情景五

跨境网络站外推广分析

子情景一　搜索引擎付费推广

知识导读

搜索引擎营销作为一种新的营销方式，通过用户使用搜索引擎检索信息的机会，尽可能地将营销信息传递给目标用户，是基于搜索引擎平台而开展的网络营销。搜索引擎营销的基础是企业网络营销信息，是一种企业网站推广的常用方法，也是一种由用户主导的网络营销方式，信任度较高。搜索引擎营销的服务方式主要包括三种：搜索引擎付费推广、搜索引擎优化、网盟推广。本章将介绍搜索引擎的基本概念、类型和工作原理，了解搜索引擎付费推广的特点、相关概念，并结合 Google Ads 了解搜索引擎付费推广的基本步骤。

学习目标

知识目标

了解搜索引擎的类型。

理解搜索引擎工作的基本原理。

理解 Google Ads 的相关基本概念。

能力目标

掌握 Google Ads 账户搭建方法。

掌握搜索引擎付费推广的基本流程。

能够制订搜索引擎付费推广计划。

素质目标

培养利用 Google Ads 进行付费推广及动态优化的系统化思维。

相关知识

1. 搜索引擎

1.1 搜索引擎概述

1.1.1 搜索引擎的概念和类型

搜索引擎(Search Engine)是指根据一定的策略,运用特定的计算机程序,自动从互联网上搜集信息,并对这些信息进行筛选、组织、整理和排序后,再按照用户搜索信息的关键词将相关信息展示给用户的系统。

随着搜索引擎技术的不断发展,搜索引擎越来越多,主要包括全文搜索引擎、目录索引、元搜索引擎、垂直搜索引擎和集合式搜索引擎等类型。

全文搜索引擎(Full Text Search Engine)是目前广泛应用的搜索引擎,是指计算机索引程序通过扫描文章中的每一个字词,对每一个字词建立一个索引,指明该字词在文章中出现的次数和位置,建立数据库。当用户查询时,检索程序根据事先建立的索引进行查找,检索与用户查询条件相匹配的记录,并将结果按一定的顺序反馈给用户。根据搜索结果来源不同,全文搜索引擎分为两类。一类拥有自己的蜘蛛程序,能够建立网页自己的数据库,搜索结果直接从其数据库中调用,如 Google、百度、360 搜索;另一类则租用其他搜索引擎的数据库,然后按照自己的规则和格式排列和显示搜索结果,如 Lycos 搜索引擎。

目录索引(Search Index/Directory)也称分类检索,是互联网上最早提供网站资源查询的服务,主要通过人工或半自动的方式搜集和整理互联网上的资源,由编辑人员查看信息后,根据搜索到的网页内容进行信息摘要,并将其分配到事先确定的分类框架中。严格意义上来说,目录搜索引擎不能称为真正的搜索引擎,只是按照目录分类的网站链接列表。用户在目录索引中查找网站时,可以使用关键词,也可以按照分类目录逐级查询。因为目录搜索引擎是依靠人工或者半自动方式搜集信息,所以信息准确、导航质量高,但是需要人工介入,维护量大、信息量少、信息更新不及时。这些信息主要面向网站,提供目录浏览服务和直接检索服务,最具代表性的目录搜索引擎为 Yahoo 和新浪分类目录搜索,搜狐目录、hao123、Dmoz 等都是目录索引。

元搜索引擎(META Search Engine)是通过一个统一的用户界面帮助用户在多个搜索引擎中选择和利用合适的搜索引擎来实现检索操作,即在接受用户查询请求后,会同时在多个搜索引擎上进行搜索,并将结果返回给用户,是对于多种检索工具的全局控制机制。在搜索结果排列方面,有的直接按来源排列,如 Dogpile;有的则按自定的规则将结果重新排列组合,如 Vivisimo。一个完整的元搜索引擎由三部分组成,即检索请求提交机制、检索接口代理机制和检索结果显示机制。检索请求提交机制负责实现用户的检索设置要求,包括调用哪些搜索引擎、检索时间限制、结果数量限制等;检索接口代理机制负责将用户的检索请求转换成满足不同搜索引擎要求的格式;检索结果显示机制负责所有元搜索引擎检索结果的去重、合并、输出处理等。

