detail/34/330483.shtml</a>	3 分，知名大站下载页面，能够下载安装
		<a href="http://sj.zol.com.cn/detail/20/19660.shtml">http://sj.zol.com.cn/detail/20/19660.shtml</a>	2 分，手机版美图秀秀下载，非主需求
		<a href="http://8.app.meitu.com/">http://8.app.meitu.com/</a>	1 分，美图官网超能力测试页面，该页面本身与 query 不太相关，页面底部有下载链接，
		<a href="http://download.digi.tech.qq.com/a/20110118/000003.htm">http://download.digi.tech.qq.com/a/20110118/000003.htm</a>	1 分，在线版美图秀秀介绍页面，有在线链接，为下载需求的替代需求，很少量满足需求，1 分
		<a href="http://bbs.waptw.com/topic/qnas/view/fid:2/tid:3227207/page:0/order:asc/">http://bbs.waptw.com/topic/qnas/view/fid:2/tid:3227207/page:0/order:asc/</a>	0 分，页面主题是美图秀秀下载了怎么不能用，不能满足用户需求

Query	用户需求意图	URL	Url 说明
兰溪地图	在线浏览地图，或下载地图	<a href="http://map.baidu.com/?newmap=1&amp;s=s%26wd%3D%25E5%2585%25B0%25E6%25BA%25AA%25E5%25B8%2582%26c%3D1462&amp;fr=alat0&amp;from=alamap">http://map.baidu.com/?newmap=1&amp;s=s%26wd%3D%25E5%2585%25B0%25E6%25BA%25AA%25E5%25B8%2582%26c%3D1462&amp;fr=alat0&amp;from=alamap</a>	知名的百度地图的地图页面，可支持各种检索 3 分
		<a href="http://www.9tour.cn/maps/detail_31_60682_3.html">http://www.9tour.cn/maps/detail_31_60682_3.html</a>	地图图片，无法支持各种检索，可保存，站点非知名，2 分
		<a href="http://zhidao.baidu.com/question/109559476.html">http://zhidao.baidu.com/question/109559476.html</a>	兰溪地图售卖相关的知道问答，极少数人需求，1 分
		<a href="http://www.ddmap.com/map/21----%C0%BC%CF%AA%C2%B7+%CA%D0%C7%F8%BC%B6%D2%BD%D4%BA----pvs/">http://www.ddmap.com/map/21----%C0%BC%CF%AA%C2%B7+%CA%D0%C7%F8%BC%B6%D2%BD%D4%BA----pvs/</a>	上海兰溪路上 XX 医院，不符合用户需求，0 分

## 2.2 query 主体确定，需求多样

即 query 主体是确定的，没有多种表意，但是该表意下需求维度多样，需求维度存在需求强度主次之分，不同需求强度打分不一样，需求强度小的打分会低。

如 query 【赵薇】，主体表意明确，就是明星“赵薇”，但是用户需求维度多样，如介绍、新闻、视频、音乐等。

Query 需求	用户需求意图	URL	Url 说明
----------	--------	-----	--------

维度中主次之分			
主要/重要需求	官方微博、博客寻址	<a href="http://weibo.com/zhaowei">http://weibo.com/zhaowei</a>	官网微博，4分
	介绍需求	<a href="http://baike.baidu.com/view/3255.htm">http://baike.baidu.com/view/3255.htm</a>	百科介绍也是重要需求，内容很优质、丰富，3分
	赵薇参演的电影、电视剧等影视在线观看需求	<a href="http://www.youku.com/star_page/uid_UNTU4MjQ=.html">http://www.youku.com/star_page/uid_UNTU4MjQ=.html</a>	电影也重要需求，优酷专题，很多电影可以在线观看，3分
次要需求	赵薇参演的某部电视剧的在线观看需求	<a href="http://www.youku.com/show_page/id_zcbfe6706962411de83b1.html">http://www.youku.com/show_page/id_zcbfe6706962411de83b1.html</a>	赵薇出演的一部电视剧的在线观看，2分
	和赵薇有关的最近的某个热度一般的新闻	<a href="http://news.jschina.com.cn/system/2012/03/07/012880473.shtml">http://news.jschina.com.cn/system/2012/03/07/012880473.shtml</a>	与赵薇有关的电影的某个近期新闻（用户搜索 query 时间为 2012.3.10），2分
冷门需求	赵薇以前的新闻	<a href="http://news.qq.com/a/20100924/000701.htm">http://news.qq.com/a/20100924/000701.htm</a>	与赵薇有关的过去发生的新闻（用户搜索 query 时间为 2012.3.10），2分
无需求	——	<a href="http://dzh.mop.com/whbm/20090827/1/3758SSI1c03d7azg.shtml">http://dzh.mop.com/whbm/20090827/1/3758SSI1c03d7azg.shtml</a>	Title 有赵薇，但是实际内容与赵薇没有任何关系，0分

如 query 【兴业银行】，主题表意明确，但是需求维度多样，包括信用卡、网上银行、股票、介绍等多方面。

Query 需求维度中主次之分	用户需求意图	URL	Url 说明
主需求	官网	<a href="http://www.cib.com.cn/n">http://www.cib.com.cn/n</a>	官网，4分
重要需求/次要需求	兴业银行信用卡	<a href="http://creditcard.cib.com.cn/card/">http://creditcard.cib.com.cn/card/</a>	信用卡也是重要需求，3分
	股票行情	<a href="http://www.xcar.com.cn/bbs">http://www.xcar.com.cn/bbs</a> <a href="http://cj.gw.com.cn/news/stockData/SH/66/SH601166/detail.shtml">http://cj.gw.com.cn/news/stockData/SH/66/SH601166/detail.shtml</a>	股票行情，走势 2分
	股票行情、公司业绩讨论	<a href="http://news.jschina.com.cn/system/2012/03/07/012880473.shtml">http://news.jschina.com.cn/system/2012/03/07/012880473.shtml</a>	股票兴业银行论坛，出自股吧 2分
	实体银行位置	<a href="http://map.baidu.com/?newmap=1&amp;s=con%26wd%3D%E5%85%B4%E4%B8%9A%E9%93%B6%E8%A1%8C%26c%3D131&amp;fr=alae0&amp;ext=1&amp;from=alamap">http://map.baidu.com/?newmap=1&amp;s=con%26wd%3D%E5%85%B4%E4%B8%9A%E9%93%B6%E8%A1%8C%26c%3D131&amp;fr=alae0&amp;ext=1&amp;from=alamap</a>	Query 查询地，兴业银行实体位置。2分

	兴业银行招聘信息	<a href="http://bbs.yingjiesheng.com/forum-247-1.html">http://bbs.yingjiesheng.com/forum-247-1.html</a>	招聘论坛，2 分
	各营业网点电话	<a href="http://www.cib.com.cn/netbank/cn/About_IB/Outlet_Locator.html">http://www.cib.com.cn/netbank/cn/About_IB/Outlet_Locator.html</a>	各营业网点电话。2 分
小众/冷门需求	关于该银行的新闻	<a href="http://finance.qq.com/a/20120425/004230.htm">http://finance.qq.com/a/20120425/004230.htm</a>	关于该银行业务的一篇新闻，1 分
	兴业银行公司介绍	<a href="http://baike.baidu.com/view/34166.htm">http://baike.baidu.com/view/34166.htm</a>	百科的词条，介绍该公司的情况，2 分
	法国兴业银行	<a href="http://news.qq.com/a/20080125/000393.htm">http://news.qq.com/a/20080125/000393.htm</a>	1 条 08 年的法国兴业银行信息，1 分

## 2.2 query 主体表意不确定，需求多样

评估者可能会碰到一些具有多种表意需求的 query，如前面一直提到的【苹果】、【土豆】的例子。打分的基本原则是：表意少的打分低，常见最主要的表意需求打分高。

主次需求	Query 范例	打分范围
主需求	Apple，主需求是苹果品牌； Windows，主需求是 windows 操作系统； 京东，主需求是京东商城电子商务； 豚鼠，主需求是介绍、与这种动物相关的信息（这个为 3-0） 天龙八部，主需求是网络游戏相关	4 到 0 分，其中部分 query 不可以拥有 4 分结果
重要需求/次要需求	Apple，稍次的需求为苹果这种水果的信息，或翻译单词等 Windows，稍次的需求为单词翻译，或者有关窗户的信息； 天龙八部，林志颖版电视剧属于需求强度一般的需求维度；	3 到 0 分，没有 4 分的情况；
小众/冷门需求	BBS，比较冷门的需求为德国的一个汽车零件生产商 星座，比较冷门的需求是某游戏中名为星座的武器 天龙八部，关于“天龙八部”的来历	2 到 0 分，需求强度比例越小，所给打分越低；
不可能的需求	荷兰猪，绝大部分是找豚鼠，而非是生长在荷兰的普通猪	0 分

下面结合 query 主体表意多样与需求维度的多样，从而确定的需求强度，来举例说明需求强度对打分的影响，以 query【苹果】为例说明：

Query 主体表意上的主次之分	需求维度	URL	Url 说明

主需求：苹果的品牌产品 (4-0 分)	苹果官网：重要需求	<a href="http://www.apple.com.cn/">http://www.apple.com.cn/</a>	4 分，苹果品牌的官网页面，满足主要表意需求好结果
	苹果产品：重要需求	<a href="http://detail.zol.com.cn/cell_phone_index/subcate57_544_list_1.html">http://detail.zol.com.cn/cell_phone_index/subcate57_544_list_1.html</a>	3 分，苹果手机相关产品的介绍评论页面
	苹果公司介绍：次要需求	<a href="http://baike.baidu.com/view/15181.htm">http://baike.baidu.com/view/15181.htm</a>	2 分
	苹果公司总部大楼：冷门需求	<a href="http://wenku.baidu.com/view/5950321214791711cc791743.html">http://wenku.baidu.com/view/5950321214791711cc791743.html</a>	1 分
重要需求/次要需求：水果、电影 (3-0 分)	苹果这种水果介绍：次要需求	<a href="http://baike.39.net/8000-8300-8301-1/">http://baike.39.net/8000-8300-8301-1/</a>	2 分
	宠物吃了苹果后的反应：基本无需求	<a href="http://zhidao.baidu.com/question/363505443.html">http://zhidao.baidu.com/question/363505443.html</a>	0 分
	电影视频：次要需求	<a href="http://www.167b.com/ysxx/151.html">http://www.167b.com/ysxx/151.html</a>	2 分
	电影介绍、演员等	<a href="http://ent.people.com.cn/GB/8222/42057/83496/index.html">http://ent.people.com.cn/GB/8222/42057/83496/index.html</a>	2 或 1 分
小众/冷门需求：苹果摄影 (2-0 分)	沈阳苹果摄影：冷门需求	<a href="http://www.applephoto.com.cn/">http://www.applephoto.com.cn/</a>	1 分

## 2.4 特殊说明：需求维度的泛化与具体

下面从另一个维度，看同一个主题的 query 的需求维度的泛化与具体，以及对打分的影响。

有的 query 是需求比较泛的查询，但有些 query 是需求描述非常具体的 query，以下是一些这些不同 query 之间比较的例子：

Query	略具体的需求	较具体的需求
椅子	餐厅椅	宜家 高后备、带扶手、软垫的椅子
数码相机	尼康 数码相机	尼康 d5000 数码相机
丰田	丰田 混合动力	丰田 普锐斯 2010
历史	哈佛历史	哈佛 人类学 历史
面试问题	教师面试问题	对于中国小学语文教师的面试问题

对于需求非常具体的 query 来说，好的页面应该很好的匹配 query 中所涵盖的要求，如果二者不是非常匹配的时候，需要认真考虑二者之间的匹配程度，来判断该页面对于用户需求的满足程度如何，从而做出合理的打分，以下是一些具体的例子。

Query	用户需求意图	URL	Url 说明
数码相机	用户对数码相机有兴趣，想了解相关的内容或者买一个相机	<a href="http://detail.zol.com.cn/digital_camera_index/subcate15_list_1.html">http://detail.zol.com.cn/digital_camera_index/subcate15_list_1.html</a>	3 分，很多机型的列表页面，信息很丰富
		<a href="http://www.360buy.com/products/652-654-831.html">http://www.360buy.com/products/652-654-831.html</a>	3 分，很多机型的列表页面，介绍详细，京东是著名的电



