

Internet 搜索引擎的安全性 - 修订版

2006 年 12 月 11 日

Ben Edelman

McAfee SiteAdvisor 顾问

Hannah Rosenbaum

研究分析家, McAfee SiteAdvisor

本文档的 Web 版包含超链接和其他图表, 请访问 http://www.siteadvisor.com/studies/search_safety_dec2006.html 进行浏览。

摘要

2006 年 5 月, McAfee 发布了一项研究, 对主要搜索引擎的安全性进行了比较。此报告以新的数据更新了该项研究, 并对搜索引擎结果所造成的特定类型的威胁进行了更深入的分析。

在此项研究中, 我们使用 McAfee SiteAdvisor 的自动网站评级对主要搜索引擎的安全性进行了比较。结果再次表明, 大多数主要搜索引擎所链接站点的安全性相似, 但是 AOL 取代了 MSN 成为最安全的引擎, 而 Yahoo! 则替代了 Ask 成为搜索结果危险性最高的引擎。在搜索引擎中我们发现, 赞助结果的安全性明显低于搜索引擎随机结果的安全性。尽管搜索引擎用户还面临着危险下载、浏览器漏洞和诈骗等的威胁, 但是令人憎恶的电子邮件方式仍是最主要的安全性风险。

主要发现

在美国, 排名前 5 位的搜索引擎 (占有所有搜索引擎使用的 91%) 都包含一些危险结果。

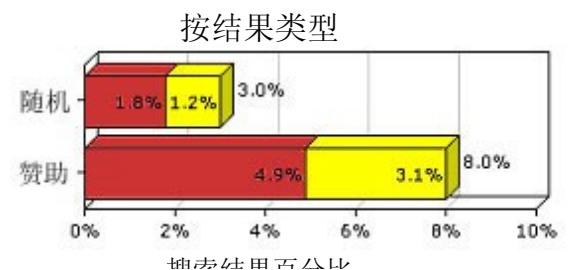
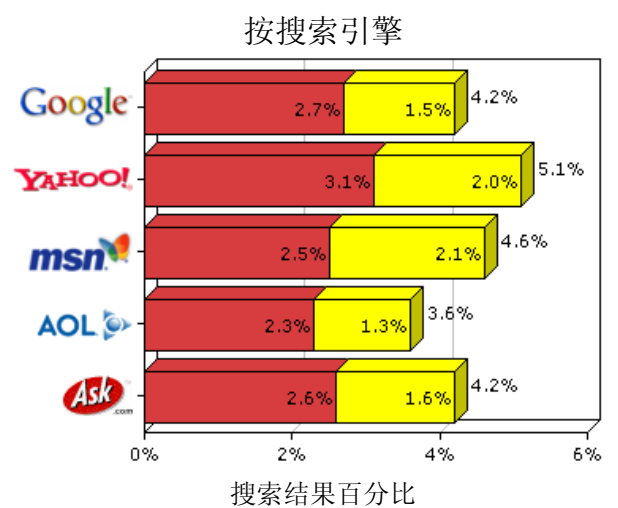
AOL 返回的结果是最安全的, 有 3.6% 的结果被 McAfee SiteAdvisor 评定为红色¹或黄色²。Yahoo! 是结果中被评定为红色或黄色最多的引擎, 百分比高达 5.1%。平均来说, 有 4.4% 的搜索结果会链接到有危险的网站。

有 8% 的赞助结果被评定为红色或黄色, 这几乎是随机结果中红色或黄色站点所占百分比的 3 倍。特别是, 赞助结果中诈骗站点的发现频率更高。

有 41.4% 的红色或黄色站点存在不当电子邮件活动、24.5% 包含危险下载、26.8% 是诈骗站点、32.3% 会链接到其他危险站点, 另有 3.0% 包含浏览器漏洞。许多站点同时包含多种危险。

成人搜索词汇的危险性是非成人搜索词汇的 2 倍。

红色和黄色²结果所占百分比...