垂直搜索引擎(Vertical Search Engine)更专注于特定的搜索领域和搜索需求,如图片搜索、视频搜索、法律搜索、专利搜索、论文搜索等,是对通用搜索内容的细分,在特定的搜索领域创造更好的用户体验,如百度文库、Google 学术等。

集合式搜索引擎与元搜索引擎类似,区别在于它并非自动同时调用多个搜索引擎进行搜索,而是由用户从提供的若干搜索引擎中进行选择,然后在其中搜索用户需要的内容。HotBot、Howsou、Duoci 等都是集合式搜索引擎。

1.1.2 国内外主要搜索引擎

目前,国内搜索引擎主要有百度、360 搜索、搜狗搜索等,国外的搜索引擎主要有 Google、Yahoo 和 Bing 等,见图 5-1 和图 5-2。

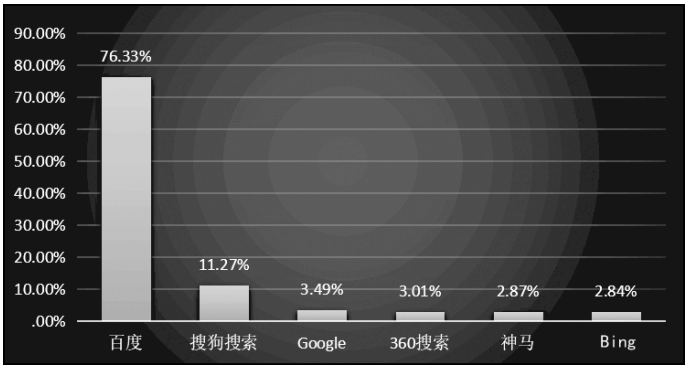

图 5-1 2020 年 10 月国内搜索引擎市场

图片来源:站长之家

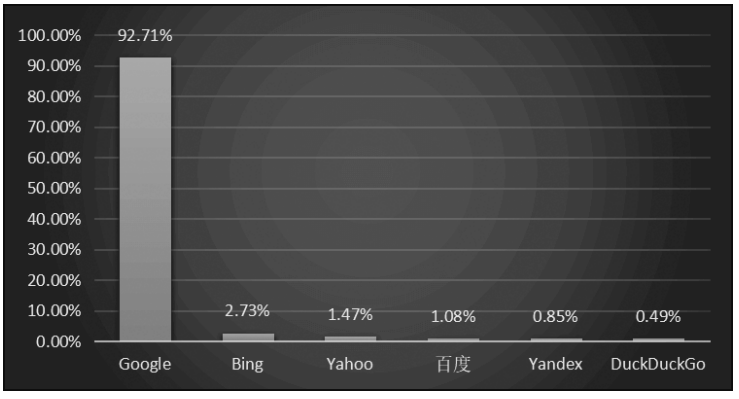

图 5-2 2020 年 10 月全球搜索引擎市场份额

图片来源:站长之家

1) 百度

百度在全球搜索引擎市场份额排名第四,是排名前五的搜索引擎中唯一的中文搜索引擎,也是国内排名第一的搜索引擎,市场份额达到 76.33%。百度拥有全球最大的中文网页库,收录的中文网页已达数百亿,且每天仍在飞速增长。百度每天处理来自 138 个国

家和地区超过数亿次的搜索请求，用户通过百度可以搜到世界上新鲜、全面的中文信息。百度目前提供网页搜索、音乐搜索、图片搜索、百度新闻、百度贴吧、百度知道、百度百科等主要产品和服务，提供地图搜索、文档搜索、视频搜索、百度翻译、百度识图等多项满足用户细分需求的搜索服务，还在个人领域提供了百度影音、百度云、百度输入法等服务。

2）搜狗

搜狗搜索是全球首个第三代互动式中文搜索引擎，也是全球首个百亿规模中文搜索引擎，致力于中文互联网信息的深度挖掘。搜狗旗下的产品有搜狗百科、搜狗问问、音乐搜索、图片搜索、新闻搜索、地图搜索、网址导航、视频搜索、知识搜索、搜狗实验室、搜狗百宝箱等。搜狗搜索的音乐搜索死链率小于2%，图片搜索有独特的组图浏览功能，新闻搜索有“看热闹”首页及时反映互联网热点事件，地图搜索有全国无缝漫游功能，极大地满足了用户的日常需求。