			子商务网站，可以方便购买
尼康 数码相机	想了解尼康数码相机的相关信息，或者买一个	<a href="http://www.nikon.com.cn/sc_CN/products.page?CategoryID=glxlvehu">http://www.nikon.com.cn/sc_CN/products.page?CategoryID=glxlvehu</a>	3 分，尼康官网上的一种类型的数码相机页面，但尼康官网上还有其他的数码相机，所以不适合打 4 分
		<a href="http://detail.zol.com.cn/digital_camera_index/subcate15_657_list_1.html">http://detail.zol.com.cn/digital_camera_index/subcate15_657_list_1.html</a>	3 分，优质站点的尼康数码相机页面
		<a href="http://www.nikon.com.cn/sc_CN/products.page?CategoryID=glxlvehv">http://www.nikon.com.cn/sc_CN/products.page?CategoryID=glxlvehv</a>	2 分，官网尼康数码相机 P 系列，比 query 需求更具体
北京 沃尔玛	找北京的沃尔玛超市相关信息，如地点、营业时间、活动等	<a href="http://www.wal-martchina.com/walmart/store/2_beijing.htm">http://www.wal-martchina.com/walmart/store/2_beijing.htm</a>	4 分，官网上的北京沃尔玛的分店信息
		<a href="http://map.baidu.com/?newmap=1&amp;s=con%26wd%3D%E6%B2%83%E5%B0%94%E7%8E%9B%26c%3D131&amp;fr=alam0&amp;ext=1&amp;from=alamap">http://map.baidu.com/?newmap=1&amp;s=con%26wd%3D%E6%B2%83%E5%B0%94%E7%8E%9B%26c%3D131&amp;fr=alam0&amp;ext=1&amp;from=alamap</a>	3 分，较官方页面给出的分布地图
		<a href="http://zhidao.baidu.com/question/291303458.html">http://zhidao.baidu.com/question/291303458.html</a>	2 分，知道回答只有少数几家店，不足以满足用户需求
		<a href="http://s.baidu.com/shop/739faf3f97d1ebf31a07f20e">http://s.baidu.com/shop/739faf3f97d1ebf31a07f20e</a>	2 分，沃尔玛的北京某具体分店，通过链接可以知道其他更多分店信息

也因为在用户需求非常具体时，容易有一些页面与 query 的匹配程度并不太高，这时候就需要评估者认真分析，看这个页面与 query 到底有多相关，到底多大程度上能够满足用户的需求，并且做出打分。

以下是一些匹配程度或好或差的例子：

Query	用户需求意图	URL	说明
教师面试问题	查找教师面试时会面临的问题	<a href="http://wenku.baidu.com/view/238091282af90242a895e591.html">http://wenku.baidu.com/view/238091282af90242a895e591.html</a>	文库结果，有很多面试问题可以参考，内容丰富 3 分
		<a href="http://dongqianxiaoxue.lingd.net/article-4102279-1.html">http://dongqianxiaoxue.lingd.net/article-4102279-1.html</a>	页面主题比 query 更具体，是小学老师面试题，但是很多是通用的问题，2 分
		<a href="http://wenku.baidu.com/view/75758019964bcf84b9d57b3c.html">http://wenku.baidu.com/view/75758019964bcf84b9d57b3c.html</a>	文库结果，页面话题更泛，更多是考试相关，面试只是其中一点，但是还是有一点价值的，1 分
		<a href="http://wenku.baidu.com/view/195d4b6d7e21af45b307a825.html">http://wenku.baidu.com/view/195d4b6d7e21af45b307a825.html</a>	有很多面试题，但是并不太适用于教师，由于教师职业本身的特殊性，这个结果几乎没有用，0 分

宝马 Mini	用户可能是想买车，希望得到这个车型的相关信息，此车分为 CABRIO、CLUBMAN、COUNTRYMAN\COUPER ROADSTER 等多个款	<a href="http://www.minichina.com.cn/mini/cn/zh/index.html">http://www.minichina.com.cn/mini/cn/zh/index.html</a>	官网，车型齐全优质，4 分
		<a href="http://db.auto.sohu.com/brand_202/index.shtml">http://db.auto.sohu.com/brand_202/index.shtml</a>	权威网站的页面，有多个系列，全面优质有报价图片等，3 分
		<a href="http://www.minichina.com.cn/mini/cn/zh/mini_coupe/index.html">http://www.minichina.com.cn/mini/cn/zh/mini_coupe/index.html</a>	官网中该系列中的一款，比 query 范围稍小些，3 分
		<a href="http://www.minichina.com.cn/mini/cn/zh/model_comparison/index.html">http://www.minichina.com.cn/mini/cn/zh/model_comparison/index.html</a>	官网上 Mini 不同款车的自助对比页面，需求满足程度稍低些，3 分
		<a href="http://drive.xcar.com.cn/201202/news_406050_1.html">http://drive.xcar.com.cn/201202/news_406050_1.html</a>	Mini 某款车 COUPE ROADSTER 的介绍，2 分
		<a href="http://www.minichina.com.cn/mini/cn/zh/accessories/xmas_promotion/index.html">http://www.minichina.com.cn/mini/cn/zh/accessories/xmas_promotion/index.html</a>	Mini 车的装饰包，需求程度很低，1 分

### 3、特定类型的 query 打分举例说明

在这部分，我们会举例介绍一些特定类型和特定 url 页面评估。例子只能覆盖一些特定的类型，不能穷举你遇到的所有情况，所以在你做评估打分的时候，永远需要记住的是用户的需求意图是什么，评估的页面用户需求满足程度如何。

#### 3.1 实体类 query

##### 3.1.1 名人类

名人类 query 即提到该人，大多数脑海里都认为的那个人，一般明星类、公众人物、体育人物等有一定知名度、被公众知晓的人物都可以列入其中。

名人一般会有 4 分权威结果，存在主要表意需求。

Query	赵薇	
Query 说明	赵薇是知名的娱乐明星	
用户需求意图	信息需求：想知道有关于赵薇相关的信息 娱乐操作需求：	
4	<a href="http://weibo.com/zhaowei">http://weibo.com/zhaowei</a>	赵薇的官方微博
3	<a href="http://baike.baidu.com/view/3255.htm">http://baike.baidu.com/view/3255.htm</a>	赵薇的百科介绍
3	<a href="http://tieba.baidu.com/f?kw=%D5%D4%DE%B1&amp;fr=ala0">http://tieba.baidu.com/f?kw=%D5%D4%DE%B1&amp;fr=ala0</a>	赵薇的贴吧
3	<a href="http://www.youku.com/star_page/uid_UNTU4MjQ=.html">http://www.youku.com/star_page/uid_UNTU4MjQ=.html</a>	赵薇的介绍、视频列表页，视频可以观看
3	<a href="http://www.baidu.com/s?tn=baidurt&amp;rtt=1&amp;bsst=1&amp;wd=%D5%D4%DE%B1">http://www.baidu.com/s?tn=baidurt&amp;rtt=1&amp;bsst=1&amp;wd=%D5%D4%DE%B1</a>	赵薇的最新新闻搜索结果页，页面内容都是相关的，质量较好

2	<a href="http://www.youku.com/show_page/id_zcbfe6706962411de83b1.html">http://www.youku.com/show_page/id_zcbfe6706962411de83b1.html</a>	赵薇出演的一部电视剧《老房有喜》的在线观看
2	<a href="http://gb.cri.cn/27564/2012/03/30/1326s3622596.htm">http://gb.cri.cn/27564/2012/03/30/1326s3622596.htm</a>	与赵薇有关的某条最近相关新闻“林心如赵薇谢娜李湘何炅 明星惨不忍睹的照片”
0	<a href="http://dzh.mop.com/whbm/20090827/1/3758SSI1c03d7azg.shtml">http://dzh.mop.com/whbm/20090827/1/3758SSI1c03d7azg.shtml</a>	Title 有赵薇,但是实际内容与赵薇没有任何关系

### 3.1.1 非名人

普通人物指非知名的人物，我们无法知道用户到底找的是哪个人。如果有多个人叫这个名字，则关于该人名的网页，不管是介绍的是同名的哪个人，一般都给 1 分，除非某个叫该名字的人略知名些，如在新闻等曾经被报道等，且页面质量较好，这时候可以打 3 分或 2 分。若结果中多个叫该名字的人，如百科有多人介绍，可以 2-3 分。

Query	李小敏	
Query 说明	有非常多叫李小敏的人，也没有非常知名的人	
用户需求意图	信息需求：想找跟叫“李小敏”相关的信息	
3 或 2	<a href="http://baike.baidu.com/view/1399700.htm">http://baike.baidu.com/view/1399700.htm</a>	百科有多义项，同一个页面上有多个同名人物的介绍，丰富度高
2	<a href="http://www.jstv.com/s/jszy/lxm/">http://www.jstv.com/s/jszy/lxm/</a>	江苏省委常委、省委政法委书记李小敏，比非常不知名的略有价值
1	<a href="http://esf.sh.soufun.com/a/lixiaominrita">http://esf.sh.soufun.com/a/lixiaominrita</a>	某个叫李小敏的房产经纪人，对大多数用户价值较低

### 3.1.2 地点名

对于一些地名，政府官方网站作为权威网站，可以给 4 分，其他结果根据需求强度和满足情况做适当打分。

Query	长沙	
Query 说明	长沙是湖南的省会城市	
用户需求意图	信息需求：长沙相关的信息，需求很广泛，如天气、地图、介绍、政府机构、旅游信息等。	
4	<a href="http://www.changsha.gov.cn/">http://www.changsha.gov.cn/</a>	长沙市政府网站
3	<a href="http://baike.baidu.com/view/7127.htm">http://baike.baidu.com/view/7127.htm</a>	百科介绍，重要需求之一且内容丰富
3	<a href="http://www.weather.com.cn/weather/101250101.shtml">http://www.weather.com.cn/weather/101250101.shtml</a>	当地天气，比较重要需求
3	<a href="http://map.baidu.com/?newmap=1&amp;s=%26wd%3D%25E9%2595%25BF%25E6%25B2%2599%25E5%25B8%2582%26c%3D158&amp;fr=alac0&amp;from=alamap">http://map.baidu.com/?newmap=1&amp;s=%26wd%3D%25E9%2595%25BF%25E6%25B2%2599%25E5%25B8%2582%26c%3D158&amp;fr=alac0&amp;from=alamap</a>	当地地图，比较重要需求

2	<a href="http://tieba.baidu.com/f?kw=%B3%A4%C9%B3&amp;fr=ala0">http://tieba.baidu.com/f?kw=%B3%A4%C9%B3&amp;fr=ala0</a>	当地贴吧，较主要需求
2	<a href="http://www.cshr.com.cn/index.asp">http://www.cshr.com.cn/index.asp</a>	当地人才网站等比较重要的网站
2	<a href="http://www.csup.gov.cn/">http://www.csup.gov.cn/</a>	长沙的各类非省政府的重要机构站点
1	<a href="http://tieba.baidu.com/p/1056453660">http://tieba.baidu.com/p/1056453660</a>	当地贴吧的一个具体帖子
1	<a href="http://zhidao.baidu.com/question/279702629.html">http://zhidao.baidu.com/question/279702629.html</a>	有部分长沙工作环境的介绍。另外，关于长沙一般的小站点、小机构的相关信息可以给 1 分。
0	<a href="http://www.iphoto9.com/contents/23/505.html">http://www.iphoto9.com/contents/23/505.html</a>	实际页面与长沙关系并不大，基本不能满足用户需求

### 3.1.3 公司名

公司类需求一般找公司官方网站、公司介绍信息、地址、联系方式、评价信息，所经营业务（如产品、服务）等相关信息。

以【东方票务】为例进行说明

打分	URL	说明	总结说明
4	<a href="http://www.ticket2010.com/">http://www.ticket2010.com/</a>	官网，4 分	官网
3	<a href="http://www.ticket2010.com/project_index.aspx">http://www.ticket2010.com/project_index.aspx</a>	重要二级频道，3 分；	官网重要频道；质量好的公司介绍页面（含介绍、地址、联系方式、网址等较丰富信息）
3	<a href="http://www.dianping.com/shop/2753518">http://www.dianping.com/shop/2753518</a>	基础信息之外还有点评、人均价格、地图等信息，页面质量好、可信，3 分	
2	<a href="http://www.ticket2010.com/aboutus.aspx">http://www.ticket2010.com/aboutus.aspx</a>	官网上的介绍我们，用户需求不大，2 分	官网上的普通二级频道、介绍页面、微博等视与公司经营等有多大关系，数量如何打分
2	<a href="http://okcard.blemall.com/merdetail_ddshh-81112.html">http://okcard.blemall.com/merdetail_ddshh-81112.html</a>	介绍页面，有官网网址，需求较少，2 分	
2	<a href="http://weibo.com/u/2051100135">http://weibo.com/u/2051100135</a>	官网微博，与票务有关的内容较多，2 分	
1	<a href="http://sports.eastday.com/eastday/tickets/">http://sports.eastday.com/eastday/tickets/</a>	东方网上共建频道，链接基本已经全部死链，内容都是 07 年之前的，仅电话等需求可以满足，1 分	页面低质、内容不全、页面过于具体、很久没更新的时效性很差的页面
1	<a href="http://www.ticket2010.com/program.aspx?id=8085">http://www.ticket2010.com/program.aspx?id=8085</a>	官网上非常具体的页面，1 分	
1	<a href="http://t.dianping.com/deal/3599">http://t.dianping.com/deal/3599</a>	东方票务产品的过期团购页面，1 分	

### 3.1.4 产品名

下面介绍一些通用的原则，但是实际评估中还是根据实际用户需求和页面对用户需求满足程度打分：

- ✓ 详尽的专题页面，至少包含介绍、报价、评论、参数，3 分；

- ✓ 专题页面仅能其中覆盖部分类别、或内容稍少、时效性略差，以及比较详细丰富的单一页面，2 分；
- ✓ 关于这个产品的相关信息，但质量较差，或者过期的产品介绍信息等，1 分；
- ✓ 仅仅是文中提及这个名词，没有任何相关介绍，例如“我明天想去买个 XX 产品”，0 分；

