

Google PageRank: What Do We Know About It?

Published on June 5th, 2007 in [Developer's Toolbox](#)

Everybody is using it, but (almost) nobody really knows how it works. **Google PageRank** is probably one of the most important algorithms ever developed for the Web. With billions of existing pages and millions of pages generated every day, the search issue in the Web is more complex than you probably think it is. PageRank, only one of hundreds of factors used by Google to determine best search results, helps to keep our search clean and efficient. But how is it actually done? How does Google PageRank work, which factors do have an impact on it and which don't? And what do we *really* know about PageRank? In this article **we put the facts straight**.

Over the last weeks we've done an extensive research and selected **dozens of facts and suggestions about PageRank, which seem to be true in practice**. Besides, we've collected academic papers related to the issue - such as scientific proposals for better search results (such as Topic-Sensitive PageRank); you'll also find references to mathematical background of PageRank as well as **16 useful PageRank tools** you can use to analyze and track the ranking of your web-projects.

Update: we'd like to apologize for some misleading facts we've initially included in this article. We've re-checked the sources and inaccurate or incomplete data. The .pdf-file won't contain any mistakes. Thanks to all the readers who've pointed us to the mistakes (particularly Dan Grossman and [Reuben Yau](#)).

- You don't have to read the whole article. Most important facts are selected in the beginning of the post as a brief summary.
- You might be interested in reading our article [Google AdSense: Facts, FAQs and Tools](#), which should provide you with the most important facts, tools and resources about Google AdSense.



Summary: How Does PageRank Work?

1. PageRank is only **one of numerous methods Google uses** to determine a page's relevance or importance.

2. Google interprets a link from page A to page B as a **vote**, by page A, for page B. Google looks not only at the sheer volume of votes; among 100 other aspects it also analyzes the page that casts the vote. However, these aspects don't count, when PageRank is calculated.
3. PageRank is **based on incoming links**, but not just on the number of them - relevance and quality are important (in terms of the PageRank of sites, which link to a given site).
4. **$PR(A) = (1-d) + d(PR(t1)/C(t1) + \dots + PR(tn)/C(tn))$** . That's the equation that calculates a page's PageRank.
5. **Not all links weight the same** when it comes to PR.
6. If you had a web page with a PR8 and had 1 link on it, the site linked to would get a fair amount of PR value. But, if you had 100 links on that page, each individual link would only get a fraction of the value.
7. Bad incoming links don't have impact on Page Rank.
8. Ranking popularity **considers site age**, backlink relevancy and backlink duration. PageRank doesn't.
9. Content is not taken into account when PageRank is calculated.
10. PageRank does not rank web sites as a whole, but is determined for each page individually.
11. **Each inbound link is important** to the overall total. Except banned sites, which don't count.
12. PageRank values don't range from 0 to 10. PageRank is a **floating-point number**.
13. Each Page Rank level is **progressively harder to reach**. PageRank is believed to be calculated on a logarithmic scale.
14. Google calculates pages PRs permanently, but we see the update **once every few months** (Google Toolbar).

Summary: Impact on Google PageRank

1. Frequent content updates don't improve Page Rank automatically. Content is not part of the PR calculation.
2. **High Page Rank doesn't mean high search ranking**.
3. DMOZ and Yahoo! Listings don't improve Page Rank automatically.
4. .edu and .gov-sites don't improve Page Rank automatically.
5. Sub-directories don't necessarily have a lower Page Rank than root-directories.
6. **Wikipedia links don't improve PageRank** automatically (*update*: but pages which extract information from Wikipedia might improve PageRank).
7. Links marked with **nofollow-attribute don't contribute** to Google PageRank.
8. Efficient internal onsite linking has an impact on PageRank.
9. Related high ranked web-sites count stronger. But: "a page with high PageRank may actually pass you less if it has more links, because it's spread too thin." [\[RY\]](#)
10. Links from and to high quality related sites have an impact on Page Rank.
11. Multiple votes to one link from the same page cost as much as a single vote.