3）360搜索

360搜索属于元搜索引擎，同时将信息聚合在一起实现网络工具化、个性化的发展需求，提升网络使用效率，目前在国内搜索引擎市场份额中排名第四。360搜索包含网页、新闻、影视等搜索产品，为用户带来安全、真实的搜索服务体验。360搜索不仅掌握通用搜索技术，而且独创PeopleRank算法、拇指计划等创新技术。

4）Google

Google是目前全球最大的搜索引擎，以简单、干净的页面设计和相关性最高的搜寻结果赢得了全球范围内无数用户的认可，在全球搜索引擎的市场份额排名第一，高达92.71%。Google提供常规搜索和高级搜索两种功能，提供多达30余种搜索语言选择，每天需要处理约2亿次搜索请求。Google目前的主要业务有谷歌搜索、谷歌广告、谷歌地图、谷歌火星、谷歌月球、YouTube、Android、Chrome、Google Play、Gmail等。

5）Microsoft Bing

Microsoft Bing（微软必应）是微软公司于2009年推出的搜索引擎服务，集成了搜索首页图片设计、崭新的搜索结果导航模式、创新的分类搜索和相关搜索用户体验模式、视频搜索结果可直接预览播放、图片搜索无须翻页等功能。作为最贴近中国用户的全球搜索引擎，微软必应致力于满足中国用户对全球搜索的需求，必应图片搜索率先实现了中文输入全球搜图。

6）Yahoo

Yahoo自1994年推出全球第一个搜索引擎，是目前全球排名第三的搜索引擎。Yahoo提供目录、网站及全文检索功能，为超过5亿的独立用户提供包括搜索引擎、电邮、新闻等在内的多元化网络服务。Yahoo是较早的分类目录搜索数据库，所收录的网站全部被人工编辑，并按照类目分类。

1.2 搜索引擎的工作原理

搜索引擎的工作大致分为四个阶段：蜘蛛爬行、抓取建库、页面处理、检索及呈现。

1.2.1 蜘蛛爬行

如何有效获取并利用互联网上的海量信息，是搜索引擎要解决的首要问题。爬行与

抓取是搜索引擎工作的第一步，包括在互联网上发现、搜集网页信息，同时对信息进行提取和建立索引库。

1）搜索引擎蜘蛛

搜索引擎蜘蛛(Spider)，又称为网络爬虫(Web Crawler)、网络蜘蛛或网络机器人，是一种按照一定的规则，自动抓取互联网信息的程序或者脚本。每个搜索引擎蜘蛛都有各自的名称，如 Googlebot、Bingbot、BaiduSpider 等。搜索引擎蜘蛛的作用就是在互联网中浏览信息，把这些信息都抓取到搜索引擎服务器上，然后建立索引库。

根据抓取目标和范围，可以将搜索引擎蜘蛛分为三类：①批量型蜘蛛，指有明确的抓取目标和范围，达到即停止；②垂直型蜘蛛，指针对某个特定领域的网络，根据主题过滤网页和信息；③增量型蜘蛛，应对网页不断更新的状态，及时反应。一般的商业搜索引擎蜘蛛都属于增量型。

2）爬行策略

为了抓取更多的页面，搜索引擎蜘蛛会跟踪页面上的链接，从一个页面爬行到下一个页面，就像蜘蛛在蜘蛛网上爬行一样。整个互联网是由相互连接的网站及页面组成，从理论上讲，搜索引擎蜘蛛从任何一个页面出发，都可以顺着链接爬行到互联网上的其他页面。由于网站及页面链接结构非常复杂，搜索引擎蜘蛛需要采取一定的策略才能爬行完互联网上所有的页面。

搜索引擎蜘蛛的爬行策略通常分为三种：深度优先策略、宽度优先策略、最佳优先策略。

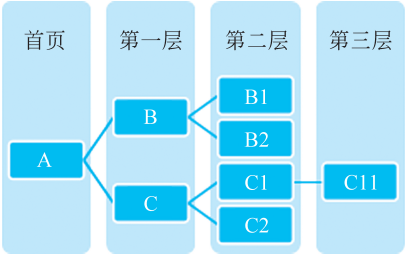

图 5-3　爬行策略示意图

(1) 深度优先。

深度优先策略是早期搜索引擎蜘蛛使用较多的一种策略，指的是在一个网页文件中，当一个链接被发现，蜘蛛沿着该链接一直向前爬行，直到再也没有其他链接，然后返回到第一个网页，再沿着另一个链接向前爬行。当不再有新的超链接可供选择时，爬行过程结束。如图 5-3 所示，搜索引擎蜘蛛爬行的顺序为 A-B-B1-A-B-B2-A-C-C1-C11-A-C-C2(访问次序并不唯一，第二个点既可以是 B，也可以是 C)。