Query	htc G16	
Query 说明	Htc G16 是一款智能手机	
用户需求意图	信息需求：用户找介绍、评论、价格等信息，使用方法等。 交互操作需求：在网上购买手机、软件下载等	
3	<a href="http://detail.zol.com.cn/cell_phone/index272619.shtml">http://detail.zol.com.cn/cell_phone/index272619.shtml</a>	知名 IT 网站中关村在线关于该款手机的专题页面
3	<a href="http://product.pconline.com.cn/mobile/htc/453778.html">http://product.pconline.com.cn/mobile/htc/453778.html</a>	同样的知名网站的专题页面信息
2	<a href="http://zhidao.baidu.com/question/340485518.html">http://zhidao.baidu.com/question/340485518.html</a>	相关的评论信息，2011.11 月的，时效性还好，回答较有价值，但不是特别丰富
2	<a href="http://baike.baidu.com/view/6227635.htm">http://baike.baidu.com/view/6227635.htm</a>	百科的介绍信息，不如好的专题好，没有评论、报价等信息
1	<a href="http://wenku.baidu.com/view/d1f3b46627d3240c8447ef8f.html">http://wenku.baidu.com/view/d1f3b46627d3240c8447ef8f.html</a>	页面介绍 HTC 从 G1 到 G16，少量信息跟 G16 相关
0	<a href="http://bbs.qdit.com/thread-635915-1-1.html">http://bbs.qdit.com/thread-635915-1-1.html</a>	页面价值很低，基本没啥价值

### 3.1.5 学校名

找学校，一般主要的需求为官网，为 4 分。重要的需求有介绍、考试招生、新闻、学院等各方面信息。根据需求强度的不同和页面对需求满足程度，可以 3-0 分。

Query	北京大学	
Query 说明	国内知名高校北京大学	
用户需求意图	信息需求：找北京大学相关的各方面信息，如介绍、招生、考试、新闻、贴吧、论坛等； 寻址需求：去北京大学的官方网站；	
4	<a href="http://www.pku.edu.cn/">http://www.pku.edu.cn/</a>	北京大学的官方网站
3	<a href="http://baike.baidu.com/view/1471.htm">http://baike.baidu.com/view/1471.htm</a>	北京大学的百科介绍
3	<a href="http://grs.pku.edu.cn/">http://grs.pku.edu.cn/</a> <a href="http://www.gotopku.cn/">http://www.gotopku.cn/</a>	北京大学校内重要的机构，如招生网站、研究生院等，给 3 分
3	<a href="http://kaoshi.edu.sina.com.cn/college/c/10001.shtml">http://kaoshi.edu.sina.com.cn/college/c/10001.shtml</a>	新浪中北大的招生专题，对于参加高考的学生家长很有价值
2	<a href="http://pkunews.pku.edu.cn/">http://pkunews.pku.edu.cn/</a>	北京大学新闻网，对部分用户有价值
2	<a href="http://news.baidu.com/ns?cl=2&amp;rn=20&amp;tn=news&amp;word=%B1%B1%BE%A9%B4%F3%D1%A7&amp;ct=1&amp;fr=ala0">http://news.baidu.com/ns?cl=2&amp;rn=20&amp;tn=news&amp;word=%B1%B1%BE%A9%B4%F3%D1%A7&amp;ct=1&amp;fr=ala0</a>	北京大学的新闻检索结果页，没有非常有名的新闻，对部分用户有价值（query 日期为 2012.3.31）
2	<a href="http://www.bjmu.edu.cn/">http://www.bjmu.edu.cn/</a>	北京大学医学部，需求强度略弱，对一些用户有价值
1	<a href="http://epaper.jinghua.cn/html/2011-01/17">http://epaper.jinghua.cn/html/2011-01/17</a>	1 月份的一篇关于北大的新闻，时效性不太好，

	/content_623924.htm	对用户价值不太高
0	<a href="http://zhidao.baidu.com/question/330356982.html">http://zhidao.baidu.com/question/330356982.html</a>	转义的结果，关于“北京有哪些大学不错”

## 3.2 信息类需求

信息类需求是搜索引擎需求中较大的一类，一般主要考虑页面提供的信息是否满足了用户的需求，信息是否丰富、真实、获取便捷等。

下面重点说明 3 种特殊的信息类需求，包括 key-value、原文、问答类。

### 3.2.1 Key-Value 需求

所谓的 key-value，即用户需求是一个 key，但是希望获取 value 的信息，如姚明身高（身高是 key，value 是 2.26 米）、北京海淀区政府地址、联想客服电话等。

所以我们需要关注页面是否提供了 value 相关信息，以及 value 信息是否可靠真实的。

Query	亚伯拉罕 林肯 生日	
Query 说明	美国总统亚伯拉罕 林肯的生日	
用户需求意图	信息需求：生日是哪年哪月哪日	
3	<a href="http://zhidao.baidu.com/question/14773657.html">http://zhidao.baidu.com/question/14773657.html</a>	较好的说明页面，页面主题很集中，3 分
2	<a href="http://zhidao.baidu.com/question/282698922.html">http://zhidao.baidu.com/question/282698922.html</a>	此网页包含所有美国总统的生日，找到林肯生日的成本较高，2 分
1	<a href="http://dongxi.net/b14tv">http://dongxi.net/b14tv</a>	该页面较高成本可得知，林肯出生于 2 月 12 日，但是无哪年，1 分
1	<a href="http://www.miniworld.com.tw/article/love2a2a2002/1530644">http://www.miniworld.com.tw/article/love2a2a2002/1530644</a>	该页面只表明哪月哪天，但是没有说明是哪年，对于大部分用户来说价值比较低,1 分

### 3.2.2 原文需求

原文需求，指用户输入的 query，是要某一篇具体的文章，通常用文章标题、文章中的某一句话等搜索。而且通过该 query，就可以定位到某篇具体文章。

原文需求常见有两类：

- ✓ 找某篇文章，那么只要找到具体文章，然后根据文章质量等给出打分；
- ✓ 找某个题目的答案，这种如果只有找到该题目文档，没有答案，那么一般只能给 1 分，有答案根据获取成本（是否便捷）给 2-3 分。

举例说明如下：

Query	痛苦抉择背后，究竟有何隐情？	
Query 说明	一篇新闻稿标题	
用户需求意图	信息需求：了解阅读这篇新闻	
3	<a href="http://news.sina.com.cn/o/2011-12-09/064023602478.shtml">http://news.sina.com.cn/o/2011-12-09/064023602478.shtml</a>	权威网站上提供的新闻稿



Query	一切导体都有阻碍电流的性质，导体的这种性质叫_____	
Query 说明	初三物理单元测试题原文	
用户需求意图	信息需求：找到这道题目的答案	
3	<a href="http://zhidao.baidu.com/question/111308844.html">http://zhidao.baidu.com/question/111308844.html</a>	直接给出了正确答案，3 分
3	<a href="http://www.jyeoo.com/physics/ques/detail/8f22759b-651e-4a7a-b6c9-4e3c175f5057">http://www.jyeoo.com/physics/ques/detail/8f22759b-651e-4a7a-b6c9-4e3c175f5057</a>	试题解析，直接给出此题答案，3 分
2	<a href="http://apps.hi.baidu.com/share/detail/30942484">http://apps.hi.baidu.com/share/detail/30942484</a>	试题集合，有答案，但查找十分困难，2 分
1	<a href="http://space.30edu.com/05384423/ReadArticle.aspx?ID=f9dc87c4-b6b8-45c1-a4f1-687f6a75e6c6">http://space.30edu.com/05384423/ReadArticle.aspx?ID=f9dc87c4-b6b8-45c1-a4f1-687f6a75e6c6</a>	找到了该试题，但是没有答案，给 1 分；

注意：原文 query 下，url 往往只要命中这句话，能够得知来源、前后文、出处等，至少是相关的，也就是非 0 分的。根据质量好坏、完整与否来判断需求满足情况，往往给 1-3 分。如果确定是始发原创，也可以给 4 分。

Query	需求	URL 与打分	易错点
镇压天地	原文 query，实际是某小说的章节名	<a href="http://www.00ks.com/Html/Book/14/14695/3712819.html">http://www.00ks.com/Html/Book/14/14695/3712819.html</a> 该章节小说阅读，3 分	可满足主需求的结果，打分一般都较高，需要先确认好主需求才打分
sqlite3 error lnk1181: 无法打开输入文件 d:	原文 query，通过这种方法来满足疑难求解需求	<a href="http://hi.baidu.com/liyy011/blog/item/7a7ff0d8937300e738012f42.html">http://hi.baidu.com/liyy011/blog/item/7a7ff0d8937300e738012f42.html</a> , 1 分，命中了该段语句，但是实际满足程度比较差	主需求，在页面较差的时候，只要满足一点需求，就不应当 0 分
大家都认为初三的甲某比较厉害. 有一天,甲某向初一的乙某借钱,乙某因惧怕甲某,被迫	原文需求，既想知道出自哪里，也想知道答案	<a href="http://zhidao.baidu.com/question/206573018.html">http://zhidao.baidu.com/question/206573018.html</a> ；只有问题集合，且找到问题原文的成本稍大，无答案，2 分	原文类 query 需要使用 ctrl+f 进行查询，以免实际包含该 query 的页面，被误当不相关
第一次握你的手	原文 query，实际是歌词中的一句，需要找到对应的歌词、歌名	<a href="http://zhidao.baidu.com/question/293223588.html">http://zhidao.baidu.com/question/293223588.html</a> <a href="http://zhidao.baidu.com/question/397707">http://zhidao.baidu.com/question/397707</a> 均有歌词，3 分	主需求就是所在歌词，只要能满足，打分不能过低
劳其筋骨饿其体肤	原文需求，了解出处、渊源，因是古文可能还有释义需求	<a href="http://blog.csdn.net/arielpx/archive/2004/08/12/72003.aspx">http://blog.csdn.net/arielpx/archive/2004/08/12/72003.aspx</a> <a href="http://v.youku.com/v_show/id_XMTkyMTk4NzMy.html">http://v.youku.com/v_show/id_XMTkyMTk4NzMy.html</a> 均为 1 分，实际只有 title 相关可知道上下文信息，页面内	即便是原文 query，也不能直接认为命中的一大段话，就一定到 2 分，需要看其多大程度的满足用户需求，是只有一小段原文命中，整体无关，还是整

		容和视频均完全无关	篇原文或者对小段原文的深入阐述
--	--	-----------	-----------------

原文 Query 有时候出现在整篇文章中的某一处，不易查找，此时可以用浏览器的查找功能进行查找。使用 Ctrl+F 热键，唤出查找界面，然后输入 Query 在网页中进行查找。

对于较长的 Query，尤其是含有标点符号的 Query，可能会因为些微的差异或者标点符号全角半角的原因而查找不到。此时建议使用 Query 的一小部分进行查找。

例如对于 Query 【天将降大任于斯人也，必先苦其心志】，可能网页中的确含有这句话，但因为标点符号全角半角不匹配而查找不到。此时可以只查找“天将降大任于斯人也”，再看找到的部分上下文是否可以匹配 Query。

### 3.2.3 问答类 query

问答类 query 是非常常见的一种类型，即用户有问题，然后希望来搜索引擎搜索获得答案。

一般衡量标准为：主要看问题是否得到回答，答案是否准确、直接、有效。过于专业的问题，参考回帖、评论等反馈来判断回答质量。如有可能，找到权威答案再判断。有问无答的页面一般给 0 分。

Query	网上银行汇款手续费怎么收	
用户需求意图	信息需求：找网上银行汇款手续费怎么收的信息，具体哪个银行不太清楚，希望知道手续费怎么收，收多少钱。	
3 or 2	<a href="http://zhidao.baidu.com/question/134840510">http://zhidao.baidu.com/question/134840510</a>	各个银行的手续费收取说明，质量好；但是时效性可能没有保障，有可能今年已经有更新了
1	<a href="http://zhidao.baidu.com/question/71194159.html">http://zhidao.baidu.com/question/71194159.html</a>	只有一个银行的说明；且时间是 2008 年的
1	<a href="http://zhidao.baidu.com/question/287116756.html">http://zhidao.baidu.com/question/287116756.html</a>	结果是关于手续费能否让收款人收，非常冷门的需求，对大多数用户都没有价值；
0	<a href="http://www.lzbs.com.cn/rb/2005-07/14/content_428393.htm">http://www.lzbs.com.cn/rb/2005-07/14/content_428393.htm</a>	网上汇款要当心，出了差错手续费照收。转义，无法得到手续费怎么收的情况；

## 3.3 交互操作需求

### 3.3.1 在线预订、购买需求

对于在线预订、购买类需求，对站点知名度、操作有效性等要求较高，用户会认为知名权威的站点更可靠。

Query	北京到哈尔滨 机票预订	
用户需求意图	交互操作需求：找“北京到哈尔滨”的机票信息，并且完成在线预订等	
3	<a href="http://flight.qunar.com/site/oneway_list.htm?searchDepartureAirport=%E5%8C%97%E4%BA%AC&amp;searchArrivalAirport=%E5%93%88%E5%B0%94%E6%B%A8&amp;searchDepartureTime=2012-03-09&amp;searchArrivalTime=2012-03-12&amp;nextNDays=0&amp;startSearch=true&amp;from=qu">http://flight.qunar.com/site/oneway_list.htm?searchDepartureAirport=%E5%8C%97%E4%BA%AC&amp;searchArrivalAirport=%E5%93%88%E5%B0%94%E6%B%A8&amp;searchDepartureTime=2012-03-09&amp;searchArrivalTime=2012-03-12&amp;nextNDays=0&amp;startSearch=true&amp;from=qu</a>	优质网站的指定线路查询页面，用户需求满足较好