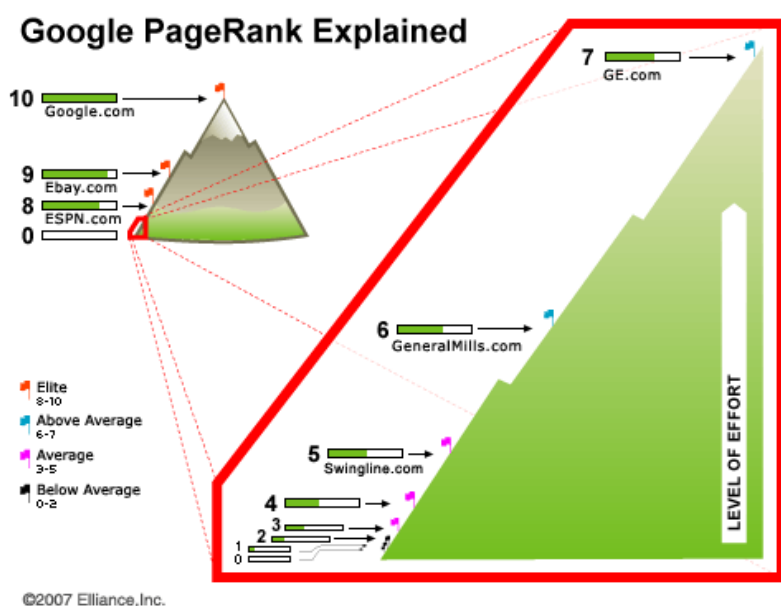
1.1. What is PageRank?

- "PageRank is [only] one of the methods Google uses to determine a page's relevance or importance." [\[PageRank Explained Correctly\]](#)
- "Google uses many factors in ranking. Of these, the PageRank algorithm might be the best known. PageRank evaluates two things: how many links there are to a web page from other pages, and the quality of the linking sites. With PageRank, five or six high-quality links from websites such as www.cnn.com and www.nytimes.com would be valued much more highly than twice as many links from less reputable or established sites." [\[Google Librarian Central\]](#)
- "PageRank has only ever been an approximation of the quality of a web page and has never had anything to do with the measuring of the topical relevance of a web page. Topical relevance is measured with link context and on-page factors such as keyword density, title tag, and everything else." [\[PageRank: An Essay\]](#)

1.2. How Does PageRank work?

- **No one really knows.** "No one knows for sure how PageRank is currently calculated by Google." [\[Google PageRank Explained\]](#)
- **$PR(A) = (1-d) + d(PR(t1)/C(t1) + \dots + PR(tn)/C(tn))$** . "That's the equation that calculates a page's PageRank. In the equation 't1 - tn' are pages linking to page A, 'C' is the number of outbound links that a page has and 'd' is a damping factor, usually set to 0.85."

- We can think of it in a simpler way: a page's PageRank = 0.15 + 0.85 * (a "share" of the PageRank of every page that links to it). "share" = the linking page's PageRank divided by the number of outbound links on the page. A page "votes" an amount of PageRank onto each page that it links to. The amount of PageRank that it has to vote with is a little less than its own PageRank value (its own value * 0.85). This value is shared equally between all the pages that it links to." [[Google's Page Rank](#)]
- "The core Google **PageRank algorithm "distributes" its established PR across all of the outbound links.** Put differently, if you had a web page with a PR8 and had 1 link on it, the site linked to would get a fair amount of PR value. But, if you had 100 links on that page, each individual link would only get a fraction of the value." [[The Importance of PageRank](#)]
- "From this, we could conclude that a link from a page with PR4 and 5 outbound links is worth more than a link from a page with PR8 and 100 outbound links. The PageRank of a page that links to yours is important but the number of links on that page is also important. The more links there are on a page, the less PageRank value your page will receive from it." [[Google's Page Rank](#)]
- "PageRank [...] **uses the link structure** as an indicator of an individual page's value. Google interprets a link from page A to page B as a vote, by page A, for page B. Google looks at considerably more than the sheer volume of votes, or links a page receives; e.g. it also analyzes the page that casts the vote. Votes cast by pages that are themselves "important" weigh more heavily and help to make other pages "important." [[Google: Technology](#)]
- "**Not all links weight the same** when it comes to PR. So an 'important' page linking to you gives you more PR than a 'less important' one. [...] A factor in PR propagation is the number of out-links the 'voting' page have. So a PR4 page with only one out-link on it might give you more weight than a PR5 page with 100 out-links on it. A typical example here would be the famous milliondollarhomepage. This page is PR7 page with hundreds of out-links therefore its weight is would contribute very little to your page PR." [[Google PageRank Explained](#)]
- **Each Page Rank level is progressively harder to reach.** "PageRank is logarithmic in its calculation. In the same way that the earthquake Richter scale is exponential in calculation, so too is the mathematics behind Google PageRank. It takes one step to move from a PR0 to a PR1, it takes a few more steps to PR3, it takes even more steps to PR4, and many more steps again to PR5, and so one." [[Google Page Rank FAQ](#)]