(2) 宽度优先。

宽度优先策略是将某个页面中的链接依次加入待爬取队列，即蜘蛛在一个页面上发现多个链接时，不是顺着一个链接一直向前，而是把页面上所有的第一层页面都爬一遍，再沿着第二层页面上发现的链接爬向第三层。如图 5-3 所示，搜索引擎蜘蛛爬行的顺序为 A-B-C(第一层链接的页面)-B1-B2-C1-C2(第二层链接的页面)-C11(第三层链接的页面)。

理论上，无论是深度优先策略，还是宽度优先策略，只要给搜索引擎蜘蛛足够的时间，都能爬行完整个网络。但在实际工作中，蜘蛛的宽带资源和时间都是有限的，不可能爬完所有的页面。因此，通常需要混合使用两种策略，既照顾到更多的网站(宽度优先)，也能照顾到一部分网站的页面(深度优先)，同时还考虑了页面权重、网站规模和外部链接等

因素。

(3) 最佳优先。

最佳优先策略是指当搜索引擎蜘蛛到达一个网页时,将其中的所有链接收集到地址库中,并对其进行分析,从中筛选出重要性较高的链接进行爬取。影响链接重要性的因素主要有 PR(Page Rank,网页排名)值、网站规模、反应速度等。其中,PR 值是根据网站的外部链接和内部链接的数量和质量来衡量网站好坏的标准,PR 值越高说明该网页越受欢迎、越重要。因此,当一个链接的 PR 值越高、网站规模越大、反应速度越快时,越会被优先抓取。

1.2.2 抓取建库

经过一定时间后,搜索引擎蜘蛛可以爬取完互联网上的所有网页,但这些网站的资源庞大,其中可能会有大量垃圾网站、重复内容等,而搜索引擎的资源有限,通常只会抓取其中的部分网页到数据库中。这就需要在搜索引擎蜘蛛到达一个网页时,对其内容进行检测,判断信息是否为垃圾信息,例如,存在大量重复内容、乱码等。检测通过后,搜索引擎蜘蛛会对有价值的网页进行收录,将网页信息存储到数据库中。为了避免蜘蛛重复爬行和抓取网址,搜索引擎会建立一个地址库,记录下已经被发现但还没有被抓取的页面,以及已经被抓取的页面。搜索引擎蜘蛛按照重要性从待访问地址库中提取 URL,访问并抓取页面,然后把这个 URL 从待访问地址库中删除,放进已访问地址库中。

1.2.3 页面处理

由于数据量过于庞大,搜索引擎蜘蛛抓取到网页数据后,不能直接用于索引服务,还需要对页面进行预处理,即建立索引,为最后的查询排名做准备。预处理主要包括结构化、分词、去停止词、降噪、去重、建立索引、链接分析、数据整合等。

1) 结构化

搜索引擎蜘蛛抓取到的网页数据中,除了用户在浏览器上可以看到的可见文字外,还包含 HTML(Hyper Text Markup Language,超文本标记语言)格式标签、JavaScript 程序、导航、友情链接等无法用于排名的内容。结构化就是从 HTML 文件中去除这些内容,提取出可以用于排名的正文文本、Meta 标签内容、锚文本、图片视频注释等内容。

2) 分词

分词是中文搜索引擎特有的工作内容,因为中文的词与词之间没有分隔符,搜索引擎必须先将一句话分解成若干词语。分词的方法主要有基于词典匹配的分词法、基于统计的分词法,以及利用人工智能并结合汉语语法和心理学知识进行的基于词义的分词法。

3) 去停止词

停止词是指在页面内容中出现的频率很高,但是对内容没有实际意义的词,如中文的“的、地、得、啊、哈、呀”等,英文的“a、an、the、to、of”等。因为它们对页面的主要意思没有什么影响,搜索引擎会去掉这些词,使索引数据的主题更突出,减少无谓的计算。