¹ 评定为“红色”表示该站点未通过 McAfee SiteAdvisor 的安全测试。例如, 那些分发广告软件、发送大量垃圾邮件或对用户计算机进行未经授权更改的站点。

² 评定为“黄色”表示此站点参与某些活动, McAfee SiteAdvisor 的安全性测试对这些活动给出了重要的安全公告信息。例如, 发送大量非垃圾电子邮件, 显示众多弹出广告或提示用户更改浏览器设置的站点。

在某些危险关键字的搜索结果中，危险站点会占到 90% 以上。特别危险的关键字包括“bearshare”、“free screensavers”和“free ringtones”。

简介

搜索引擎已成为用户与 Internet 的主要通道。几乎有 80% 的网站页面访问源自于搜索引擎。虽然搜索引擎通常是浏览 Web 最快捷的方式，但它们未必最安全。

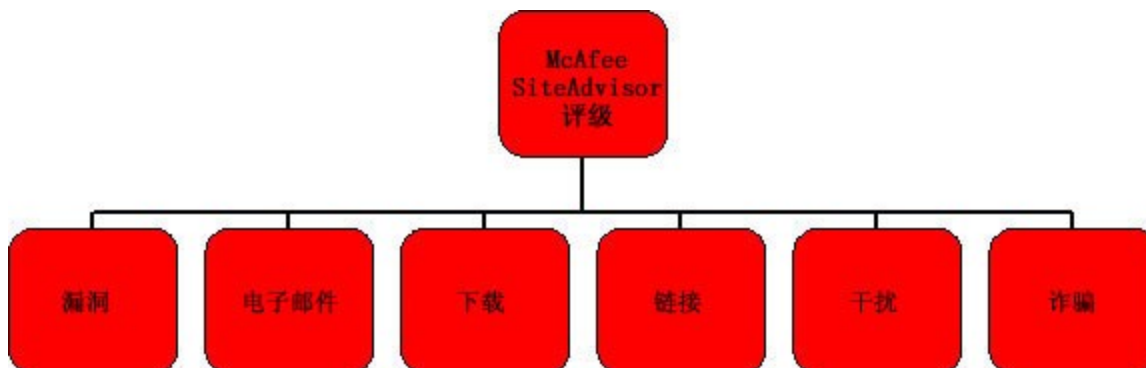
通常，搜索引擎排名不会明确地以网站安全为中心。但用户却经常会依据站点的搜索排名假定其安全性排名。1/3 以上的搜索者会认为在搜索结果中排名较前的公司同样是其领域内的佼佼者，这表示用户可能会错误地将可信赖性与搜索结果的排名关联起来。但事实上搜索排名并不会反映出安全性。除非用户是依据之前的经验选择信任的熟悉站点，否则就可能让自己面临众多威胁。

方法

比较以下 5 个搜索引擎搜索结果的安全性：Google、Yahoo!、MSN、AOL 和 Ask。首先，收集一份包含大约 2,500 个常用关键字的列表，这些常用关键字来自 Google Zeitgeist、Yahoo!、AOL、Ask、Lycos、Wordtracker、Hitwise 和其他行业来源的常见搜索列表。我们分别对 5 个搜索引擎的每个关键字搜索结果的前 5 页进行评估。然后根据结果位置和结果类型（随机与赞助）分析站点安全性。

我们的站点安全性评估来自 McAfee SiteAdvisor 790 万个网站的安全性数据库，占全部 Web 通讯的 95% 以上。自先前于 2006 年 5 月作出的搜索引擎分析之后，McAfee SiteAdvisor 在网站、站点下载和浏览器漏洞的覆盖范围大大增长，从 300 万个站点增加到 790 万个。

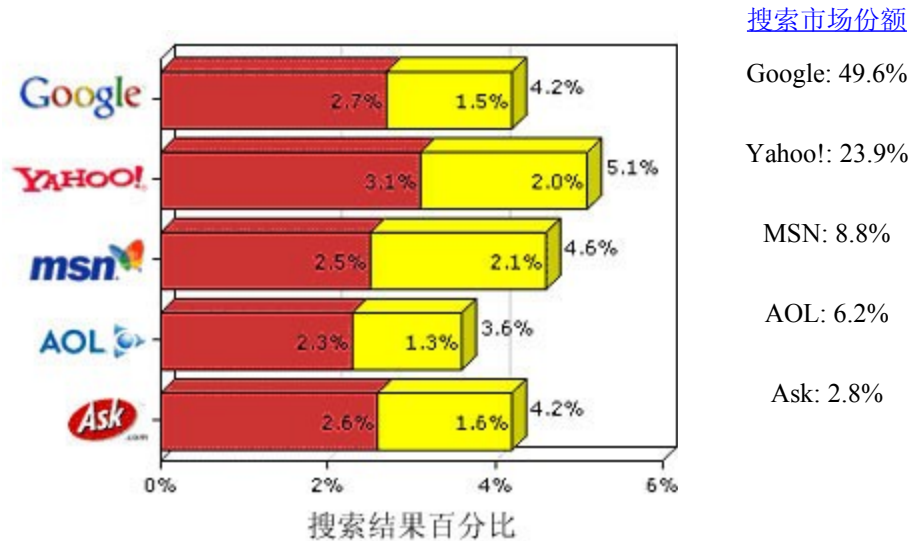
我们使用全面的 McAfee SiteAdvisor 评级以及对下列特定行为的组成评级，分析网站的安全性：浏览器漏洞、电子邮件、下载、诈骗、干扰（如弹出窗口）和到其他此类站点的链接。



