

---

# How does Google Decide Which Web Page to Display First?

Ilse Ipsen

Joint work with: Teresa Selee & Rebecca Wills



**Web** [Images](#) [Groups](#) [News](#) [Froogle](#) [Local](#) [Scholar](#) **[more »](#)**

[Advanced Search](#)  
[Preferences](#)  
[Language Tools](#)

[Advertising Programs](#) - [Business Solutions](#) - [About Google](#)

©2006 Google

**Web**Results **1 - 10** of about **1,370,000** for **ipсен**. (0.11 seconds)

### [Ipsen, spécialiste en oncologie, désordres neuromusculaires et ...](#)

Groupe pharmaceutique européen spécialisé dans l'oncologie, l'endocrinologie et les désordres neuromusculaires, avec plus de 20 médicaments commercialisés ...

[www.ipсен.com/](#) - 40k - [Cached](#) - [Similar pages](#)

[The Group](#) - [Products](#) - [Career Area](#) - [Research & Development](#)

[More results from www.ipсен.com »](#)

### [Ipsen Limited, UK](#)

**Ipsen** Limited: Pharmaceutical company specialising in controlled release peptides and botulinum toxin.

[www.ipсен.ltd.uk/](#) - [Similar pages](#)

### [IPSEN Pharmaceuticals Limited, Ireland](#)

About **IPSEN** · Medical Information · Patient Information · Contact Us · RCSI Anthony Walsh Travelling Fellowship in Urology.

[www.ipсен.ie/](#) - 5k - [Cached](#) - [Similar pages](#)

### [Ipsen](#)

member of the **Ipsen** Group, the largest manufacturer of heat treating equipment in the world.

[www.ipсен-intl.com/](#) - 2k - [Cached](#) - [Similar pages](#)

### [Abar Ipsen Vacuum Heat Treating Equipment](#)

**Ipsen** is dedicated to developing furnace technology that fulfills all of industry's processing needs, capabilities, and environmental requirements with ...

[www.ipсен-intl.com/index.asp](#) - 25k - [Cached](#) - [Similar pages](#)

### [BEAUFOUR IPSEN Pharma : Bienvenue](#) - [ [Translate this page](#) ]

[www.bipmed.com/](#) - 2k - [Cached](#) - [Similar pages](#)

### [Iise Ipsen](#)

**ipсен** at ncsu dot edu Department of Mathematics North Carolina State University Raleigh, NC 27695-8205, USA. Research. numerical linear algebra, matrix ...

[www4.ncsu.edu/~ipсен/](#) - 2k - [Cached](#) - [Similar pages](#)

? ? ? ? ? ?

---

- Why is my web page **ALWAYS** in 7th place?
- Why does my page not appear **first**?

Is there a system to how Google displays pages?

**Web**Results **1 - 10** of about **2,560,000,000** for **google**. (**0.19** seconds)

## [Google](#)

Enables users to search the Web, Usenet, and images. Features include **PageRank**, caching and translation of results, and an option to find similar pages.

[www.google.com/](http://www.google.com/) - 4k - [Cached](#) - [Similar pages](#)

## [Google Talk](#)

Convenience: Your Gmail contacts are pre-loaded into **Google** Talk so inviting or ... **Google** Talk is in beta and requires a Gmail username and password. ...

[www.google.com/talk/](http://www.google.com/talk/) - 6k - [Cached](#) - [Similar pages](#)

## [Google Analytics](#)

Log Analysis/Web Statistics. Aimed at ISPs and large sites. Fully browser based reports, links to revenue, Multilingual functionality (including ...

[www.urchin.com/](http://www.urchin.com/) - 7k - [Cached](#) - [Similar pages](#)

## [Google Local](#)

Provides directions, interactive maps, and satellite/aerial imagery of the United States. Can also search by keyword such as type of business.

[maps.google.com/](http://maps.google.com/) - 23k - [Cached](#) - [Similar pages](#)

## [Official Google Blog](#)

Official weblog, with news of new products, events and glimpses of life inside the Googleplex.