[via [einfach-persoehnlich](#)]

- "PageRank does not rank web sites as a whole, but is determined for each page individually. Further, the PageRank of page A is recursively defined by the PageRanks of those pages which link to page A." [[The Page Rank algorithm](#)]
- "Google combines PageRank with sophisticated text-matching techniques to find pages that are both important and relevant to user's search. Google examines all aspects of the page's content (and the content of the pages linking to it) to determine if it's a good match for user's queries." [[What Is Google PageRank?](#)]
- "Google calculates pages PRs once every few months (PR update). After a PR update is done, all pages are assigned a new PR by Google and you will have this PR until a new PR update is done. New sites that were just launched will have a PR of 0 until an update is done by Google so that they are assigned an appropriate PR." [[Google PageRank Explained](#)]

- “Google PageRank is calculated all the time, but what we see in the [Google Toolbar](#) (or other online PR tools) is a snapshot in time which is updated every 3 months or so.” [\[Reuben Yau\]](#)
- PageRank values don’t range from 0 to 10. **PageRank is a floating-point number.** “It’s more accurate to think of it as a floating-point number. Certainly our internal PageRank computations have many more degrees of resolution than the 0-10 values shown in the toolbar.” [\[Matt Cutts\]](#)
- “We’re sure that their curve is similar to an **exponential curve** with each new “plateau” being harder to reach than the last. I have personally done [some research into this](#), and so far the results point to an exponential base of 4. So a PR of 6 is 4 times as difficult to attain as a PR of 5. [...] The difference between a high PR of 6, and a low PR of 6, could be hundreds or thousands of links.” [\[Top 10 Google Myths Revealed\]](#)
- “PageRank is believed to be calculated on a **logarithmic scale**. What this roughly means is that the difference between PR4 and PR5 is likely 5-10 times than the difference between PR3 and PR4. So, there are likely over a 100 times as many web pages with a PageRank of 2 than there are with a PageRank of 4. This means that if you get to a PageRank of 6 or so, you’re likely well into the top 0.1% of all websites out there. If most of your peer group is straggling around with a PR2 or PR3, you’re way ahead of the game.” [\[Importance of Google PageRank\]](#)
- “The fact is that PageRank is based on incoming links, but not just on the number of them. Instead PageRank is based on the value of your incoming links. To find the value of an incoming link look at the PR of the source page, and divide it by the number of links on that page. It’s very possible to get a PR of 6 or 7 from only a handful of incoming links if your links are “weighty” enough.” [\[Top 10 Google Myths Revealed\]](#)
- **“Google tries to find pages that are both reputable and relevant.** If two pages appear to have roughly the same amount of information matching a given query, we’ll usually try to pick the page that more trusted websites have chosen to link to. Still, we’ll often elevate a page with fewer links or lower PageRank if other signals suggest that the page is more relevant. For example, a web page dedicated entirely to the civil war is often more useful than an article that mentions the civil war in passing, even if the article is part of a reputable site such as Time.com.” [\[Google Librarian Central\]](#)
- **Links don’t give PR away, they are votes.** “When a page votes its PageRank value to other pages, its own PageRank is not reduced by the value that it is voting. The page doing the voting doesn’t give away its PageRank and end up with nothing. It isn’t a transfer of PageRank. It is simply a vote according to the page’s PageRank value.” [\[Page Rank Explained\]](#)
- “We know from the paper “The Anatomy of a Large-Scale Hypertextual Web Search Engine” ([Paper](#)) that the PageRank of a Web page is a number calculated using a recursive algorithm in which the page receives a share of the PageRank of each page that links to it.” [\[Google PageRank\]](#)
- **Crawlers don’t analyze web-sites permanently.** “It often takes two full monthly updates for all of your incoming links to be discovered, counted, calculated and displayed as backlinks.” [\[Google FAQ\]](#)