分析

搜索引擎比较

虽然所有搜索引擎都会返回一些不安全的结果，但是 Yahoo! 结果中的危险性特别高。在我们的分析中，有 5.1% 的 Yahoo! 结果被评定为红色或黄色。AOL 返回的结果最安全，有 3.6% 被评定为红色或黄色。总体来讲，有 4.4% 的搜索结果被评定为红色 (2.6%) 或黄色 (1.7%)。



红色和黄色结果所占百分比（按搜索引擎）

自 2006 年 5 月发布搜索引擎研究之后，搜索结果中的红色和黄色站点所占百分比已由 5.0% 下降到 4.4%。目前，Google、AOL 和 Ask 返回的结果较安全，而 Yahoo! 和 MSN 返回结果的危险性较高。导致这些变化的因素可能有许多。例如，某些搜索引擎更改了其排名算法。另有一些搜索引擎完全更改了广告平台。同时，自上次分析后我们还更新了关键字列表，以期能更好地反映出最新和最常用的搜索词汇。

搜索引擎的相对排名也有所改变。AOL 取代了 MSN 成为返回结果最安全的搜索引擎，而 Yahoo! 则替代 Ask 成为搜索结果最危险的搜索引擎。MSN 安全性的下降可能归因于其付费搜索程序的扩展，而在我们早先文章收集数据期间，其正处于转换中（显示极少的广告）。在 MSN 的赞助结果中，红色和黄色站点的百分比五月份为 6.3%，到十一月激增至 10.7%。

自五月份以来的变化：

	2006 年 11 月	2006 年 5 月
Google (↓)	4.2%	5.3%
Yahoo (↑)	5.1%	4.3%
MSN (↑)	4.6%	3.9%
AOL (↓)	3.6%	5.3%
Ask (↓)	4.2%	6.1%
Organic (↓)	3.0%	3.1%
Sponsored (↓)	8.0%	8.5%

↓ 自上次研究以来安全性增加

↑ 自上次研究以来安全性降低

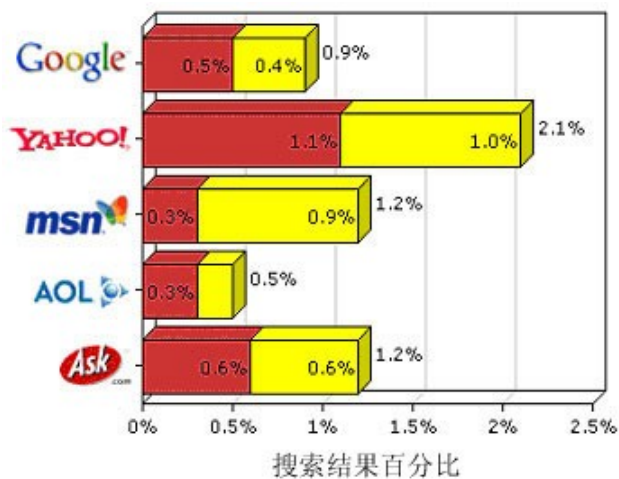
结果（按安全性方面）：漏洞、电子邮件、下载、诈骗、链接

搜索引擎用户面临多种类型的网站安全性风险。在评定为红色或黄色的搜索结果中，有 24.5% 是因为危险下载，41.4% 是因为不当电子邮件活动，26.8% 是因为诈骗，另有 31.2% 是因为链接到其他站点。3.0% 的红色和黄色搜索结果包含浏览器漏洞，这种威胁相当严重，即使用户只是浏览站点也可能会损坏用户的计算机。Google 最近针对特定漏洞站点采取了[插入式警告](#)页面措施，使 Google 的漏洞结果的安全性略有提升。此近期发展表明 Google 正在努力帮助用户远离危险站点，但是我们发现 Google 仅会对包含浏览器漏洞的 Google 搜索结果中的 18% 发出警告，所以用户不能依赖这些警告来保证安全。