	<a href="#">narindex</a>	
2	<a href="http://www.airchina.com.cn/?cid=bptopt0301">http://www.airchina.com.cn/?cid=bptopt0301</a>	中国国航的官网，可以输入城市名来预订
2	<a href="http://www.gz020hotel.com/jipiao/cityto-city-PEK-HRB.html">http://www.gz020hotel.com/jipiao/cityto-city-PEK-HRB.html</a>	也是北京到哈尔滨的机票查询，但网站知名度低，认为权威性比较差，不一定准确
1	<a href="http://zhidao.baidu.com/question/286195840.html?an=2&amp;si=1">http://zhidao.baidu.com/question/286195840.html?an=2&amp;si=1</a>	用户需求很明确，就是预订需求，此页面是关于预订的一些参考信息，对少数人来说有用
0	<a href="http://jipiao.quna.com/flightList/CA1640.html">http://jipiao.quna.com/flightList/CA1640.html</a>	是指定线路的某一个航班的具体信息，但是无法预订

### 3.3.2 游戏需求

游戏类需求，主要为在线玩儿游戏、游戏下载，以及攻略、介绍、视频等需求。

Query	洛克王国	
用户需求意图	寻址需求：去官网玩儿游戏； 交互操作需求：在某个站点玩儿游戏； 信息需求：找洛克王国攻略等相关信息	
4	<a href="http://17roco.qq.com/">http://17roco.qq.com/</a>	官方在线玩页面，4 分
3	<a href="http://www.4399.com/flash/32979.htm">http://www.4399.com/flash/32979.htm</a>	知名小游戏网站 4399 的洛克王国游戏，用户可在线玩儿，以及通过链接去官网玩儿
2	<a href="http://news.4399.com/luoke/xinwengonggao/">http://news.4399.com/luoke/xinwengonggao/</a>	洛克王国最新资讯信息专题
2	<a href="http://news.4399.com/luoke/gonglue/">http://news.4399.com/luoke/gonglue/</a>	洛克王国攻略专题
1	<a href="http://jingyan.baidu.com/article/6766299710b1ed54d51b84a0.html">http://jingyan.baidu.com/article/6766299710b1ed54d51b84a0.html</a>	洛克王国某一方面的具体攻略

### 3.3.4 图片浏览

图片的需求，一般越丰富、越清晰等越好，如果图片资源失效，无法查看，则没有价值。

以【赵薇 图片】为例：

URL	说明
<a href="http://gb.cri.cn/9964/2006/02/16/Zt114@902338.htm">http://gb.cri.cn/9964/2006/02/16/Zt114@902338.htm</a>	很多类型的图片，非常丰富，3 分
<a href="http://www.haibao.cn/star/1689/">http://www.haibao.cn/star/1689/</a>	虽然图片也不少，但是毕竟凌乱无分类，而且页面有很多其他的主题，并非是图片专门页面，2 分
<a href="http://zhidao.baidu.com/question/142833250.html">http://zhidao.baidu.com/question/142833250.html</a>	图片较少还算有特点，满足部分人需求，2 分
<a href="http://zhidao.baidu.com/question/142444328.html">http://zhidao.baidu.com/question/142444328.html</a>	主要是文字描述，最佳答案所给链接不能满足，底部有个链接点击后可看到较单一的图片，1 分
<a href="http://zhidao.baidu.com/question/13949263.html">http://zhidao.baidu.com/question/13949263.html</a>	无图片，无任何参考价值，0 分

### 3.3.5 小说阅读

小说类需求主要为在线阅读，其次为下载等需求。一般评估原则如下：

- ✓ 原发官方书籍主页、目录页，4 分，即使部分章节需要购买权限等也是给 4 分；

- ✓ 其他站点小说主页、目录页，更新到最新章节（或完结的有全部章节），都能免费在线看，根据页面质量 3-2 分，如广告太多，排版略差，给 2 分；
- ✓ 章节不太全，缺少少量章节，根据页面质量，一般 2-1 分；
- ✓ 只有少量章节可在线看，其他缺少或需要权限阅读等，一般 1 分；

Query	斗破苍穹	
用户需求意图	交互操作需求：主要是在线看斗破苍穹小说，其次为下载小说 信息需求：小说介绍信息、小说评论信息等	
4	<a href="http://www.qidian.com/BookReader/1209977.aspx">http://www.qidian.com/BookReader/1209977.aspx</a>	官网目录页
4	<a href="http://www.qidian.com/Book/1209977.aspx">http://www.qidian.com/Book/1209977.aspx</a>	官网该书主页
3	<a href="http://www.laishu.com/doupocangqiong/">http://www.laishu.com/doupocangqiong/</a>	其他小站转载的免费、全面文章，3 分
1	<a href="http://www.morui.com/book/11/11386/index.html">http://www.morui.com/book/11/11386/index.html</a>	虽然看起来章节是齐全的，但是后面有些章节无法阅读
1	<a href="http://www.shuoshuo520.com/Html/0/391/3242188.html">http://www.shuoshuo520.com/Html/0/391/3242188.html</a>	其中的某一章节页面，点击可以回到目录页
0	<a href="http://www.xkzw.org/xkzw3232/">http://www.xkzw.org/xkzw3232/</a>	为续集，非原需求的书

### 3.3.6 软件下载需求

官方网站上的下载页 4 分，其它最高给 3。

软件需求满足的关键因素能否下载使用，这是最重要的。若无法下载，对用户价值很低。软件 query 的打分，还要考虑软件的类型，是不是免费软件，要不要花钱。一些重要的参考原则如下：

- ✓ 如果是免费软件或不影响使用的共享软件，知名下载站的结果，可打 3 分。如搜【迅雷下载】，天空、华军、太平洋的下载页都能打 3 分；
- ✓ 若软件要花钱注册，不注册限制多难以正常使用，则受限版本的结果，用户信息获取的成本较高，给 1 分；能方便使用的版本，给 2 分或 3 分，如【photoshop 下载】、【autocad 下载】。
- ✓ 结果限定了要使用某种下载软件才能下载，要考虑下载软件的普及率。如很多用户都电脑都安装了迅雷，若大型的软件、网络游戏只能用迅雷下载，可默认用户安装了，不影响打分。如果要用网际快车（FlashGet）、才能下载，那就要减分，因为 flashget 普及程度相对低，很多用户没装。

Query	迅雷下载	
用户需求意图	交互操作需求：下载迅雷软件	
4	<a href="http://dl.xunlei.com/xl7.html">http://dl.xunlei.com/xl7.html</a>	官网下载页
4	<a href="http://dl.xunlei.com/">http://dl.xunlei.com/</a>	官网，软件下载列表页
3	<a href="http://www.skycn.com/soft/14857.html">http://www.skycn.com/soft/14857.html</a>	知名网站下载页面，3 分

Query	acdsee 下载	
3	<a href="http://www.crsky.com/soft/9483.html">http://www.crsky.com/soft/9483.html</a>	破解版，不注册可使用，3 分
2	<a href="http://www.51chongdian.net/soft/36/45/2007/20070606628.html">http://www.51chongdian.net/soft/36/45/2007/20070606628.html</a>	破解版，但需要多次点击才能找到下载链接，便捷性较差，2 分
1	<a href="http://www.skycn.com/soft/6612.html">http://www.skycn.com/soft/6612.html</a>	需要注册后才能正常使用，1 分

0	<a href="http://www.yajyw.com/Soft/qtrj/200609/117.html">http://www.yajyw.com/Soft/qtrj/200609/117.html</a>	不能下载，0 分
---	---	----------

### 3.3.7 视频观看需求

搜索电视剧、电影、电视节目，用户主要是想在线观看，其次是下载、介绍、影评、演员等。视频可分为单个和连续的。单个如电影、某期电视节目，连续的如电视剧、多期的综艺节目。

评估时，主要考虑是否视频能否观看、清晰程度等。如果是连续性的，还要考虑排列是否有序、是否完整、有没有最新剧集。

有些结果要安装软件或插件才能播放，这增加了成本，评分要比直接能看的低。

视频类的搜索很少有 4 分，因为没有官方网站。一般评估参考原则如下：

- ✓ 剧集完整（连载更新到最新），排列有序，可以有效观看，3 分；
- ✓ 丰富的百科介绍，根据需求情况，如果需求也较高，3 分；
- ✓ 剧集不太完整（如缺少少数集数），排列不太有序，2 分；
- ✓ 单独某一集，但能方便获取到完整剧集列表，2 分；
- ✓ 需要安装 Qvod、百度影音，但剧集更新到最新，根据质量情况，一般 2 分，如果页面质量很差，1 分；
- ✓ 电视剧简单介绍、花絮和预告、单独具体某一集，一般 1 分；

Query	失恋 33 天	
用户需求意图	交互操作需求：主要是在线观看《失恋 33 天》这部电影；购买电影票 信息需求：了解影院信息、介绍信息、演员信息、评论信息等	
3	<a href="http://www.tudou.com/playlist/album/id69874.html">http://www.tudou.com/playlist/album/id69874.html</a>	完整版高质结果在线观看
3	<a href="http://www.letv.com/ptv/pplay/54534/1.html?ref=baiduopen">http://www.letv.com/ptv/pplay/54534/1.html?ref=baiduopen</a>	完整版高质结果在线观看
3	<a href="http://baike.baidu.com/view/3158523.htm">http://baike.baidu.com/view/3158523.htm</a>	失恋 33 天的百科介绍，很丰富
2	<a href="http://app.ent.ifeng.com/movie/detail/4655">http://app.ent.ifeng.com/movie/detail/4655</a>	该电影的介绍，海报等，次需求
1	<a href="http://tv.sohu.com/20111124/n326756810.shtml">http://tv.sohu.com/20111124/n326756810.shtml</a>	只有一小部分，无法观看正常影片
1	<a href="http://www.cnipr.com/news/dailykeyword/201111/t20111123_138699.html">http://www.cnipr.com/news/dailykeyword/201111/t20111123_138699.html</a>	简短介绍，主要是因此而扩展出来的版权等相关思考
0	<a href="http://wenda.tianya.cn/wenda/thread?tid=16f10Ins843agjv9lpulgj8d8ckebyqfr86g4v">http://wenda.tianya.cn/wenda/thread?tid=16f10Ins843agjv9lpulgj8d8ckebyqfr86g4v</a>	Title 中有失恋 33 天，但是实际页面是关于二维码的使用的

### 3.4 寻址类型

寻址类 query，一般官网给 4 分，如新浪、4399、淘宝、北京大学等。。在 4 分介绍中说明已经较多，所以这里就不详细解说了。只列举一个例子，供大家参考。



Query	1 号店	
用户需求意图	寻址需求：找 1 号店官网 信息需求：了解 1 号店这个网站相关信息等	
4	<a href="http://www.yihaodian.com/">http://www.yihaodian.com/</a>	官网，4 分
3	<a href="http://e.weibo.com/yihaodian">http://e.weibo.com/yihaodian</a>	官网微博，页面基本上都是在介绍一号店的产品，3 分
3	<a href="http://www.yihaodian.com/2/">http://www.yihaodian.com/2/</a>	1 号点北京分站页面，3 分；如果确定 query 是来自北京用户，4 分
2	<a href="http://baike.baidu.com/view/2164590.htm">http://baike.baidu.com/view/2164590.htm</a>	百科介绍，非主需求，2 分
2	<a href="http://b2b.yihaodian.com/index.do">http://b2b.yihaodian.com/index.do</a>	企业频道，需求比例比个人低，2 分
2	<a href="http://www.51fanli.com/product-633.html">http://www.51fanli.com/product-633.html</a>	返利网上一号店的介绍、评论、优惠券等，2 分
1	<a href="http://zhidao.baidu.com/question/355464402.html">http://zhidao.baidu.com/question/355464402.html</a>	评论，相关但非主需求，1 分
1	<a href="http://www.yihaodian.com/cmsPage/show.do?pageId=1186&amp;merchant=2">http://www.yihaodian.com/cmsPage/show.do?pageId=1186&amp;merchant=2</a>	非常具体的官网内页，1 分
0	<a href="http://www.21cbh.com/HTML/2012-3-6/5MMDQwXzQwNzI5MQ.html">http://www.21cbh.com/HTML/2012-3-6/5MMDQwXzQwNzI5MQ.html</a>	页面是“它是下一个 1 号店”，只有 title 好，实际是讲京东的，0 分

## 3.5 其他特定类型 query

### 3.5.1 色情需求

当需求是寻找色情资源时，需要点到最后一步，确定能否真正获取到相关资源再打分。不要看到“看起来诱人”的词语就中止，因为存在很多虚假的色情网站，让人频繁点击，最后却只有广告，没有相关资源。