1.3. Which factors do have an impact on PageRank?

- **Each inbound link is important to the overall total. Except banned sites.** “PageRank is a form of a voting system. A link to a page is a vote for that page. Higher PageRank pages are viewed by Google as more important. Their votes are given more value by Google — much more value, in some cases. In general, the more voting links, the stronger the PageRank.” [\[Google PageRank FAQ\]](#)
- **Adding new pages can decrease Page Rank.** “The effect is that, whilst the total PageRank in the site is increased, one or more of the existing pages will suffer a PageRank loss due to the new page making gains. Up to a point, the more new pages that are added, the greater is the loss to the existing pages. With large sites, this effect is unlikely to be noticed but, with smaller ones, it probably would.” [\[PageRank Explained\]](#)
- **Page Rank can decrease.** “You can lose some important links that are no longer linking to your site. PR loss can also occur if some of your linking partners also experience a drop in their own PR, possibly setting off a chain reaction of lower PageRank all through the immediate linking network.” [\[Google PageRank FAQ\]](#)
- **Links from and to high quality related sites are important.** “The more closely related the pages, the higher the PageRank amount transferred.” “Linking to high quality sites shows the search engines your site is very useful to your visitors. Unless your site has been around for years and is well established and trusted by Google, this factor will have an adverse effect on your site’s overall ranking. Linking only to high quality content sites will give your site an edge over your competition.” [\[Let Google’s Algorithm Show You The Traffic, FAQ\]](#)
- **Incoming Links from popular sites are important.** If pages linking to you have a high PageRank then your

page gains some part of their reputation.

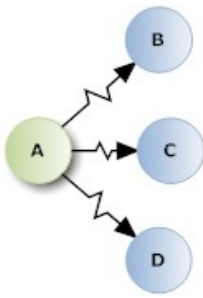
- Site can be banned if it links to banned sites. “Be extremely careful of any out-going links from your site. Don't link to bad neighborhoods (link farms, banned sites, etc.) Google will penalize you for bad links so always check the PageRank of the sites you're linking to from your site.” [\[SiteProNews\]](#)
- **Illegal activities will penalize your PageRank** and possibly ban your site from Google. “Hidden text, deceptive redirects, cloaking, automated link exchanges, or anything else against [Google's quality guidelines](#)” can ban your site from Google.
- **Myth: the higher your google PageRank, the better the results.** “While pages with a higher PageRank do tend to rank better, it is perfectly normal for a site to appear higher in the results listings even though it has a lower PageRank than competing pages. [...] Google examines the context of your incoming links, and only those links that relate to the specific keyword being searched on will help you achieve a higher ranking for that keyword.” [\[Top 10 Google Myths Revealed\]](#)
- **Related high ranked web-sites count stronger (or don't they?).** “One-way inbound links from websites with topics that are related to your website's topic will help you gain a higher Page Rank.” Other one-way inbound links from pages with high page rank but unrelated topics do help a little, but not nearly as much. [\[What Is Page Rank?\]](#)
- Different pages from a site can have different Page Rank. “Search engines crawl and index webpages not websites, that is why your page rank may vary from page to page within your website.” [\[What Is Page Rank?\]](#)