[googleblog.blogspot.com/](http://googleblog.blogspot.com/) - 49k - [Cached](#) - [Similar pages](#)

## [Google News](#)

### Sponsored Links

#### [\\$1000.00 A Day](#)

Earn up to \$3000+ per day  
Working Only 30 Mins a day Now!  
[dataiye.com](#)

#### [Make \\$500+ Per Day](#)

Work P/T at Home, Simple Data Entry  
Start Earning in 30 Mins from Now!  
[Onlinejobcorp.com](#)

#### [Earn Extra \\$3,500+ /mo.](#)

Make Money While in Your Pajamas -  
Get Paid for Your Opinions Today!  
[Paidsurveysonline.com](#)

#### [Make \\$750+ /Day Online](#)

Top 2006 Home Business Opportunity  
No More Excuses, Start Earning Now!  
[Domaincashvault.com](#)

Enter url PageRank	Your
-----------------------	------

MyGoogle PageRank was created to enable webmasters to easily know and post PageRank on their pages without any Toolbar.

PageRank is owned by [Google Inc.](#)

This site is not affiliated with Google Inc. Trademarks remain trademarks of their respective companies.

[Article World](#)

Enter url   
PageRank

The Google™ - PageRank™  
of http://www.ipsen.com is: 6



**Here is the HTML code to add on your website:**

Which will show on your site:



**Note:** To avoid an error, please keep the HTML code intact.

Enter url Your  
PageRank

The Google™ - PageRank™  
of <http://www4.ncsu.edu/~ipsen> is: 5



**Here is the HTML code to add on your website:**

Which will show on your site:



**Note:** To avoid an error, please keep the HTML code intact.





[Home](#)

[About Google](#)

[Help Central](#)

[Google Features](#)

[Services & Tools](#)

Our Technology

► [Why Use Google](#)

[Benefits of Google](#)

*Find on this site:*

**Google searches more sites more quickly, delivering the most relevant results.**

## Introduction

Google runs on a unique combination of advanced hardware and software. The speed you experience can be attributed in part to the efficiency of our search algorithm and partly to the thousands of low cost PC's we've networked together to create a superfast search engine.

The heart of our software is PageRank™, a system for ranking web pages developed by our founders [Larry Page](#) and [Sergey Brin](#) at Stanford University. And while we have dozens of engineers working to improve every aspect of Google on a daily basis, PageRank continues to provide the basis for all of our web search tools.

## PageRank Explained

PageRank relies on the uniquely democratic nature of the web by using its vast link structure as an indicator of an individual page's value. In essence, Google interprets a link from page A to page B as a vote, by page A, for page B. But, Google looks at more than the sheer volume of votes, or links a page receives; 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."

Important, high-quality sites receive a higher PageRank, which Google remembers each time it conducts a search. Of course, important pages mean nothing to you if they don't match your query. So, Google combines PageRank with sophisticated text-matching techniques to find pages that are both important and relevant to your search. Google goes far beyond the number of times a term appears on a page and 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 your query.