一般没有提供真实色情资源的，即没有满足用户的需求，给 0 分。

### 3.5.2 Url 查询的 query

有些 query 直接就是某个具体的 url，或者看起来和 url 非常相似，称之为“url 查询”；

url 查询一般有下面两种形式：一是规范的典型 url 查询，二是非规范的 url 查询，但同样我们都认为是 url 查询的需求。

#### (1) 典型的 url 查询：

即输入浏览器地址栏中，能直接打开某个网页，包括：

- ✓ Query 本身就是一个完整的 url，范例：[www.youku.com](http://www.youku.com)
- ✓ 有些虽然没有前缀，但确实也是完整 url 的：[xunlei.com](http://xunlei.com)

#### (2) 非规范的 url 查询

有一些 query 看上去直观印象就是 url 查询，但是实际将这个 url 复制到地址栏时，却不能打开。

虽然如此，我们仍然认为只要用户输入了一个明确的 url，那么在用户心目中应当是有一个明确的页面的，一般包括：



- ✓ 跟正确的 url 一样，但是无法打开页面，比如该站点域名过期导致无法访问；
- ✓ url 中存在明显输入错误的情况，如 url 完整，但是有拼错部分，范例：www.mcriosoft.com，明显可知是微软官网拼错了；
- ✓ url 查询中包括了其他 term，如【佳能 D50 360buy.com】

### (3) 非 url 查询

- ✓ Query 本身没有 www,com,cn 之类的常见 url 前缀后缀，只有主域名，范例：youku
- ✓ Query 本身直接是正常的中文 query，范例：优酷、京东

### (4) 评估注意事项

对于 url 查询的 query，请仔细查看评估页面的 url 地址，看跟 query 的 url 地址是否匹配。url 查询的 query 很容易输入错，所以对于输入不完整以及不正确的 url，需要仔细分析用户需求意图。

4 分的情况：

- ✓ 当 url query 跟页面的 Url 地址完全匹配时。注意：url 页面的地址有时候可能会带有更长的一些字符串或者看起来跟 query 不一样，但是实际页面是一样的。如 t.sina.com.cn 与 weibo.com 即是同一个页面，但 url 不一样。所以对于 query 【t.sina.com.cn】，weibo.com 应该给 4 分。
- ✓ 如果 query 的 url 不是完整的正确的，需要根据 query 分析用户需求意图，如果比较确信用户输入 query 所找的 url 页面跟评估的页面一致，才给 4 分。如【www.mcriosoft.com】，显然是找微软官网，而 query 拼写错误了，官网为 <http://www.microsoft.com/en-us/default.aspx> 或者中文官网 <http://www.microsoft.com/zh-cn/default.aspx>。

当用户输入 url query 的时候，用户的需求目的很明确，往往是要去具体的页面。这些直接匹配的页面应该给 4 分。

很少有直接匹配 url query 的结果，但是却不打 4 分的情况。有时候，直接匹配的 4 分页面是唯一对用户有价值的结果，但有时候，同时也存在一些其他对用户有价值的页面。如一些关于网站介绍信息的页面，这些页面打分可以从 0-3 档。大致规则如下：

- ✓ 目标 url，4 分；
- ✓ 频道页，3 或 2 分；
- ✓ 网站统计信息，1 or 0 分；
- ✓ 介绍信息，但是网站的介绍评论说明等页面，根据质量情况，可以给 2-0 分，如作弊广告很多，1 或 0 分；内容较旧，一般给 1 或者 0 分

### 3.5.3 输入错误的 query

在评估中，我们会遇到一些用户输入错误的 query。一般评估原则为：

- ✓ 对于明显输入错误的 query，你可以根据用户需求意图来进行打分，而不是用户实际输入的错误 query；
- ✓ 对于非明显输入错误的 query，可以就根据用户实际输入的 query 本身，来判断用户的需求是什么；
- ✓ 人名类 query，如果发现结果 url 页面有直接跟 query 匹配的人物，那么一般不要猜测用户输入了错误的 query，直接以用户输入的原 query 进行对应打分；如果一个人名下，同时可能有比较知名的人物跟该人名表述类似，可以根据你的判断，看是否对名人也有需求，从而对名人进行需求满足打分。

下面是一些典型用户输入错误 query 的例子：

Query	Query 真正表意	url	url 页面说明	打分
奇异	主要 query 表意为视频网站“奇艺”	<a href="http://www.iqiyi.com/">http://www.iqiyi.com/</a>	奇艺视频网站的官网	4
投书韵达快递	Query 表意需求为“投诉韵达快递”	<a href="http://www.kuaidi100.com/all/yd.shtml">http://www.kuaidi100.com/all/yd.shtml</a>	韵达快递全国网点电话	3

人名的 query 比较难以判断，下面是一些例子。第一个 query 不应该认为是绝对输入 query 错误，第 2 个 query 是明显的 query 输入错误。

Query	Query 真正表意	url	url 页面说明	打分
陈格儿	可能就是找“陈格儿”，而不要臆测用户找“陈格尔”	<a href="http://t.qq.com/huaxianzi7">http://t.qq.com/huaxianzi7</a>	一个叫“陈格儿”的人的微博	1
失传零花	主要是找“石川铃华”这个 AV 女优	<a href="http://baike.baidu.com/view/1751582.htm">http://baike.baidu.com/view/1751582.htm</a>	该女优的百科介绍	2

### 3.5.4 列表页需求 query

有些 query，看上去用户是期望获取较多信息，而不是单个结果信息能满足需求，常见的好结果为一个相关列表，所以称之为列表页类需求 query。这类 query 打分一般考虑结果提供信息的丰富程度。

对于这类 query 来说，大致的判断原则为：

- ✓ Query 所求的是一个列表时，一个页面上有多个相关结果，是列表页面时，往往比单条结果好；
- ✓ Query 所求的列表比较短的时候，仅有一两个有效信息条目的结果，也还算有帮助，但价值并不一定高于单条优质的结果。

Query	需求	URL 与打分	说明
鸡肉食谱	有关鸡肉的食谱，用户希望从中挑选从而做菜，有列表需求，对丰富程度有要求	<a href="http://www.meishichina.com/YuanLiao/JiRou/">http://www.meishichina.com/YuanLiao/JiRou/</a> , 3 分，优质页面，具有多种以鸡肉为主要食材的菜谱	对于丰富度有较高的要求，根据其条目多少，各条目的质量好坏来决定 0-4 之间的打分
		<a href="http://www.meishichina.com/Eat/Magic/200711/24321.html">http://www.meishichina.com/Eat/Magic/200711/24321.html</a> , 2 分，做法不是特别多，页面质量相对一般	
		<a href="http://www.meishichina.com/Eat/RMenu/200803/32028.html">http://www.meishichina.com/Eat/RMenu/200803/32028.html</a> , 1 分，只有一种菜的做法，丰富度比较差	
		<a href="http://wenwen.soso.com/z/q125507923.htm">http://wenwen.soso.com/z/q125507923.htm</a> , 0 分，仅仅提到了做鸡鸭的饭店，但是完全没有做法，无法满足用户需求	
婴儿玩具	寻找婴儿玩具的相关信息或购买婴儿玩具	<a href="http://category.dangdang.com/all/?category_id=4002067">http://category.dangdang.com/all/?category_id=4002067</a> , 3 分，知名购物网站上的婴儿玩具专题，量大且相对权威	根据页面实际的条目多少，来进行打分，条目越多、质量越高，则得分越高；任一者不够好，则需要适当降档打分
		<a href="http://www.baobeiwanju.com/">http://www.baobeiwanju.com/</a> , 2 分，多个品牌 logo，点击可去往对应的淘宝商城，但是	

		<p>页面本身是非权威知名的</p> <p><a href="http://www.amazon.cn/PLAYGO-%E8%B4%9D%E4%B9%90%E9%AB%98-%E5%A4%9A%E5%8A%9F%E8%83%BD%E4%B8%89%E8%A7%92%E6%B8%B8%E6%88%8F%E6%9E%B6-2249/dp/B003Z0CAQY/ref=sr_1_5?s=toys-and-games&amp;ie=UTF8&amp;qid=1331138950&amp;sr=1-5">http://www.amazon.cn/PLAYGO-%E8%B4%9D%E4%B9%90%E9%AB%98-%E5%A4%9A%E5%8A%9F%E8%83%BD%E4%B8%89%E8%A7%92%E6%B8%B8%E6%88%8F%E6%9E%B6-2249/dp/B003Z0CAQY/ref=sr_1_5?s=toys-and-games&amp;ie=UTF8&amp;qid=1331138950&amp;sr=1-5</a>; 1 分，具体的某个婴儿玩具，需求满足程度低</p> <p><a href="http://search.yihaodian.com/s/c0-k%25E9%2581%25A5%25E6%258E%25A7%25E9%25A3%259E%25E6%259C%25BA">http://search.yihaodian.com/s/c0-k%25E9%2581%25A5%25E6%258E%25A7%25E9%25A3%259E%25E6%259C%25BA</a>, 0 分，遥控飞机实际不适合婴儿</p>	
旅馆	<p>用户可能是想要旅行，因而需要订酒店，即使这个需求很泛，没有明确的具体限定，我们也可以判断出哪个结果好，哪个结果差</p>	<p><a href="http://hotel.qunar.com/?ex_track=bd_aladding_hotelsearch_title">http://hotel.qunar.com/?ex_track=bd_aladding_hotelsearch_title</a>;</p> <p><a href="http://www.ctrip.com/">http://www.ctrip.com/</a>;</p> <p>两个 3 分结果，去哪儿，携程是比较有名的网站，可以方便的查询到各处的旅馆信息</p> <p><a href="http://www.homeinns.com">http://www.homeinns.com</a>; 2 分，知名的连锁旅店官网，可以方便的查询预订</p> <p><a href="http://www.7daysinn.cn/innpage.php?id=1151">http://www.7daysinn.cn/innpage.php?id=1151</a>, 1 分，7 天官网的某个非常具体的旅店，虽然也可以查询，但是整体稍差</p> <p><a href="http://www.kudipetshotel.com/">http://www.kudipetshotel.com/</a>;</p> <p>旅店所指的是人所需要的旅店，宠物旅店的需求几乎没有，0 分即可</p>	<p>一些聚合信息的网站，例如去哪儿、携程等类似的网站，提供的信息非常多，往往打分会比较高，可以达到 3 分；对于一些比较知名的连锁旅店，给出官网也比较受认可，可达 2 分；</p> <p>对于非常具体的某个地方的某个旅馆，需求满足程度很低但至少是相关的，可以给 1 分；</p> <p>字面含义基本就是人所需要的旅馆，宠物旅馆符合用户需求的比例非常低，可以给 0</p>

### 3.5.5 时效性需求 query

有些 query，对时效性要求很高。如对于近期或突发的一些事件，用户希望看到最新的内容。如用户输入一个 query 来找体育赛事、每年的考试等，一般来说都是在最近的相关信息，而不是几年前的信息，除非 query 明确表述找过去某个时间的信息。

#### (1) 时效性类别：

根据时效性特点，以及好的结果的特征，具有时效性需求的 query 大概可以分为以下几个类别：

类型	需求特征	好的结果特征	举例
一般泛时效性	主要需求维度上一直有时效性要求	资源随时间一直有变化，能够较好地满足用户需求的资源一般都是一年以内的	<p>A. 商品名（未明确询价）且主需求是购买类，市场行情，商品对比，生活服务类查询（比如找保洁、搬家等服务信息）</p> <p>B. 名人（不包括已故名人）</p>

			<p>C. 最新政策、法规等（考虑更新频率，如几年更新一次法律不在此范畴，而关注是否是最新的法律条文）</p> <p>D. 连载中的小说（因为大多更新周期不固定，不属于周期），热播电视剧</p> <p>E. 媒体，如《南方周末》</p>
强泛时效性	主要需求维度上一直有时效性要求	资源更新频率高，只有最新资源（一天或几天以内）才能较好地满足用户需求，非时效性资源对用户伤害大	<p>A. 天气预报，汇率，股票行情——每日更新</p> <p>B. 各种价格，二手商品，房屋租赁，招聘，彩票，影院排片表，dnf 封号查询，违章查询——过期信息基本无价值</p> <p>C. 外挂、私服、激活码——失效快</p> <p>D. query 中包含特定 term，如：最新，新闻，XX 小说最新章节，现在几点；</p> <p>E. 无固定周期的排行榜，如金曲排行</p> <p>F. 团购——过期团购信息价值低</p> <p>* 特例（资源更新频率低，但过期信息对用户伤害大）：列车时刻表，火车票、机票查询</p> <p>* 注意：违章/火车票/机票/列车时刻/影讯查询：如是明确的查询寻址需求，不是此范围</p>
周期时效性	主要需求维度上有时效性要求	资源周期性更新或突增，但最新一个周期内的资源才是用户最想要的	<p>A. 定期更新的漫画、动漫、电视剧</p> <p>B. 定期播出的比赛、电视节目</p> <p>C. 年度、季度、月度、甚至每周发布的排行榜</p> <p>D. 节日、节气</p>
突发时效性	主要需求维度上有强烈的时效性要求	资源突增，非时效性资源不能满足用户需求	即一般的社会事件，如最近发生的有一定影响力的地震、大火、洪水等都属于此范畴

**（2）评估中以下几个方面需要说明：**

✓ 突发时效性：

典型具有时效性要求的 query 如突发新闻事件等，如突然发生的地震、大火等，一旦时效性很差，没有给出最新的结果，那么有可能是不能满足用户需求，打分时可以因为时效性差，而打分比较低。