评定为红色/黄色的搜索结果百分比（按安全性风险）					
	下载	电子邮件	漏洞	诈骗	链接
Google	1.1%	1.9%	0.13%	0.9%	1.6%
Yahoo!	1.0%	2.0%	0.09%	2.1%	1.6%
MSN	1.3%	1.6%	0.09%	1.2%	1.3%
AOL	1.1%	2.0%	0.17%	0.5%	1.2%
Ask	1.0%	1.6%	0.17%	1.2%	1.4%
整体	1.1%	1.8%	0.13%	1.2%	1.4%

红色/黄色搜索结果百分比（按安全性风险）					
	下载	电子邮件	漏洞	诈骗	链接
Google	26.2%	45.2%	3.1%	21.4%	38.1%
Yahoo!	19.6%	39.2%	1.8%	41.2%	31.4%
MSN	28.3%	34.8%	2.0%	26.1%	28.3%
AOL	30.6%	55.6%	4.7%	13.9%	33.3%
Ask	23.8%	38.1%	4.0%	28.6%	33.3%
整体	25.0%	41.4%	3.0%	26.8%	32.3%

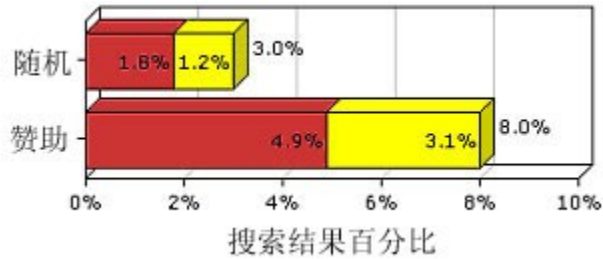
这些表格证明用户使用不同的搜索引擎会面临不同的危险。Yahoo 搜索结果中包含的诈骗结果是其他搜索引擎的 2 倍，AOL 搜索结果包含的诈骗站点最少（SiteAdvisor 诈骗评级会警告用户有关误导性站点内容，如欺骗性声明、优惠或付费活动）。但是 AOL 也并非在各个方面均安全：漏洞和大量电子邮件散播站点在 AOL 列出的结果中更为普遍。



红色和黄色诈骗结果百分比（按搜索引擎）

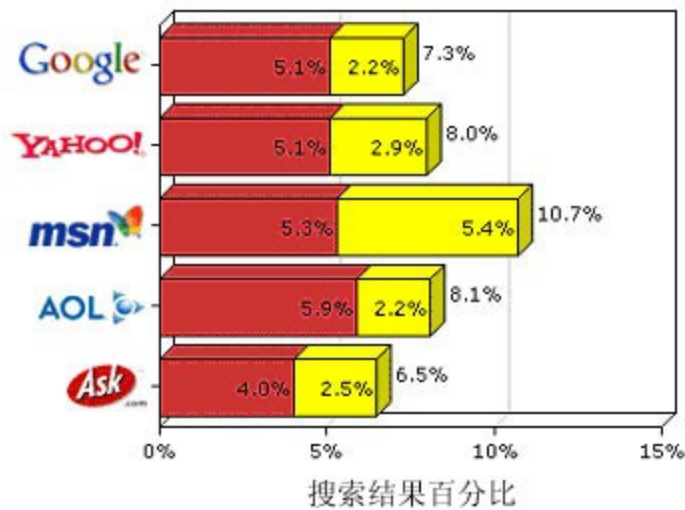
随机与赞助结果

在我们的分析中，由搜索引擎的赞助（付费广告）结果引向红色和黄色站点的可能性约为随机（非付费）结果的 3 倍。此结果表明随机搜索排名算法安全性更佳，它还表明搜索引擎从危险站点获取了实际的报酬。通过允许危险网站购买搜索结果的靠前位置，搜索引擎帮助这些站点获得（并可能会伤害）信任用户。2006 年全年搜索广告收入将超过 [160 亿美元](#)。假设不安全站点的广告价格与其他广告商的价格相近（猜想一个合理的近似值），搜索引擎通过将用户引向危险站点每年可从其美国营业收入中赚取大约 12.8 亿美元。