# Ranking Web Pages

---

Popular pages: high PageRank  
Unpopular pages: low PageRank

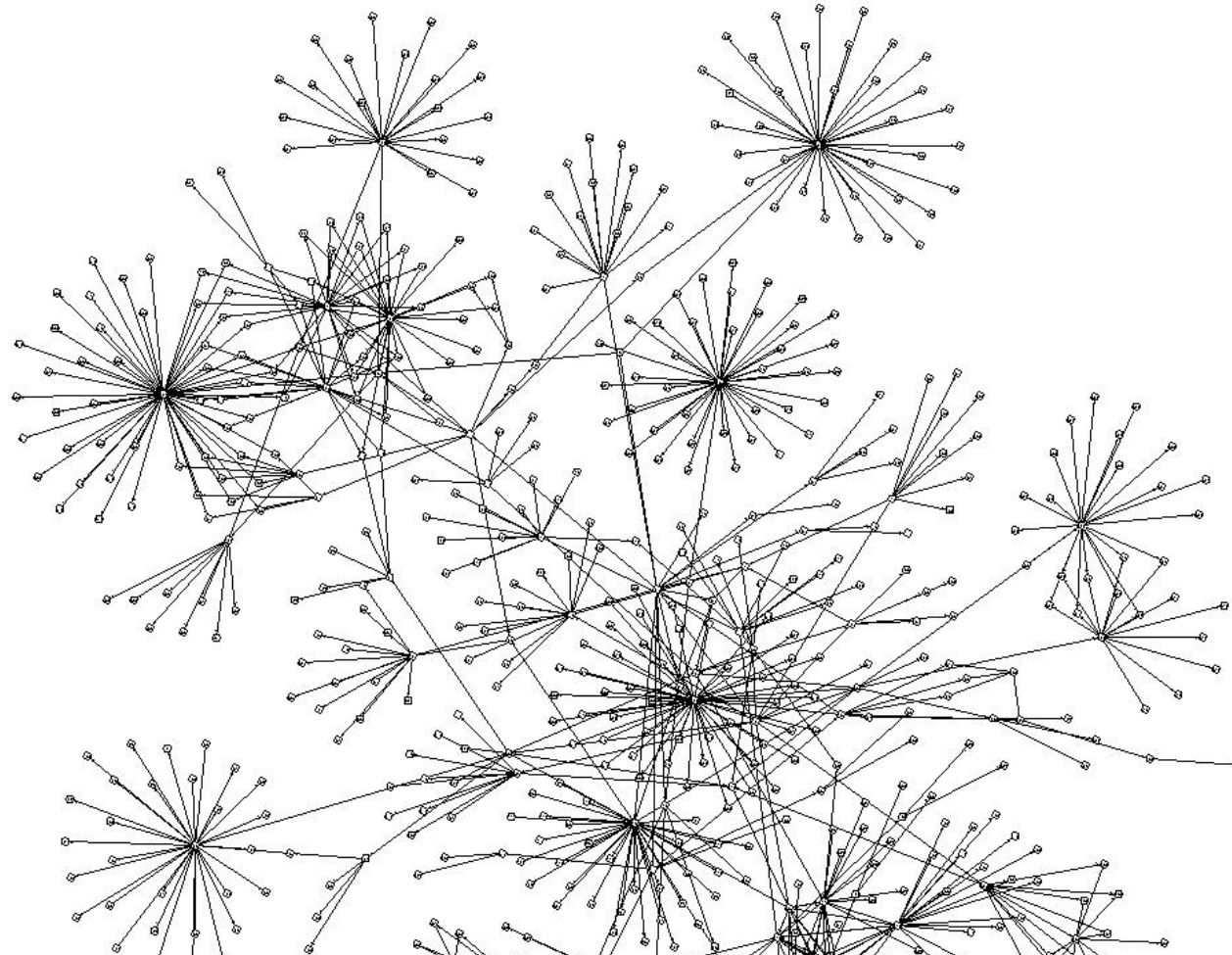
A web page is **popular** if **many popular** web pages link to it