#### ✓ 周期性事件的时效性：

有一些 query，是跟周期性有关，例如高考、奥运、世界杯等，是周期往复发生的，在这类 query 下，判断什么样的页面最有用之前，要考虑抓取结果当时该事件处于什么阶段，如果是一场比赛刚发生在几个月前，那么最有价值的页面一般都是刚过去比赛的相关信息，如果一场比赛在几个月后举行，那么最有价值的页面是有关于即将举办的比赛有关信息，这方面需要做一下判断。

如果页面是事件的官网，那么可以给 4 分，不管事件是过去还是即将到来的。

#### ✓ Query 有明确时间：

“中国人口”，往往是需要最近的统计数据，比如最新一次数据产出于 2012 年，之前的结果就都不如 2012 的结果好，在同样质量、同样相关的情况下，时效性差的结果应该适当降档。

如果 query 是【1980 年中国人口】，那么就不存在内容是否最新的，而是是否提供了 query 表述的信息。有时候，旧的页面是唯一的满足“过去事件”的好资源。旧的页面并不一定就是差的，这主要取决于用户的需求和页面内容匹配程度。

#### ✓ 页面显示时间与页面内容时间

大多数时候，你需要看页面实际内容描述的时间，而不仅是页面上显示的时间。

#### ✓ 时效性有时不需考虑

有时候，时效性差的网页反而是唯一的关于某事的好结果。时效性差的结果，未必一定是差结果，需要根据实际情况来分析。

Query	需求	URL 与打分	易错点
鸦片战争	较久以前的鸦片战争的信息需求	<a href="http://book.360buy.com/10880231.html">http://book.360buy.com/10880231.html</a> ，2 分，京东上某本关于该战争的书，虽然已经比较久了，最近一次修订已经有些年，但是本身是历史事件，对于时效性要求不高	对于比较久远的事情，时效性需求往往并不高，即使页面内容时效性稍差，但往往也不需要特意降档
奥黛丽 赫本 逝世	查找关于某人逝世相关的信息	<a href="http://info.pharmacy.hc360.com/2009/03/17125676822.shtml">http://info.pharmacy.hc360.com/2009/03/17125676822.shtml</a> <a href="http://www.80kuku.com/a/234/87995.html">http://www.80kuku.com/a/234/87995.html</a> ; <a href="http://www.people.com.cn/digest/200005/10/wy051006.html">http://www.people.com.cn/digest/200005/10/wy051006.html</a> 以上多个结果都可以给 2 分	赫本逝世已经很久，关于当时逝世情况、人物历史回顾等内容，都可以给到 2 分，并不因为有些网页是 08 年、09 年或者 10 年而导致某个页面认为会比较差。
迈克尔杰克逊 死讯	信息需求	<a href="http://book.360buy.com/10880231.html">http://book.360buy.com/10880231.html</a> <a href="http://news.163.com/special/00013FR0/MichaelJosephJackson.html">http://news.163.com/special/00013FR0/MichaelJosephJackson.html</a> 2 个时效性都比较新的专题页面，维护及时，3 分 <a href="http://ent.qq.com/zt/2009/mjfore">http://ent.qq.com/zt/2009/mjfore</a>	迈克尔逝世当时的报道，后续已经有了更多的进展，对时效性就有了要求；所以内容更新不够及时（时效性差）的页面，需要适当降档打分

		<a href="#">ver/</a> 虽然也是专题，但是最新消息都没有，只更新到 09 年，时效性比较差，但相当完善，2 分	
--	--	---	--

### 3.5.6 地域需求 query

在某些 query 下，query 没有明确地点，但是根据 query 描述，知道用户需求具有地域性特点，即只有当地的结果才能更好的满足需求，如天气预报、移动营业厅、团购、餐厅、电影院等。外包评估中根据标注类型，会分为下面两种情况：

#### (1) 需要考虑地域性

对于线上搜索效果的评估，如持续评估、监控评估、策略评估等，因为我们评估的结果都来自北京地区，所以评估大致原则如下：

基本原则是官网 4 分，符合北京地域需求的结果也 4 分，其他地域结果 1 分。如果所求地域与结果的分站地域有上下位关系，比如北京-海淀，那么给 2 分

Query	需求	URL 与打分	说明
赶集	寻址需求，限定用户所在地是北京	<a href="http://www.ganji.com/">http://www.ganji.com/</a> ，官网，4 分 <a href="http://bj.ganji.com/">http://bj.ganji.com/</a> ，北京分站，4 分 <a href="http://wh.ganji.com/">http://wh.ganji.com/</a> ，武汉分站，1 分	根据评估类型，看是否需要考虑地域需求

#### (2) 不需考虑地域性

对于某些标注类型，不需要考虑地域性需求特点，那么基本原则是官网 4 分，其他各分站不认为有明显区别，全都统一到 3 分。

Query	需求	URL 与打分	说明
赶集	寻址需求，并不限定用户是在北京，无地域要求	<a href="http://www.ganji.com/">http://www.ganji.com/</a> ，官网，4 分 <a href="http://bj.ganji.com/">http://bj.ganji.com/</a> ，北京分站，3 分 <a href="http://wh.ganji.com/">http://wh.ganji.com/</a> ，武汉分站，3 分	根据评估类型，看是否需要考虑地域需求

### 3.5.7 纯英文 query

当 query 是纯英文，给出英文结果页时，不能因为是英文页面就扣分，仍需要按照需求来判断此页面满足情况如何。比如主需求是翻译，则给出全英文的页面不能满足主需求，不好；但比如主需求是介绍，给出英文的介绍页面，也是满足需求的。



## 4、特定 url 页面打分说明

### 4.1 百科类结果

百科结果是评估中大家经常会碰到的页面，这些百科结果往往对于用户是有帮助的，但是在当用户已经了解百科中所介绍的内容、所寻找的是其他方面的信息时，那么对于用户的帮助未必很大，下面举例说明：

Query	需求	URL 与打分	说明
光合作用	介绍需求	<a href="http://baike.baidu.com/view/8885.htm">http://baike.baidu.com/view/8885.htm</a>	百科是非常优质丰富的好结果，3 分
e. g.	找到 e. g. 的翻译，是个信息需求	<a href="http://hi.baidu.com/changcaibai/blog/item/3234319510428b0d7bf48015.html">http://hi.baidu.com/changcaibai/blog/item/3234319510428b0d7bf48015.html</a>	这个类似百科结果，是 e. g. 的相关介绍等，内容比较优质全面，3 分
银行	查找一个银行等	<a href="http://baike.baidu.com/view/20233.htm">http://baike.baidu.com/view/20233.htm</a>	介绍需求很低，满足需求会很少，1 分
电子贺卡	在线制作、发送电子贺卡	<a href="http://baike.baidu.com/view/1202827.htm">http://baike.baidu.com/view/1202827.htm</a>	1 分，用户主需求是在线进行贺卡操作，而非看介绍

### 4.2 搜索结果页

评估中有时会碰到结果是搜索引擎结果页的，这部分结果应该像普通的页面一样去评估，因为搜索结果页的特殊性，可能这部分结果上不太好评估，所以特别给出这部分的说明。

#### 4.2.1 什么是搜索结果页

搜索结果页，除了各个搜索引擎的搜索结果页面外，也包括垂直搜索等范围更小的搜索结果

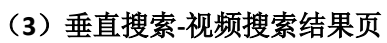
##### (1) 普通搜索结果页

如百度、google 等搜索引擎结果页面



##### (2) 垂直搜索页面-图片搜索结果页

<http://cn.bing.com/images/search?q=%e6%a8%b1%e8%8a%b1&qpv=%e6%a8%b1%e8%8a%b1&FORM=Z7FD>



56

舞林大会

找到约 4,180,000 条结果 (用时 0.06 秒)



### 舞林大会2011 - 高清在线观看- 优酷视频

youku.com

《舞林大会》是由东方卫视和星空卫视联合制作的2011年全明星版的一档周日晚黄金档大型综艺秀 ...

▶ 第20111009期 - 舞林大会 111009 - 2011 - 综艺 ... - 130 分钟

▶ 第20111002期 - 舞林大会 111002 - 2011 - 综艺 ... - 96 分钟

▶ 第20110925期 - 舞林大会 110925 - 2011 - 综艺 ... - 88 分钟



### 舞林大会2012 - 第20120415期- 舞林大会120415 - 2012 - 综 ...

v.youku.com

新增 1 天前 - 91 分钟

舞林大会2012 《舞林大会》是由东方卫视和星空卫视联合制作的全明星版的一档周日 晚黄金档大型综艺秀 ...

(4) 某些站点的站内搜索结果页

比如下面的京东商城的搜索结果页，

<http://search.360buy.com/Search?keyword=%B4%F7%B6%FB%B5%E7%C4%D4>

“戴尔电脑”搜索结果(189)

相关搜索： 戴尔电脑主机 | 戴尔电脑键盘保护膜 | 戴尔电脑适配器 | 戴尔电脑键盘膜

排序： 相关性 销量 价格 评论数 上架时间

1/8 上一页 下一页

库存 选择省份 只显示有货

共189个商品



戴尔 (Dell) V230SR-826C台式电脑 (双核E5700 2G内存 500G硬盘 DVD光驱)  
¥2999.00

已有638人评论 返券

购买 关注 对比



戴尔 (Dell) I620SR-356B台式电脑 (双核i3-2120 2G内存 500G硬盘 1G独显)  
¥3899.00

已有381人评论 返券

购买 关注 对比



戴尔 (Dell) I560SR-826M-8台式电脑 (双核E6600 2G内存 500G硬盘 DVD光驱)  
¥3399.00

已有69人评论 返券

购买 关注 对比



戴尔 (Dell) V260SR-526台式电脑 (双核G630 2G内存 500G硬盘 DVD 三年上)  
¥2999.00

已有98人评论 返券

购买 关注 对比

## 4.2.2 搜索结果页的打分

如果搜索引擎页实际为空页面，或者实际没有任何搜索结果，那么应该是 0 分；

如果搜索引擎页具有一组搜索结果，那么其评分应在 0-3 分之间，具体根据搜索结果页实际有多少内容对用户需求满足打分。

注意：对于百度视频、google 视频、图片等搜索结果页，不能不查看实际页面需求满足就给出高分。

下面给出一些具体评估范例参考：

Query	需求	URL 与打分	说明
北京的必胜客	查找北京必胜客地址	<a href="http://map.baidu.com/?newmap=1&amp;s=con%26wd%3D%E5%BF%85%E8%83%9C%E5%AE%A2%26c%3D131&amp;fr=alaeg&amp;ext=1&amp;from=alamap">http://map.baidu.com/?newmap=1&amp;s=con%26wd%3D%E5%BF%85%E8%83%9C%E5%AE%A2%26c%3D131&amp;fr=alaeg&amp;ext=1&amp;from=alamap</a> ; 3 分，地图上多家必胜客地址信息，方便丰富	地图、地址很类似，地图搜索页往往能够很好的满足地址寻找需求
		<a href="http://image.baidu.com/i?tn=baiduimage&amp;ct=201326592&amp;cl=2&amp;lm=-1&amp;st=-1&amp;fm=result&amp;fr=&amp;sf=1&amp;fmq=1331708406997_R&amp;pv=&amp;ic=0&amp;z=&amp;se=1&amp;showtab=0&amp;fb=0&amp;width=&amp;height=&amp;face=0&amp;istype=2&amp;word=%B1%B1%BE%A9%B1%D8%CA%A4%BF%CD&amp;s=0">http://image.baidu.com/i?tn=baiduimage&amp;ct=201326592&amp;cl=2&amp;lm=-1&amp;st=-1&amp;fm=result&amp;fr=&amp;sf=1&amp;fmq=1331708406997_R&amp;pv=&amp;ic=0&amp;z=&amp;se=1&amp;showtab=0&amp;fb=0&amp;width=&amp;height=&amp;face=0&amp;istype=2&amp;word=%B1%B1%BE%A9%B1%D8%CA%A4%BF%CD&amp;s=0</a> 0 分	Query 主需求是地址，某个店的地址大概是 1 分，某个店的图片更难以满足用户需求，给 0 分较好
吸尘器	查找吸尘器的相关信息，购买吸尘器	<a href="http://www.amazon.cn/s/ref=nb_sb_noss_1?_mk_zh_CN=%E4%BA%9A%E9%A9%AC%E9%80%8A%E7%BD%91%E7%AB%99&amp;url=search-alias%3Daps&amp;field-keywords=%E5%90%B8%E5%B0%98%E5%99%A8&amp;x=0&amp;y=0">http://www.amazon.cn/s/ref=nb_sb_noss_1?_mk_zh_CN=%E4%BA%9A%E9%A9%AC%E9%80%8A%E7%BD%91%E7%AB%99&amp;url=search-alias%3Daps&amp;field-keywords=%E5%90%B8%E5%B0%98%E5%99%A8&amp;x=0&amp;y=0</a> ; 3 分，购物网站上的检索信息，方便丰富	商品类，各优质 b2c 网站上的分类页面、检索结果页面只要确实相关，且比较丰富，一般较好
尼桑骐达新车是否适合上高速	信息需求，有疑难求解需求	<a href="http://www.52che.com/z/guanzhu_364911.html">http://www.52che.com/z/guanzhu_364911.html</a> ; 3 分，网站骐达类目下的检索信息	注意页面上虽然没有很多骐达的字样，但是本身是在骐达论坛下的检索结果，也是符合用户需求的，搜索结果页要特别注意实际结果的满