4) 降噪

噪声指页面上对页面主题没有贡献的内容,如版权声明、导航条、广告等。搜索引擎

需要识别并消除噪声。降噪的基本方法是根据 HTML 标签对页面分块，区分出页头、导航、正文、页脚、广告等区域，剔除大量重复出现的无关区块。

5）去重

由于网站之间相互转载，以及使用网页模板，互联网中存在大量的重复内容，搜索引擎需要在进行索引前对重复内容进行识别和删除，这个过程就称为去重。去重的基本方法是从页面主题内容中选取出现频率最高的一部分关键词，作为特征关键词，如果页面的特征关键词相同，会被判定为内容重复，不予收录。

由此可知，通过在文章中简单地添加“的、地、得”，以及调整语句和段落顺序等产生的伪原创内容，不能逃过搜索引擎的去重算法，因为这些操作并不能改变文章内容中的特征关键词，不能使转载和抄袭变成原创内容。

6）建立索引

搜索引擎会记录每个关键词在页面上出现的频率、次数、格式、位置等信息，并计算每个关键词的重要性，按照重要性对关键词进行排序，然后将页面及其对应的关键词构建为正向索引并存储进索引库。正向索引可以快速找到一个页面中包含哪些关键词，但是实际搜索是通过关键词寻找包含它的页面。这就需要扫描每一个页面，判断它是否包含该关键词，如果计算量太大，则无法满足实时返回搜索及排名结果的要求。因此，正向索引不能直接用于排名，还需要重新构建为倒排索引，把文档对应关键词的映射转换为关键词到文档的映射，根据关键词快速获取包含这个关键词的文档列表。

7）链接分析

搜索引擎对页面进行排序时，除了需要考虑网页本身的关键词密度和位置，还需要引入其他标准来衡量网页。链接分析是一种通过分析链接关系从而获取网页重要性的方法。搜索引擎会分析页面上有哪些链接指向哪些页面，每个页面有哪些导入链接，链接使用的锚文本等。复杂的链接关系能够反映网站和页面的质量及其与关键词的相关度。由于网页和链接数量巨大，且网页之间的链接关系不断更新，链接分析需要耗费很多时间。目前主流的搜索引擎广告质量分析因素中，都包含链接结构信息，很多搜索引擎在进行倒排索引前就要先进行链接分析，这对索引的排序会产生影响。

8）数据整合

除了 HTML 文件外，搜索引擎还能抓取和索引以文字为基础的多种文件类型，包括 PDF、Word、WPS、XLS、PPT、TXT 文件等，但不能直接处理图片、视频、Flash 等非文字内容，只能通过说明性文本对其进行处理。

1.2.4 检索及呈现

当用户输入查询关键词后，搜索引擎会对搜索请求进行详细分析。对于有多个关键词的搜索请求，搜索引擎的默认处理方式是在关键词之间使用“与”的逻辑。用户在搜索时，如果输入了明显的错别字，或者英文单词有明显的拼写错误，搜索引擎会提示用户正确的用字或拼写。某些搜索词还会触发整合搜索，如热门话题容易触发资讯内容。

搜索词分析完成后，搜索引擎需要从索引库中将与搜索词匹配的页面提取出来。由于包含搜索关键词的页面可能有成千上万甚至上百万，搜索引擎会通过相关性、权威性、时效性、丰富度等不同的维度，对页面得分进行综合排序，选择综合权重较高的页面进行匹配。

将搜索结果反馈给用户之前，需要对网页进行排序。相关性计算是排名中最重要的一步。影响相关性的因素主要包括关键词常用度、链接的使用和页面权重、关键词出现的位置、关键词密度和字频、关键词的距离等。

网页排名基本确定后，搜索引擎还会根据用户搜索数据（IP 地址、搜索时间、搜索记录、网页浏览记录等）对结果进行优化，并通过一些过滤算法对排名进行微调，得到最终的搜索结果。

搜索结果确定后，排名程序会调用原始页面的标题标签、说明标签等数据显示在页面上。有时搜索引擎需要动态生成页面摘要。目前搜索引擎搜索结果的展现形式非常丰富，包括摘要式（图 5-4）、图片式、视频式（图 5-5）、软件下载式、步骤式、新闻资讯式（图 5-6）等。