赞助结果中显示的红色和黄色站点百分比最多可达随机结果的 3 倍

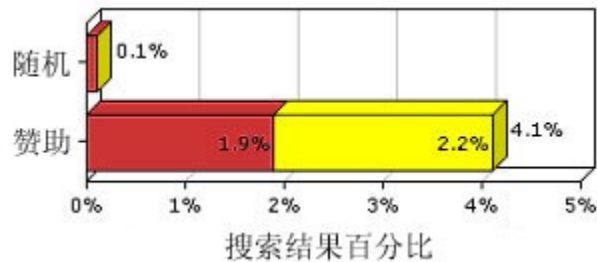
MSN 的危险赞助结果的百分比最高 (10.7%)，Ask 的赞助结果最安全 (6.5%)。



红色和黄色赞助结果百分比（按搜索引擎）

赞助结果会促进诈骗站点发展

特别值得注意的是，赞助结果包含的诈骗站点的百分比要远大于随机结果。4.1% 的赞助结果会因诈骗站点被评定为红色或黄色，相比之下随机结果只有 0.1%。诈骗站点包含误导性站点内容，如欺骗性声明或促销优惠和诱饵行为付费活动。例如，诈骗站点可能会对[免费下载收费](#)、[对进入免费绿卡彩券的用户收费](#)，或[将宝塔式方案伪装成有利的在家办公机会](#)。尽管搜索引擎的广告准则禁止欺骗性内容，但诈骗站点仍在赞助结果中频繁出现。另请参阅 Ben Edelman 的 [False and Deceptive Pay-Per-Click Ads](#)。



红色和黄色诈骗结果几乎全部出现在赞助列表中。

免费一词的使用

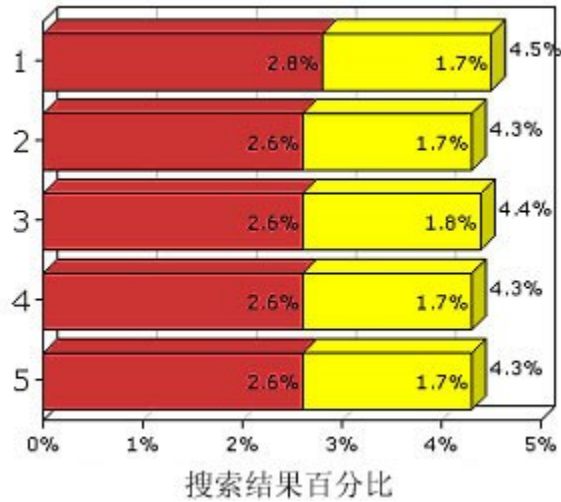


搜索词汇中包含“免费”一词的红色和黄色结果百分比

包含“免费”一词的常见搜索尤其可能将用户引向具有令人憎恶活动的站点。在包含“免费”一词的常见搜索中，有 14.5% 的搜索结果会链接到评定为黄色或红色的站点。购买付费搜索词“免费”的网站会使用各种策略诱惑用户。“免费”下载常常与有害的程序捆绑在一起，特别是广告软件、工具栏或跟踪软件。提供“免费”服务（如铃声或信用检查）的站点常会提供误导性声明，将收费说明隐藏在不引人注意的缩印中，或强加其他欺骗性付费活动。另有一些站点会[对可以在其他地方轻松取得的产品收费](#)，欺骗没有经验的客户。例如，某些站点会对免费浏览器 Firefox 收费，并[含糊承诺](#)提供技术支持或客户服务来证明其收费的合理性。最后，声明提供免费产品（如 iPod 或平面电视）的站点常常会将用户的电子邮件地址泄漏给第三方，让用户饱受每星期数百封电子邮件的骚扰。

按页面的安全性

搜索结果前 5 页的安全性基本没有差别，这表明搜索结果排名和站点安全性之间没有相关性。[62%](#) 的搜索者会选择第一页的结果，但这一页的结果并不比后续几页更为安全。所以仅依靠选择排名靠前的搜索结果并不能保证用户的安全。