1.4. Which factors don't have an impact on PageRank?

- **Frequent content updates don't improve PR automatically.** “Although Google might send crawlers more frequently to analyze your site, what is more significant are links pointing to you.
- “Content is not taken into account when PageRank is calculated. Content is taken into account when you actually perform a search for specific search terms.” [\[Google PageRank\]](#)
- “High PageRank does NOT guarantee a high search ranking for any particular term. If it did, then PR10 sites like Adobe would always show up for any search you do. They don't.” [\[What Is Google PageRank?\]](#)
- Google considers **site age**, backlink relevancy and backlink duration. **PageRank doesn't.** If backlink isn't relevant, it won't weight much.
- **Wikipedia Links don't improve Page Rank.** “Wikipedia implemented a no-follow rule, indicating that outbound links should not be followed by search engine spiders.” [\[A Survival Guide to SEO & Wikipedia\]](#)
- **Listing in DMOZ and Yahoo! doesn't give your site a special PR Bonus.** “Google uses Open Directory Project (DMOZ.org), to power its directory. Coupling that fact with the observation that sites listed in DMOZ often get decent and inexplicable PageRank boosts, has lead many to conclude that Google gives a special bonus to sites listed in DMOZ. This is simply not true. The only bonus gained from being in DMOZ is the same bonus a site would achieve from being linked to by any other site.” However, DMOZ data is used by hundreds of sites.” [\[Top 10 Google Myths Revealed\]](#)
- **Sub-directories don't necessarily have a lower Page Rank than root-directories.** Depending on the popularity of a web-site your subdirectories can have a higher PageRank than the root pages.
- Meta-Tags don't improve PageRank. “Google can sometimes use the meta description tag to create an abstract for your site, so it may be useful to you if your home page is primarily composed of graphics. However, do not expect it to increase your rank.” [\[10 Google Myths Revealed\]](#)
- **.edu and .gov-sites do not provide higher PageRank (or do they?).** “We don't really have much in the way to say “Oh this is a link from the ODP, or .gov, or .edu, so give that some sort of special boost.” Its just those sites tend to have higher PageRank because-because more people link to them and reputable people link to them.” [\[A Google Myth Busted\]](#)

No Follow Treatment

Search Engine Comparison



no follow	Google	Yahoo	Ask
Linked page indexed?	<input type="radio"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Trust/ attribution passed?	<input type="radio"/>	<input type="radio"/>	<input checked="" type="checkbox"/>
Traffic passed?	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

- **Links marked with nofollow-attribute don't contribute to Google PageRank.** "Google implemented a new value, "nofollow", for the rel attribute of HTML link and anchor elements, so that website builders and bloggers can make links that Google will not consider for the purposes of PageRank – they are links that no longer constitute a "vote" in the PageRank system." [\[Wikipedia: PageRank\]](#)
- Multiple votes to one link from the same page cost as much as a single vote. "It is reasonable to assume that a page can cast only one vote for another page, and that additional votes for the same page are not counted." [\[PageRank FAQ\]](#)
- Links from one page to itself don't improve Page Rank. "It is reasonable to assume that a page can't vote for itself, and that such links are not counted." [\[PageRank Explained\]](#)
- **Bad incoming links don't have impact on Page Rank.** "Where the links come from doesn't matter. Sites are not penalized because of where the links come from." [\[Google PageRank\]](#)
- **Dangling links don't have impact on Page Rank.** "Dangling links are simply links that point to any page with no outgoing links. They affect the model because it is not clear where their weight should be distributed, and there are a large number of them. Because dangling links do not affect the ranking of any other page directly, we simply remove them from the system until all the PageRanks are calculated. After all the PageRanks are calculated they can be added back in without affecting things significantly." [\[PageRank Paper\]](#)