图 5-4　摘要式搜索结果

图 5-5　视频式搜索结果

图 5-6　图片和新闻资讯式搜索结果

2. 搜索引擎付费推广

2.1　搜索引擎付费推广概述

搜索引擎付费推广，也叫搜索引擎广告、搜索引擎竞价推广，是一种将企业网站或产品服务以关键词的形式在搜索引擎平台上推广的一种按效果付费的网络广告。企业通过购买一定数量的相关关键词，让企业网站和广告信息出现在搜索引擎的搜索结果中，还可以通过调整每次单击付费价格，控制推广信息在特定关键词结果中的排名，并通过设定不同的关键字，捕捉到不同类型的目标访问者，用少量的投入带来大量潜在客户，让网络营销活动效果在短期内立竿见影。

2.1.1　搜索引擎付费推广的优点

1）准确定位，精准投放

搜索引擎付费推广可以通过关键词锁定有需求的用户。选择好要推广的关键词之后，企业网站和广告只会出现在搜索这些关键词的用户面前。此外，还可以通过设置投放地域和投放时间，将广告精准覆盖到特定地区、特定时段的用户，大大提高了营销效果。

2）预算灵活控制

搜索引擎付费推广虽然是一种付费广告，但推广账户的预算费用可以通过设置和调整关键词的数量、位置、投放地域和时间等来灵活管理。

3）效果实时查看

企业登录账户后可以随时查看推广效果统计报告，以及推广效果实时信息，及时了解账户推广计划的展现、单击及转化的流量，详细地了解广告效果。

2.1.2　搜索引擎付费推广的基本概念

1）单击付费

一种广告费用定价模式。当用户单击广告时，根据广告被单击的次数，向企业的广告账户收取费用。如果广告针对某个搜索查询显示但用户并未单击它，则不会产生任何费用。例如，有 100 人搜索到了广告并有 3 个人单击了广告，企业只需要为这 3 次单击付费。

(1) 最高每次单击费用。

最高每次单击费用是指当广告展示在搜索引擎广告平台时,企业愿意为其得到的每次单击而支付的最高费用,也即出价金额。

(2) 实际每次单击费用。

搜索引擎推广平台综合考量企业的最高每次单击费用和质量得分等因素,对每次单击向企业的广告账户实际收取的费用。在质量得分比较理想的情况下,实际每次单击费用要低于最高每次单击费用。

(3) 首页出价估算值。

首页出价估算值是指在搜索查询与设置的关键词完全匹配的情况下,要让广告展示在搜索结果首页,需要支付的最高每次单击费用。如果企业的出价低于首页出价估算值,广告仍会正常展示,但不太可能出现在搜索结果的首页。一般而言,更优质的广告能获得更低的首页出价估算值。

(4) 最低展现价格。

最低展现价格也称为起价,是指为确保带有关键词的推广信息能够发布展示在搜索结果页,企业为每次单击所必须支付的最低金额。如果出价低于最低展现价格,关键词状态将显示“搜索无效”而无法获得展现机会。不同关键词的最低展现价格各不相同,企业在推广账户后台看到的最低展现价格较高,可能意味着关键词的质量太低或者关键词商业价值较大。此外,同一个关键词,在不同的推广账户中的最低展现价格也可能不同,这与推广账户的质量得分有关。

2) 质量得分

搜索引擎付费推广平台会对推广账户和关键词质量进行评估,以此决定最低出价并影响实际每次单击费用。账户中每个关键词都会获得一个质量得分,得分越高,代表平台认为广告和网站着陆页对于看到推广信息的搜索用户来说更具有相关性。同等条件下,潜在用户关注与认可的可能性更高,相应的最低出价和为每次单击支付的实际费用就越低。

影响质量得分的主要因素如下。

(1) 点击率。

点击率是搜索广告的点击量占展现量的比例,即广告的单击次数除以展示次数,是用于表明广告效果的指标。较高的点击率表示搜索用户对广告的关注度和认可度较高。广告与所在页面内容的相关性越强,用户单击广告的频率就越高,点击率也会随之增长。点击率是直接影响质量得分的重要因素,相关性和创意水平会通过影响点击率来影响关键词质量度。

(2) 创意内容和关键词的相关性。

由于语言的复杂性,搜索引擎付费推广平台不能做到完全智能化地判断关键词和创意内容的相关性。推广人员应该围绕关键词来撰写广告创意,用户在搜索关键词信息时才能引发创意标题和描述的飘红,达到系统对于创意内容和关键词相关性的要求,最终获得较高的质量得分。