红色和黄色结果百分比（按搜索结果页面）

按位置的安全性

随机搜索结果的安全性在位置排名中无变化，而赞助搜索结果的安全性却会因排名的不同而有所不同。在我们的测试中，Google 页面顶端的第一个赞助结果与第 2 和第 3 个相比似乎更安全。在页面右侧列出的 Google 赞助结果中，前两个结果的危险性比其后面结果的危险性要低。我们注意到许多排名第一的 Google 广告实际上是非盈利的，是通过 [Google.org](#) 收到的免费广告。这些非盈利广告的特殊性质大概能够说明 Google 最先位置相对安全的原因：这些广告的目的是促进安全站点（无间谍软件、诈骗等），而且由于这些广告的创建者不需要向 Google 付费，所以常常排名靠前。

按排名的 Google 结果

随机	
位置	红色/黄色百分比
1	3.1%
2	3.2%
3	3.5%
4	2.9%
5	3.2%
6	2.9%
7	3.0%
8	3.1%
9	3.1%
10	3.2%

赞助：页面顶端	
位置	红色/黄色百分比
1	6.2%
2	9.5%
3	9.0%

赞助：页面右侧	
位置	红色/黄色百分比
1	5.2%
2	6.2%
3	8.4%
4	9.1%
5	8.9%
6	8.3%
7	7.4%
8	7.7%

类别分析

在所分析的 Google Zeitgeist 搜索词汇中，最危险的 Zeitgeist 类别是“科技玩具” - 其中有 23.3% 的结果被 McAfee SiteAdvisor 评定为红色或黄色。“科技玩具”搜索词汇包括“iPod nano”、“mp3 music downloads”和“winmx” - 这些词汇通常会将用户引向诈骗和其他可疑活动。总体来讲，与音乐和技术相关的 Google Zeitgeist 类别返回的危险结果百分比最高。这些类别的安全性危害主要包括危险下载和对免费软件收费的诈骗。

最危险的五个 Google Zeitgeist 类别包括：

Google Zeitgeist 类别	红色/黄色结果百分比
科技玩具	23.3%
流行软件	20.0%
数字音乐	19.7%
技术查询	18.6%
流行男歌手	15.3%

名人搜索经常会包含许多危险站点。在搜索“流行男歌手”（如“Usher”和“Nelly”）以及“最受欢迎女明星”（如“Lindsay Lohan”和“Paris Hilton”）时，我们发现许多危险站点，前者为 15.3%，后者为 10.5%。危险搜索结果在“流行运动查询”中也普遍存在（11.3%），如“皇家马德里队”和“板球”，同样情况还有“儿时最爱”（6.7%），这包括“小熊维尼”和“崔弟”等关键字。

相反，在 Google Zeitgeist 类别“Web 2.0”中搜索“wikipedia”和“youtube”等关键字却不包含任何危险搜索结果。其他相对安全的 Google Zeitgeist 类别有“夏季旅行”（0.8%）（包括“Expedia”和“Travelocity”等搜索词汇）和“周末差事”（0.4%）（包括“USPS”和“Home Depot”）。

个别关键字分析

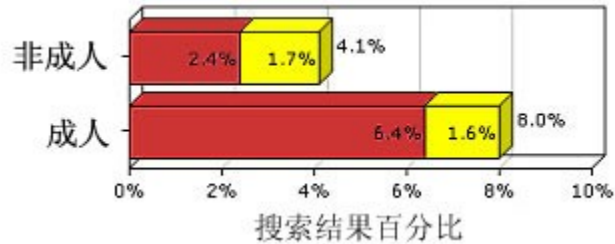
我们的测试表明，在所有搜索引擎中都是最危险的单个关键字是“bearshare”，其中有 53.3% 的搜索结果被评定为红色或黄色。Bearshare 是一个免费文件共享程序，在我们的测试中，程序的安装绑定有广告软件。“Bearshare”还是一个危险的关键字，因为许多站点会对获得此免费程序的用户收取费用。在 MSN 上搜寻“rotten.com”时，红色和黄色搜索结果的百分比最高可达 90.9%。

最危险的十个搜索词汇：

	搜索词汇	红色/黄色结果百分比
1	bearshare	53.3%
2	rotten.com	53.1%
3	free screensavers	53.0%
4	winmx	52.0%
5	screensavers	48.2%
6	limewire	46.2%
7	kazaa	45.8%
8	free ringtones	44.3%
9	ringtones	43.8%
10	lime wire	42.7%