PageRank does (almost) **not** depend on the contents of a web page



# A Different View of the Internet

---

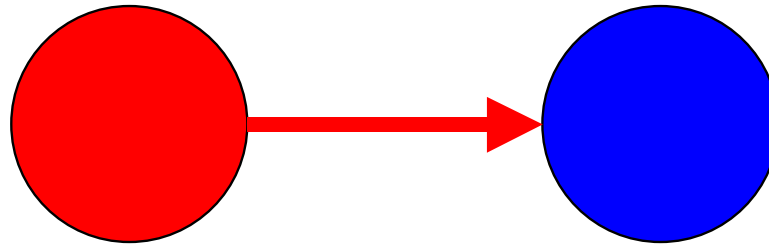


This is a graph

# The Internet as a Graph

---

Link from web page  $i$  to web page  $k$

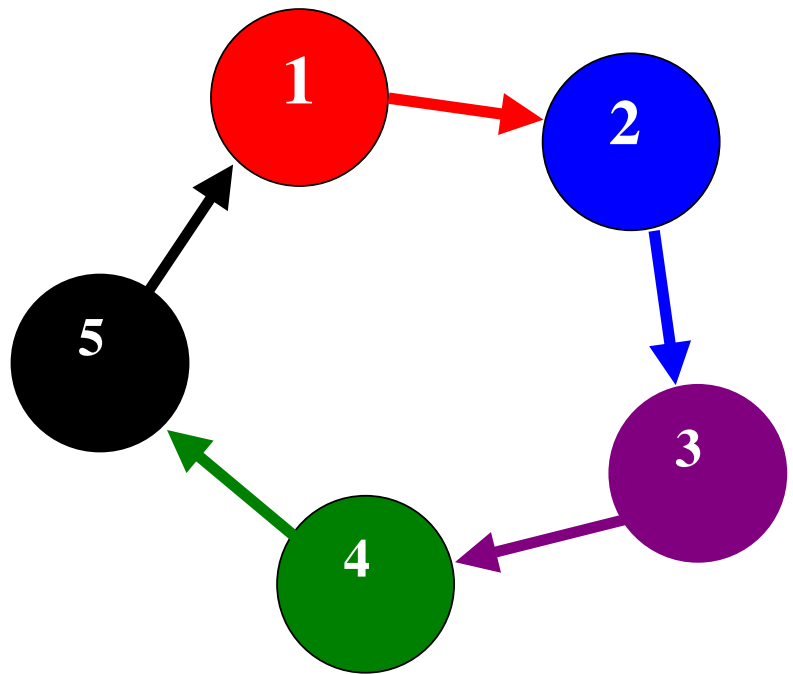


Web graph:

Web pages = nodes

Links = edges

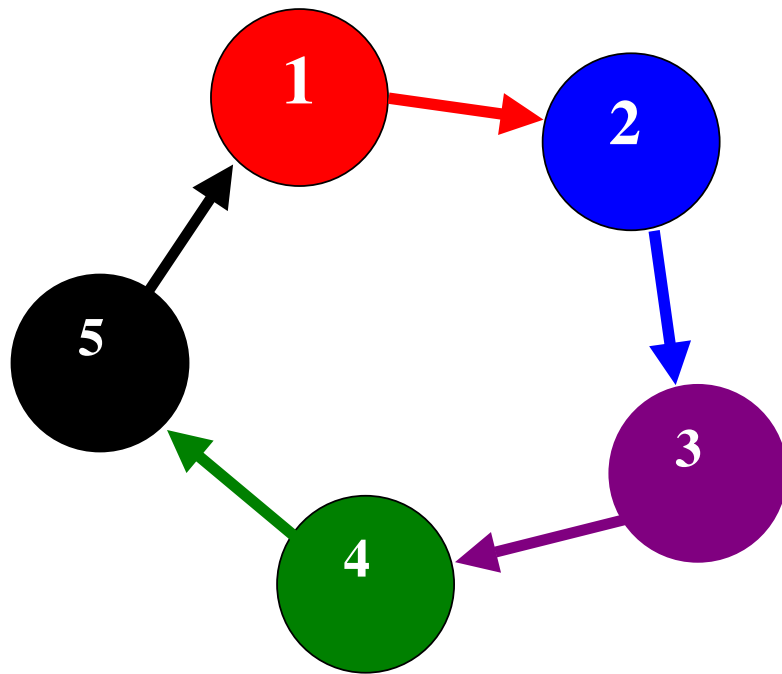
# The Web Graph as a Matrix



$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Links = nonzero elements in matrix

# PageRank is a Vector



$$p = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \end{pmatrix}$$

$p_i$  is PageRank of page  $i$

# Computing PageRank

---

Google matrix  $G$   
PageRank vector  $p$

A page has high PageRank if many pages with high PageRank link to it:

$$G p = \lambda p, \quad \lambda = 1$$

PageRank is an eigenvector



# How Google Ranks Web Pages

---

- Model:  
Internet  $\rightarrow$  web graph  $\rightarrow$  matrix  $G$
- Computation:  
PageRank  $p$  is eigenvector of  $G$   
 $p_i$  is PageRank of page  $i$
- Ranking:  
If  $p_i > p_k$  then  
page  $i$  is displayed before page  $k$

# It is Difficult to Compute PageRank

---

$$G p = p$$

- Matrix  $G$  is **large**: 11 billion static web pages
- Computing the exact  $p$  takes **too long**  
Iterative method:  $p^{(k)} = G p^{(k-1)}$   
When should we **stop**?  
How do we know that  $p^{(k)}$  is **accurate** enough?
- What happens to the PageRanks, if a link is added?
- How **often** should we re-compute  $p$ ?  
250,000 new domain names every day

# Numerical Analysis Tools

---

- Properties of  $p$  (Teresa):  
 $G$  is convex combination of two stochastic matrices  
Markov chain theory, Jordan canonical form
- Computing  $p$ :  
As eigenvector: Krylov space methods  
As linear system solution: iterative methods
- Accuracy (Rebecca):  
How to measure error: absolute/relative errors, norm/component-wise, ranking
- Adding/deleting links: Perturbation theory

# Google's

## PageRank and Beyond

The Science of

Search Engine Rankings

Amy N. Langville

Carl D. Meyer