百丽雅文	百丽雅文女鞋	<a href="http://www.smarter.com.cn/_la_njmapgnbmfmomemfkonakm-se-ch-1018-c-11/">http://www.smarter.com.cn/_la_njmapgnbmfmomemfkonakm-se-ch-1018-c-11/</a> ; 0 分, title 虽然是百丽雅文, 但是实际均是百丽的鞋子, 和用户需求不符合	足好坏 检索结果页中, 要特别注意有一些结果, 是看上去很相关, title 和 query 很匹配的, 但是实际上无真正意义上的检索结果, 给出一堆其他的信息的, 需要给 0 分
现在白银的价格	具有高时效性要求的某商品价格需求	<a href="http://www.baidu.com/s?tn=baidurt&amp;rtt=1&amp;wd=%CF%D6%D4%DA%B0%D7%D2%F8%B5%C4%BC%DB%B8%F1">http://www.baidu.com/s?tn=baidurt&amp;rtt=1&amp;wd=%CF%D6%D4%DA%B0%D7%D2%F8%B5%C4%BC%DB%B8%F1</a> ; 1 分, 检索页面中多个白银相关结果, 但也有黄金查询结果, 本身并非所有搜索结果都是相关的, 且实际多个结果并不能满足用户需求, 内容实际没价格	搜索结果页, 要看相关结果条数、在整体搜索结果中所占的比例, 也要考虑实际结果的满足情况如何; 例如是不是全都作弊的、看似相关实际不满足的

### 4.3 权限页面

有些网页具有权限限制, 即需要注册、登陆该网站后, 才能下载或者浏览有用信息。在评估有权限限制的网页时, 要根据权限限制强度有多大, 进行适当的降档打分。

权限描述	权限限制程度	帮助程度
可以在线浏览全部内容, 仅对下载有权限。	弱	如: 百度文库、豆丁网。对用户有较好的帮助, 因为浏览即可以满足部分用户的需求。
不可在线浏览, 下载需要登录, 是知名网站。	较强	如新浪爱问、CSDN 资料下载。因为是知名网站, 所以应该有更多的用户已经有账户。而且注册知名网站的风险较小。所以这类结果对用户有些帮助。
不可在线浏览, 下载需要登录, 不是知名网站。	强	如书香门第网。因为不是知名网站, 所以用户已有账号的可能性不大。注册账户的话, 有收到垃圾邮件的风险。所以这类结果对用户帮助较小。
不可在线浏览, 下载需要大量积分或付费。	很强	如中国知网。对这类网站, 即使用户注册, 也需要付费或者花费大量时间累积积分才能获得所需额的信息, 对用户的帮助较小。

下面举例说明:

Query	需求	URL 与打分	说明
八年级上学期历史重点总结	信息需求。用户想浏览或者下载这一总结资料。	<a href="http://lishi.ht88.com/downinfo/366553.html">http://lishi.ht88.com/downinfo/366553.html</a> ; 1 分, 不知名网站, 强权限,	尽管网站有权限限制, 但内容是相关的, 可满足部分愿意注册的用户的需求, 因此

		可预览少部分内容，但不足以满足用户需求。	不可以打 0 分。
通讯录管理系统系统设计	浏览或下载这些资料的需求	<a href="http://www.cnki.com.cn/Article/CJFDTotal-KJXX201011377.htm">http://www.cnki.com.cn/Article/CJFDTotal-KJXX201011377.htm</a> ; 1 分，虽然是大站，但是需要付费，也算强权限	知网等虽然是知名大站，但是资料大部分都需要付费，算是强权限
新视野大学英语三读写教程答案	信息需求。用户想浏览或下载这部分教程的答案。	<a href="http://wenku.baidu.com/view/6dffb23482fb4daa58d4ba0.html">http://wenku.baidu.com/view/6dffb23482fb4daa58d4ba0.html</a> ; 2 分	百度文库结果可以浏览、打印，积分要求比较低，属弱权限，对用户有帮助。

#### 4.4 垃圾、作弊等低质页面

有些网页在只看搜索结果页中的展现时，或只看 title 时，往往会被认为相关性比较好，但实际页面却往往因为质量低劣而价值很低，或者完全是广告垃圾作弊等页面。这类结果往往视其低质程度，给 1 或者 0 分。下面举一些典型例子说明：

类型	Query 与 url 页面	打分与说明
有问无答页面	<a href="http://wenwen.soso.com/z/q317634928.htm?sp=1510">http://wenwen.soso.com/z/q317634928.htm?sp=1510</a>	因无答案导致的用户需求难以满足，给 0 分
	<a href="http://wenwen.soso.com/z/q287727880.htm">http://wenwen.soso.com/z/q287727880.htm</a>	因无答案导致的用户需求难以满足，给 0 分
	<a href="http://www.16k.net.cn/book/88545/8087828.shtml">http://www.16k.net.cn/book/88545/8087828.shtml</a>	有 Title，有正常页面的所属成分，只是完全没有该小说章节的正文，0 分
链接、图片失效，实际无法浏览、下载	【麻生舞图片】 <a href="http://hi.baidu.com/zhouyh09/blog/item/6f8d55a7636c0a90d043586a.html">http://hi.baidu.com/zhouyh09/blog/item/6f8d55a7636c0a90d043586a.html</a>	实际页面上图片已无法观看，对于图片需求来说，该页面完全不能满足，0 分；
	【月光传奇补丁】 <a href="http://download.pcpop.com/Down/239080.html#">http://download.pcpop.com/Down/239080.html#</a>	实际下载链接无法点击，因此对于下载需求来说完全不能满足，0 分；
	【月光传奇补丁】 <a href="http://down.it168.com/228/231/34314/comment.shtml">http://down.it168.com/228/231/34314/comment.shtml</a>	实际下载链接虽然可以点击，但是全部是死链的，同样给 0 分
虚假视频	【小沈阳】 <a href="http://www.tudou.com/programs/view/UEMWLitg-oE/">http://www.tudou.com/programs/view/UEMWLitg-oE/</a>	Title 是“小沈阳彪高音”，实际是猫咪吃东西，完全无法满足用户需求，0 分；
	【金星舞蹈】 <a href="http://v.ku6.com/show/-V6ssKcVveWdycTNfO8dZA...html">http://v.ku6.com/show/-V6ssKcVveWdycTNfO8dZA...html</a>	Title 是“金星老师舞蹈视频”，实际是路边随便乱拍的，无法满足任何需求，0 分；



虚假站点	<b>【qq2012】</b> <a href="http://www.up620.com/">http://www.up620.com/</a>	Title 与首屏看上去不错，但首屏后大量不相关链接，点击首屏表面好的链接，一直是广告，始终无法满足用户下载需求，给 0 分；
低质搜索结果页	<b>【虐待凌辱女人】</b> <a href="http://www.41717.com/game/nuedailingrunvren_838368/">http://www.41717.com/game/nuedailingrunvren_838368/</a>	搜索结果全都不相关，对于 <b>【虐待凌辱女人】</b> 来说，0 分
	<b>【小学六年级奥数补习班】</b> <a href="http://cn.coovee.net/3888742/shop.html">http://cn.coovee.net/3888742/shop.html</a>	搜索结果是篮球架、CAD，完全不相关，0 分
文不对题	<b>【北京双眼皮医院】</b> <a href="http://zhidao.baidu.com/question/218643062.html">http://zhidao.baidu.com/question/218643062.html</a>	整个页面都没有关于北京医院的推荐；完全没有价值
	<b>【怎么注销 qq 空间网址】</b> <a href="http://www.91gupiao.com/info/20110523/345743.html">http://www.91gupiao.com/info/20110523/345743.html</a>	Title 看起来相关，实际完全不相关，给 0 分
页面不清晰，阅读困难	<b>【掌事】</b> <a href="http://www.yueduge.com/html/8/8484/2054418.html">http://www.yueduge.com/html/8/8484/2054418.html</a>	小说阅读页面中，有些是图片格式，其中模糊不清、阅读成本很大的，为低质结果； 但有些虽然是图片，但是非常清晰，完全不影响阅读的，不算低质，如： <a href="http://www.uczw.com/html/10/10531/6128806.html">http://www.uczw.com/html/10/10531/6128806.html</a>
低质采集，无法阅读	<b>【达内培训】</b> <a href="http://www.17u.net/wd/showguide/1455126">http://www.17u.net/wd/showguide/1455126</a>	内容是从其他网页上直接采集、复制过来的，由于没有排版，非常难以阅读，对于用户来说基本没有帮助，可以直接给 0； 采集的页面要看采集后是否阅读方便，能否满足用户需求，好的采集结果也可以根据需求满足打分。
广告泛滥	<a href="http://www.baiwandu.com/Html/Book/39/39610/7783653.html">http://www.baiwandu.com/Html/Book/39/39610/7783653.html</a> <a href="http://www.wzyy.org/guochanju/AAzhishenghuo/player.html?14210-0-2">http://www.wzyy.org/guochanju/AAzhishenghuo/player.html?14210-0-2</a> <a href="http://www.ranwen.com/html/44/44218/7293048.html">http://www.ranwen.com/html/44/44218/7293048.html</a>	广告有普通、悬浮、弹窗等多种，少量不影响阅读的广告不认为很差，但如广告太多影响浏览阅读，或者频繁的弹窗广告，则认为是低质结果，即使能够满足，也只能最高 1 分；
内容堆砌、穿插	<a href="http://www.94mfk.cn/viewnews-14836.html">http://www.94mfk.cn/viewnews-14836.html</a>	页面完全无排版，且语句完全无法连起来阅读，往往无法满足任何需求，可以直接给 0
	<a href="http://se.huangse228.com/dde/3574.html">http://se.huangse228.com/dde/3574.html</a>	网页上有很多“90 后 mm 全裸”的关键词，但是同样无法满足任何需求，直接给 0
	<a href="http://www.xici.net/d159503568.htm">http://www.xici.net/d159503568.htm</a>	网页上有很多《av 电影》的关键词，将原来的正常文章严重打散，直接给 0

乱码页面	<a href="http://launch.groups.yahoo.com/group/dove/r/messages/24278?xm=1&amp;o=1&amp;m=p&amp;tidx=1">http://launch.groups.yahoo.com/group/dove/r/messages/24278?xm=1&amp;o=1&amp;m=p&amp;tidx=1</a>	页面基本乱码，无法满足任何需求，直接给 0
------	---	-----------------------

## Part6 Query-Url 质量评估快速导读

本部分是对 Query-Url 评估标准的简要总结，主要介绍评估思路与打分的衡量方法，也希望能帮助大家快速了解本文档。

### 1、评估必备软件

一般评估常用的必备软件，包括以下一些：

- Firefox 浏览器、ie 浏览器
- 杀毒软件
- adobe flash player、windows media、qvod、百度影音等播放器
- 迅雷、QQ 旋风、快车（flashget）等下载软件

### 2、评估必备基础知识

#### ✓ Query

也可称查询词、关键词、搜索词，指用户输入到搜索框想要查找的字、词、符号等。以后作为 Query 出现的词，都会用“【”、“】”括起来。

#### ✓ Term

即组成一个 query 的字、词，如 query【新浪微博】，拆开后的“新浪”、“微博”，我们称为 Term。

#### ✓ 搜索结果

指搜索引擎（如百度、google 等）根据 Query 返回的搜索结果，可以指具体某一条，也可以指所有的结果。

#### ✓ URL

也称为网页地址，简称网址。也就是你将要评估的页面。

#### ✓ 用户需求/用户需求意图

简称“需求”，用户键入搜索词是希望获得相应的满足，比如获得相应的信息、网上购物等，我们把用户 query 背后的这种诉求称为用户需求。

### 3、query-url 质量评估概述

#### 3.1 什么是 query-url 质量评估

Query-URL 打分是最重要的一种打分方式，是评价各搜索引擎好坏的基本方式之一。这种评估是根据 Query 背后的需求，去评价某条结果（即 URL 页面）对用户需求的满足程度。

### 3.2 Query-Url 质量评估考虑的维度

- ✓ **相关性**：是指结果主题和用户需求的主题匹配程度，两者说的是不是一回事，有没有跑题。
- ✓ **需求强度**：搜同一个 Query，不同用户的需求可能不同。此时，要站在多数用户的角度来判断，越多人查找的需求越重要。主要需求是最多人找的，满足主要需求的结果，是最好的结果；满足冷门需求，则结果差些。
- ✓ **丰富程度**：页面内容越详细越好。
- ✓ **有效性**：能否真正满足，如是否可以真实下载、在线试听等。
- ✓ **时效性**：有些查询，对时间有着较高的要求，期望得到最新的信息。常见的是突发新闻的搜索，哪怕没有明确指定什么新闻，给出最新的内容也是很好的。
- ✓ **便捷性**：找到有用信息要花多少时间和成本。
- ✓ **权威度**：通常出自官方网站、行业知名网站的信息比一般网站要好。