成人与非成人搜索词汇

我们的测试显示，成人搜索词汇引向不安全结果的可能性是非成人搜索词汇的 2 倍。为了调查成人搜索词汇的安全性，我们将前 142 个成人关键字组成一组。这些成人搜索词汇的结果中有 8.0% 被评定为红色或黄色，而非成人词汇仅为 4.1%。成人关键字返回结果中包含浏览器漏洞和具有不当电子邮件活动的站点的百分比也较高。



成人搜索词汇引向危险搜索结果的可能性是非成人搜索词汇的 2 倍。

讨论

搜索引擎的排名最终是否会结合站点安全性仍有待观察。Google 的广告排名已经[结合](#)了“目标页面质量”的观念，而且 Google 的[广告原则](#)中论及的几个因素对我们也有实际意义，如避免向用户发送过多电子邮件，以及避免用户因取得所承诺的免费产品之前要面对“众多障碍”而离开。而且仍令我们感到忧虑的是，有那么多的 Google 广告商在继续促销服务，而这些服务对我们而言似乎完全是诈骗 - 承诺产品“免费”而不免费，或对随处都可以免费得到的产品进行收费。总体来说，搜索引擎确实从诈骗者的赞助结果中获取利润 - 这一因素往往会阻碍搜索引擎积极地从其结果中去除不安全站点。严格的广告准则和强制过滤可能会在短期内减少搜索引擎的收入，但此类活动最终定能让搜索引擎受益。如果赞助结果变得比随机结果更安全，用户可能会更频繁的点击赞助链接，而广告商也会相应增加其支付的搜索费用。

由于 Web 匿名的灵活性而产生的特殊问题，搜索引擎恰好可以解决。如果搜索引擎对此视而不见，诈骗者就可以操纵市场、欺骗客户，而且似乎可以轻松逃脱处罚。但是搜索引擎是有可能通过限制可以成为广告商的客户并禁止欺骗和恶意广告，帮助用户避开危险站点。同时，感兴趣的用户可以通过安装 [McAfee SiteAdvisor](#) 浏览器插件来取得这些好处，该插件可以自动标注搜索结果，以报告 McAfee SiteAdvisor 对每个搜索结果的安全性评估。

许多用户仍对 Web 安全一无所知。他们随意点击、提交电子邮件地址并进行下载，丝毫意识不到后果，直到受到损害。作为 Web 安全的倡导者，我们将此铭记心中，因为这意味着如果用户不是在安全的环境中，那么他们的网上体验将会相当糟糕。另有一些用户通过限制自己的线上活动，来设法避免 Web 的危险：他们仅访问熟悉的站点，从不提供其电子邮件地址，而且从不下载新的程序。这也非常不好：他们的谨慎使他们无法探索整个 Web，而且会让他们无法享受许多有用站点和娱乐站点提供的服务。用户应该能够安全地浏览 Web 而无需过度担心。我们希望搜索引擎能够采取措施改进其搜索结果的安全性，让用户可以更轻松地进行浏览。

其他图表和表格

[随机结果的分析（按搜索引擎）和赞助结果的分析（按搜索引擎）](#)

按 [Zeitgeist 关键字组分析](#)（最安全和最危险的关键字类别）

[按个别关键字分析](#)

资源

[Forrester to Marketers: Let Age Define Your Buy](#) - ClickZ

[False and Deceptive Pay-Per-Click Ads](#) - Ben Edelman

[Google Revenue to Exceed \\$10 Billion in 2007](#) - eMarketer

[How Consumers Find Web Sites In 2006](#) - Forrester

[The Philanthropic Arm of Google](#) - Google

[Search Engine User Behavior Study](#) - iProspect

[Search Engine Market Share](#) - Nielsen/NetRatings

[Google AdWords Landing Page Quality Ranking Initiated](#) - Search Engine Journal

搜索引擎广告准则

[Google AdWords 编辑指南](#)

[Google AdWords 目标网页和网站质量指南](#)

[Inside AdWords: Landing Page Quality Update](#)

[Yahoo! - Sponsored Search Listing Guidelines](#)

[MSN - AdCenter Search Ads Content Guidelines](#)

[AOL - Ad Specs - Policies and Guidelines](#)

[Ask - Sponsored Listings - Guidelines](#)

SiteAdvisor Web 安全工具

siteadvisor.com