1.5. Ranking Factors (related to PageRank)

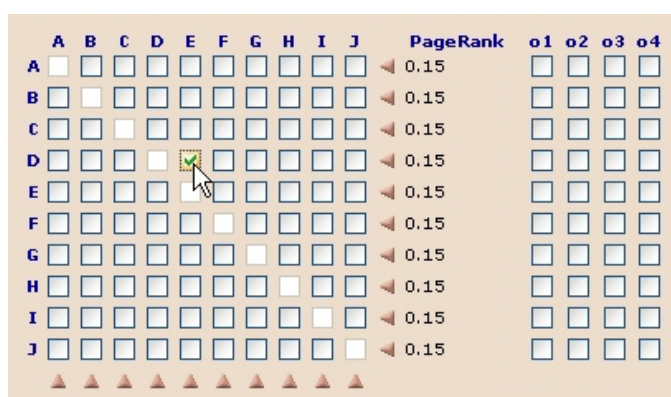
- **Efficient internal onsite linking is important.** "Internal linking is important to your overall ranking. Make sure your linking structure is easy for the spiders to crawl. Most suggest a simple hierarchy with links no more than three clicks away from your home/index page. Creating traffic modes or clusters of related links within a section on your site has proven very effective." [\[Let Google's Algorithm Show You The Traffic\]](#)
- **Anchor text is important.** The more specific is the reference, the better Google can evaluate it and consider it in relates search queries.
- Google penalizes link farms. "Google is only concerned with pages of over 100 outgoing links. Google considers overly linked pages to be link farms, and they are penalized as such." [\[Google FAQ\]](#)
- Headers (h1, ... ,h6), strong tags and **semantic content** are important. (*Update:* But it doesn't improve PageRank.) "Place it in the description and meta tags, place it in bold/strong tags, but keep your content readable and useful. Be aware of the text surrounding your keywords, search engines will become more semantic in the coming years so context is important." [\[Let Google's Algorithm Show You The Traffic\]](#)
- "The anchor text of a link is often far more important than whether it's on a high PageRank page." [\[What Is Google PageRank?\]](#)
- "If you really want to know what are the most important, relevant pages to get links from, forget PageRank. Think search rank. Search for the words you'd like to rank for. See what pages come up tops in Google. Those are the most important and relevant pages you want to seek links from. That's because Google is explicitly telling you that on the topic you searched for, these are the best." [\[What Is Google PageRank?\]](#)

2.1. Google PageRank: Theory & Scientific Background

- [A Survey of Google's PageRank](#)
 Calculation of Page Rank, Page Rank Implementation, Inbound Links, Outbound Links, Number of Pages, PageRank Distribution, Additional Factors and more.
- [The Lineal Algebra Behind Google](#)
 The \$25,000,000,000 Eigenvector - The Linear Algebra Behind Google. Google's success derives in large part from its PageRank algorithm, which ranks the importance of webpages according to an eigenvector of a weighted link matrix. Analysis of the PageRank formula provides a wonderful applied topic for a linear algebra course.
- [The Intelligent Surfer: Probabilistic Combination of Link and Content Information in PageRank](#)
 We propose to improve Page-Rank by using a more intelligent surfer, one that is guided by a probabilistic model of the relevance of a page to a query. Efficient execution of our algorithm at query time is made possible by precomputing at crawl time (and thus once for all queries) the necessary terms.
- [Topic-Sensitive PageRank](#)
 To yield more accurate search results, we propose computing a set of PageRank vectors, biased using a set of representative topics, to capture more accurately the notion of importance with respect to a particular topic. By using these (precomputed) biased PageRank vectors to generate query-specific importance scores for pages at query time, we show that we can generate more accurate rankings than with a single, generic PageRank vector.
- [Method for node ranking in a linked database](#)
 A method assigns importance ranks to nodes in a linked database, such as any database of documents containing citations, the world wide web or any other hypermedia database. The rank assigned to a document is calculated from the ranks of documents citing it. In addition, the rank of a document is calculated from a constant representing the probability that a browser through the database will randomly jump to the document. By Page and Lawrence.
- [How Google Finds Your Needle in the Web's Haystack](#)
 Mathematical Background of Google PageRank. By David Austin, Grand Valley State University
- [A Large-Scale Hypertextual Web Search Engine](#)
 Original Slides, by Larry Page.
- [Wikipedia: PageRank](#)
 Mathematical Theory Behind Google PageRank

3.1. Google PageRank Tools & Services

- [PageRank Search](#)
 Showing search results in order of PageRank.
- [Google PageRank Inspector.](#)
 Google PageRank inspector is PHP scripts that can seek all of your website, include out linked page or not, and display Pagerank value for each of your website pages. New pages linked by high pagerank pages can be indexed in google quickly and have higher keyword rank in google search.
- [Google's PageRank - Calculator](#)
 The results produced by the calculator indicate each page's PageRank share and are not equivalent to the values in the Google toolbar.



- [Webmastereyes, Visual PageRank View](#)

The results will show the page given along with the PageRank of each link on that page. You also have the option to show “nofollow” and external links.

- [Smart PageRank](#)

Checks PageRank from multiple datacenters and sends emails automatically if PageRank is updated.

- [Google PageRank Notifier](#)

“This script will send you an email whenever the PageRank of the given page changes. PageRank is taken from the Google Toolbar “API” and is updated once an hour.”