**总结**：不同维度适用 query 范围

这 7 个维度在不同特征的 query 下，关注的重点有不一样，不同维度适用的 query 范围说明如下：

维度	适用哪些 query	举例说明
相关性	所有	不相关，如丢掉了核心词肯定是 0； 相关，也有可能是 0，如软件虚假，视频不能播放。
需求比例	需求有多个	需求分析时，重要的是分析主要需求是什么； 非主要需求，最高只能 2 分。
丰富程度	需要详细的内容	搜作文范文、报告总结、人物资料等，丰富度越高越好，如【实习总结】、 【杨幂资料】 如一句话、数字等简短的内容即可满足，则丰富程度要求不高，如【刘备字什么】、【横看成岭侧成峰的下一句】
有效性	大多为交互操作类 query，如下载、视频、试听、文档阅读等，只有资源有效才能满足用户需求的	下载，一定要尝试能否下载，是否真实 如视频，要看内容是否查看，是否真实 如物流查询，要真能查，才有价值。
时效性	明确或隐含着时间要求的查询	近期事件（如地震）、招聘、交易信息等，对时效性要求高； 高考等周期性事件也有时效性要求； 中国人口等隐含最新数据的时效性要求；
便捷性	所有	用户要找的位于显著位置，第一眼就能看到为便捷性好；当比较难以在页面方便的找到时，即为获取信息满足成本高，会稍差；
权威度	软件下载、商品交易、医疗健康等对权威有要求	来自知名网站、专业网站，让人放心。软件下载、医疗健康、银行等，对权威方面的要求高；

### 3.3 Query-Url 打分分档介绍

打分类别	含义	描述
<b>4 (Vital)</b>	官方	官网权威页面
<b>3 (Useful)</b>	很好，对大多数人都非常有用	该 url 页面能满足大多数用户需求，很有价值
<b>2 (Relevant)</b>	比较有价值，对部分人较有帮助	该 url 页面对较多或一些用户较有价值

<b>1 (Slightly Relevant)</b>	稍沾边，略微有帮助	该 url 页面对大多数用户价值不高，但是跟 query 相关，少数用户觉得该页面有点帮助
<b>0 (Off-topic OR Useless)</b>	不相关或无用资源	该 url 页面毫无价值或价值很低
<b>死链</b>	——	死链导致无法被评估的页面

### 3.4 Query-Url 评估步骤

整个评估过程可划分为三个步骤：

- ✓ 需求判断：根据 query 判断用户需求，确定用户在找什么
- ✓ 页面质量判断：查看实际 url 页面，判断结果能否以及多大程度上满足用户需求；
- ✓ 打分：根据本评估指南，选择合适的分值。

## 4、理解用户需求

打分的第一步是要先理解 query 含义，弄明白用户搜这个 query 要找什么。

**要了解 query 背后的真实需求意图：**除了理解 query 本身的含义之外，您还应该考虑用户输入这个 query 时，他实际的需求是什么，并以此来给所对应的 url 页面打分。如【俄罗斯方块】，用户需求很可能是在线玩这个游戏，也有可能是要下载它。

### 4.1 Query 需求的主次之分：

搜索同一个 Query，不同人可能有着不同的需求，站在大多数用户的角度来衡量，把握需求主次之分，主次需求会影响打分的高低。

根据 query 不同表意下的需求强度，以及同样表意下不同需求维度的需求强度，需求强弱可能有以下几种情况：

**主需求：**即输入这个 query 的用户中，大多数用户心目中的需求。如【windows】，当大多数用户输入这个 query 时，需求是微软的操作系统，而不是在墙上的玻璃窗户。

**重要需求/次要需求：**即输入这个 query 的用户中，有许多或部分用户存在的需求。如【apple】，水果的需求；【亚马逊】，南美洲亚马逊河流的需求。

**小众/冷门需求：**在某些 query 下，有的表意和需求维度比较冷僻，用户需求的比例非常低。

**泛需求：**在某些 query 下，几乎没有占绝对主导地位的表意或同义表意需求维度上绝对强弱之分，都有一定比例的用户需求，这样情况下根据需求强弱的差别打分，如【长沙】、【书籍】。

### 4.2 query 需求意图类型：

根据用户输入 query 时的需求意图不同，我们将用户需求大致分为 3 类：寻址需求、信息需求、交互操作需求。

**寻址需求：**寻址需求就是指，用户输入这个 query，就是为了去这个特定的页面或网站。如【京东商城】、【北京大学】。

**信息需求：**用户的需求是获取信息，想了解某信息、某知识，搜索引擎的目标是让用户找到所求信息，给出相关的优质、权威、全面的信息。如：【怎样开车可以省油】、【西夏王陵】

**交互操作需求：**用户是要从事某项操作和活动，往往是需要某个 url 上，与页面进行一定的交互，从而达到用户的目标、完成相应的操作。比如想下载某款软件、想在网上购买机票，如：【迅雷下载】、【英文翻译】；这

类需求很看重结果页面是否进行操作、达成目标。

**Query 需求辅助判断：**当遇到一个您不熟悉的 query，或 query 的需求很难确定时，您可以参考搜索引擎结果(百度、谷歌)、参考搜索提示或相关搜索、借助百科词条理解。

**需要注意的是：**当采用通过搜索引擎来确认用户需求的方法时，千万不要直接将排在前面的结果当做好结果，将排在后面的结果当做差结果。这应当只作为您确认用户需求的方式之一，而非直接根据现有排序情况进行打分。

## 5、五档打分标准

### 5.1 4分——官方

#### ➤ 4分结果的基本要求

Query 表意清晰，主需求明确；Url 为官方或者首发地址；二者缺一不可。

#### ➤ 4分的典型情况

- ✓ 寻址 query 下，给出官网或指定站内页面；
- ✓ 实体 query 下，给出对应实体的官网；
- ✓ 人名 query 下，著名人物的官方微博 4 分；
- ✓ Url 类查询下，仅 url 查询本身对应 url 页面可以给 4 分；
- ✓ 文章、小说名 query 下，原创首发可以给 4 分（报刊发表的，该报网站上该文 4 分）

#### ➤ 典型不能有 4 分结果的情况

##### ✓ Query 角度

没有主需求，需求比较分散：如【DDD】，指代多种实体缩写

虽有主需求，但非寻址需求，而是信息需求：如【iphone 评论】、【北大建校多少年】、【王超英勇事迹】

##### ✓ Url 角度

没有对应的官方网站或网页存在：如【谈南路小学】，这个小学目前没官方网站

没有任何人或实体可以对应并完全覆盖这个 query，如【蘑菇】、【心脏病】

#### ➤ 典型可有多个 4 分结果的情况

- ✓ 多个 url 对应的实际页面完全相同，例如有无 www 前缀，index 后缀等区别；
- ✓ 多个 url 具有相互跳转关系，例如 urlA 会自动跳转到 urlB；
- ✓ 多个 url 都可以快捷满足用户需求，例如：下载需求：软件官网首页与下载内页

### 5.2 3分——很好，对大多数人都非常有用

#### ➤ 3分定义

3 分结果，认为应当是在该 query 下，对于大多数用户都非常有用的好结果。非常有用的 3 分结果一般应该具备的特征为：符合 query 需求、页面质量很高。

注意：一个 query，不止一个 url 页面可以打 3 分。

#### ➤ 3分结果特征

这些页面往往具有以下全部或者部分的典型特征：需求满足程度高、权威性、页面资源真实可用（如可以播放、收听）、页面看起来整洁组织排版很好、页面有用信息容易找到，在时效性需求的 query 下（如突然爆发的新闻事件），时效性也很好。

### 5.3 2分——比较有价值，对部分人较有帮助

#### ➤ 2分定义

2分，意味着该页面对于搜索这个 query 的部分用户是有帮助的。

#### ➤ 2分结果特征

可以认为，2分结果对于用户来说，在需求满足方面要稍微差于3分结果，比如丰富全面程度、时效性程度、权威程度（来自不太知名站点）等方面略差；或者只能涵盖 query 用户需求的一部分，如常见的缺乏明确主需求的 query，往往最高只能到2分。

2分的网页虽然比3分差，但对于用户来说也必须是有价值帮助的，即这个结果对于用户来说，与 query 主题相关，页面质量不垃圾

### 5.4 1分——稍沾边，略微有帮助

#### ➤ 1分定义

1分结果可以理解为，页面跟 query 是相关的，但是对于大多数用户来说这个页面都是没有用的，可能只对少量用户略有价值。“表面看起来相关的差结果”，也会是1分。

#### ➤ 1分结果特征：

- ✓ 页面相关，但页面质量较低劣，如页面是打广告目的、采集排版较差等；
- ✓ 页面相关，但不适合 PC 端查看浏览，如 wap 页面
- ✓ 页面相关，但包含的有用信息很少；
- ✓ 页面信息因过时、过于具体、过于笼统等而难以达到2分的要求；
- ✓ 页面信息因满足的需求过于冷僻，而难以达到2分；

### 5.5 0分——不相关或无用资源

#### ➤ 0分定义

- ✓ 0分代表对于用户来说，这个结果对于搜该 query 的所有用户完全没有帮助的

#### ➤ 0分特征

- ✓ 不相关，即搜 A 给 B，完全不能满足用户需求
- ✓ 页面无价值，页面即使看起来相关，但实质上页面完全无法提供有价值信息，是一些非常低质页面。如只有链接和广告、虚假视频、虚假站点、有问无答、文不对题、无用内容堆砌等，这些页面较多是作弊的，对用户没有价值。

### 5.6 死链

如页面死链而无法给出打分时，请选择“死链”，死链主要包括以下几种：

- ✓ 页面抓取不成功，且点击 url 打开也无法加载，无实际内容的空白页；
- ✓ 页面报错提示页面：除了报错信息外，页面上没有其他内容；
- ✓ 标准报错页面：网站直接返回的标准出错页面，如 404、500 等标准报错代码
- ✓ 网页恶意软件提示，如“请注意，访问此软件可能会损害您的计算机”；



## 6、特定类型打分说明

### ➤ 输入错误的 query:

在评估中，我们会遇到一些用户输入错误的 query。对于明显输入错误的 query，可以根据用户正确需求意图来进行打分，而不是对错误的 query 进行理解打分；对于非明显输入错误的 query，可以根据用户实际输入的 query 本身来判断用户的需求是什么，从而进行打分。

### ➤ url 查询 query

即用户想去对应的网站，规范的 url 查询，如：[www.youku.com](http://www.youku.com)、360buy.com。不规范的 url 查询。如：[www.mcriosoft.com.query](http://www.mcriosoft.com.query) 拼错了；[www.360buy.com](http://www.360buy.com) 多输入了一个‘w’。只要确定是找某个 url 页面，即使输入错误的 url，给出对应网站或网页也可以给 4 分。

### ➤ 列表页需求 query:

有些 query，用户期望得到很多信息做参考，如：鸡肉食谱、10-15 万的 suv。Query 所需是一个列表时，一个页面上有多个结果往往比单条结果好，一般为页面上多个优质结果>单条优质结果≈普通少量结果>普通单个结果。

### ➤ 时效性需求的 query:

有些 query，对时效性要求很高，如近期或突发的事件，用户希望看到最新进展。如输入的 query 是找体育赛事、高考等，一般来说都是想得到最近的相关信息，而不是几年前的信息，除非 query 明确表明要找过去某个时间的信息。

### ➤ 地域性需求 query

定义：有些 query 具有地域性需求，即当地的结果对于本地用户更能满足需求。如餐馆、电影院等信息，当地信息更好，如【58】，给当地分站较好。

打分：

- ✓ 如不知用户所在地，仅官网可给 4 分，各地分站结果给 3 分；
- ✓ 如已知用户所在地，一般都认为是北京，即官网与北京分站结果都可以给 4；外地各分站结果给 1 分

### ➤ 纯英文 query

纯英文 query 时，给出英文页面不一定就降分，仍然需要根据用户的主次需求来判断页面价值，从而打分。

### ➤ 百科结果页面

百科结果虽然往往质量比较好，但不能直接给予百科结果较高的打分；需要首先判断用户是否具有较强的介绍需求，如百科本身不能满足主需求，最高 2 分；

### ➤ 搜索结果页

打分：如果搜索引擎页实际为空页面，或者实际没有任何搜索结果，那么应该是 0 分。如果搜索引擎页具有一组搜索结果，那么其评分应在 0-3 分之间，具体根据搜索结果页实际有多少内容对用户需求满足打分。

注意：对于百度视频、google 视频、图片等搜索结果页，不能不查看实际需求满足就给出高分。

### ➤ 权限结果页面

有些网页具有权限限制，即需要注册、登陆后，才能下载或者浏览有用信息。打分一般要根据权限限制强度有多大，进行适当的降档打分。

注意：

- ✓ 可浏览仅下载有权限的，认为权限弱，浏览下载均具有权限的，认为权限强
- ✓ 大站点需登录权限，认为权限弱，不知名小站点需登录权限，认为权限强
- ✓ 仅需注册登录权限，认为权限弱，需大量积分或付费的，认为权限强；

感谢您的认真阅读，如在打分过程中或任务领取上有疑问请登录[评测论坛](http://wse1.baidu.com/wsebbs/)进行交流讨论。

评测论坛地址：<http://wse1.baidu.com/wsebbs/>