- [Google PageRankâ„¢ Checker](#) (registration required)

You can monitor site’s PageRank via RSS and you can also be notified via e-mail when the PageRank has been changed.

Google PageRank™ Checker

Check the Google Page Rank of Any Website

- [Dig PageRank](#)

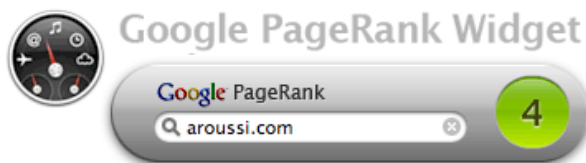
Checks the current Page Rank of a page in over 100 Google data centers.

- [Live PageRank Check](#)

The Live PageRank value may be used as an indicator of what will show when Google decides to export the PageRank values to the Google Toolbar. The Live PageRank calculator gives you the current PageRank value in the Google index, not just the snapshot that is displayed in the toolbar. Google updates its internal PageRank value continuously as the web changes and their index is updated. Only once every third month or so this value is exported to be displayed in the Google Toolbar.

- [Page Rank Widget](#) for Mac OS X.

Little Widget finds the Google Page Rank for any URL by calculating the checksum and requesting the PR from Google’s servers.



- [Google PageRank Prediction](#)

The tool analyzes the popularity of a given web-site and tries to predict its future Google PageRank. [More Page Rank Tools](#).

- [PageRank Checker](#)

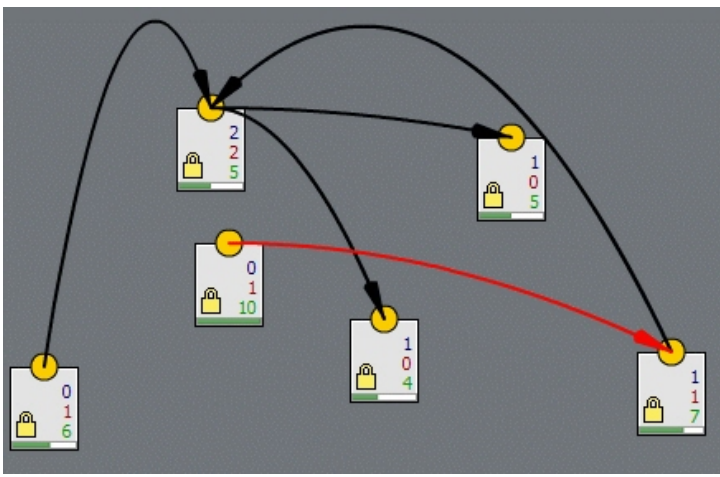
Shows PageRank of your backlinks.

- [PageRank Overlay \(PR Mapper\)](#) (both currently offline)

Browse your competitors website and view the Google PR of all the links at once. Also available as [Firefox Extension](#).

- [PageRank Decoder \(Demo\)](#)

“This little tool is not too much different then a tool that tells you your PageRank, however it allows you to organize your sites (with PR information) in a visual network and then correspondingly connect them with arrows. You can move them around like cards, connect them or not, and even delete them by throwing them in a trash can.” [[Search Engine Roundtable](#)]



- [Page Rank Export List History](#)
This Page Rank Update/Export List History contains the dates that Google Toolbar Pagerank (PR) was exported.
- [Google Ranking Factors](#)
Alleged positive and negative SEO Google Ranking Factors

3.2. Google Tools & Services

- [Google Quality Guidelines](#)
These quality guidelines cover the most common forms of deceptive or manipulative behavior, but Google may respond negatively to other misleading practices not listed here (e.g. tricking users by registering misspellings of well-known websites).
- [Check if your site is in Google database](#)
- [Reinclusion request form](#)
Request reinclusion of a site that has violated the webmaster guidelines
- [Google Tools](#)
A comprehensive overview on Dmoz.org.

Visit Smashingmagazine.com for more!

we smash you with information, which will make your life easier. really.

by Vitaly Friedman, Sven Lennartz, www.smashingmagazine.com, 04.08.2007

about: <http://www.smashingmagazine.com/about/>

e-mail: office@smashingmagazine.com

advertise with us: advertising@smashingmagazine.com (Michael Dobler)